MICHIGAN ENGLISH LANGUAGE PROFICIENCY ASSESSMENT (ELPA)

FINAL TECHNICAL REPORT 2009 ADMINISTRATION

KINDERGARTEN THROUGH GRADE 12

DRAFT

PREPARED BY
Strategic Measurement and Evaluation, Inc.

PREPARED FOR
Michigan Department of Education
Office of Educational Assessment and Accountability

TABLE OF CONTENTS

APPENDICES ..... **iv**
FIGURES AND TABLES ..... **vii**
OVERVIEW **1**
SECTION 1. INTRODUCTION **3**
1.1 Background 3
1.2 Rationale and Purpose 4
1.3 Recommended Test Use 4

SECTION 2. TEST DESIGN AND DEVELOPMENT **6**
2.1 Overview 6
2.2 General Design 6
2.3 Test Blueprints According to Michigan Learning Standards by Assessment Level and Domain 7
2.4 Item Development and Review Processes 7
2.5 Differential Item Functioning (DIF) 9
DIF for Dichotomous Items 9
DIF for Polytomous Items 11
2.6 Test Construction 14
Psychometric Guidelines for Operational Item Selection and Form Construction 15
Testing Written Language 16
Testing Oral Language 16
2.7 Test Accommodations 16

SECTION 3. SCORING OPEN-ENDED ITEMS **17**
3.1 ELPA Range-finding 17
3.2 Reader Training 18
Scoring Guides 19
Training and Qualifying Sets 20
3.3 Inter-Rater Agreement 21
Retraining Sets for Monitoring Room Drift 21
3.4 Monitoring Reports 21
3.5 Retraining 22

SECTION 4. DESCRIPTIVE STATISTICS AT ITEM AND TEST LEVELS **24**
4.1 Item-Level Descriptive Statistics 24
4.2 Higher-Level Descriptive Statistics 25

SECTION 5. RELIABILITY **29**
5.1 Classical True Score Theory 29
5.1a Internal Consistency Reliability 29
5.1b Standard Error of Measurement (SEM) within CTT Framework 30
5.2 Conditional SEM within IRT Framework 33
5.3 Inter-Rater Reliability 34
5.4 Quality (or Reliability) of Classificaiton Decision at Proficient Cut 34

SECTION 6. CALIBRATION, EQUATING, AND SCALING 37
6.1 The Unidimensional Rasch and Partial Credit Models 37
6.2 Calibration of the Spring 2009 ELPA 39
Stability of Anchor Items 40
6.3 Scale Scores for the ELPA 41
6.4 Test Characteristic Curves for the ELPA by Assessment Level 42

SECTION 7. IRT STATISTICS 44
7.1 Rasch Statistics 44
7.2 Evidence of Model Fit 45
7.3 Item Information 47

SECTION 8. VALIDITY 49
8.1 Validity Evidence of ELPA Test Content 49
Relation between ELPA and Michigan Content Standards 49
Validity of the Presentation of the Listening Stimulus 51
8.2 Validity Evidence of the Internal Structure of the ELPA 51
8.3 Validity Evidence of the External Structure of the ELPA 56
Relation between the ELPA and the MEAP Subject Tests 56
Relation between the ELPA and the MME Subject Tests 58
Distribution of Student Classification at Performence Levels 59
Summary 60

SECTION 9. SUMMARY OF OPERATIONAL TEST RESULTS 61

REFERENCES 68

APPENDICES

APPENDIX A: ELPA TEST BLUEPRINT BY ASSESSMENT LEVEL AND DOMAIN 71
A.1: Level I-Operational Items.....71
A.2: Level II-Operational Items 72
A.3: Level III-Operational Items.....73
A.4: Level IV-Operational Items 74
A.5: Level V-Operational Items.....75
A.6: Level I-Field Test Items.....76
A.7: Level II-Field Test Items.....77
A.8: Level III-Field Test Items.....78
A.9: Level IV-Field Test Items.....79
A.10: Level V-Field Test Items.....80

APPENDIX B: READER AGREEMENT AND SCORE POINT DISTRIBUTION SUMMARY 81

APPENDIX C: ELPA ITEM-LEVEL STATISTICS BY ASSESSMENT LEVEL 83
C.1: Level I-Operational Items.....83
C.2: Level II-Operational Items 85
C.3: Level III-Operational Items 87
C.4: Level IV-Operational Items.....89
C.5: Level V-Operational Items 91
C.6: Level I-Field Test Items.....93
C.7: Level II-Field Test Items.....95
C.8: Level III-Field Test Items 97
C.9: Level IV-Field Test Items.....98
C.10: Level V-Field Test Items.....99
C.11.a: DIF Analysis for Level I Ethnicity Comparison (Asians and Whites) 100
C.11.b: DIF Analysis for Level I Ethnicity Comparison (African Americans and Whites).....101
C.11.c: DIF Analysis for Level I Ethnicity Comparison (Hispanics and Whites).....102
C.11.d: DIF Analysis for Level I Gender Comparison.....103
C.12.a: DIF Analysis for Level II Ethnicity Comparison (Asians and Whites) 104
C.12.b: DIF Analysis for Level II Ethnicity Comparison (African Americans and Whites).....106
C.12.c: DIF Analysis for Level II Ethnicity Comparison (Hispanics and Whites).....108
C.12.d: DIF Analysis for Level II Gender Comparison 110
C.13.a: DIF Analysis for Level III Ethnicity Comparison (Asians and Whites) 112
C.13.b: DIF Analysis for Level III Ethnicity Comparison (African Americans and Whites) 113
C.13.c: DIF Analysis for Level III Ethnicity Comparison (Hispanics and Whites).....114
C.13.d: DIF Analysis for Level III Gender Comparison.....115
C.14.a: DIF Analysis for Level IV Ethnicity Comparison (Asians and Whites) 116
C.14.b: DIF Analysis for Level IV Ethnicity Comparison (African Americans and Whites).....117
C.14.c: DIF Analysis for Level IV Ethnicity Comparison (Hispanics and Whites).....118
C.14.d: DIF Analysis for Level IV Gender Comparison 119
C.15.a: DIF Analysis for Level V Ethnicity Comparison (Asians and Whites) 120
C.15.b: DIF Analysis for Level V Ethnicity Comparison (African Americans and Whites).....121
C.15.c: DIF Analysis for Level V Ethnicity Comparison (Hispanics and Whites).....122
C.15.d: DIF Analysis for Level V Gender Comparison.....123

APPENDIX D: RAW SCORE TO SCALE SCORE CONVERSION TABLES FOR THE ELPA 124
D.1.a: Level I –Total ELPA 124
D.1.b: Level I ELPA by Domain-Listening 126
D.1.c: Level I ELPA by Domain-Reading 127
D.1.d: Level I ELPA by Domain-Writing 128
D.1.e: Level I ELPA by Domain-Speaking 129
D.1.f: Level I ELPA by Domain-Comprehension 130
D.2.a: Level II –Total ELPA 131
D.2.b: Level II ELPA by Domain-Listening 133
D.2.c: Level II ELPA by Domain-Reading 134
D.2.d: Level II ELPA by Domain-Writing 135
D.2.e: Level II ELPA by Domain-Speaking 136
D.2.f: Level II ELPA by Domain-Comprehension 137
D.3.a: Level III –Total ELPA 138
D.3.b: Level III ELPA by Domain-Listening 140
D.3.c: Level III ELPA by Domain-Reading 141
D.3.d: Level III ELPA by Domain-Writing 142
D.3.e: Level III ELPA by Domain-Speaking 143
D.3.f: Level III ELPA by Domain-Comprehension 144
D.4.a: Level IV –Total ELPA 145
D.4.b: Level IV ELPA by Domain-Listening 147
D.4.c: Level IV ELPA by Domain-Reading 148
D.4.d: Level IV ELPA by Domain-Writing 149
D.4.e: Level IV ELPA by Domain-Speaking 150
D.4.f: Level IV ELPA by Domain-Comprehension 151
D.5.a: Level V –Total ELPA 152
D.5.b: Level V ELPA by Domain-Listening 154
D.5.c: Level V ELPA by Domain-Reading 155
D.5.d: Level V ELPA by Domain-Writing 156
D.5.e: Level V ELPA by Domain-Speaking 157
D.5.f: Level V ELPA by Domain-Comprehension 158

APPENDIX E: IRT STATISTICS 159
E.1: Level I Items 159
E.2: Level II Items 162
E.3: Level III Items 165
E.4: Level IV Items 168
E.5: Level V Items 171

APPENDIX F: ITEM INFORMATION AT DIFFERENT CUT SCORES BY GRADE 174
F.1: Grade K 174
F.2: Grade 1 176
F.3: Grade 2 178
F.4: Grade 3 180
F.5: Grade 4 182
F.6: Grade 5 184
F.7: Grade 6 186
F.8: Grade 7 188
F.9: Grade 8 190
F.10: Grade 9 192
F.11: Grade 10 194
F.12: Grade 11 196
F.13: Grade 12 198

APPENDIX G: CLASSIFICATION DISTRIBUTION ACROSS PERFORMANCE LEVELS FOR THE ELPA, MEAP, AND MME 200

G.1: Grade 3-Percentage of ELLs Scoring at Each MEAP ELA and Math Performance Level by Each ELPA Performance Level 200

G.2: Grade 4-Percentage of ELLs Scoring at Each MEAP ELA and Math Performance Level by Each ELPA Performance Level 201

G.3: Grade 5-Percentage of ELLs Scoring at Each MEAP ELA, Math, and Science Performance Level by Each ELPA Performance Level 202

G.4: Grade 6-Percentage of ELLs Scoring at Each MEAP ELA, Math, and Social Studies Performance Level by Each ELPA Performance Level 204

G.5: Grade 7-Percentage of ELLs Scoring at Each MEAP ELA and Math Performance Level by Each ELPA Performance Level 206

G.6: Grade 8-Percentage of ELLs Scoring at Each MEAP ELA, Math, and Science Performance Level by Each ELPA Performance Level 207

G.7: Grade 11-Percentage of ELLs Scoring at Each MME ELA, Math, Science, and Social Studies Performance Level by Each ELPA Performance Level 209

FIGURES AND TABLES

Table 2.1: Test Specifications by Domain and Assessment Level 7
Table 2.2: Maximum Number of Points by Domain and Assessment Level 7
Table 2.3: DIF Classification for Multiple-Choice Items 11
Table 2.4: DIF Classification for Constructed-Response Items 14
Table 4.1: Summary Statistics of the ELPA Domains by Assessment Level 26
Table 4.2: Summary Statistics of the ELPA Domains by Grade 27
Table 5.1: Raw Score Descriptive Statistics and Reliability by Grade and Domain 31
Figure 5.1: Classification Accuracy 34
Figure 5.2: Classification Consistency 35
Table 5.2: Decision Accuracy and Consistency Table by Grade 36
Figure 6.1: Category Response Curves for a Two-Step Item Using the PCM 38
Table 6.1: Number of Students in the Calibration Sample by Assessment Level and Form 40
Table 6.2: Number of Anchor Items by Domain and Assessment Level 41
Table 6.3: Scale Score Transformation Equations for Total Test and Domain 42
Figure 6.2: ELPA Test Characteristic Curves (TCCs) by Assessment Level 43
Table 7.1: Summary of Rasch Difficulty Values by Assessment Level and Domain 44
Table 7.2: General Criteria to Evaluate Mean-Square Fit Statistics 47
Table 8.1: Inter-correlations among ELPA Domains and Total Test by Grade 54
Table 8.2: Inter-correlations among ELPA Total Scale Score and MEAP Subject Test Scale Scores by Grade 58
Table 8.3: Inter-correlations among ELPA Total Scale Score and MME Subject Test Scale Scores for Grade 11 59
Table 8.4: Percentage of Grade 3 ELLs Scoring at Each MEAP ELA Performance Level by Each ELPA Performance Level 60
Table 9.1: Raw Score Summary by Grade, Domain, and Total Test 61
Table 9.2a: Scale Score Summary by Grade and Domain 63
Table 9.2b: Scale Score Summary of the Total ELPA Test by Grade 65
Table 9.3: Percent of Students in Each Performance Level by Grade 65
Table 9.4: Percent of Students Meeting Achievement Target by Grade and Domain 66

OVERVIEW

This Michigan English Language Proficiency Assessment (ELPA) Technical Report for the 2009 administration is organized around nine major sections—Introduction; Test Design and Development; Scoring Open-Ended Items; Descriptive Statistics at Item and Test Levels; Reliability; Calibration, Equating, and Scaling; Item Response Theory (IRT) Statistics; Validity; and Summary of Operational Test Results. A brief overview of this report is provided below.

Section 1: Introduction

This section presents the background, rationale, purpose, and recommended test use.

Section 2: Test Design and Development

This section describes the test development process of the ELPA. It includes the test design and specifications, the item development and review processes, the test construction, and the test accommodations of the ELPA.

Section 3: Scoring Open-Ended Items

This section provides a description of the scoring process for open-ended items. It includes a description of the range-finding meetings that were held in Lansing, Michigan, and information about reader training and monitoring.

Section 4: Descriptive Statistics at Item and Test Levels

This section provides item-level and test-level descriptive statistics based on Classical Test Theory (CTT).

Section 5: Reliability

This section explains internal consistency reliability, classical Standard Error of Measurement (SEM), and conditional SEM based on IRT. It also provides information on inter-rater reliability (or agreement), and the reliability of classification decisions at the proficient cut.

Section 6: Calibration, Equating, and Scaling

This section explains the unidimensional Rasch and Partial Credit models. It also includes a description of the calibration process and reports the corresponding results for the spring 2009 ELPA test items. The test characteristic curves (TCCs) are presented for each assessment level.

DRAFT

Section 7: Item Response Theory (IRT) Statistics

This section presents the IRT statistics, including a summary of the minimum, maximum, and average Rasch difficulties by assessment level and domain. This section also includes a discussion of the IRT model fit statistics, and the item information at each cut score.

Section 8: Validity

This section presents the evidence for the validity of the ELPA. It includes evidence of validity based on test content, internal structure, and external structure.

Section 9: Summary of Operational Test Results

This section presents the raw score summary, scale score summary, and percentage of students in each performance category by grade for the 2009 administration of the ELPA.

DRAFT

SECTION 1. INTRODUCTION

1.1 Background

Title III of the Federal *No Child Left Behind* (*NCLB*) *Act* of 2001 requires annual assessment of the English language skills of limited English proficient (LEP) students. Section 3121(d)(1) explains that each state must use evaluation measures designed to assess "the progress of children in attaining English proficiency, including a child's level of comprehension, speaking, listening, reading, and writing skills in English."

Moreover, the *NCLB* requires that the annual assessment of LEP students be based on specific student achievement objectives. Section 3122(a)(1) states that "Each State educational agency or specially qualified agency receiving a grant under subpart 1 shall develop annual measurable achievement objectives for limited English proficient children served under this part that relate to such children's development and attainment of English proficiency while meeting challenging State academic content and student academic achievement standards."

The Annual Measurable Achievement Objectives (AMAOs) are described in the *NCLB* law to provide targets for performance and annual progress that these students must make in their acquisition of the English language. These AMAOs also describe targets for performance and annual progress on state content and achievement assessments that must be met by the English language learner (ELL) subgroup.

Federal guidelines require that states develop assessments of English language proficiency (ELP) that are aligned to state-approved ELP curricular and instructional standards. In addition, these assessments must measure and report scores for five areas of language acquisition. These areas, also called domains, are as follows: Listening, Reading, Writing, Speaking, and Comprehension.

With these requirements guiding its development, the English Language Proficiency Assessment, or ELPA, was implemented beginning in the spring of 2006 by the Office of Educational Assessment and Accountability (OEAA) of Michigan's Department of Education (MDE). Beginning in the spring of 2007, the ELPA was designed for five assessment levels designated by the Roman numerals I through V and corresponding to the kindergarten (K), 1–2, 3–5, 6–8, and 9–12 grade levels. Each of these assessment levels feature domain subtests for Listening, Reading, Writing, and Speaking. The fifth domain, Comprehension, is assessed through a composite of test items selected from the Listening and Reading domains that contribute to comprehension. Each cycle of the ELPA is designed to be accessible for students with disabilities (SWD) at each assessment level along with federal testing requirements described in the *Individuals with Disabilities Act* (*IDEA*) of 2004 (see Section 2.7 for more information).

To design the spring 2009 ELPA, OEAA continued its supervision of the creation of test forms that featured embedded field test (EFT) items. A minimum of 25 percent of the items at each assessment level were items that also appeared in the spring 2008 forms.

These items are called "linking" or "anchor" items. The EFT items that appeared in the spring 2009 forms may be added into a pool from which operational items for future administrations of the ELPA can be drawn. This procedure ensures that anchor items can be cyclically retired. Only those items that pass the data review process, however, can be included in the operational item pool.

1.2 Rationale and Purpose

Besides meeting the measurement and reporting requirements of English acquisition and progress described in the *NCLB* law, the ELPA also provides reports to Michigan's schools, districts, and intermediate school districts (ISDs) that help those agencies determine academic placement policies and guide instructional interventions for Michigan's English language learner (ELL) subgroup. While creation of policy remains within the purview of individual Michigan districts and decisions governing classroom instruction rest with educators, ELPA raw and scale scores for each of its language domains as well as overall student performance are used as one set among several criteria that provide such direction at the district and school levels.

Although it is not detailed within this report, OEAA also developed the ELPA Initial Screening for use as a diagnostic tool of English acquisition at the time of enrollment for all K–12 students who are potential ELLs. Since its inception in the fall of 2006, the items selected for the ELPA Initial Screening have come from a pool of items that appeared as operational items on the previous spring ELPA. Since the scale score tables and the performance level cut scores are the same between the ELPA Initial Screening and the spring ELPA, the ELPA Initial Screening can be used by Michigan's educators as one factor to determine student placement into ELL services. The ELPA Initial Screening also aids the State in gathering a count of students who are eligible to receive ELL services for Title III funding purposes.

In 2007, a new performance level, Advanced Proficient, was added. Cut scores for Advanced Proficient were determined from a logistic regression model that was based on an equal-access comparison between students who were assessed on the Michigan Educational Assessment Program (MEAP) English Language Arts (ELA) test and the ELPA. This performance level provides districts with an additional tool for making decisions for student placement in or exit from ELL services.

1.3 Recommended Test Use

The Michigan ELPA is designed to assess students at all performance levels within each assessment level. This vertical development of the language tested allows the test to differentiate more finely among students at different stages of language acquisition. Because test results provide students, teachers, and parents with an objective report of each student's strengths and weaknesses in the English language skills of speaking, listening, reading, writing, and comprehension, the ELPA helps determine whether these students are making adequate progress towards English language proficiency.

DRAFT

The test results can also help schools focus on ways to improve instruction so that English language learners become proficient in English, thereby allowing more time for content-based materials, such as mathematics and science.

DRAFT

SECTION 2. TEST DESIGN AND DEVELOPMENT

2.1 Overview

The spring 2009 ELPA forms were developed to be administered at one of five assessment levels. Level I forms were designed to be administered to kindergarten students; Level II forms were designed for students in grades 1 and 2; Level III forms were designed for students in grades 3 to 5; Level IV forms were designed for students in grades 6 to 8; and Level V forms were designed for students in grades 9 to 12. Each assessment level featured domain subtests for Listening, Reading, Writing, and Speaking. The fifth domain, Comprehension, was assessed through a composite of test items from selected Listening and Reading standards that contribute to comprehension.

2.2 General Design

As previously noted, the Michigan ELPA was designed to assess four domains (Listening, Reading, Writing, and Speaking) and one composite domain (Comprehension) within each assessment level. The test forms included multiple-choice (MC) and constructed-response (CR) items that required either a short-response or an extended-response.

The total number of items per assessment level varied. Level I had 52 items; Level II had 62 items; Level III had 62 items; Level IV had 66 items; and Level V had 65 items.

The Listening and Reading domains were assessed with MC items only. The number of items for the Listening and Reading domains varied from 16–22 across the different assessment levels. The Speaking domain was assessed with constructed-response (CR) items only. The number of items for the Speaking domain varied from 8–10 items across the different assessment levels. The number of items for the Writing domain ranged from 12–14 across the various assessment levels. The Writing domain was assessed with both MC and CR items and included the following three skill sets:

- MC items that assess ELLs' understanding of the principles of written English at the phoneme, word, and sentence levels
- Developmental writing items (Levels I and II only)
- Sentence-writing items, paragraph-writing items, and items requiring an extended response to a graphics-based prompt (numbers and types of items vary by assessment level)

The test design for the 2009 administration of the ELPA is shown in Table 2.1, while Table 2.2 provides the maximum number of points by domain and assessment level.

DRAFT

Table 2.1: Test Specifications by Domain and Assessment Level

	Listening	Speaking	Reading	Writing		Comprehension	Total Number of Core Items per Level
Level	MC	CR	MC	MC	CR	MC	(MC + CR)
I	16	8	16	6	6	21	52
II	20	9	20	8	5	33	62
III	19	9	22	8	4	29	62
IV	21	10	21	10	4	28	66
V	20	10	21	10	4	28	65

Note. Comprehension consists of items selected from the Listening and Reading domains and is not included in the column titled "Total Number of Core Items per Level."

Table 2.2: Maximum Number of Points by Domain and Assessment Level

	Listening	Speaking	Reading	Writing		Comprehension	Total Number of Points for Core Items per Level
Level	MC	CR	MC	MC	CR	MC	(MC + CR)
I	16	18	16	6	12	21	68
II	20	20	20	8	12	33	80
III	19	20	22	8	12	29	81
IV	21	21	21	10	12	28	85
V	20	21	21	10	12	28	84

Note. Comprehension consists of items selected from the Listening and Reading domains and is not included in the column titled "Total Number of Points for Core Items per Level."

2.3 Test Blueprints According to Michigan Learning Standards by Assessment Level and Domain

Appendices A.1 to A.5 provide detailed test blueprints for the operational items (which appeared on all forms) administered at each assessment level. Appendices A.6 to A.10 provide a summary of the field test items administered at each assessment level and across all forms. Within each appendix, the table is sorted by domain and Michigan Learning Standard. The number of items and the points associated with each learning standard and test form (for field test items) are also provided.

2.4 Item Development and Review Processes

The operational items on the 2009 ELPA forms were drawn from 2008 operational items and items that were field tested as part of the ELPA program in 2008 or earlier. The item development process for these items was managed by OEAA and the previous ELPA contractor.

In accordance with the item specifications, the educators and assessment specialists emphasize the following criteria when creating and reviewing passages and items:

- Absence of bias and sensitive topics in passages
- Item soundness
- Absence of bias in items

DRAFT

- Appropriateness of topic, vocabulary, and language structure for each assessment level
- Match to the intended Michigan State English as a second language (ESL) standard

The passage and item development process for the ELPA program includes several steps. The first step is the development of appropriate passages to serve as stimulus material for the domain subtests. Those passages meeting OEAA approval are presented to an item writing committee. The item writers for the ELPA program are Michigan educators with background and experience with ELL populations. These educators are also familiar with the Michigan ELPA content standards. The item writing committee meets twice. At the first meeting, the focus is on training the educators to be item writers and providing initial feedback regarding the quality of their items.

Following the initial training and feedback, the educators continue to write items and then bring their work to a second item review meeting. At this second meeting, the teachers and OEAA staff review and rate each item. Again, participants are provided with specific checklists related to item construction and test content to guide their ratings. During this meeting, all new items also undergo a separate bias and sensitivity review before being eligible for inclusion in the item bank or on a field test form.

The next step in the development process is to select new field test items to be embedded into an operational test form. Following the administration of these items, responses to the new CR items are reviewed and sample student responses are selected to serve as training papers. Training papers are used to illustrate the range of student responses associated with each point on the corresponding scoring rubric for the item.

Following the scoring and analysis of the field test items, the items are submitted to a final data review. This review is done with panels of Michigan educators serving as reviewers. Data review committee members are presented with item statistics[1] and differential item functioning (DIF)[2] analysis results for each item and are asked to review the alignment to ELP standards and content, and absence of bias for each item. Only those field test items that are judged to be of acceptable quality and to be fair to students coming from all over the world are approved for inclusion in the item bank as possible operational items.

Although the final tests were challenging for students, the questions, graphics, and stories engaged students and reflected the kinds of activities in which they were involved on a daily basis. This helped to ensure that the tests would measure the learning of each individual student and provide meaningful information about his or her English language proficiency.

---

[1] These item statistics are computed within the framework of Classical Test Theory (CTT), which is described in Section 4.

[2] DIF is described in detail below in Section 2.5.

2.5 Differential Item Functioning (DIF)

The 2009 ELPA had newly constructed field test items embedded within the operational forms. In addition to the bias-related judgments provided by Michigan educators as part of the item development review process (see previous discussion in Section 2.4), the field test items were also checked for statistical indicators of bias. Differential Item Functioning (DIF) refers to "the differential impact of an item on the performance of one subgroup when compared to that of another subgroup" (Welch, 2006, p. 320).

All of the field test items were eligible for DIF testing. However, because of small N-counts, the DIF procedures only compared students for cases in which enough student responses (N > 200 per subgroup) were available. For the 2009 ELPA, the DIF procedures were conducted between male and female students, White and Hispanic students, White and Asian students, and White and African American students. In these comparisons, White and male students were considered reference groups for the comparisons of ethnicity and gender, respectively. Penfield's (2005) DIFAS software was used to complete the DIF analysis for both the dichotomous and polytomous items.

DIF for Dichotomous Items

For dichotomous items, the Mantel-Haenszel (MH) procedure was used (Mantel & Haenszel, 1959). Specifically, the MH chi-square (χ^2) statistic was used in conjunction with the MH log-odds ratio. These two statistics are discussed in more detail below.

The MH χ^2 indicates "whether there is a relationship between performance on an item and group membership, after taking into account performance on the instrument" (de Ayala, 2009, p. 329). This statistic is calculated as follows (Camilli & Shepard, 1994, p. 120):

$$MH\chi^2 = \frac{\left\{\left|\sum_{j=1}^{S}[A_j - E(A_j)]\right| - 1/2\right\}^2}{\sum_{j=1}^{S} VAR(A_j)} \qquad (2.1)$$

where $VAR(A_j) = \frac{n_{Rj}\, n_{Fj}\, m_{1j}\, m_{0j}}{T_j^2 (T_j - 1)}$

and $E(A_j) = \frac{n_{Rj}\, m_{1j}}{T_j}$.

In Equation (2.1), $A_j - E(A_j)$ represents the discrepancy between the observed number and the expected number of correct responses on the item by the reference group members who have *j*th score on the matching variable[3]; n_{Rj} and n_{Fj} represent the number of examinees in the reference and focal groups, respectively, for the *j*th score on the matching variable; m_{1j} represents the total number of examinees (both reference and focal) with *j*th score on the matching variable and with a correct response on the current

[3] Total observed score is used as the matching variable for DIF analysis here.

item; m_{0j} represents the total number of examinees with the *j*th score on the matching variable and with an incorrect response on the current item. The MH χ^2 is evaluated against the standard χ^2 critical values using one degree of freedom.

The MH χ^2 does not indicate the strength of association of the relationship between item performance and group membership. The MH odds ratio can be computed to estimate the strength of this association. The resulting estimate represents the relative likelihood of success on a particular item for members of two different groups of examinees (Camilli, 2006). This odds ratio thus provides an estimate of effect size with a value of 1.0 indicating no DIF. A value greater than 1.0 indicates that, on average, the reference group members performed better than comparable focal group members and a value less than 1.0 indicates that, on average, the reference group members performed worse than comparable focal group members.

The odds of a correct response (proportion passing divided by proportion failing) is *P/Q* (i.e., *P/[1-P]*). The MH odds ratio is simply the odds of a correct response of the reference group divided by the odds of a correct response of the focal group. The formula for its estimation is as follows (Camilli & Shepard, 1994, p. 116):

$$\hat{\alpha}_{MH} = \frac{\sum_{j=1}^{S} A_j D_j / T_j}{\sum_{j=1}^{S} B_j C_j / T_j} \qquad (2.2)$$

where $S = K-1$ and represents the actual number of 2 x 2 contingency tables (assuming the tables have at least 1 person in each cell); *K* represents the number of items on the test; *j* signifies the *j*th score on the matching variable and runs from 0 to *K*[4]. For *j*th score category, A_j represents the number of reference group members with a correct response, B_j represents the number of reference group members with an incorrect response, C_j represents the number of focal group members with a correct response, and D_j represents the number of focal group members with an incorrect response. T_j represents the total number of examinees who have the *j*th score on the matching variable.

The corresponding null hypothesis is that the odds of getting the item correct are equal for the two groups (i.e., the odds ratio is equal to 1):

$$H_0: \alpha_{MH} = 1 \qquad (2.3)$$

In order to make the odds ratio symmetrical around zero with its range located in the interval $-\infty$ to $+\infty$, the odds ratio is transformed into a log-odds ratio as follows (Camilli & Shepard, 1994, p.116):

$$\beta_{MH} = \log(\alpha_{MH}) \qquad (2.4)$$

The natural logarithm transformation of this odds ratio is symmetrical around zero (where 0 indicates no DIF). This DIF measure is a signed index where a positive value

[4] Although the value of the matching variable runs from 0 to *K*, the all correct (*K*) and all incorrect (0) score categories are not included in the DIF analysis in order to avoid a denominator equal to 0.

represents DIF in favor of the reference group and a negative value indicates DIF in favor of the focal group.

The variance of the log-odds ratio estimate (V_β) is computed as follows (Camilli & Shepard, 1994, p. 121):

$$V_\beta = \frac{\sum_{j=1}^{S} T_j^{-2}\left(A_j D_j + \alpha_{MH} B_j C_j\right)\left[A_j + D_j + \alpha_{MH}\left(B_j + C_j\right)\right]}{2\left(\sum_{j=1}^{S} A_j D_j / T_j\right)^2}. \quad (2.5)$$

The terms included in Equation (2.5) correspond to those presented for Equation (2.2). In practice, a standardized MH log-odds ratio is computed by dividing the estimate $\hat{\beta}_{MH}$ by the estimated standard error. A value greater than 2.0 or less than -2.0 may be considered evidence of the presence of DIF (Penfield, 2005).

Educational Testing Service (ETS) utilizes the fact that β_{MH} is amenable to linear transformations to other interval scale metrics and transforms β_{MH} to their delta scale metric (D) via (Camilli & Shepard, 1994, p. 121):

$$D = -2.35(\beta_{MH}) \quad (2.6)$$

The quantity D represents the difference in the difficulty of the items in the delta metric for the two groups whose performance is being compared. Table 2.3 depicts DIF classifications for multiple-choice items based on the MH χ^2 and the item delta scale difference value (D).

Table 2.3 DIF Classification for Multiple-Choice Items

Category	Description	Criterion
A	Negligible DIF	Non-significant MH χ^2 or |D| < 1.0
B	Moderate DIF	Neither A nor C
C	Large DIF	Significant MH χ^2 and |D| ≥ 1.5

Note. Table adapted from Zieky (1993).

DIF for Polytomous Items

Since the ELPA includes CR items that are polytomously scored, the MH procedure cannot be used as a DIF index for all the items on a form. Instead, an extension of the MH χ^2 procedure was computed (see Mantel, 1963) and a generalization of the MH odds ratio for ordered categories, the Liu-Agresti Cumulative Common Odds Ratio (Liu & Agresti, 1996), was used to assess DIF for the polytomous items.

Mantel (1963) proposed an extension of the MH procedure for ordered response categories that involves comparing the means for the reference and focal groups,

conditional on a matching variable. The statistic is computed as follows (Zwick, Donaghue, & Grima, 1993, p. 239):

$$\text{Mantel } \chi^2 = \frac{(\sum_k F_k - \sum_k E(F_k))^2}{\sum_k VAR(F_k)} \qquad (2.7)$$

where F_k is the sum of scores for the focal group at the *k*th level of the matching variable and is defined as

$$F_k = \sum_t y_t n_{Ftk},$$

and the expectation of F_k under the hypothesis of no association is

$$E(F_k) = \frac{n_{F+k}}{n_{++k}} \sum_t y_t \, n_{+tk} \, ,$$

and the variance of F_k under the assumption of no association is

$$\text{Var}(F_k) = \frac{n_{R+k} n_{F+k}}{n_{++k}^2 (n_{++k} - 1)} \left\{ \left(n_{++k} \sum_t y_t^2 n_{+tk} \right) - \left(\sum_t y_t \, n_{+tk} \right)^2 \right\}. \qquad (2.8)$$

Using the Mantel approach for ordered categories, the data are organized into a 2 x *T* x *K* contingency table, where *T* is the number of response categories and *K* is the number of levels of the matching variable. $y_1, y_2, \ldots, y_T$ represent the *T* scores that can be obtained on the item; n_{Rtk} and n_{Ftk} represent the number of examinees in the reference and focal groups, respectively, who are at the *k*th level of the matching variable and received an item score of y_T on the item. The "+" denotes summation over a particular index (e.g., n_{R+k} denotes the total number of reference group members at the *k*th level of the matching variable). Under the null hypothesis of no association, the Mantel statistic has a chi-square distribution with one degree of freedom. For dichotomous items, the Mantel statistic reduces to the MH statistic (without the continuity correction).

The Liu-Agresti estimator of the common odds ratio has the following mathematical formulation (Liu & Agresti, 1996; Penfield & Algina, 2003, p. 357):

$$\hat{\psi}_{LA} = \frac{\sum_{k=1}^{K} \sum_{j=1}^{J-1} A_{jk} \, D_{jk} \, / N_k}{\sum_{k=1}^{K} \sum_{j=1}^{J-1} B_{jk} \, C_{jk} \, / N_k}, \qquad (2.9)$$

where each item has *J* levels of response scores (*j* = 1, …, *J*, and thus $(J-1)$ dichotomizations), the matching variable has *K* levels (*k* = 1, …, *K*), and N_k is the total number of individuals at level *k*. The cumulative frequency for response level *j* at stratum *k* for the reference group is denoted as $N^*_{Rjk} = N_{R1k} + \cdots + N_{Rjk}$ and that of the

focal group is denoted as $N^*_{Fjk} = N_{F1k} + \ldots + N_{Fjk}$. Moreover, $A_{jk} = N^*_{Rjk}$, $B_{jk} = N_{Rjk} - N^*_{Rjk}$, $C_{jk} = N^*_{Fjk}$, and $D_{jk} = N_{Fjk} - N^*_{Fjk}$. These values are analogous to A_j, B_j, C_j and D_j in Equation (2.2), but it is important to note that in the dichotomous case A and C are associated with the correct responses but for the cumulative polytomous case A and C are associated with the lower portion of the response variable dichotomization. Thus, for dichotomous items, the Liu-Agresti statistic reduces to the inverse of $\hat{\alpha}_{MH}$. To match the scale of the Liu-Agresti estimator to the scale for dichotomous items, the following computation is necessary (Penfield & Algina, 2003, p. 357):

$$\hat{\alpha}_{LA} = \frac{1}{\hat{\psi}_{LA}} \qquad (2.10)$$

Values equal to 1.0 indicate no DIF while values less than 1.0 indicate DIF in favor of the focal group. Similar to the MH procedure, the natural log of $\hat{\alpha}_{LA}$ can be computed, where $\log(\hat{\alpha}_{LA})$ of 0 indicates no DIF, $\log(\hat{\alpha}_{LA})$ greater than 0 indicates DIF against the focal group, and $\log(\hat{\alpha}_{LA})$ less than 0 indicates DIF against the reference group. In practice, a standardized Liu-Agresti cumulative common log-odds ratio is computed by dividing $\log(\hat{\alpha}_{LA})$ by the estimated standard error. A value greater than 2.0 or less than -2.0 may be considered evidence of the presence of DIF (Penfield, 2005).

The Liu-Agresti estimate of the variance of $\log(\hat{\psi}_{LA})$ is expressed as follows (Penfield & Algina, 2003, p. 357):

$$\mathrm{Var}\left[\log\left(\hat{\psi}_{LA}\right)\right] = \frac{\sum_{k=1}^{K} \hat{\xi}_k(\hat{\psi})}{\hat{\psi}_{LA}^2 \left[\sum_{k=1}^{K} \sum_{j=1}^{J-1} B_{jk}\, C_{jk} / N_k\right]^2}, \qquad (2.11)$$

where $\hat{\xi}_k(\hat{\psi}) = \sum_{j=j'=1}^{J-1} \hat{\phi}_{jj'k}(\hat{\psi}) + 2\sum_{j<j'}^{J-1} \hat{\phi}_{jj'k}(\hat{\psi})$,

and $\hat{\phi}_{jj'k}(\hat{\psi}) = \frac{n_{Rk}\, n_{Fk}}{N_k^2} \left\{ \frac{\hat{\psi} B_{jk}\, C_{j'k}}{n_{Rk}} \left[1 + \left(\hat{\psi} - 1\right) \frac{C_{j'k}}{n_{Rk}}\right] + \frac{A_{jk}\, D_{j'k}}{n_{Fk}} \left[\hat{\psi} - \left(\hat{\psi} - 1\right) \frac{A_{jk}}{n_{Rk}}\right] \right\}$.

$$j \le j' = 1, \ldots, J-1$$

This variance is a generalization of the variance estimator for the MH estimate of the common odds ratio for dichotomous items. Because $\log(\hat{\alpha}_{MH}) = -\log\left(\hat{\psi}_{LA}\right)$, the statistics have the same asymptotic variance (Penfield & Algina, 2003).

Because the Liu-Agresti estimator is theoretically equivalent to the MH odds ratio, a classification scheme can be constructed that parallels the ETS scheme for dichotomous items (Penfield, 2007). Table 2.4 presents the DIF classification for CR items.

Table 2.4 DIF Classification for Constructed-Response Items

Category	Description	Criterion
A	Negligible DIF	Non-significant $log(\hat{\psi}_{LA})$ or $\lvert log(\hat{\psi}_{LA})\rvert < 0.43$
B	Moderate DIF	Significant $log(\hat{\psi}_{LA})$ and $\lvert log(\hat{\psi}_{LA})\rvert \geq 0.43$, and either: (a) $\lvert log(\hat{\psi}_{LA})\rvert < 0.64$, or (b) $\lvert log(\hat{\psi}_{LA})\rvert$ not significantly greater than 0.43
C	Large DIF	$\lvert log(\hat{\psi}_{LA})\rvert$ significantly greater than 0.43 and $\lvert log(\hat{\psi}_{LA})\rvert \geq 0.64$

Note.Table adapted from Penfield (2007).

A list of all the embedded field test items with DIF classification based on gender and ethnicity is shown in Appendices C.11.a through C.15.d. These appendices include the following information:

- Mantel Chi-Square (Note that this is the Mantel-Haenszel Chi-Square for MC items and the Mantel Chi-Square for CR items)
- Mantel-Haenszel (for MC items) or Liu-Agresti (for CR items) Log-Odds Ratio
- Log-Odds Ratio Standard Error (SE)
- Standardized Log-Odds Ratio (Note that this is the standardized Mantel-Haenszel common log-odds ratio for MC items or the standardized Liu-Agresti cumulative common log-odds ratio for CR items)
- ETS Criteria for Classification of DIF

All field test items flagged for DIF will be carefully reviewed and only those items that pass the reviews will be included in future operational tests.

2.6 Test Construction

A fundamental consideration in constructing the ELPA is the language that is being tested. Rigorous methods in language choice must be applied to provide consistency across the forms of the five assessment levels and to create an integrated structure within each form. The use of an integrated structure means that language ranges from the most simple, which is first acquired by non-native speakers, to advanced language that would indicate a level of English proficiency sufficient for participation in regular academic classes. Being able to accurately identify students at different levels of language development provides better information to classroom teachers who must find the most effective way to help their students reach English proficiency. It also provides very important evidence of students' progress toward English proficiency that is required by the *NCLB* legislation.

Items selected for the 2009 ELPA represent a complete range of difficulty at all grade levels from K–12. Items range from simple ones (primarily aimed at students with limited ability in English) to difficult ones (aimed at students with advanced ability in English).

Psychometric Guidelines for Operational Item Selection and Form Construction

Several criteria guided the construction of the 2009 ELPA. An important statistical consideration was the item difficulty, or *p*-value[5], which is a proportion of examinees who correctly answered an item.

Another statistical consideration for the construction of the 2009 ELPA was the adjusted point-biserial correlation coefficient (for MC items) and the adjusted item-total correlation coefficient (for CR items). These correlation coefficients indicate whether people who did well on a particular item also did well on the overall test.[6]

In addition to selecting items with appropriate *p*-values and correlations, staff members were given the following instructions during the item selection and form construction process:

A. Item Statistics
   1. Check range of item difficulties (i.e., *p*-values): flag items if *p*-values are < 0.30 or > 0.95.
   2. Check adjusted point-biserial correlation ranges (for MC items) and adjusted item-total correlation ranges (for CR items): avoid items with values < 0.30.
   3. Check invalid response rate[7]: watch items with an invalid rate > 5%.
   4. Avoid items with DIF bias flag (see discussion in Section 2.5). If it is necessary to select an item with a flag, it needs to be carefully reviewed during the operational test construction.

B. Typically, no changes are made to items once an item is field tested and passes the data review. If changes are needed for an item, that particular item (after revision) should then be field tested and subjected to the data review process again.

C. Operational items are selected according to several statistical targets including, but not limited to:

   1. Similar average *p*-value comparisons between spring 2009 forms and spring 2008 forms.

[5] See Section 4.1 for a detailed discussion of the *p*-value.
[6] See Section 4.1 for a detailed discussion of the correlation coefficients.
[7] Invalid responses include blanks, illegible responses, foreign language, off topic or off task, copy of prompt, incomprehensible responses, and refusal to respond.

2. Similar average point-biserial and item-total correlation comparisons between spring 2009 and spring 2008 forms.
3. Similar average Rasch item difficulty values by domain and for the total test across the 2008 and 2009 assessments.

D. The total number of items at each assessment level and the number of items for each standard must follow the test blueprint.

Testing Written Language

Readability measures are primarily based on factors such as the number of words per sentence and the number of letters or syllables per word. Additionally, ESL assessment specialists evaluate the coherence of a passage, vocabulary difficulty, sentence and text structure, and concreteness and abstractness. It is the combination of these factors that determines the appropriateness of the language of a passage.

Testing Oral Language

Recognizing that oral language structure and vocabulary of English differ vastly from those of written language, issues of oral language assessment among ELL students in kindergarten through grade 12 have received special attention. Additionally, for the oral language components of the ELPA to be relevant, the Listening and Speaking subtests must have predictive validity for academic achievement; therefore, academic language and social language are both included in the Listening and Speaking subtests of the ELPA.

2.7 Test Accommodations

In accordance with the guidelines mandated by the *IDEA* of 2004, test items were developed for the spring 2009 administration of the ELPA which adhered to standards of accessibility, reliability, and validity for the widest range of students possible. This range included SWDs with visual and auditory impairments. Specific to these populations, enlarged-print versions of the ELPA were available at all assessment levels, and Braille versions were available to students at Levels III through V. For students with auditory impairments, an accommodated CD was available; schools were also able to use amplification devices for administration of the ELPA. In addition to these design features, OEAA also provides routine and consistent guidance to testing coordinators for the purpose of proper administration of the ELPA to SWDs with other kinds of impairments.

During form construction, Questar Assessment, Inc. (Questar), the primary assessment development contractor, utilized in-house content and fairness experts to ensure that general universal design concepts were preserved. These experts also reviewed the forms for fairness, reliability, and accessibility to this special population of students.

DRAFT

SECTION 3. SCORING OPEN-ENDED ITEMS

This section describes the scoring process for both operational and field test open-ended items for the spring 2009 ELPA. The scoring team for the operational items included a scoring director, six team leaders, and fifty readers. The scoring team for the field test items included a scoring director, two team leaders, and twenty-five readers.

To prepare for the scoring training, Questar and OEAA staff reviewed training materials used in previous administration cycles, and reviewed and scored samples of existing student responses to the operational items. Questar then developed annotated anchor papers (responses that demonstrate the features of each score point), training sets, and qualifying sets. These materials were combined to form a training manual for each item.

The details of the scoring process for the operational and field test items are presented below.

3.1 ELPA Range-finding

Questar facilitated two range-finding sessions. The purpose of the first session, held on January 12-13, 2009, was to score responses to operational prompts in order to create training materials for the spring 2009 operational scoring. The participants included three Questar staff members, three OEAA staff members, and fourteen Michigan teachers. The participants were divided into two groups, based on background and teaching experience. One group was assigned to work with student responses to open-ended items appearing on the Level I, II, and III forms (grades K-5), and the other group was assigned to work with responses to the open-ended items appearing on the Level IV and V forms (grades 6-12).

Questar staff selected approximately 50 responses per item to present at the range-finding meeting. The responses for each item were selected to capture the range of student answers to each item. To ensure continuity with scoring interpretations developed during previous administrations of the ELPA tests, Questar also included scored responses from the 2008 range-finding session for each item (if available) to use as anchor papers.

The range-finding process for an individual item began with a review of the rubric for that item. Following the presentation of the rubric, the Questar meeting facilitators reviewed the existing anchor sets for the item. Next, participants were given packets of sample responses for each item. The sample responses for each item were ordered randomly within the set.

Each sample response was read aloud. After reading the response, participants were asked to assign a score using the rubric and the anchor papers as reference points. Scores assigned by the participants were reviewed and discussed by the group. The facilitators worked to identify a consensus score for each sample response that was consistent with the rubric and the existing anchor papers. If questions arose during scoring, the problem was generally discussed by the group to identify the elements that led to the assignment of different scores. After discussion, the facilitators recorded the final score point

assigned to each sample paper. The group made note of papers that should be used as anchors for the 2009 scoring. Papers for which a score could not be agreed upon were noted as "do not use." Questar facilitators documented discussions and decisions that were made. The notes taken during the range-finding session were used by the scoring director during future reader and team leader training.

After the range-finding session, Questar created prompt-specific scoring guides as well as prompt-specific training, qualifying, and retraining sets. These were used to train the readers to ensure that they were well prepared for operational scoring. The selected papers were carefully reviewed and compared throughout this process to ensure consistency.

The purpose of the second range-finding session, held on June 15-16, 2009, was to score selected responses to field test items from the spring 2009 administration. A team consisting of Questar's ELPA performance assessment manager, the scoring director, and an expert team leader pulled sample responses from several hundred randomly selected responses for each field test item. As with the operational items, Questar selected approximately 50 responses for each of the prompts to score at the range-finding session. These sample papers covered a range of score points and represented a variety of issues that might arise when scoring student responses.

There were three range-finding committees for field test items with one committee being responsible for Levels I and II; one committee for Level III (with a few items for Levels IV and V); and one committee for Levels IV and V. Members on each of these committees reviewed and scored the papers in a manner similar to the procedure described above for the operational items. From these scored responses, anchor sets were created to train readers to score the remainder of the field test responses. (Training and qualifying sets were not created for field test scoring.) Participants included three Questar staff, three OEAA staff, and nineteen Michigan teachers.

3.2 Reader Training

All scoring staff had, at a minimum, a B.S. or B.A. degree and were experienced in scoring other ELL or writing assessments. They had expertise in English Language Arts and were familiar with the nuances of rubric application and scoring.

Prior to scoring the student responses, the scoring director conducted team leader training. These team leaders are expert readers who have experience in all facets of scoring. They are responsible for carefully monitoring the scoring accuracy of readers on their team. These team leaders are the next level experts regarding the items and the requirements and procedures for the project.

Team leaders completed the same training received by the readers. They actually completed the training twice – once during their own session and then again with the readers. During the team leader training sessions, the logistics of the scoring sessions and scoring routines were discussed. This included the criterion by which readers would

qualify, procedures for monitoring accuracy and reliability, and procedures for retraining and evaluating readers on their teams.

The training of readers was conducted by the scoring director for all prompts within each assessment level and was based on the anchors and training sets developed after the first range-finding session. After training, each reader was given a qualifying set to score in order to apply the criteria they had learned. Any reader who failed to meet the criterion of 70% exact agreement with the expert-provided score for a qualifying set was either retrained and given a second chance to meet the scoring criteria, or removed from the project.

The scoring director, in concert with the ELPA performance assessment manager, developed the training (and qualifying) sets for the 2009 forms. The training materials were grouped by assessment level. Training materials were constructed to progress from Level I to Level V and from the simpler 0-1 point range rubrics to the most complex 0-4 point range rubrics. The details for the operational training materials (scoring guide, training and qualifying sets), along with item type and score point ranges, are reported below. (For field test scoring, scoring guides or anchor sets were not used to train the readers.)

Scoring Guides

There were 7 Scoring Guides (SGs) covering all 23 operational prompts across the 5 assessment levels as follows:

SG 1: *Levels I & II:*
Letter Writing (LW) [0-1 point range]
Word Writing with Visual Stimulus (WWV) [0-2 point range]
Word Writing with Aural Stimulus (WWA) [0-2 point range]

SG 2: *Levels I & II:*
Sentence Writing (SW) [0-2 point range]

SG 3: *Levels I & II:*
Extended Response (ER) [0-4 point range]

SG 4: *Level III:*
Sentence Writing with Short Reading (SWSR) [0-2 point range]

SG 5: *Level III:*
Paragraph Writing (PW) [0-4 point range]
Extended Response (ER) [0-4 point range]

SG 6: *Levels IV & V:*
Sentence Combination (SC) [0-2 point range]

DRAFT

SG 7: *Levels IV & V:*
Paragraph Writing (PW) [0-4 point range]
Extended Response (ER) [0-4 point range]

Training and Qualifying Sets

There were a total of 10 Training (T) Sets and 10 Qualifying (Q) Sets. These covered all 23 operational prompts across the 5 assessment levels as follows:

Training Sets 1 & 2: *Levels I & II:*
Letter Writing (LW) [0-1 point range]
Word Writing with Visual Stimulus (WWV) [0-2 point range]
Word Writing with Aural Stimulus (WWA) [0-2 point range]

Training Sets 3 & 4: *Levels I & II:*
Sentence Writing (SW) [0-2 point range]
Extended Response (ER) [0-4 point range]

Training Sets 5 & 6: *Level III:*
Sentence Writing with Short Reading (SWSR) [0-2 point range]
Paragraph Writing (PW) [0-4 point range]
Extended Response (ER) [0-4 point range]

Training Sets 7 & 8: *Levels IV & V:*
Sentence Combination (SC) [0-2 point range]

Training Sets 9 & 10: *Levels IV & V:*
Paragraph Writing (PW) [0-4 point range]
Extended Response (ER) [0-4 point range]

Qualifying Sets 1 & 2: *Levels I & II:*
Letter Writing (LW) [0-1 point range]
Word Writing with Visual Stimulus (WWV) [0-2 point range]
Word Writing with Aural Stimulus (WWA) [0-2 point range]

Qualifying Sets 3 & 4: *Levels I & II:*
Sentence Writing (SW) [0-2 point range]
Extended Response (ER) [0-4 point range]

Qualifying Sets 5 & 6: *Level III:*
Sentence Writing with Short Reading (SWSR) [0-2 point range]
Paragraph Writing (PW) [0-4 point range]
Extended Response (ER) [0-4 point range]

Qualifying Sets 7 & 8: *Levels IV & V:*
Sentence Combination (SC) [0-2 point range]

DRAFT

Qualifying Sets 9 & 10: *Levels IV & V:*
Paragraph Writing (PW) [0-4 point range]
Extended Response (ER) [0-4 point range]

3.3 Inter-Rater Agreement

All readers were trained to score according to the same rubric to ensure accurate, consistent, reliable results. Questar adhered to stringent criteria in its general screening, training, and qualifying procedures as preliminary measures for obtaining high levels of consistency and reliability.

Each response was scored once independently and 20% of the responses receiving a second reading for reliability purposes. The first reader's scores were used as the score of record as long as the second reader had more than 90% agreement with the first reader. If the second reader's scores differed more than 10% (not within a one point difference) from those of the first reader, a third reader scored all of the first reader's responses and the scores assigned by the third reader became the score of record. This did not occur, however, because the scores of all readers were less than 10% nonadjacent.

Team leaders conducted "read-behinds" as an additional monitoring method to ensure that individual readers continued to maintain their scoring accuracy and did not change the way they were scoring over time. When conducting read-behinds, the team leader reviewed student responses and the scores assigned by an individual reader. The team leader could agree with and confirm the score, or disagree with the reader's score and send the paper back for review, citing specific anchor papers as guides.

The evaluation of inter-rater agreement is described in Section 5.3. All papers receiving invalid codes were also read by either a team leader or a scoring director.

Retraining Sets for Monitoring Room Drift

In addition to scoring the student responses, readers also periodically scored a set of retraining responses during the scoring process. This process is used to detect room drift, and it is also used to make sure individual readers are maintaining accuracy. Retraining sets consisted of student responses of mixed quality in a random order that were pre-scored by the scoring director or the range-finding committee. The retraining sets covered all levels and prompts. Questar developed six retraining sets but only had the opportunity to administer three. The retraining sets varied in the number of responses per set based on the item type. All readers met the 70% or higher exact agreement with the master scorers during the scoring of the 2009 ELPA responses.

3.4 Monitoring Reports

Questar's online scoring system generated many different types of internal monitoring reports that enabled Questar to monitor the accuracy of scoring. These reports, computed by individual reader and by team, listed all of a team's readers and provided the results of their scoring on an ongoing basis. The information in these reports included the number

of responses read by the readers during a specific period, reader agreement (or reliability) rate, and score point distribution by prompt.

Reader Reliability Report: Every student response is read once by a randomly selected reader. Twenty percent of the responses received a second independent reading by a randomly selected reader. This provided data for reporting the number of instances and percent of exact agreement, adjacent agreement, and non-adjacent scores between two different readers. Adjacent agreement occurs when the second reader's score was a point higher or lower than the first reader's score. Non-adjacent scores occur when two readers' scores differ by more than one point.

Score Point Distribution Report: This report included the percentage of responses each reader awarded each score point and indicated whether the reader tended to distribute scores in a manner similar to other readers.

In addition to the reports described above, other reports, such as the Project Summary Report, were generated each day to monitor the progress of the project. This report indicated the number and the percent of responses for which first and second readings were required and completed.

Reader Reliability Reports and Score Point Distribution Reports were printed and posted in the scoring room daily. These reports were examined daily by the scoring director and team leaders. In addition, the readers were instructed on how to monitor themselves by reviewing these reports.

3.5 Retraining

The scoring director conducted group retraining while waiting for more responses to become available for scoring. In addition, individual readers received retraining during the scoring process as deemed necessary by team leader observations and by the results of various reports. The need for retraining was identified in different ways: a large number of non-adjacent scores requiring resolution by a team leader, low or irregular scores for the retraining sets, or unsatisfactory exact agreement rates or anomalies detected via the read-behind monitoring.

Retraining involved several techniques:

- Discussion of student response(s) and the scores involved in a resolution
- Discussion of specific papers identified by the read-behind process
- Review of anchors

Appendix B provides the following information for each item per level:

- Item ID
- Item name
- Item type

- Reader agreement summary (including percentage of exact agreement, adjacent agreement, and non-adjacent agreement)
- Score point distribution

DRAFT

SECTION 4. DESCRIPTIVE STATISTICS AT ITEM AND TEST LEVELS

This section includes descriptive statistics at both the item and test levels. The statistics reported at the item level are different from those reported at the test (and domain) level. Therefore, these results are described below separately. Note, however, that some statistics reported at the item level (e.g., *p*-value and item-total correlation) are from the framework of Classical Test Theory (CTT). In this section, only the concepts relating to the *p*-value and the item-total correlation are introduced; other important concepts in the framework of CTT are introduced in Section 5.

4.1 Item-Level Descriptive Statistics

This section presents various item-level statistics for all items (both operational and field test) on the 2009 ELPA. First, item difficulty in the framework of CTT is reported. For dichotomously scored items, the *p*-value is computed, which equals the proportion of students who answer an item correctly. A high *p*-value means that an item is easy; a low *p*-value means that an item is difficult. For polytomously scored items, an adjusted *p*-value is reported. It is computed by dividing the average score points of a CR item by the item's maximum possible points.

Besides item difficulty, an adjusted item-total correlation is also reported for each item. The adjusted item-total correlation is an index of the association between students' performance on the item and their overall performance on all other test items (i.e., with the item of interest excluded from the total score). A high adjusted item-total correlation is desired, as high correlations indicate that students with high scores on all other test items (i.e., high ability) tend to get a particular item correct (or obtain higher score points if it is a CR item), and the students with low scores on all other test items (i.e., low ability) tend to get this particular item incorrect (or obtain lower score points if it is a CR item). Note that for dichotomously scored items, an adjusted point-biserial correlation is computed, and for polytomously scored items, an adjusted Pearson product-moment correlation is computed.

The item-level statistics for the operational items and the embedded field test items for the 2009 ELPA are presented in Appendix C for each assessment level. The statistics for the operational items appear in appendices C.1 to C.5, and the statistics for the field test items are presented in C.6 to C.10. The tables are grouped by domain (i.e., Listening, Reading, Writing, and Speaking). The following item information and statistics are presented for each item in Appendix C:

- Item ID
- Item Sequence (the sequential appearance on the test by domain)
- Domain (Listening, Reading, Writing, or Speaking)
- Type (item type: multiple-choice (MC) or constructed-response (CR))
- Points (maximum number of possible points)
- N (number of students)
- 0, 1/A, 2/B, 3/C, 4/D (percentage of students obtaining each response option for MC items or score point for CR items)

DRAFT

- Omits/Invalids (percentage of students omitting an item or providing an invalid response)[8]
- P-value (*p*-value for MC items and adjusted *p*-value for CR items)
- Item-Total Correlation (correlation between the item score and the adjusted total-test score; for MC items, an adjusted point-biserial correlation is computed, and for CR items, an adjusted Pearson product-moment correlation is computed.)

4.2 Higher-Level Descriptive Statistics

Descriptive statistics for the ELPA scores are also reported by domain and for the overall test. The test directly assesses the following domains: Listening, Reading, Writing, and Speaking. The sum of these four domain scores results in the overall total test score. The fifth domain, Comprehension, is composed of selected items from the Listening and Reading domains, and is not included in the computation of the overall total test score.

The measures of central tendency, variability, and score precision of raw scores are presented in Table 4.1 for each of the five domains as well as for the total test for each assessment level. Table 4.2 presents the same statistics by grade.

The tables include the following information:

- Level or Grade
- Domain
- Number of items
- Max points (maximum number of possible points)
- N-Count
- Raw Score Mean
- Raw Score Standard Deviation (SD)
- Raw Score Standard Error of the Sample Mean (SE)

The standard error of the sample mean is an estimate of the magnitude of sampling error associated with the estimation of the population mean. It is mathematically defined as follows (Hays, 1994, p. 215):

$$SE(\bar{X}) = \frac{\sigma}{\sqrt{n}}, \qquad (4.1)$$

where $SE(\bar{X})$ represents the standard error of the sample mean, σ represents the standard deviation of the population, and n represents the number of examinees.

[8] Note that omits, blanks, and invalids were scored as "0" and are included in the final N-count that appears in all subsequent tables and appendices.

Equation (4.1) expresses $SE(\bar{X})$ in terms of the population standard deviation σ. The practical version of this equation is expressed in terms of the sample standard deviation as follows (Hays, 1994, p. 218):

$$SE(\bar{X}) = \frac{S_X}{\sqrt{n}}, \qquad (4.2)$$

where S_X represents the standard deviation of the sample, and n represents the number of examinees in the sample.

Table 4.1: Summary Statistics of the ELPA Domains by Assessment Level

Level	Domain	Number of Items	Max Points	N-Count	Raw Score		
					Mean	SD	SE
I	Listening	16	16	9141	9.37	3.26	0.03
	Reading	16	16	9141	9.12	3.26	0.03
	Writing	12	18	9141	6.91	4.01	0.04
	Speaking	8	18	9141	10.00	4.64	0.05
	Comprehension	21	21	9141	11.91	3.90	0.04
	Total Test	52	68	9141	35.39	12.01	0.13
II	Listening	20	20	16003	13.55	3.82	0.03
	Reading	20	20	16003	11.93	4.09	0.03
	Writing	13	20	16003	12.89	4.62	0.04
	Speaking	9	20	16003	13.92	4.61	0.04
	Comprehension	33	33	16003	20.85	6.15	0.05
	Total Test	62	80	16003	52.28	13.99	0.11
III	Listening	19	19	15081	12.46	3.65	0.03
	Reading	22	22	15081	11.00	4.37	0.04
	Writing	12	20	15081	11.82	3.86	0.03
	Speaking	9	20	15081	15.35	4.35	0.04
	Comprehension	29	29	15081	17.03	5.35	0.04
	Total Test	62	81	15081	50.62	13.30	0.11
IV	Listening	21	21	10947	12.43	3.70	0.04
	Reading	21	21	10947	12.03	4.25	0.04
	Writing	14	22	10947	14.43	4.34	0.04
	Speaking	10	21	10947	16.40	4.87	0.05
	Comprehension	28	28	10947	16.07	4.88	0.05
	Total Test	66	85	10947	55.29	14.21	0.14
V	Listening	20	20	10908	12.32	3.79	0.04
	Reading	21	21	10908	12.25	4.57	0.04
	Writing	14	22	10908	13.80	4.56	0.04
	Speaking	10	21	10908	17.31	5.09	0.05
	Comprehension	28	28	10908	17.41	5.52	0.05
	Total Test	65	84	10908	55.69	15.39	0.15

Note. 1. Total Test does not include the Comprehension composite.
2. Omits, blanks, and invalid responses were scored as "0."

DRAFT

Table 4.2: Summary Statistics of the ELPA Domains by Grade

Grade	Domain	Number of Items	Max Points	N-Count	Raw Score		
					Mean	SD	SE
K	Listening	16	16	9141	9.37	3.26	0.03
	Reading	16	16	9141	9.12	3.26	0.03
	Writing	12	18	9141	6.91	4.01	0.04
	Speaking	8	18	9141	10.00	4.64	0.05
	Comprehension	21	21	9141	11.91	3.90	0.04
	Total Test	52	68	9141	35.39	12.01	0.13
1	Listening	20	20	8292	12.28	3.79	0.04
	Reading	20	20	8292	10.22	3.78	0.04
	Writing	13	20	8292	11.00	4.48	0.05
	Speaking	9	20	8292	12.82	4.73	0.05
	Comprehension	33	33	8292	18.28	5.69	0.06
	Total Test	62	80	8292	46.32	13.04	0.14
2	Listening	20	20	7711	14.90	3.37	0.04
	Reading	20	20	7711	13.77	3.59	0.04
	Writing	13	20	7711	14.91	3.86	0.04
	Speaking	9	20	7711	15.10	4.17	0.05
	Comprehension	33	33	7711	23.61	5.39	0.06
	Total Test	62	80	7711	58.69	12.00	0.14
3	Listening	19	19	5786	11.25	3.61	0.05
	Reading	22	22	5786	9.49	3.86	0.05
	Writing	12	20	5786	10.51	3.75	0.05
	Speaking	9	20	5786	14.65	4.41	0.06
	Comprehension	29	29	5786	15.12	4.99	0.07
	Total Test	62	81	5786	45.90	12.40	0.16
4	Listening	19	19	4978	12.80	3.42	0.05
	Reading	22	22	4978	11.26	4.25	0.06
	Writing	12	20	4978	12.14	3.67	0.05
	Speaking	9	20	4978	15.58	4.23	0.06
	Comprehension	29	29	4978	17.49	5.11	0.07
	Total Test	62	81	4978	51.78	12.64	0.18
5	Listening	19	19	4317	13.67	3.45	0.05
	Reading	22	22	4317	12.70	4.46	0.07
	Writing	12	20	4317	13.22	3.65	0.06
	Speaking	9	20	4317	16.02	4.25	0.06
	Comprehension	29	29	4317	19.06	5.22	0.08
	Total Test	62	81	4317	55.62	13.07	0.20
6	Listening	21	21	3709	11.61	3.50	0.06
	Reading	21	21	3709	11.03	4.02	0.07
	Writing	14	22	3709	13.53	4.23	0.07
	Speaking	10	21	3709	15.96	4.83	0.08
	Comprehension	28	28	3709	14.93	4.57	0.08
	Total Test	66	85	3709	52.13	13.42	0.22

DRAFT

Table 4.2: Summary Statistics of the ELPA Domains by Grade (Continued)

Grade	Domain	Number of Items	Max Points	N-Count	Raw Score		
					Mean	SD	SE
7	Listening	21	21	3567	12.53	3.72	0.06
	Reading	21	21	3567	12.04	4.23	0.07
	Writing	14	22	3567	14.42	4.32	0.07
	Speaking	10	21	3567	16.39	4.91	0.08
	Comprehension	28	28	3567	16.14	4.88	0.08
	Total Test	66	85	3567	55.38	14.17	0.24
8	Listening	21	21	3671	13.14	3.73	0.06
	Reading	21	21	3671	13.04	4.25	0.07
	Writing	14	22	3671	15.35	4.29	0.07
	Speaking	10	21	3671	16.85	4.83	0.08
	Comprehension	28	28	3671	17.15	4.94	0.08
	Total Test	66	85	3671	58.39	14.35	0.24
9	Listening	20	20	3544	11.70	3.74	0.06
	Reading	21	21	3544	11.53	4.47	0.08
	Writing	14	22	3544	13.08	4.65	0.08
	Speaking	10	21	3544	16.73	5.68	0.10
	Comprehension	28	28	3544	16.47	5.41	0.09
	Total Test	65	84	3544	53.03	15.99	0.27
10	Listening	20	20	3150	12.30	3.85	0.07
	Reading	21	21	3150	12.19	4.59	0.08
	Writing	14	22	3150	13.67	4.65	0.08
	Speaking	10	21	3150	17.12	5.25	0.09
	Comprehension	28	28	3150	17.34	5.59	0.10
	Total Test	65	84	3150	55.28	15.79	0.28
11	Listening	20	20	2265	12.97	3.71	0.08
	Reading	21	21	2265	12.89	4.51	0.09
	Writing	14	22	2265	14.54	4.34	0.09
	Speaking	10	21	2265	17.92	4.36	0.09
	Comprehension	28	28	2265	18.32	5.39	0.11
	Total Test	65	84	2265	58.32	14.11	0.30
12	Listening	20	20	1949	12.74	3.73	0.08
	Reading	21	21	1949	12.93	4.58	0.10
	Writing	14	22	1949	14.48	4.28	0.10
	Speaking	10	21	1949	17.98	4.27	0.10
	Comprehension	28	28	1949	18.18	5.48	0.12
	Total Test	65	84	1949	58.13	14.13	0.32

Note. 1. Total Test does not include the Comprehension composite.
2. Omits, blanks, and invalid responses were scored as "0."

DRAFT

SECTION 5. RELIABILITY

In this section, classical true score theory is introduced, along with two concepts that are directly derived from this theory: reliability and standard error of measurement. The concept of standard error of measurement can also be expressed within the framework of item response theory (IRT), and it is also described and reported in this section.

In practical terms, reliability is understood as the degree of consistency over repeated administrations of the same or alternative tests (Crocker & Algina, 1986). Such consistency can also be broadened to situations where the same test item is scored by different raters (Haertel, 2006). Therefore, the fourth topic covered in this section is inter-rater reliability.

In addition, students can be categorized into different performance levels based on their observed test scores. It is informative to examine the accuracy or consistency of such categorization, especially the cut score that separates students into proficient and non-proficient categories. The classification accuracy and classification consistency at the proficient level cut is presented in the final section.

5.1 Classical True Score Theory

According to classical true score theory (which is a fundamental component of Classical Test Theory or CTT), an observed score consists of the sum of two parts—a random component of true score (T) and a random component of error score (E), or mathematically, $X = T + E$ (McDonald, 1999). This model has the following properties (Lord & Novick, 1968): (1) the expected error score is zero, (2) the correlation between true and error scores is zero, and (3) the correlation between the error scores on different but parallel forms is zero.

According to the model presented above, a student's observed test score is an imprecise estimate of their actual ability because a portion of that score is attributable to random error. A fundamental theoretical quantity in test theory, the *reliability coefficient* of the observed scores, is defined as the ratio of the variance of true scores to the variance of observed scores. Tests are thus most reliable when the proportion of observed score variance that may be attributed to error variance is minimal. There are three main recognized methods for estimating the reliability coefficients (see McDonald, 1999), but only internal consistency reliability is described here.

5.1a Internal Consistency Reliability

Estimates of internal consistency reliability involve "dividing the test into two or more constituent parts and in some way estimating reliability from the consistency of performance across these part-tests" (Haertel, 2006, p. 71).

Various internal consistency reliability estimates have been proposed, but the most widely used is Cronbach's (1951) coefficient alpha (Haertel, 2006). Using sample statistics, it is computed as follows (adapted from Haertel, 2006, p. 74):

$$\alpha = \frac{I}{I-1}\left(1-\frac{\sum_i^I S_i^2}{S_X^2}\right), \qquad (5.1)$$

where I represents the number of items on the test, S_i^2 represents the sample variance of item i, and S_X^2 represents the sample variance of the total test score.

The use of coefficient alpha has several theoretical advantages (Haertel, 2006). First, since it equals the mean of all possible split-half reliability coefficients (i.e., another estimate of internal consistency reliability which involves the division of the total test into two "parallel" sub-tests), the use of coefficient alpha avoids the arbitrary choice of a split or division. Second, it is mathematically equivalent to one of the lower bounds of the theoretical reliability coefficient. The implication of this is that the theoretical reliability coefficient is higher than the observed coefficient alpha.

5.1b Standard Error of Measurement (SEM) within CTT Framework

As mentioned above, another related concept that is directly derived from the classical true score theory is SEM. It is calculated with sample statistics as follows (Hays, 1994, p. 617):

$$SEM(X) = S_X\sqrt{1-r_{XX'}}\,, \qquad (5.2)$$

where $SEM(X)$ represents the estimated standard error of measurement of the observed test score X, S_X denotes the estimated standard deviation (i.e., sample standard deviation) of the observed score, $r_{XX'}$ represents the estimated reliability coefficient of a test. In this document, the observed coefficient alpha is used as the estimated reliability coefficient.

According to Equation (5.2), it is obvious that the SEM is inversely related to the reliability of a test; as for any standard deviation of the observed score, the SEM decreases when the reliability coefficient increases. Thus, when the SEM is small, one has more confidence in the accuracy, or precision, of observed test scores.

Table 5.1 provides the raw score descriptive statistics and reliabilities by grade and domain. It includes the following information:

- Grade
- Domain
- Number of items
- Max Points (maximum number of possible points)
- N-Count
- Mean (Raw Score Mean)
- SD (Raw Score Standard Deviation)
- Reliability (estimated Cronbach's alpha coefficient)
- Classical SEM (estimated SEM within CTT framework)

Table 5.1: Raw Score Descriptive Statistics and Reliability by Grade and Domain

Grade	Domain	Number of Items	Max Points	N-Count	Mean	SD	Reliability	Classical SEM
K	Listening	16	16	9141	9.37	3.26	0.70	1.79
	Reading	16	16	9141	9.12	3.26	0.73	1.69
	Writing	12	18	9141	6.91	4.01	0.79	1.84
	Speaking	8	18	9141	10.00	4.64	0.82	1.97
	Comprehension	21	21	9141	11.91	3.90	0.73	2.03
	Total Test	52	68	9141	35.39	12.01	0.90	3.80
1	Listening	20	20	8292	12.28	3.79	0.73	1.97
	Reading	20	20	8292	10.22	3.78	0.72	2.00
	Writing	13	20	8292	11.00	4.48	0.82	1.90
	Speaking	9	20	8292	12.82	4.73	0.84	1.89
	Comprehension	33	33	8292	18.28	5.69	0.79	2.61
	Total Test	62	80	8292	46.32	13.04	0.90	4.12
2	Listening	20	20	7711	14.90	3.37	0.72	1.78
	Reading	20	20	7711	13.77	3.59	0.75	1.80
	Writing	13	20	7711	14.91	3.86	0.80	1.73
	Speaking	9	20	7711	15.10	4.17	0.82	1.77
	Comprehension	33	33	7711	23.61	5.39	0.81	2.35
	Total Test	62	80	7711	58.69	12.00	0.91	3.60
3	Listening	19	19	5786	11.25	3.61	0.71	1.94
	Reading	22	22	5786	9.49	3.86	0.69	2.15
	Writing	12	20	5786	10.51	3.75	0.77	1.80
	Speaking	9	20	5786	14.65	4.41	0.84	1.76
	Comprehension	29	29	5786	15.12	4.99	0.76	2.44
	Total Test	62	81	5786	45.90	12.40	0.90	3.92
4	Listening	19	19	4978	12.80	3.42	0.71	1.84
	Reading	22	22	4978	11.26	4.25	0.76	2.08
	Writing	12	20	4978	12.14	3.67	0.78	1.72
	Speaking	9	20	4978	15.58	4.23	0.84	1.69
	Comprehension	29	29	4978	17.49	5.11	0.79	2.34
	Total Test	62	81	4978	51.78	12.64	0.91	3.79
5	Listening	19	19	4317	13.67	3.45	0.74	1.76
	Reading	22	22	4317	12.70	4.46	0.79	2.04
	Writing	12	20	4317	13.22	3.65	0.77	1.75
	Speaking	9	20	4317	16.02	4.25	0.86	1.59
	Comprehension	29	29	4317	19.06	5.22	0.81	2.28
	Total Test	62	81	4317	55.62	13.07	0.92	3.70
6	Listening	21	21	3709	11.61	3.50	0.64	2.10
	Reading	21	21	3709	11.03	4.02	0.73	2.09
	Writing	14	22	3709	13.53	4.23	0.76	2.07
	Speaking	10	21	3709	15.96	4.83	0.87	1.74
	Comprehension	28	28	3709	14.93	4.57	0.72	2.42
	Total Test	66	85	3709	52.13	13.42	0.90	4.24

DRAFT

Table 5.1: Raw Score Descriptive Statistics and Reliability by Grade and Domain (Continued)

Grade	Domain	Number of Items	Max Points	N-Count	Mean	SD	Reliability	Classical SEM
7	Listening	21	21	3567	12.53	3.72	0.70	2.04
	Reading	21	21	3567	12.04	4.23	0.76	2.07
	Writing	14	22	3567	14.42	4.32	0.78	2.03
	Speaking	10	21	3567	16.39	4.91	0.88	1.70
	Comprehension	28	28	3567	16.14	4.88	0.76	2.39
	Total Test	66	85	3567	55.38	14.17	0.92	4.01
8	Listening	21	21	3671	13.14	3.73	0.72	1.97
	Reading	21	21	3671	13.04	4.25	0.78	1.99
	Writing	14	22	3671	15.35	4.29	0.78	2.01
	Speaking	10	21	3671	16.85	4.83	0.88	1.71
	Comprehension	28	28	3671	17.15	4.94	0.78	2.32
	Total Test	66	85	3671	58.39	14.35	0.92	4.06
9	Listening	20	20	3544	11.70	3.74	0.73	1.94
	Reading	21	21	3544	11.53	4.47	0.80	2.00
	Writing	14	22	3544	13.08	4.65	0.81	2.03
	Speaking	10	21	3544	16.73	5.68	0.92	1.61
	Comprehension	28	28	3544	16.47	5.41	0.82	2.30
	Total Test	65	84	3544	53.03	15.99	0.94	3.92
10	Listening	20	20	3150	12.30	3.85	0.75	1.93
	Reading	21	21	3150	12.19	4.59	0.82	1.95
	Writing	14	22	3150	13.67	4.65	0.81	2.03
	Speaking	10	21	3150	17.12	5.25	0.91	1.58
	Comprehension	28	28	3150	17.34	5.59	0.84	2.24
	Total Test	65	84	3150	55.28	15.79	0.94	3.87
11	Listening	20	20	2265	12.97	3.71	0.74	1.89
	Reading	21	21	2265	12.89	4.51	0.82	1.91
	Writing	14	22	2265	14.54	4.34	0.80	1.94
	Speaking	10	21	2265	17.92	4.36	0.89	1.45
	Comprehension	28	28	2265	18.32	5.39	0.83	2.22
	Total Test	65	84	2265	58.32	14.11	0.93	3.73
12	Listening	20	20	1949	12.74	3.73	0.74	1.90
	Reading	21	21	1949	12.93	4.58	0.82	1.94
	Writing	14	22	1949	14.48	4.28	0.79	1.96
	Speaking	10	21	1949	17.98	4.27	0.89	1.42
	Comprehension	28	28	1949	18.18	5.48	0.84	2.19
	Total Test	65	84	1949	58.13	14.13	0.93	3.74

Note. 1. Total Test does not include the Comprehension domain.
2. Omits, blanks, and invalid responses were scored as "0."

DRAFT

5.2 Conditional SEM within IRT Framework

Unlike the SEM based on CTT, the SEM based on IRT is not the same for all persons. Under the unidimensional Rasch model, the SEM for each person is as follows (adapted from Wright & Stone, 1979, p. 135):

$$\sigma_\theta = \frac{1}{\sqrt{\sum_{i=1}^{L} p_{vi}(1-p_{vi})}}, \qquad (5.3)$$

where v is the subscript for a person, i is the subscript for an item, L represents the length of a test, θ represents the ability estimate, and p_{vi} represents the probability that person v answers item i correctly. For the unidimensional Rasch model, p_{vi} is defined as follows (adapted from Wright & Stone, 1979, p. 15):

$$p_{vi} = \frac{e^{\theta_v - \delta_i}}{1+e^{\theta_v - \delta_i}}, \qquad (5.4)$$

where θ represents person v's ability and δ_i represents item i's difficulty.

Under the unidimensional PCM, the probability that a person with ability θ obtains a score of x on item i is expressed as follows (adapted from Masters & Wright, 1997, p. 103):

$$p_{ix}(\theta) = \frac{\exp \sum_{k=0}^{x}(\theta_v - \delta_{ik})}{\sum_{h=0}^{m_i} \exp \sum_{k=0}^{h}(\theta_v - \delta_{ik})}, x = 0, 1, \ldots, m_i, \qquad (5.5)$$

where, for notational convenience,

$$\sum_{k=0}^{0}(\theta_v - \delta_{ik}) \equiv 0 \text{ and } \sum_{k=0}^{h}(\theta_v - \delta_{ik}) \equiv \sum_{k=1}^{h}(\theta_v - \delta_{ik}).$$

In Equation (5.5), θ_v is the ability of person v, and δ_{ik} is an item parameter governing the probability of scoring k rather than $k-1$ on item i. In other words, δ_{ik} parameters represent "the relative difficulty of each step" (Embretson & Reise, 2000, p. 106). When $m_i = 1$ (i.e., an item with only two response categories), Equation (5.5) simplifies to Equation (5.4).

Note that the SEM for item difficulty is smallest when the probability of passing is close to the probability of failing. That is, when an item is near the threshold level for many persons in the sample, the SEM is small (Embretson & Reise, 2000).

The conditional SEMs for the total test and each domain are presented in Appendix D.

DRAFT

5.3 Inter-Rater Reliability

Another source of measurement error occurs during the evaluation of student work. As first discussed in Section 3.3, inter-rater reliability investigates the consistency of the evaluation given by different raters (Crocker & Algina, 1986). One way to obtain evidence of inter-rater reliability is to calculate the percent agreement between raters. If raters always agree in their assignment of scores, there is 100% agreement. If raters never agree in their assignment of scores, there is 0% agreement. Appendix B presents the results of inter-rater agreement for the 2009 ELPA CR items. As shown in Appendix B, the exact agreement for operational CR items scored by two readers across all assessment levels ranged from 87% to 97%, and the exact agreement for field test CR items across all assessment levels ranged from 73% to 97%.

5.4 Quality (or Reliability) of Classification Decision at Proficient Cut

Based on the ELPA scale scores, student performance is classified into one of the five performance levels. While it is always important to know the reliability of student scores in any examination, it is also important to assess the quality of the decisions based on these scores. Evaluation of the quality of classification decisions is performed through estimation of the probabilities of correct and consistent classification of student performance. Procedures from Livingston and Lewis (1995) were applied to derive measures of the accuracy and consistency of the classifications. A brief description of the procedures used and the results obtained are presented below.

The accuracy of decisions is the extent to which decisions would agree with those that would be made if each student could somehow be tested with all possible forms of an examination. The consistency of decisions is the extent to which decisions would agree with the decision that would be made if the students had taken a parallel form of the ELPA, equal in difficulty and covering the same content as the form they actually took. These ideas are shown schematically in Figures 5.1 and 5.2. In both figures, "Achieves Proficient Status" refers to the proficient and above category on the total test score, and "Does Not Achieve Proficient Status" refers to all categories below the proficient cut.

Figure 5.1: Classification Accuracy

		Decision made on a form actually taken	
		Does Not Achieve Proficient Status	Achieves Proficient Status
True status made on all-forms average	Does Not Achieve Proficient Status	Correct Classification	Misclassification
	Achieves Proficient Status	Misclassification	Correct Classification

Note: Adapted from Young and Yoon (1998).

Figure 5.2: Classification Consistency

		Decision made on the 2nd form taken	
		Does Not Achieve Proficient Status	Achieves Proficient Status
Decision made on the 1st form taken	Does Not Achieve Proficient Status	Correct Classification	Misclassification
	Achieves Proficient Status	Misclassification	Correct Classification

Note: Adapted from Young and Yoon (1998).

In Figure 5.1, accurate classification occurs when the decision made on the basis of the form actually taken agrees with the decision made on the basis of the theoretical "all-forms" average. Misclassification occurs, for example, when a student who actually "Does Not Achieve Proficient Status" based on his or her "all-forms" average is classified incorrectly as "Achieves Proficient Status."

Consistent classification occurs (see Figure 5.2) when two possible alternate forms agree on the classification of a student as either "Achieves Proficient Status" or "Does Not Achieve Proficient Status," whereas inconsistent classification occurs when the decisions made by the forms differ.

These analyses make use of the techniques outlined and implemented by Hanson (1991), Livingston and Lewis (1995), and Brennan (2004). In this method, the reliability of the total test and the frequency distribution of the scores at each grade level are used to build models for the theoretical true score and alternate form classifications. The proportional true score is estimated by fitting a 4-parameter beta distribution. The conditional distribution of scores on an alternate form given the true score is estimated from a binomial distribution based on the estimated effective test length. The BB-CLASS software was used to complete these analyses (Brennan, 2004).

Table 5.2 presents the analysis results of decision accuracy and consistency for classifying students at each grade level as "Achieves Proficient Status" or "Does Not Achieve Proficient Status" based on their total ELPA score. In addition, Table 5.2 provides information on the proportion of false positives and false negatives (i.e., the two types of misclassification). The false positive is the type of misclassification where students should be classified in the "Does Not Achieve Proficient Status" category based on their "all-forms" average but end up in the "Achieves Proficient Status" category based on the actual form. The false negative is just the opposite—students who should be in "Achieves Proficient Status" category based on their "all-forms" average end up in the "Does Not Achieve Proficient Status" category based on the actual form. The sum of the proportion values for accuracy, false positives, and false negatives should be equal to 1.00. Due to rounding, however, the sum of these values in the table may not be equal to 1.00.

DRAFT

Table 5.2: Decision Accuracy and Consistency Table by Grade

Grade	Test	Accuracy	False Positives	False Negatives	Consistency
K	Total ELPA	0.91	0.05	0.04	0.87
1	Total ELPA	0.91	0.05	0.04	0.88
2	Total ELPA	0.88	0.05	0.07	0.83
3	Total ELPA	0.90	0.05	0.05	0.86
4	Total ELPA	0.90	0.05	0.05	0.86
5	Total ELPA	0.91	0.04	0.05	0.87
6	Total ELPA	0.89	0.06	0.05	0.85
7	Total ELPA	0.89	0.05	0.06	0.84
8	Total ELPA	0.87	0.06	0.07	0.82
9	Total ELPA	0.89	0.05	0.06	0.85
10	Total ELPA	0.90	0.05	0.06	0.85
11	Total ELPA	0.90	0.05	0.05	0.86
12	Total ELPA	0.91	0.05	0.04	0.88

NOTE. The sum of Accuracy, False Positives, and False Negatives may not add up to 1.00 because of rounding.

According to Table 5.2, across all grade levels, the accuracy value ranged from .87 (87%) at grade 8 to .91 (91%) at grades K, 1, 5, and 12.

The proportion of false positives (i.e., labeling a student as proficient when he or she should be categorized as not proficient) ranged from .04 (4%) at grade 5 to .06 (6%) at grades 6 and 8. The proportion of false negatives (i.e., labeling a student as not proficient when he or she should be categorized as proficient) ranged from .04 (4%) at grades K, 1, and 12 to .07 (7%) at grades 2 and 8.

The last column in Table 5.2 reports the proportion of students predicted by the model that would be assigned to the same category (i.e., proficient or not proficient) if an alternate form of the ELPA had been administered with similar content coverage and average item difficulty as the actual form. These values ranged from .82 (82%) at grade 8 to .88 (88%) at grades 1 and 12.

DRAFT

SECTION 6. CALIBRATION, EQUATING, AND SCALING

In addition to the reporting of raw score summary statistics and item-level statistics computed within the framework of CTT, the items on the ELPA were also analyzed within the framework of IRT. CTT is limited in that examinee characteristics cannot be separated from the test characteristics. Theoretically, IRT overcomes this limitation because of the following (Yen & Fitzpatrick, 2006): (1) *person-free item measurement* (i.e., the stability of an item's parameters across different groups of examinees) and (2) *item-free person measurement* (i.e., the stability of an examinee's trait level across different items for the same construct). These two terms are "exaggerations relative to what is found with real tests, but they do convey the ideals that motivate the use of IRT" (Yen & Fitzpatrick, 2006, p. 113).

The unidimensional Rasch model (Rasch, 1960) for dichotomous items and the unidimensional Partial Credit Model (PCM) (Masters, 1982) for polytomous items were used to calibrate, equate, and scale the 2009 ELPA. The related details are provided below.

6.1 The Unidimensional Rasch and Partial Credit Models

Equation (5.4) in Section 5 defines the unidimensional Rasch model, and it expresses the probability of obtaining a correct response given the examinee's ability. The expression of such item response probability is called the *item response function* (IRF). From this equation, it is clear that an item response in the unidimensional Rasch model is assumed to be "governed by the difference between the ability of the person and the difficulty of the item" (Wright & Stone, 1979, p. xiii).

The graphical representation of the IRF is called the *item response curve* (IRC). While the IRF "describes the probability of a given response to an item, the *item characteristic function* (ICF), and its graphical representation, the *item characteristic curve* (ICC), defines the expected item score as a function of ability" (Yen & Fitzpatrick, 2006, p. 113). The IRF and the ICF are equal (as are IRC and ICC) for dichotomous items, but are different (and thus cannot be used interchangeably) for polytomous items.

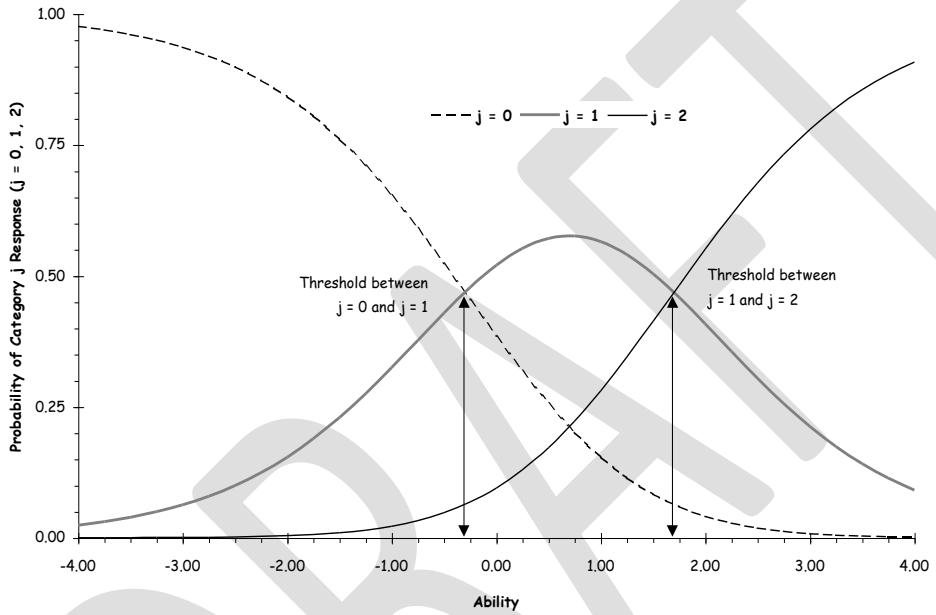
The probability of a correct response in an IRF is bounded by 1 (certainty of a correct response) and 0 (certainty of an incorrect response). The ability scale is, in theory, unbounded. In practice, however, the ability scale ranges from - 4 to + 4 logits (log-odds) for heterogeneous ability groups.

The unidimensional PCM is a direct extension of the unidimensional Rasch model for ordered response categories (≥ 3 categories). Equation (5.5) in Section 5 defines the unidimensional PCM and expresses the probability of an examinee obtaining a particular score on a specific item.

Thissen and Steinberg (1986) referred to this model as a divide-by-total model because "the probability of responding in a particular category [is] written directly as an

exponential divided by a sum of exponentials" (Embretson & Reise, 2000, p. 105). When there are only two response categories (i.e., 0 and 1), Equation (5.5) is equivalent to Equation (5.4). Therefore, the key difference between Equation (5.5) and Equation (5.4) is to assume one or more additional response categories. Suppose that, rather than scoring items as completely wrong or completely right, a category is added. This category represents answers that, though not totally correct, are still clearly not totally incorrect. These relationships are shown in Figure 6.1.

Figure 6.1: Category Response Curves for a Two-Step Item Using the PCM



As shown in Figure 6.1, the category response curve for j=0 represents the probability of observing a score of "0" (completely incorrect) on the item, given an examinee's ability. Those of very low ability (e.g., below -2) are very likely to be in this category and, in fact, are more likely to be in this category than the other two. Those receiving a "1" (partial credit) tend to fall in the middle range of abilities. The category response curve for j=2 represents the probability of observing a score of "2" (completely correct) on the item, given an examinee's ability. Very high-ability people are clearly more likely to be in this category, but there are still some examinees with average or low ability who get full credit on the item.

Similar to that of the unidimensional Rasch model, the separation of item parameters from person measures for the unidimensional PCM "results in sufficient statistics for person ability and step difficulty. For a person's ability the sufficient statistic is the count of the total number of steps the person completes, and for an item's step difficulties the sufficient statistics are counts of the number of persons completing each step" (Masters, 1982, p. 173).

It is important to note that WINSTEPS has a slightly different parameterization of the PCM (see Linacre, 2008, p. 453). Instead of estimating δ_{ik} directly (see Equation (5.5)), WINSTEPS estimates δ_i and F_{ik} where $\text{sum}(F_{ik}) = 0$ and $\delta_{ik} = \delta_i + F_{ik}$. Moreover, δ_i is the mean δ_{ik} of a particular item.

The item difficulty and step difficulty parameters estimated by the unidimensional Rasch model and PCM are provided in Appendix E.

6.2 Calibration of the Spring 2009 ELPA

A brief outline of the calibration, equating, and scaling procedure for the 2009 ELPA is presented below followed by a more detailed description:

- The 2009 operational items with known calibration values were used as anchor items to calibrate the remaining operational items onto the existing ELPA scale.
- Based on the early return data, items with known calibration values were checked for the appropriateness of these "historic" values and the remaining items were calibrated based on these "historic" values. The resulting set of item calibration values for all operational items were used to produce raw score to scale score conversion tables for the total test and all domains for each assessment level.
- For each assessment level, the embedded field test items were then calibrated to the same metric as those operational items with known calibration values.

The first step was to identify the anchor items in the spring 2009 ELPA with known calibration values. A file containing these calibrations was created so that these items could serve as a reference point for the calibration of the non-anchored operational items. It should be noted that the non-anchored operational items were selected from previous field test administrations and met the content and statistical criteria for use as operational items. Due to various reasons such as items being administered in a stand-alone field test administration, these items had never been calibrated to the existing ELPA scale.

The items were calibrated for each assessment level with the use of WINSTEPS 3.67 (Linacre, 2008). The data set used for the item calibrations was from early return data that included all assessment levels and test forms. Table 6.1 provides the number of students in the early-return data for each assessment level and test form[9]. Although the operational items appeared in different positions across the 2009 ELPA forms, their positions on Form 1 were used to calibrate the items regardless of their positions on other forms.

Once all of the operational items were calibrated, the resulting item parameter values were used to obtain the raw score to scale score tables for each of the five assessment levels and the individual domains within each assessment level. The final step was to

[9] The calibration process for the non-anchored operational items was repeated after the full data set was available. The calibrations established with the early-return data set were consistent with the calibrations established with the full data set.

calibrate the embedded field test items for the 2009 administration to the existing ELPA scale.

Table 6.1: Number of Students in the Calibration Sample by Assessment Level and Form

Level	Form 1	Form 2	Form 3	Form 4	Form 5	Total	Percent of Total Population
I	1217	1113	1053	1030	1071	5484	0.60
II	2386	2116	2735	2109	N/A	9346	0.58
III	2641	3037	3018	N/A	N/A	8696	0.58
IV	2023	1824	2011	N/A	N/A	5858	0.54
V	1651	1231	1580	N/A	N/A	4462	0.41

Stability of Anchor Items

There are various methods for evaluating anchor item stability. For the spring 2009 ELPA, however, anchor item stability was checked using only the displacement values from WINSTEPS. Each displacement value indicates the difference between the anchor value and the estimation based on the current data for each item. According to Linacre (2008), if a displacement value is greater than or equal to |.50|, the item should be unanchored. Table 6.2 presents the number of anchor items included in the calibration at each assessment level.

Across all five assessment levels, the following numbers of items were excluded because they exceeded the displacement value of |.50|: 3 items at Level I, 2 items at Level II, 1 item at Level III, and 1 item at Level IV. None of the Level V items exceeded the established criteria. The items with high displacement values for Levels I through IV were un-anchored and re-calibrated

DRAFT

Table 6.2: Number of Anchor Items by Domain and Assessment Level

Level	Domain	Total Number of Items	Number of Anchor Items	Percentage of Total Items by Domain
I	Listening	16	13	0.81
	Reading	16	11	0.69
	Writing	12	9	0.75
	Speaking	8	8	1.00
	Total	52	41	0.79
II	Listening	20	11	0.55
	Reading	20	13	0.65
	Writing	13	8	0.62
	Speaking	9	5	0.56
	Total	62	37	0.60
III	Listening	19	12	0.63
	Reading	22	11	0.50
	Writing	12	5	0.42
	Speaking	9	6	0.67
	Total	62	34	0.55
IV	Listening	21	11	0.52
	Reading	21	14	0.67
	Writing	14	8	0.57
	Speaking	10	6	0.60
	Total	66	39	0.59
V	Listening	20	9	0.45
	Reading	21	14	0.67
	Writing	14	8	0.57
	Speaking	10	6	0.60
	Total	65	37	0.57

6.3 Scale Scores for the ELPA

Once the spring 2009 ELPA items were calibrated to the existing ELPA scale, the 2009 ELPA raw scores were transformed to scale scores that ranged from 300 to 804 on the total test, and from 30 to 83 on each of the domains (i.e., Listening, Reading, Writing, Speaking, and Comprehension).

The ELPA scaling procedure involves linear transformations of the theta (ability) values into scale score points. These transformations do not give more weight to particular domain subtests, and they do not change the rank ordering of students. The performance level classification is also unaffected because the cut scores are defined on the theta metric. Linear transformation constants are utilized. The equations used to establish each total test and domain scale score are presented in Table 6.3.

DRAFT

Table 6.3: Scale Score Transformation Equations for Total Test and Domains

Domain	*Equation*
Total Test	31.25 * (theta) + 550
Listening	3.85 * (theta) + 57
Speaking	4.20 * (theta) + 57
Reading	3.75 * (theta) + 55
Writing	4.00 * (theta) + 52
Comprehension	3.55 * (theta) + 56

Appendices D.1.a–D.5.f provide the raw score to scale score conversion tables by assessment level for the total test as well as for the Listening, Reading, Writing, Speaking, and Comprehension domains. Appendix D includes the following information:

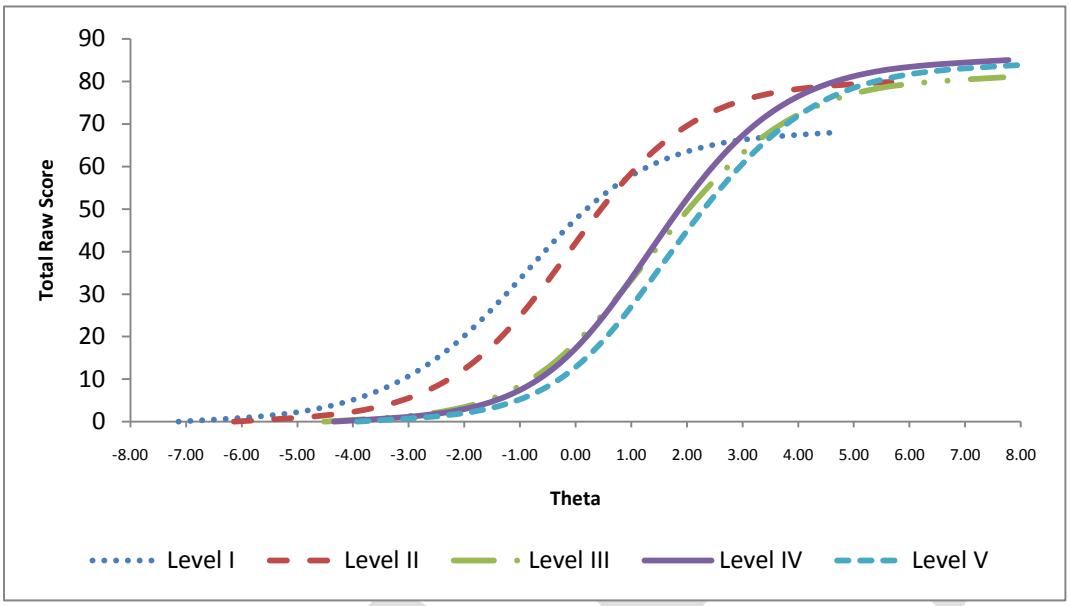
- Raw Score
- Theta (i.e., ability)
- SE (Theta) (Standard Error of Theta)
- Scaled Score
- SE (Scale Score) (Standard Error of the Scale Score)

6.4 Test Characteristic Curves for the ELPA by Assessment Level

Figure 6.2 displays the test characteristic curves (TCCs) for the spring 2009 ELPA by assessment level. The TCCs are the sum of the item characteristic functions conditioned on ability. Note that only the Rasch model and the PCM were used. Therefore, as shown in Figure 6.2, each score on the total score continuum of the test maps to a unique estimate of a student's ability. Thus, an increase in ability estimates of the examinees corresponds to an increase in their total scores. Moreover, as shown in Figure 6.2, the relation between an examinee's score and his/her ability estimate is non-linear. In general, the TCCs shift to the right as one progresses to the next higher assessment level, indicating the relative increase in student ability required as one advances through the assessment levels.

DRAFT

Figure 6.2: ELPA Test Characteristic Curves (TCCs) by Assessment Level



DRAFT

SECTION 7. IRT STATISTICS

7.1 Rasch Statistics

As discussed in Section 6, the unidimensional Rasch model and the unidimensional PCM were used to calibrate the spring 2009 ELPA within the framework of IRT. IRT emphasizes the statistical properties of individual items and their unique contributions to a test and so yields statistics that differ from those produced within the framework of CTT (Yen & Fitzpatrick, 2006).

A summary of the Rasch difficulty values (i.e., minimum, maximum, and mean) by assessment level and domain is presented in Table 7.1.

Table 7.1: Summary of Rasch Difficulty Values by Assessment Level and Domain

Level	Domain	Number of Items	Max Points	Minimum Rasch Difficulty	Maximum Rasch Difficulty	Mean Rasch Difficulty
I	Listening	16	16	-3.33	-0.22	-1.35
	Reading	16	16	-4.38	0.14	-1.27
	Writing	12	18	-2.05	1.06	-0.63
	Speaking	8	18	-3.58	0.00	-1.96
	Comprehension	21	21	-4.38	0.14	-1.29
	Total Test	52	68	-4.38	1.06	-1.25
II	Listening	20	20	-1.67	0.77	-0.25
	Reading	20	20	-3.29	1.70	0.10
	Writing	13	20	-1.06	0.90	-0.13
	Speaking	9	20	-2.51	0.50	-0.85
	Comprehension	33	33	-1.47	1.70	0.01
	Total Test	62	80	-3.29	1.70	-0.20
III	Listening	19	19	-0.26	2.54	1.34
	Reading	22	22	0.25	3.13	2.10
	Writing	12	20	0.33	2.73	1.50
	Speaking	9	20	-1.00	1.51	0.33
	Comprehension	29	29	-0.26	3.13	1.68
	Total Test	62	81	-1.00	3.13	1.49
IV	Listening	21	21	0.77	4.22	1.83
	Reading	21	21	1.08	2.89	1.95
	Writing	14	22	0.34	2.65	1.35
	Speaking	10	21	-0.78	1.43	0.49
	Comprehension	28	28	0.77	4.22	1.93
	Total Test	66	85	-0.78	4.22	1.57
V	Listening	20	20	0.73	3.94	2.19
	Reading	21	21	1.20	3.66	2.38
	Writing	14	22	0.82	3.17	1.99
	Speaking	10	21	0.13	1.67	0.83
	Comprehension	28	28	0.73	3.94	2.16
	Total Test	65	84	0.13	3.94	2.00

Note. Total Test does not include the Comprehension composite.

DRAFT

The ELPA IRT statistics for each item are presented in Appendix E and include the following information for each operational and field test item by assessment level:

- Item ID
- Type (i.e., OP (operational) or FT (field test))
- N-Count (i.e., number of students)
- Rasch Difficulty
- SE of Rasch Difficulty (i.e., standard error of Rasch difficulty)
- INFIT (INFIT mean-square)
- OUTFIT (OUTFIT mean-square)
- Step Values (for CR items only)

Note that the INFIT and OUTFIT columns report the fit statistics provided by WINSTEPS. The details of these two statistics are provided in the next section.

7.2 Evidence of Model Fit

The assessment of item fit, which is the micro-level model-data fit evaluation, typically involves the comparison of examinees' observed and model predicted performance at the item level (Yen & Fitzpatrick, 2006). As mentioned above, WINSTEPS provides two mean-square fit statistics (OUTFIT and INFIT) which indicate the amount of distortion of the measurement system.

The estimation of these two fit statistics begins with the calculation of an expected (predicted) value $\mathrm{E}(\mathrm{X}_{vi})$ which is calculated as follows (adapted from Wright & Stone, 1979, p. 23):

$$\mathrm{E}(\mathrm{X}_{vi}|\theta) = \sum_{j=1}^{m_i} x_{vij} P(X_{vi} = x_{vij}|\theta) \qquad (7.1)$$

and a model variance which is calculated as:

$$V(X_{vi}|\theta) = \sum_{j=1}^{m_i} \Big(x_{vij} - \mathrm{E}(X_{vi}|\theta)\Big)^2 P(X_{vi} = x_{vij}|\theta)\Big]. \qquad (7.2)$$

For both Equation (7.1) and (7.2), $P(X_{vi} = x_{vij} \mid \theta)$ is equivalent to Equation (5.4) for dichotomous items and Equation (5.5) for polytomous items (see Section 5). Note that when there are only two response categories for an item, Equation (5.5) (i.e., the unidimensional PCM) simplifies to the unidimensional Rasch model.

Next, a response residual (Y_{vi}) for each person *v* for each item *i* is calculated. The residual indicates how far an observed response (X_{vi}) deviates from the model expectations ($\mathrm{E}(\mathrm{X}_{vi})$) and is computed as follows (adapted from Wright & Stone, 1979, p. 23):

$$Y_{vi} = X_{vi} - E(x_{vi}) \qquad (7.3)$$

The OUTFIT mean-square and INFIT mean-square are then computed as follows (adapted from Linacre, 2008, p.438):

$$\text{OUTFIT mean-square} = \frac{\sum_{n=1}^{N} Y_{vi}^2 / V(X_{vi})}{N}, \quad (7.4)$$

and

$$\text{INFIT mean-square} = \frac{\sum_{n=1}^{N} Y_{vi}^2}{\sum_{n=1}^{N} V(X_{vi})}. \qquad (7.5)$$

The OUTFIT mean-square for each item is in fact the unweighted average of the sum of squared standardized residuals[10], and thus is more sensitive to outliers or unexpected observations on items far from a person's ability level (i.e., a large logit difference between item difficulty and person ability). Under the unidimensional Rasch model or PCM, such outlying responses can often be guesses and careless mistakes, so these outlying characteristics of respondent behavior can make a "good" item look "bad." To diminish the effect of off-target responses, more attention is often directed towards evaluating the INFIT statistic (Bond & Fox, 2007).

INFIT is in fact a weighted standardized residual. The weights are applied to down-weight the outliers and put more emphasis on the "response string close to the item difficulty (or person ability)" (Linacre, 2008, p. 438). Specifically, the squared residuals are weighted by their model variance ($V(X_{vi})$) to reduce the influence of less informative, off-target responses.

In general, mean-squares are expected to be close to 1.0 when the data fit the model (Smith, 2004). According to Linacre (2008), values less than 1.0 indicate model over-fitting (i.e., predicts the data too well), while values greater than 1.0 indicate model under-fitting (i.e., cannot predict the data well). Linacre (2008) provides some general guidelines for evaluating mean-square fit statistics (see Table 7.2).

[10] The concept of standardized residual is clear when Equation (7.3) is inserted into Equation (7.4):
$\frac{Y_{vi}^2}{V(X_{vi})} = \frac{(X_{vi} - E(X_{vi}))^2}{V(X_{vi})} = \left(\frac{X_{vi} - E(X_{vi})}{\sqrt{V(X_{vi})}}\right)^2$

Table 7.2: General Criteria to Evaluate Mean-Square Fit Statistics

Mean-Square Value	Interpretation
> 2.0	Distorts or degrades the measurement system
1.5–2.0	Unproductive for construction of measurement, but not degrading
0.5–1.5	Productive for measurement
< 0.5	Less productive for measurement, but not degrading. May produce misleadingly good reliabilities and separations

Note. Adapted from Linacre (2008, p. 251).

Besides INFIT and OUTFIT mean-square values, INFIT and OUTFIT can also be reported in a Z-Standardized (ZSTD) format (i.e., INFIT ZSTD and OUTFIT ZSTD, respectively). The ZSTD values are *t*-tests of the hypotheses about whether the data fit the model (perfectly). According to Linacre (2008), mean-squares are not sample size dependent but ZSTD values are. Therefore, only the INFIT and OUTFIT mean-squares were examined and reported for the spring 2009 ELPA.

Note that items were only flagged if they distorted or degraded the measurement system (i.e., if they were > 2.0 logits). The OUTFIT and INFIT mean-squares are presented by assessment level in the IRT item statistics tables in Appendix E. As shown in Appendix E, the majority of the items were in the productive range (i.e., between 0.5 and 1.5) when judged on both INFIT and OUTFIT mean-square values. A few operational items (i.e., 1 at Level II, 1 at Level III, 2 at Level IV, and 2 at Level V) had OUTFIT mean-square values less than 0.50 or greater than 1.5. In addition, a few field test items (e.g., 2 at Level II, 1 at Level IV, and 1 at Level V) had OUTFIT mean-square values outside of the productive range of 0.5 to 1.5. Though unproductive for the construction of measurement, these items were deemed to not degrade the measurement system. One Level IV field test item, however, had an OUTFIT value greater than 2.0.

7.3 Item Information

The item information function, $I_i(\theta)$, represents the information provided by a specific item *i* across the range of θ. Items that provide the most information at the cuts should be considered for inclusion in form building. The information provided by item *i* at any point on the latent trait scale (theta) under the unidimensional Rasch model is defined mathematically as follows (Yen & Fitzpatrick, 2006, p. 127):

$$I_i(\theta) = Q_i(\theta)P_i(\theta), \qquad (7.6)$$

where $P_i(\theta)$ is the probability of a correct response on item *i* (see Equation (5.4) for the mathematical formula), and $Q_i(\theta)$ is the probability of an incorrect response or $1 - P_i(\theta)$.

The information provided by item *i* under the unidimensional PCM is computed as follows (de Ayala, 2009, p. 200):

$$I_i(\theta) = \sum_{k=1}^{m_i} k^2 p_{xi} - \left[\sum_{k=0}^{m_i} k p_{xi}\right]^2 \qquad (7.7)$$

Where p_{xi} (see Equation (5.5) for the mathematical formula) is the probability that a person with ability θ will score x on item i, where x has possible score of each k value which ranges from 0 to m_i.

Appendix F provides the following item information for the operational items at each of the four cut scores by grade:

- Item ID
- Rasch difficulty
- Item Information at Low Intermediate Cut
- Item Information at High Intermediate Cut
- Item Information at Proficient Cut
- Item Information at Advanced Proficient Cut

DRAFT

SECTION 8. VALIDITY

In the broadest sense, a test or instrument is valid if it measures what it was intended to measure (Nunnally, 1978). As stated in *Standards for Educational and Psychological Testing*, "validation involves accumulating evidence to provide a sound scientific basis for the proposed score interpretations" (AERA, APA, NCME, 1999, p. 9). Validity is thus a property of the interpretations or uses of a particular test and not of the test itself. Test validation is a continuous process that begins during test construction and continues for the duration of a test's use.

The ELPA was constructed to adhere to the validity-related standards set forth in *Standards for Educational and Psychological Testing*. The following sources of evidence can be used to evaluate the validity of the proposed interpretations and use of the ELPA scores (AERA et al., 1999):

- Test content—"relationship between a test's content and the construct it is intended to measure" (p. 11)
- Internal structure—"the degree to which the relationships among test items and test components conform to the construct on which the proposed test score interpretations are based" (p. 13)
- Relationships to other variables—"relationship of test scores to variables external to the test" (p. 13)

8.1 Validity Evidence of the ELPA Test Content

Evidence of validity based on test content is revealed by the extent to which the test material represents an adequate sampling of the skills and knowledge underlying the overall construct being measured. The ELPA was developed to measure the level of English language proficiency demonstrated by ELL students and to gauge whether students possess the requisite English language skills needed to participate effectively in Michigan classrooms. The specific English language skills to be assessed by the test (i.e., Listening, Reading, Writing, Speaking, and Comprehension) are identified in the related Michigan English Language Proficiency Standards.

Relation between ELPA and Michigan Content Standards

As discussed in *Standards for Educational and Psychological Testing* (AERA et al., 1999), expert judgments about the relationship between the test content and the intended assessment construct provide evidence of test content validity.

During 2009, OEAA assembled the Content Advisory Committee (CAC). This committee was composed of Michigan educators and members of the Michigan ELL community to serve as content experts. These experts were selected based on a combination of the following criteria established by OEAA:

DRAFT

- Information contained in the ELPA Educator Committee Application Form (e.g., gender, ethnicity, education, teaching experience, grade levels taught, location of school, and current role as an educator)
- Additional information provided by phone interviews with the applicants
- Recommendations made by experienced ELL administrators in the field
- Prior committee experience
- Knowledge gained from interactions with committee members both in the actual committee sessions as well as other ELL related meetings and conferences sponsored by OEAA or meetings that OEAA was invited to attend

OEAA relied heavily on the first three criteria listed above when selecting new committee members to participate in the committee. The last two criteria were more important for selecting veteran committee members.

The CAC session was held in August after the passages and items had been written. The CAC committee members were given two tasks. First, they were asked to review the content of the Michigan ELPA newly written passages and associated MC items, and newly written stand-alone MC and CR items. Second, they were asked to verify that each item was aligned with the appropriate individual Michigan English Language Proficiency (ELP) standard.

In order to accomplish the first task, CAC members used their field expertise as well as other resources to guide them in their content review of passages and items. These resources included:

- Michigan's K-8 Grade Level Content Expectations (GLCEs) and High School Content Expectations (HSCEs)
- Education Development Laboratory (EDL) Core Vocabularies in Reading, Mathematics, Science, and Social Studies
- Flesch-Kincaid Readability tests which were designed to indicate comprehension difficulty when reading a passage of contemporary academic English. (There are two tests involved: the Flesch Reading Ease, and the Flesch–Kincaid Grade Level Test.)
- ELPA Item Format Descriptions
- ELPA Style Guide
- ELPA Item Content Checklist designed by the OEAA development team
- Copies of the previous year's spring ELPA materials including all test booklets and Directions for Administration manuals (will be referred to as Test Administrator Manuals beginning spring 2010)

In order to accomplish the second task, the CAC members systematically examined the items against the Michigan English Language Proficiency (ELP) Standards. Members examined each assessment level and verified that the ELPA items accurately measured or aligned to a specific ELP standard. Only those items that were appropriately aligned to a specific ELP standard (see the following URL: http://www.michigan.gov/mde/0,1607,7-

140-22709_40192---,00.html for the "Michigan English Language Proficiency Standards K-12 Full Checklist") were considered as possible field test items.

Taken together, the rigorous scrutiny of the item development process and the alignment of items to particular learning standards provide evidence for the content validity of the spring 2009 ELPA.

Validity of the Presentation of the Listening Stimulus

Another important source of evidence for the content validity of the ELPA test concerns the extent to which the presentation of the stimulus to the test taker is appropriate for the construct being measured. This issue is particularly important for the Listening domain because items in the Listening domain are not exclusively presented in an oral format. In other words, the text of the question that the administrator reads to the students is also printed out in the test booklet.

The items for each Listening passage are printed in the test booklet in order to preserve the construct being measured. The purpose of each Listening item is to assess whether or not a student can accurately respond to the item based on what he or she hears when the prompt is read, not whether they can understand the individual item options. Stated differently, a barrier preventing an accurate response is removed when the answer options and item stems appear in the test booklet.

The Listening prompts, however, which are the true basis of the assessment, are not printed in the test booklet. If the Listening passages appeared in print, the domain being assessed would be transformed into Reading, as students would simply read while hearing an aural version of the passage in the background. Instead, a line-art graphic is included in the booklet to enhance the overall comprehension of the main topic of the passage without providing any clues for the Listening items. Such a design increases the chance that the students will respond to the item based on what they hear in the prompt. This is because they won't be distracted by the additional challenge of having to listen to the item stem and its options, and keeping all of these in short-term memory while simultaneously trying to recall information of the prompt. Thus, this design preserves the Listening construct being measured.

8.2 Validity Evidence of the Internal Structure of the ELPA

In order to investigate the ELPA's internal structure, Pearson product-moment correlations (r) were calculated among each of the five ELPA domains (i.e., Listening, Reading, Writing, Speaking, and Comprehension) and between the domains and the ELPA total score. The purpose of calculating these correlations was to examine the degree to which the domains were related to each other and the degree to which each domain was related to the total score. The direction and magnitude of these correlations provide validity evidence for the internal structure of the test. In particular, strong positive correlations provide more evidence for the validity of the internal structure than weaker or negative correlations. In order to account for measurement error, the correction for attenuation ($\bar{r}$; see Nunnally, 1978) was also computed for each

correlation. The correction for attenuation is computed as follows (Nunnally, 1978, p. 237):

$$\bar{r}_{12} = \frac{r_{12}}{\sqrt{r_{11} r_{22}}}, \qquad (8.1)$$

where r_{12} represents the uncorrected correlation coefficient and r_{11} and r_{22} are the reliability coefficients (reported as Cronbach's alpha reliability coefficients in this report). Table 8.1 reports both the Pearson product-moment correlation coefficients (below the diagonal) and the disattenuated correlation coefficients (above the diagonal), as well as the reliability coefficients for each domain or total test (on the diagonal).

The correlations between Listening and the ELPA total score were relatively high across grades. These correlations ranged from 0.76 for grade 1 to 0.85 for grades 9, 10, and 11 ($\bar{r}s$ ranged from 0.94 for grade 1 to 1.04 for grade 6[11]). Similar correlations were observed between Reading and the ELPA total score; these correlations ranged from 0.77 at grade 1 to 0.86 at grades 10 and 12 ($\bar{r}s$ ranged from 0.96 for grade 1 to 1.00 for grades 2, 3, and 6). The correlations between Writing and the ELPA total score ranged from 0.82 at kindergarten to 0.91 at grades 9 and 10 ($\bar{r}s$ ranged from 0.97 for grades K, 1, and 2 to 1.04 for grades 9, 10, and 12). For Speaking, the correlations with the ELPA total score ranged from 0.75 at grade 1 to 0.86 at grade 9 ($\bar{r}s$ ranged from 0.84 for grade 12 to 0.92 for grade 9). The correlations between Comprehension and the ELPA total score ranged from 0.84 at kindergarten to 0.91 at grade 10 ($\bar{r}s$ ranged from 1.02 for grades 10, 11, and 12 to 1.06 for grades 1 and 6).

In looking at the inter-correlations among the domains themselves[12], there is some evidence of a pattern across grade levels. The highest correlations were observed between Reading and Writing for 11 of the 13 grades (i.e., excluding grade 4 and 11). These correlations ranged from 0.60 at grade 3 to 0.72 at grade 10 ($\bar{r}s$ ranged from 0.79 for grade 4 to 0.90 for kindergarten). For grades 4 and 11, the highest correlations were observed between Listening and Reading (rs = 0.62 and 0.69, respectively; $\bar{r}s$ = 0.84 and 0.89, respectively). For grade 5, the highest correlation (r = 0.64) was found between Listening and Reading, as well as between Reading and Writing ($\bar{r}s$ = 0.84 and 0.82, respectively). For all but one grade (i.e., kindergarten), the lowest correlation was between Speaking and Reading (rs ranged from 0.35 at grade 1 to 0.54 at grades 9 and 10; $\bar{r}s$ ranged from 0.45 for grade 1 to 0.63 for grades 8, 9, and 10).

[11] Note that disattenuated correlations greater than 1.0 are theoretically impossible because it implies that valid variance would exceed obtained variance (Muchinsky, 1996). It is possible, however, to obtain a coefficient greater than 1.0 using the Equation (8.1) when the correlation between true scores is high and reliability is low (Nunnally, 1978).

[12] Comprehension was excluded from this investigation, as it is composed of selected items from the Listening and Reading domains. High correlations between Comprehension and each of those two domains are thus expected.

Because the Comprehension domain is a composite of Listening and Reading items, the correlations between Comprehension and Listening, and between Comprehension and Reading are not informative regarding internal structure. However, in looking at the correlations between Comprehension and Writing (*r*s ranged from 0.60 at kindergarten to 0.76 at grades 9 and 10; $\bar{r}s$ ranged from 0.79 for kindergarten to 0.93 for grade 9) and Comprehension and Speaking (*r*s ranged from 0.46 at grade 3 to 0.63 at grade 9; $\bar{r}s$ ranged from 0.58 for grades 1 and 3 to 0.73 for grade 9), we see correlations similar in magnitude to those that were observed among the other domains.

Point-biserial correlations (for MC items) and item-total correlations (for CR items) also provide validity evidence for the internal structure of an assessment. Because an English language proficiency test should be able to detect performance and proficiency differences among students, it is important to examine how well each item functions consistently with the overall intent of the test. As first introduced in Section 4.1, point-biserial correlation coefficients and item-total correlation coefficients reveal how well an item discriminates between high- and low-achieving students. These statistics are presented in Appendices C.1 to C.10.

As previously mentioned, the unidimensional Rasch model and the unidimensional PCM were used to fit the data. The fit statistics reported in this technical report (i.e., INFIT and OUTFIT mean-squares) help to determine whether the data fit the model. According to Linacre (2008), high INFIT mean-squares are a bigger threat to validity than high OUTFIT mean-squares, as high INFIT mean-squares "indicate that the items are mis-performing for the people on whom the items are targeted" (p. 439). Based on the results shown in Appendix E, no such validity threat was found with the spring 2009 ELPA data.

DRAFT

Table 8.1: Inter-correlations among ELPA Domains and Total Test by Grade

		Correlation Coefficient					
Grade	**Domain**	**Listening**	**Reading**	**Writing**	**Speaking**	**Comprehension**	**Total Test**
K	Listening	(0.70)	0.74	0.70	0.65	1.30	0.98
	Reading	0.53	(0.73)	0.90	0.54	1.00	0.99
	Writing	0.52	0.68	(0.79)	0.51	0.79	0.97
	Speaking	0.49	0.42	0.41	(0.82)	0.65	0.90
	Comprehension	0.93	0.73	0.60	0.50	(0.73)	1.04
	Total Test	0.78	0.80	0.82	0.77	0.84	(0.90)
1	Listening	(0.73)	0.65	0.65	0.59	1.13	0.94
	Reading	0.47	(0.72)	0.82	0.45	1.09	0.96
	Writing	0.50	0.63	(0.82)	0.52	0.80	0.97
	Speaking	0.46	0.35	0.43	(0.84)	0.58	0.86
	Comprehension	0.86	0.82	0.64	0.47	(0.79)	1.06
	Total Test	0.76	0.77	0.83	0.75	0.89	(0.90)
2	Listening	(0.72)	0.75	0.67	0.62	1.15	0.96
	Reading	0.55	(0.75)	0.85	0.59	1.10	1.00
	Writing	0.51	0.66	(0.80)	0.59	0.81	0.97
	Speaking	0.48	0.46	0.48	(0.82)	0.65	0.90
	Comprehension	0.88	0.86	0.65	0.53	(0.81)	1.04
	Total Test	0.78	0.83	0.83	0.78	0.89	(0.91)
3	Listening	(0.71)	0.83	0.77	0.61	1.23	1.01
	Reading	0.58	(0.69)	0.82	0.49	1.12	1.00
	Writing	0.57	0.60	(0.77)	0.61	0.82	1.00
	Speaking	0.47	0.37	0.49	(0.84)	0.58	0.87
	Comprehension	0.90	0.81	0.63	0.46	(0.76)	1.05
	Total Test	0.81	0.79	0.83	0.76	0.87	(0.90)
4	Listening	(0.71)	0.84	0.78	0.67	1.20	1.02
	Reading	0.62	(0.76)	0.79	0.54	1.10	0.99
	Writing	0.58	0.61	(0.78)	0.67	0.82	0.99
	Speaking	0.52	0.43	0.54	(0.84)	0.63	0.89
	Comprehension	0.90	0.85	0.64	0.51	(0.79)	1.05
	Total Test	0.82	0.82	0.83	0.78	0.89	(0.91)
5	Listening	(0.74)	0.84	0.82	0.69	1.16	1.01
	Reading	0.64	(0.79)	0.82	0.57	1.08	0.99
	Writing	0.62	0.64	(0.77)	0.71	0.85	1.01
	Speaking	0.55	0.47	0.58	(0.86)	0.66	0.89
	Comprehension	0.90	0.86	0.67	0.55	(0.81)	1.04
	Total Test	0.83	0.84	0.85	0.79	0.90	(0.92)
6	Listening	(0.64)	0.83	0.80	0.67	1.27	1.04
	Reading	0.57	(0.73)	0.85	0.56	1.13	1.00
	Writing	0.56	0.63	(0.76)	0.68	0.84	1.03
	Speaking	0.50	0.45	0.55	(0.87)	0.63	0.90
	Comprehension	0.86	0.82	0.62	0.50	(0.72)	1.06
	Total Test	0.79	0.81	0.85	0.80	0.85	(0.90)

DRAFT

Table 8.1: Inter-correlations among ELPA Domains and Total Test by Grade (Continued)

		Correlation Coefficient					
Grade	**Domain**	**Listening**	**Reading**	**Writing**	**Speaking**	**Comprehension**	**Total Test**
7	Listening	(0.70)	0.84	0.83	0.68	1.21	1.01
	Reading	0.61	(0.76)	0.83	0.59	1.12	0.98
	Writing	0.61	0.64	(0.78)	0.70	0.86	1.02
	Speaking	0.53	0.48	0.58	(0.88)	0.64	0.90
	Comprehension	0.88	0.85	0.66	0.52	(0.76)	1.03
	Total Test	0.81	0.82	0.86	0.81	0.86	(0.92)
8	Listening	(0.72)	0.85	0.83	0.72	1.17	1.02
	Reading	0.64	(0.78)	0.87	0.63	1.10	0.99
	Writing	0.62	0.68	(0.78)	0.75	0.87	1.03
	Speaking	0.57	0.52	0.62	(0.88)	0.68	0.91
	Comprehension	0.88	0.86	0.68	0.56	(0.78)	1.04
	Total Test	0.83	0.84	0.87	0.82	0.88	(0.92)
9	Listening	(0.73)	0.88	0.91	0.76	1.16	1.03
	Reading	0.67	(0.80)	0.88	0.63	1.09	0.97
	Writing	0.70	0.71	(0.81)	0.82	0.93	1.04
	Speaking	0.62	0.54	0.71	(0.92)	0.73	0.92
	Comprehension	0.90	0.88	0.76	0.63	(0.82)	1.03
	Total Test	0.85	0.84	0.91	0.86	0.90	(0.94)
10	Listening	(0.75)	0.89	0.90	0.71	1.15	1.01
	Reading	0.70	(0.82)	0.88	0.63	1.07	0.98
	Writing	0.70	0.72	(0.81)	0.80	0.92	1.04
	Speaking	0.59	0.54	0.69	(0.91)	0.70	0.91
	Comprehension	0.91	0.89	0.76	0.61	(0.84)	1.02
	Total Test	0.85	0.86	0.91	0.84	0.91	(0.94)
11	Listening	(0.74)	0.89	0.84	0.64	1.14	1.02
	Reading	0.69	(0.82)	0.81	0.54	1.08	0.97
	Writing	0.65	0.66	(0.80)	0.70	0.86	1.01
	Speaking	0.52	0.46	0.59	(0.89)	0.61	0.85
	Comprehension	0.89	0.89	0.70	0.52	(0.83)	1.02
	Total Test	0.85	0.85	0.87	0.77	0.90	(0.93)
12	Listening	(0.74)	0.89	0.89	0.60	1.14	1.01
	Reading	0.69	(0.82)	0.87	0.55	1.07	0.98
	Writing	0.68	0.70	(0.79)	0.70	0.91	1.04
	Speaking	0.49	0.47	0.59	(0.89)	0.59	0.84
	Comprehension	0.90	0.89	0.74	0.51	(0.84)	1.02
	Total Test	0.84	0.86	0.89	0.76	0.90	(0.93)

Note. 1. Comprehension consists of items selected from the Listening and Reading Domains.
2. Total Test does not include the Comprehension composite.
3. The disattenuated correlation coefficients appear above the diagonal, and the Pearson product-moment correlations appear below the diagonal.
4. The reliability coefficients for each test appear on the diagonal.

DRAFT

8.3. Validity Evidence of the External Structure of the ELPA

An examination of the relationship of the ELPA total score to other external variables provides another source of validity evidence. As indicated in *Standards for Educational and Psychological Testing*, "evidence based on relationships with other variables addresses questions about the degree to which these relationships are consistent with the construct underlying the proposed test interpretations" (AERA et al., 1999, p. 13). Correlations between theoretically similar constructs should generally be "high" while correlations between theoretically dissimilar constructs should be relatively "low." Analyses that investigate the relation between theoretically similar constructs provide evidence of convergent validity while analyses of supposedly different constructs provide evidence of discriminant validity.

In order to ascertain the convergent and discriminant validity of the ELPA vis-à-vis an external criterion, grades 3–8 ELLs' performance on the spring 2009 ELPA were compared to their performance on the fall 2008 Michigan Educational Assessment Program (MEAP). Similarly, grade 11 ELLs' performance on the spring 2009 ELPA were compared to their performance on the spring 2009 Michigan Merit Examination (MME).

Relation between the ELPA and the MEAP Subject Tests

In Michigan, all ELL students, except those who have been in U.S. schools for 12 months or less, are required to take both the MEAP and the ELPA. The MEAP is a statewide assessment composed of the following subject tests: Mathematics, English Language Arts (ELA; includes reading and writing), Science, and Social Studies. The MEAP assessment is an important measure of student learning and school success for the state of Michigan.

By contrast, the purpose of the ELPA is to measure ELLs' English language ability. Scores on the ELPA may be used to aid in placement decisions for students for whom English is a non-native language. It is expected that ELLs who score proficient on the ELPA will be able to function successfully in mainstream education programs and have the same chance as their English proficient counterparts to pass the required state examinations. Given that the ELPA and the MEAP ELA test assess similar constructs, students who perform well on the ELPA should generally perform well on the MEAP ELA test. The existence of such a relationship provides evidence of convergent validity.

Overall performance on the ELPA will also likely be related to students' performance on the other MEAP tests (i.e., Math, Science, and Social Studies) as students who are more proficient in English language are likely to be more successful in the classroom than their less proficient peers. However, given that the items assessing these subject areas assess different (but related) types of skills, the relationship may not be as strong as that between the ELPA and the MEAP ELA test. The existence of such a pattern provides evidence of discriminant validity.

Table 8.2 provides the correlations between the ELPA total scale score and the individual MEAP scale scores across subject tests (e.g., ELA, Math, Science, and Social Studies) for grades 3 through 8. These analyses were conducted after removing all duplicate records from the final data file. Moreover, only cases that had valid data across all MEAP subject tests were included in the analyses. The Pearson product-moment correlation coefficients appear below the diagonal in Table 8.2, the disattenuated correlation coefficients appear above the diagonal, and the reliability coefficients appear on the diagonal.

Table 8.2 indicates that the ELPA total scale score had the strongest linear relation with the MEAP ELA scale score (i.e., rs ranged from 0.69 to 0.71; $\bar{r}$s ranged from 0.76 to 0.79). These correlations were positive and moderately high across all grade levels. Thus, higher scores on the ELPA were associated with higher scores on the MEAP ELA test for each of the six grade levels. The ELPA was also linearly related to the other MEAP subject tests (i.e., Math, Science, and Social Studies), though these correlations were not as strong as those between the ELPA and the MEAP ELA test. Specifically, correlations between the ELPA and the other MEAP subject tests were found to be as follows: rs ranged from 0.46 to 0.67; $\bar{r}$s ranged from 0.51 to 0.75.

DRAFT

Table 8.2: Inter-correlations among ELPA Total Scale Score and MEAP Subject Test Scale Scores by Grade

Grade	Domain	Correlation Coefficient				
		ELPA	MEAP ELA	MEAP Math	MEAP Science	MEAP Soc. Studies
3	ELPA	(0.90)	0.79	0.69	NA	NA
	MEAP ELA	0.70	(0.88)	0.72	NA	NA
N = 5225	MEAP Math	0.62	0.64	(0.89)	NA	NA
4	ELPA	(0.91)	0.76	0.68	NA	NA
	MEAP ELA	0.69	(0.90)	0.69	NA	NA
N = 4498	MEAP Math	0.61	0.61	(0.88)	NA	NA
5	ELPA	(0.92)	0.78	0.62	0.75	NA
	MEAP ELA	0.71	(0.90)	0.62	0.77	NA
N = 3771	MEAP Math	0.56	0.56	(0.90)	0.68	NA
	MEAP Science	0.67	0.68	0.60	(0.87)	NA
6	ELPA	(0.90)	0.77	0.60	NA	0.73
	MEAP ELA	0.69	(0.89)	0.60	NA	0.77
N = 3291	MEAP Math	0.53	0.53	(0.88)	NA	0.65
	MEAP Soc. Studies	0.65	0.68	0.57	NA	(0.87)
7	ELPA	(0.92)	0.79	0.64	NA	NA
	MEAP ELA	0.71	(0.88)	0.68	NA	NA
N = 3113	MEAP Math	0.58	0.60	(0.88)	NA	NA
	ELPA	(0.92)	0.76	0.51	0.69	NA
8	MEAP ELA	0.69	(0.90)	0.57	0.78	NA
	MEAP Math	0.46	0.50	(0.87)	0.66	NA
N = 3214	MEAP Science	0.63	0.70	0.58	(0.90)	NA

Note. 1. NA indicates that a particular subject test was not administered at that grade level.
2. The disattenuated correlation coefficients appear above the diagonal, and the Pearson product-moment correlations appear below the diagonal.
3. The reliability coefficients for each test appear on the diagonal.

Relation between the ELPA and the MME Subject Tests

The MME is a statewide assessment for students in grade 11. In addition to the two ACT assessments: the ACT Plus Writing exam and WorkKeys job skills assessments, the MME also includes Michigan-developed content to assess standards not covered by the ACT assessments for mathematics, science, and social studies. Scores for the following subject tests are computed: ELA, Math, Science, and Social Studies. Because the ELPA and the MME ELA test assess similar constructs, they should be more strongly related to each other than to other subject tests (i.e., Math, Science, and Social Studies).

Table 8.3 provides the correlations between the ELPA total scale score and the individual MME subject test scale scores for ELA, Math, Science, and Social Studies. Similar to Table 8.2, the Pearson product-moment correlation coefficients appear below the diagonal in Table 8.3, the disattenuated correlation coefficients appear above the diagonal, and the reliability coefficients appear on the diagonal.

DRAFT

Again, the ELPA total scale score was found to have the strongest linear relation with the MME ELA scale score (r = 0.73; $\bar{r}$ = 0.78), providing evidence for convergent validity. The ELPA scale score was only moderately (linearly) related to the other MME subject test scale scores (i.e., Math, Science, and Social Studies), providing evidence of discriminant validity (rs ranged from 0.54 to 0.58; $\bar{r}$s ranged from 0.60 to 0.63).

Table 8.3: Inter-correlations among ELPA Total Scale Score and MME Subject Test Scale Scores for Grade 11

		Correlation Coefficient				
Grade	**Domain**	**ELPA**	**MME ELA**	**MME Math**	**MME Science**	**MME Soc. Studies**
11	ELPA	(0.93)	0.78	0.63	0.60	0.61
	MME ELA	0.73	(0.93)	0.66	0.62	0.63
	MME Math	0.58	0.61	(0.92)	0.58	0.61
N = 1944	MME Science	0.54	0.56	0.52	(0.87)	0.64
	MME Soc. Studies	0.54	0.56	0.54	0.55	(0.85)

Note. 1. The disattenuated correlation coefficients appear above the diagonal, and the Pearson product-moment correlations appear below the diagonal.
2. The reliability coefficients for each test appear on the diagonal.

In general, the correlations between the theoretically similar constructs (e.g., the ELPA and the MEAP ELA or MME ELA tests) were stronger than the correlations between theoretically dissimilar constructs (e.g., the ELPA and the MEAP or MME Mathematics, Science, and Social Studies tests). Such a pattern provides evidence for convergent and discriminant validity of the external structure of the spring 2009 ELPA.

Distribution of Student Classification at Performance Levels

In addition to correlations, the distribution of student classification across performance levels on two related tests is another important piece of evidence for validity. Thus, the classification distribution for each of the MEAP (or MME) subject tests given each performance level of the spring 2009 ELPA was examined at each appropriate grade (i.e., grades 3-8 for MEAP and grade 11 for MME). Related results can be found in Appendices G.1 to G.7. The first table in Appendix G.1 is presented below as Table 8.4. The contents of this table are discussed below.

DRAFT

Table 8.4: Percentage of Grade 3 ELLs Scoring at Each MEAP ELA Performance Level by Each ELPA Performance Level

ELA			ELPA Performance Levels				
			1 Advanced Proficient (N=491)	2 Proficient (N=1260)	3 High Intermediate (N=3191)	4 Low Intermediate (N=277)	5 Basic (N=6)
MEAP Performance Levels	1 Exceeded Standards	%	28.90	5.20	0.50	0.00	0.00
	2 Met Standards	%	69.70	88.90	58.00	15.50	0.00
	3 Basic	%	1.40	5.90	40.20	71.50	83.30
	4 Apprentice	%	0.00	0.10	1.30	13.00	16.70

As demonstrated in Table 8.4, the MEAP test has four performance levels: Exceeded Standards, Met Standards, Basic, and Apprentice, while the ELPA test has five performance levels: Advanced Proficient, Proficient, High Intermediate, Low Intermediate, and Basic. Note that the percentage values in each column add up to 100%. Therefore, a value of 28.90 means that among the grade 3 students who were categorized as Advanced Proficient based on the spring 2009 ELPA (N=491), 28.90% of them were found to be in the Exceeded Standards category based on their performance on the fall 2008 MEAP ELA. Other values in this table (and in other tables in Appendices G.1 to G.7) can be interpreted in the same fashion.

Summary

The positive correlations found between the ELPA and the MEAP ELA as well as between the ELPA and the MME ELA indicate that higher scores on the ELPA were associated with higher scores on either the MEAP ELA or the MME ELA tests. It is important to note, however, that the ELA test and the ELPA tests were developed to serve different populations of students. In addition, the Listening and Speaking domains were not tested as part of either the MEAP ELA or MME ELA tests. Thus, it is not surprising that we see only a moderate to high positive correlation between the two tests.

The analysis of the distribution of student classification across performance levels provides evidence to support the validity of the ELPA cut scores in that about 64% to 94% of grades 3 through 8 students classified as Proficient on the ELPA were also classified as Met or Exceeded Standards on the MEAP ELA test. For grade 11 students, about 94% of those classified as Proficient on the ELPA were also classified as Partially Proficient or Proficient on the MME ELA test.

The majority of students in grades 3 through 8 (from 92-99%) who were classified as Advanced Proficient on the ELPA were also classified as Met or Exceeded Standards on the MEAP ELA test. About 59% of those grade 11 students who were classified as Advanced Proficient on the ELPA were classified as Proficient on the MME ELA test.

DRAFT

SECTION 9. SUMMARY OF OPERATIONAL TEST RESULTS

This section presents both the raw score and scale score summaries by grade for the spring 2009 ELPA total test and each domain subtest. Table 9.1 presents the raw score summary by grade and domain (or the total test). Table 9.2a presents the scale score summary by grade and domain, and Table 9.2b presents the scale score summary of the total test by grade. These tables include the sample size (N-Count), mean, median, interquartile range (IQR), and standard deviation (SD) of either the raw score or the scale score.

Table 9.1: Raw Score Summary by Grade and Domain (or the Total Test)

Grade	Domain (or the Total Test)	N-Count	Mean	Median	IQR	SD
K	Listening	9141	9.37	9	5	3.26
	Reading	9141	9.12	9	4	3.26
	Writing	9141	6.91	6	6	4.01
	Speaking	9141	10.00	10	7	4.64
	Comprehension	9141	11.91	12	6	3.90
	Total Test	9141	35.39	35	17	12.01
1	Listening	8292	12.28	13	5	3.79
	Reading	8292	10.22	10	5	3.78
	Writing	8292	11.00	11	7	4.48
	Speaking	8292	12.82	13	7	4.73
	Comprehension	8292	18.28	18	8	5.69
	Total Test	8292	46.32	47	18	13.04
2	Listening	7711	14.90	16	4	3.37
	Reading	7711	13.77	14	4	3.59
	Writing	7711	14.91	16	5	3.86
	Speaking	7711	15.10	16	5	4.17
	Comprehension	7711	23.61	25	7	5.39
	Total Test	7711	58.69	61	14	12.00
3	Listening	5786	11.25	12	5	3.61
	Reading	5786	9.49	9	5	3.86
	Writing	5786	10.51	11	5	3.75
	Speaking	5786	14.65	15	6	4.41
	Comprehension	5786	15.12	15	7	4.99
	Total Test	5786	45.90	47	17	12.40
4	Listening	4978	12.80	13	4	3.42
	Reading	4978	11.26	11	6	4.25
	Writing	4978	12.14	13	5	3.67
	Speaking	4978	15.58	17	6	4.23
	Comprehension	4978	17.49	18	7	5.11
	Total Test	4978	51.78	54	16	12.64
5	Listening	4317	13.67	14	4	3.45
	Reading	4317	12.70	13	7	4.46
	Writing	4317	13.22	14	4	3.65
	Speaking	4317	16.02	17	5	4.25
	Comprehension	4317	19.06	20	7	5.22
	Total Test	4317	55.62	58	16	13.07

DRAFT

Table 9.1: Raw Score Summary by Grade and Domain (or the Total Test) (Continued)

Grade	Domain (or the Total Test)	N-Count	Mean	Median	IQR	SD
6	Listening	3709	11.61	12	5	3.50
	Reading	3709	11.03	11	6	4.02
	Writing	3709	13.53	14	6	4.23
	Speaking	3709	15.96	17	6	4.83
	Comprehension	3709	14.93	15	6	4.57
	Total Test	3709	52.13	54	18	13.42
7	Listening	3567	12.53	13	5	3.72
	Reading	3567	12.04	12	6	4.23
	Writing	3567	14.42	15	6	4.32
	Speaking	3567	16.39	18	6	4.91
	Comprehension	3567	16.14	16	7	4.88
	Total Test	3567	55.38	58	18	14.17
8	Listening	3671	13.14	14	5	3.73
	Reading	3671	13.04	14	6	4.25
	Writing	3671	15.35	16	5	4.29
	Speaking	3671	16.85	18	5	4.83
	Comprehension	3671	17.15	18	7	4.94
	Total Test	3671	58.39	62	17	14.35
9	Listening	3544	11.70	12	5	3.74
	Reading	3544	11.53	12	7	4.47
	Writing	3544	13.08	14	6.75	4.65
	Speaking	3544	16.73	19	6	5.68
	Comprehension	3544	16.47	17	8	5.41
	Total Test	3544	53.03	57	20	15.99
10	Listening	3150	12.30	13	5	3.85
	Reading	3150	12.19	13	7	4.59
	Writing	3150	13.67	15	6	4.65
	Speaking	3150	17.12	19	5	5.25
	Comprehension	3150	17.34	18	9	5.59
	Total Test	3150	55.28	59	20	15.79
11	Listening	2265	12.97	13	6	3.71
	Reading	2265	12.89	13	7	4.51
	Writing	2265	14.54	16	6	4.34
	Speaking	2265	17.92	20	4	4.36
	Comprehension	2265	18.32	19	8	5.39
	Total Test	2265	58.32	61	18	14.11
12	Listening	1949	12.74	13	5.5	3.73
	Reading	1949	12.93	14	7.5	4.58
	Writing	1949	14.48	15	6	4.28
	Speaking	1949	17.98	20	4	4.27
	Comprehension	1949	18.18	19	8	5.48
	Total Test	1949	58.13	61	19	14.13

Note.
1. Omits, blanks, and invalid responses were scored as "0".
2. Total Test does not include Comprehension, as Comprehension is composed of selected items from Listening and Reading.

DRAFT

Table 9.2a: Scale Score Summary by Grade and Domain

Grade	Domain	N-Count	Mean	Median	IQR	SD
K	Listening	9141	53.82	53	6	4.42
	Reading	9141	52.17	52	5	4.95
	Writing	9141	48.35	48	7	5.28
	Speaking	9141	52.95	53	8	7.24
	Comprehension	9141	52.86	53	5	3.62
1	Listening	8292	58.44	59	5	4.20
	Reading	8292	55.80	56	4	4.03
	Writing	8292	53.24	53	7	5.01
	Speaking	8292	59.00	58	7	6.63
	Comprehension	8292	57.05	57	4	3.30
2	Listening	7711	61.47	62	4	4.30
	Reading	7711	59.56	59	5	4.31
	Writing	7711	58.06	58	7	5.27
	Speaking	7711	62.28	62	7	6.67
	Comprehension	7711	60.32	61	5	3.71
3	Listening	5786	64.14	64	5	3.98
	Reading	5786	61.64	61	4	3.42
	Writing	5786	59.16	60	6	4.90
	Speaking	5786	66.08	65	8	6.73
	Comprehension	5786	62.41	62	4	3.20
4	Listening	4978	65.87	65	4	4.10
	Reading	4978	63.16	63	4	3.74
	Writing	4978	61.30	62	7	4.99
	Speaking	4978	67.71	68	10	6.90
	Comprehension	4978	63.92	64	4	3.43
5	Listening	4317	67.00	67	5	4.50
	Reading	4317	64.48	64	6	4.22
	Writing	4317	62.89	63	6	5.24
	Speaking	4317	68.49	68	9	7.11
	Comprehension	4317	65.03	65	4	3.77

DRAFT

Table 9.2a: Scale Score Summary by Grade and Domain (Continued)

Grade	Domain	N-Count	Mean	Median	IQR	SD
6	Listening	3709	65.01	65	4	3.33
	Reading	3709	62.75	63	5	3.53
	Writing	3709	60.13	60	5	4.71
	Speaking	3709	67.06	66	10	7.14
	Comprehension	3709	63.40	63	3	2.92
7	Listening	3567	65.90	66	4	3.67
	Reading	3567	63.69	63	5	3.88
	Writing	3567	61.26	61	7	5.16
	Speaking	3567	67.86	68	10	7.36
	Comprehension	3567	64.19	64	4	3.24
8	Listening	3671	66.54	67	5	3.75
	Reading	3671	64.62	65	5	4.05
	Writing	3671	62.54	62	6	5.56
	Speaking	3671	68.72	68	9	7.46
	Comprehension	3671	64.89	65	4	3.37
9	Listening	3544	67.15	67	4	4.01
	Reading	3544	64.99	65	6	4.34
	Writing	3544	62.33	63	6.75	5.33
	Speaking	3544	69.53	70	13	8.12
	Comprehension	3544	65.33	65	5	3.82
10	Listening	3150	67.83	68	5	4.23
	Reading	3150	65.66	66	6	4.61
	Writing	3150	63.05	64	6	5.48
	Speaking	3150	70.16	70	12	7.62
	Comprehension	3150	65.98	66	6	4.06
11	Listening	2265	68.62	68	7	4.28
	Reading	2265	66.39	66	6	4.58
	Writing	2265	64.16	65	7	5.37
	Speaking	2265	71.24	73	11	6.79
	Comprehension	2265	66.74	67	6	4.11
12	Listening	1949	68.33	68	6	4.16
	Reading	1949	66.45	67	6.5	4.68
	Writing	1949	64.09	64	7	5.27
	Speaking	1949	71.38	73	11	6.71
	Comprehension	1949	66.62	67	5	4.09

Note.
1. Omits, blanks, and invalid responses were scored as “0”.
2. Total Test does not include Comprehension, as Comprehension is composed of selected items from Listening and Reading.

DRAFT

Table 9.2b: Scale Score Summary of the Total ELPA Test by Grade

Grade	Test	N-Count	Mean	Median	IQR	SD
K	ELPA	9141	522.85	522	38	30.21
1	ELPA	8292	558.94	559	33	26.73
2	ELPA	7711	586.28	587	34	28.94
3	ELPA	5786	606.20	607	34	26.70
4	ELPA	4978	619.30	622	34	28.57
5	ELPA	4317	628.87	631	38	31.47
6	ELPA	3709	614.00	616	33	26.18
7	ELPA	3567	620.95	623	36	28.79
8	ELPA	3671	627.77	631	37	30.45
9	ELPA	3544	630.91	636	41	33.62
10	ELPA	3150	636.16	640	44	34.04
11	ELPA	2265	643.12	645	41	32.19
12	ELPA	1949	642.81	645	43	32.24

Note. The mean for each grade should generally increase from one grade to the next but due to artifacts of the population whereby some grades may have a greater percentage of higher scoring students than the next higher grade, the mean for the lower grade can be higher than the next higher grade(s).

Table 9.3 presents the percentage of students at each performance level by grade. Note that the performance level categorization is based on the total ELPA test score only.

Table 9.3: Percent of Students in Each Performance Level by Grade

Grade	Test	N-Count	Performance Levels				
			5	4	3	2	1
K	ELPA	9141	15.20	26.70	17.90	23.00	17.20
1	ELPA	8292	5.70	16.70	51.60	16.60	9.40
2	ELPA	7711	2.70	9.40	46.90	27.90	13.10
3	ELPA	5786	0.80	8.20	59.20	22.80	9.00
4	ELPA	4978	0.60	7.00	48.80	26.50	17.00
5	ELPA	4317	1.40	6.50	43.90	28.90	19.30
6	ELPA	3709	2.30	20.50	57.40	15.80	4.00
7	ELPA	3567	3.20	22.60	49.10	18.00	7.10
8	ELPA	3671	3.30	19.70	51.70	17.40	7.90
9	ELPA	3544	11.00	19.90	50.00	13.50	5.60
10	ELPA	3150	10.40	18.10	48.30	16.80	6.40
11	ELPA	2265	5.90	26.40	40.80	19.50	7.40
12	ELPA	1949	7.10	32.50	41.40	13.30	5.60

Note.
Performance levels are defined as follows:
1—advanced proficient
2—proficient
3—high intermediate
4—low intermediate
5—basic

Table 9.4 presents the percentage of students meeting the achievement target in each domain by grade. Note that a practical decision was made by OEAA to adopt the exact same achievement targets reported in the 2008 ELPA technical report for the spring 2009 ELPA.

DRAFT

Table 9.4: Percent of Students Meeting Achievement Target by Grade and Domain

Grade	Domain	N-count	% of Students Meeting Achievement Target for Each Domain
K	Listening	9141	29.00
	Reading	9141	43.90
	Writing	9141	48.50
	Speaking	9141	25.80
	Comprehension	9141	27.20
1	Listening	8292	41.20
	Reading	8292	20.70
	Writing	8292	33.30
	Speaking	8292	25.40
	Comprehension	8292	29.70
2	Listening	7711	51.00
	Reading	7711	36.50
	Writing	7711	41.50
	Speaking	7711	34.80
	Comprehension	7711	34.80
3	Listening	5786	29.50
	Reading	5786	29.10
	Writing	5786	33.10
	Speaking	5786	41.10
	Comprehension	5786	33.00
4	Listening	4978	47.20
	Reading	4978	31.70
	Writing	4978	40.70
	Speaking	4978	41.50
	Comprehension	4978	45.10
5	Listening	4317	47.10
	Reading	4317	37.20
	Writing	4317	41.40
	Speaking	4317	48.00
	Comprehension	4317	43.50
6	Listening	3709	21.40
	Reading	3709	21.40
	Writing	3709	16.10
	Speaking	3709	26.20
	Comprehension	3709	23.50
7	Listening	3567	23.40
	Reading	3567	23.90
	Writing	3567	25.50
	Speaking	3567	32.00
	Comprehension	3567	20.30

DRAFT

Table 9.4: Percent of Students Meeting Achievement Target by Grade and Domain (Continued)

Grade	Domain	N-count	% of Students Meeting Achievement Target for Each Domain
8	Listening	3671	20.00
	Reading	3671	34.10
	Writing	3671	35.90
	Speaking	3671	24.30
	Comprehension	3671	29.00
9	Listening	3544	24.00
	Reading	3544	28.80
	Writing	3544	15.20
	Speaking	3544	31.90
	Comprehension	3544	26.40
10	Listening	3150	22.60
	Reading	3150	28.30
	Writing	3150	20.10
	Speaking	3150	35.00
	Comprehension	3150	26.90
11	Listening	2265	28.6
	Reading	2265	33.80
	Writing	2265	26.40
	Speaking	2265	38.80
	Comprehension	2265	33.80
12	Listening	1949	7.60
	Reading	1949	26.20
	Writing	1949	26.40
	Speaking	1949	40.00
	Comprehension	1949	17.10

Note. Omits, blanks, and invalid responses were scored as “0”.

DRAFT

REFERENCES

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (AERA, APA, NCME). (1999). *Standards for educational and psychological testing.* Washington, D.C.: Author.

Bond, T. G., & Fox C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Brennan, R. L. (2004). Manual for BB-CLASS: A computer program that uses the beta-binomial model for classification consistency and accuracy (Version 1.1). Iowa City, IA: Center for Advanced Studies in Measurement and Assessment.

Camilli, G. (2006). Test fairness. In R. L. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 221-256). Westport, CT: American Council on Education and Praeger.

Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items*. Thousand Oaks, CA: SAGE Publications.

Crocker, L., & Algina, J. (1986). *An introduction to classical and modern test theory*. New York: Holt, Rinehart & Winston.

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

de Ayala, R. J. (2009). *The theory and practice of Item Response Theory*. New York: Guilford Press.

Embretson, S., & Reise, S. (2000). *Item Response Theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates.

Haertel, E. H. (2006). Reliability. In R. L. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 65-110). Westport, CT: American Council on Education and Praeger.

Hanson, B. A. (1991). *Method of moments estimates for the four-parameter beta compound binomial model and the calculation of classification consistency indexes*. ACT Research Report 91–5. Iowa City, IA: American College Testing.

Hays, W. L. (1994). *Statistics* (5th ed.). Belmont, CA: Wadsworth.

Linacre, J. M. (2008). WINSTEPS (Version 3.67.0) [Computer Software]. Chicago: Winsteps.com.

Liu, I-M., & Agresti, A. (1996). Mantel-Haenszel-type inference for cumulative odds ratios with a stratified ordinal response. *Biometrics, 52*, 1223-1234.

DRAFT

Livingston, S. A. & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement, 32*, 179–1987.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Mantel, N. (1963). Chi-square tests with one degree of freedom: Extensions of the Mantel-Haenszel procedure. *American Statistical Association Journal, 58*, 690-700.

Mantel, N., & Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of National Cancer Institute, 22*, 719-748.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*, 149–174.

Masters, G. N., & Wright, B. D. (1997). The Partial Credit Model. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 101-121). New York: Springer-Verlag.

McDonald, R. P. (1999). *Test theory: A unified treatment.* Hillsdale, NJ: Erlbaum.

Muchinsky, P. M. (1996). The correction for attenuation. *Educational and Psychological Measurement, 56*, 63-75.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill Book Company.

Penfield, R. D. (2005). DIFAS: Differential item functioning analysis system. *Applied Psychological Measurement, 29*, 150-151.

Penfield, R. D. (2007). An approach for categorizing DIF in polytomous items. *Applied Measurement in Education, 20*, 335-355.

Penfield, R. D., & Algina, J. (2003). Applying the Liu-Agresti estimator of the cumulative common odds ratio to DIF detection in polytomous items. *Journal of Educational Measurement, 40*, 353-370.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.

DRAFT

Smith, E. V. (2004). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal components analysis of residuals. In E. V. Smith & R. M. Smith (Eds.), *Introduction to Rasch measurement* (pp. 575-600). Maple Grove, MN: JAM Press.

Thissen, D., & Steinberg, L. (1986). A taxonomy of item response models. *Psychometrika, 51*(4), 567-577.

Welch, C. (2006). Item and prompt development in performance testing. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 303-327). Mahwah, NJ: Lawrence Erlbaum Associates.

Wright, B. D., & Stone, M. H. (1979). *Best test design: Rasch measurement*. Chicago, IL: MESA Press.

Yen, W. M., & Fitzpatrick, A. R. (2006). Item response theory. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 111-153). Westport, Ct: American Council on Education and Praeger.

Young, M. J. & Yoon, B. (1998, April). Estimating the consistency and accuracy of classifications in a standards-referenced assessment. (CSE Technical Report 475). Center for the Study of Evaluation, Standards, and Student Testing. Los Angeles, CA: University of California, Los Angeles.

Zieky, M. (1993). Practical questions in the use of DIF statistics in test development. In P. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 337-347). Hillsdale, NJ: Erlbaum.

Zwick, R., Donoghue, J. R., & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement, 30*, 233-251.

DRAFT

APPENDIX A: ELPA TEST BLUEPRINT BY ASSESSMENT LEVEL AND DOMAIN

A.1: Level I – Operational Items

DOMAIN	STANDARD	NO. OF ITEMS	NO. OF POINTS
Listening	L.1 Follow simple and complex directions	1	1
	L.2 Understand spoken English to participate in social contexts	1	1
	L.3 Identify main ideas and supporting details from spoken English	8	8
	L.4 Identify meaning of vocabulary in the content areas	2	2
	L.5 Identify speaker attitude and point of view	2	2
	L.6 Make inferences and predictions	2	2
	SUBTOTAL	16	16
Reading	R.1 Recognize concepts of print literacy	3	3
	R.2 Demonstrate phonological awareness and the relationship of listening/speaking to decoding	3	3
	R.3 Build vocabulary to develop concepts	6	6
	R.5 Read and demonstrate comprehension of main idea and supporting details	1	1
	R.6 Apply reading skills in social and academic context	2	2
	R.8 Make inferences, predictions, and conclusions from reading	1	1
	SUBTOTAL	16	16
Writing	W.1 Use conventions and formats of written English	6	8
	W.2 Use grammatical conventions of English	4	4
	W.4 Construct sentences and develop paragraphs to organize writing supporting a central idea	1	4
	W.6 Use various types of writing for specific purposes	1	2
	SUBTOTAL	12	18
Speaking	S.2 Engage in conversations for personal expression and enjoyment	1	1
	S.4 Use English to interact in the classroom	2	8
	S.5 Provide and obtain information; express and exchange opinions	1	2
	S.6 Demonstrate comprehensible pronunciation and intonation for clarity in oral communication	3	3
	S.7 Present information, concepts, and ideas to an audience of listeners on a variety of topics	1	4
	SUBTOTAL	8	18
	TOTAL	52	68

DRAFT

A.2: Level II – Operational Items

DOMAIN	STANDARD	NO. OF ITEMS	NO. OF POINTS
Listening	L.1 Follow simple and complex directions	1	1
	L.2 Understand spoken English to participate in social contexts	3	3
	L.3 Identify main ideas and supporting details from spoken English	8	8
	L.4 Identify meaning of vocabulary in the content areas	2	2
	L.5 Identify speaker attitude and point of view	3	3
	L.6 Make inferences and predictions	3	3
	SUBTOTAL	20	20
Reading	R.2 Demonstrate phonological awareness and the relationship of listening/speaking to decoding	2	2
	R.5 Read and demonstrate comprehension of main idea and supporting details	7	7
	R.6 Apply reading skills in social and academic context	2	2
	R.8 Make inferences, predictions, and conclusions from reading	6	5
	R.9 Analyze style and form of various genre	3	3
	SUBTOTAL	20	20
Writing	W.1 Use conventions and formats of written English	3	6
	W.2 Use grammatical conventions of English	9	9
	W.6 Use various types of writing for specific purposes	1	4
	SUBTOTAL	13	19
Speaking	S.2 Engage in conversations for personal expression and enjoyment	1	2
	S.4 Use English to interact in the classroom	2	8
	S.5 Provide and obtain information; express and exchange opinions	1	2
	S.6 Demonstrate comprehensible pronunciation and intonation for clarity in oral communication	4	4
	S.7 Present information, concepts, and ideas to an audience of listeners on a variety of topics	1	4
	SUBTOTAL	9	20
	TOTAL	62	80

DRAFT

A.3: Level III – Operational Items

DOMAIN	STANDARD		NO. OF ITEMS	NO. OF POINTS
Listening	L.1	Follow simple and complex directions	3	3
	L.3	Identify main ideas and supporting details from spoken English	6	6
	L.4	Identify meaning of vocabulary in the content areas	2	2
	L.5	Identify speaker attitude and point of view	4	4
	L.6	Make inferences and predictions	4	4
		SUBTOTAL	19	19
Reading	R.3	Build vocabulary to develop concepts	6	6
	R.4	Understand and use grammatical structures of English to improve reading comprehension	1	1
	R.5	Read and demonstrate comprehension of main idea and supporting details	9	9
	R.8	Make inferences, predictions, and conclusions from reading	1	1
	R.9	Analyze style and form of various genre	3	3
	R.10	Identify author's voice, attitude and point of view	2	2
		SUBTOTAL	22	22
Writing	W.2	Use grammatical conventions of English	8	8
	W.4	Construct sentences and develop paragraphs to organize writing supporting a central idea	2	8
	W.6	Use various types of writing for specific purposes	2	4
		SUBTOTAL	12	20
Speaking	S.2	Engage in conversations for personal expression and enjoyment	1	2
	S.4	Use English to interact in the classroom	1	4
	S.5	Provide and obtain information; express and exchange opinions	1	2
	S.6	Demonstrate comprehensible pronunciation and intonation for clarity in oral communication	4	4
	S.7	Present information, concepts, and ideas to an audience of listeners on a variety of topics	1	4
	S.8	Use strategies to extend communicative competence	1	4
		SUBTOTAL	9	20
		TOTAL	62	81

DRAFT

A.4: Level IV – Operational Items

DOMAIN	STANDARD	NO. OF ITEMS	NO. OF POINTS
Listening	L.1 Follow simple and complex directions	2	2
	L.2 Understand spoken English to participate in social contexts	3	3
	L.3 Identify main ideas and supporting details from spoken English	3	3
	L.4 Identify meaning of vocabulary in the content areas	5	5
	L.5 Identify speaker attitude and point of view	6	6
	L.6 Make inferences and predictions	2	2
	SUBTOTAL	21	21
Reading	R.3 Build vocabulary to develop concepts	8	8
	R.5 Read and demonstrate comprehension of main idea and supporting details	4	4
	R.7 Read for research purposes	1	1
	R.8 Make inferences, predictions, and conclusions from reading	6	6
	R.10 Identify author's voice, attitude and point of view	2	2
	SUBTOTAL	21	21
Writing	W.2 Use grammatical conventions of English	12	14
	W.4 Construct sentences and develop paragraphs to organize writing supporting a central idea	1	4
	W.8 Use tone and voice to engage specific audiences	1	4
	SUBTOTAL	14	22
Speaking	S.4 Use English to interact in the classroom	1	4
	S.5 Provide and obtain information; express and exchange opinions	2	4
	S.6 Demonstrate comprehensible pronunciation and intonation for clarity in oral communication	5	5
	S.7 Present information, concepts, and ideas to an audience of listeners on a variety of topics	1	4
	S.8 Use strategies to extend communicative competence	1	4
	SUBTOTAL	10	21
	TOTAL	66	85

DRAFT

A.5: Level V – Operational Items

DOMAIN	STANDARD		NO. OF ITEMS	NO. OF POINTS
Listening	L.1	Follow simple and complex directions	3	3
	L.2	Understand spoken English to participate in social contexts	2	2
	L.3	Identify main ideas and supporting details from spoken English	4	4
	L.4	Identify meaning of vocabulary in the content areas	4	4
	L.5	Identify speaker attitude and point of view	3	3
	L.6	Make inferences and predictions	4	4
		SUBTOTAL	20	20
Reading	R.3	Build vocabulary to develop concepts	6	6
	R.4	Understand and use grammatical structures of English to improve reading comprehension	1	1
	R.5	Read and demonstrate comprehension of main idea and supporting details	3	3
	R.6	Apply reading skills in social and academic context	1	1
	R.8	Make inferences, predictions, and conclusions from reading	4	4
	R.9	Analyze style and form of various genre	2	2
	R.10	Identify author's voice, attitude and point of view	4	4
		SUBTOTAL	21	21
Writing	W.2	Use grammatical conventions of English	12	14
	W.6	Use various types of writing for specific purposes	1	4
	W.8	Use tone and voice to engage specific audiences	1	4
		SUBTOTAL	14	22
Speaking	S.2	Engage in conversations for personal expression and enjoyment	2	4
	S.4	Use English to interact in the classroom	2	8
	S.6	Demonstrate comprehensible pronunciation and intonation for clarity in oral communication	5	5
	S.7	Present information, concepts, and ideas to an audience of listeners on a variety of topics	1	4
		SUBTOTAL	10	21
		TOTAL	65	84

DRAFT

A.6: Level I – Field Test Items

DOMAIN	STANDARD		NO. OF ITEMS F1	F2	F3	F4	F5	NO. OF POINTS F1	F2	F3	F4	F5
Listening	L.1	Follow simple and complex directions				1					1	
	L.2	Understand spoken English to participate in social contexts				1					1	
	L.3	Identify main ideas and supporting details from spoken English	1	1	1		1		1	1		1
	L.4	Identify meaning of vocabulary in the content areas	1			1	1	1			1	1
	L.5	Identify speaker attitude and point of view	1	1	1		1	1	1	1		1
	L.6	Make inferences and predictions		1	1			1	1	1		
		SUBTOTAL	3	3	3	3	3	3	3	3	3	3
Reading	R.1	Recognize concepts of print literacy	2	1	1	1		2	1	1	1	
	R.2	Demonstrate phonological awareness and the relationship of listening/speaking to decoding		1					1			
	R.3	Build vocabulary to develop concepts	1	1	2		1	1	1	2		1
	R.5	Read and demonstrate comprehension of main idea and supporting details	1	1	1	1	2	1	1	1	1	2
	R.8	Make inferences, predictions, and conclusions from reading				2	1				2	1
		SUBTOTAL	4	4	4	4	4	4	4	4	4	4
Writing	W.1	Use conventions and formats of written English			2	2	3			2	2	4
	W.2	Use grammatical conventions of English			1					2		
	W.3	Write using appropriate vocabulary choice and variation				1					2	
	W.4	Construct sentences and develop paragraphs to organize writing supporting a central idea	1					4				
	W.6	Use various types of writing for specific purposes		1					4			
		SUBTOTAL	1	1	3	3	3	4	4	4	4	4
Speaking	S.1	Use spoken language for daily activities within and beyond the school setting	1					1				
	S.2	Engage in conversations for personal expression and enjoyment	2					3				
	S.4	Use English to interact in the classroom					1					4
	S.5	Provide and obtain information; express and exchange opinions		1					2			
	S.6	Demonstrate comprehensible pronunciation and intonation for clarity in oral communication		2					2			
	S.7	Present information, concepts, and ideas to an audience of listeners on a variety of topics			1					4		
	S.8	Use strategies to extend communicative competence				1					4	
		SUBTOTAL	3	3	1	1	1	4	4	4	4	4
		TOTAL*	11	11	11	11	11	15	15	15	15	15

DRAFT

* Some field test items were repeated across forms. The actual number of unique field test items by domain is: 15 Listening, 16 Reading, 11 Writing, and 9 Speaking.

A.7: Level II – Field Test Items

DOMAIN	STANDARD		NO. OF ITEMS F1	F2	F3	F4	NO. OF POINTS F1	F2	F3	F4
Listening	L.1	Follow simple and complex directions	1	2			1	2		
	L.2	Understand spoken English to participate in social contexts	1	1			1	1		
	L.3	Identify main ideas and supporting details from spoken English	2		1	2	2		1	2
	L.4	Identify meaning of vocabulary in the content areas	1				1			
	L.5	Identify speaker attitude and point of view		1	3	1		1	3	1
	L.6	Make inferences and predictions		1	1	2		1	1	2
		SUBTOTAL	5	5	5	5	5	5	5	5
Reading	R.1	Recognize concepts of print literacy	1	1			1	1		
	R.2	Demonstrate phonological awareness and the relationship of listening/speaking to decoding	1	1			1	1		
	R.3	Build vocabulary to develop concepts		1	1	1		1	1	1
	R.5	Read and demonstrate comprehension of main idea and supporting details	1	1	2	2	1	1	2	2
	R.6	Apply reading skills in social and academic context	1				1			
	R.8	Make inferences, predictions, and conclusions from reading	1	1	2	1	1	1	2	1
	R.9	Analyze style and form of various genre				1				1
		SUBTOTAL	5	5	5	5	5	5	5	5
Writing	W.1	Use conventions and formats of written English	1	1	1		2	2	1	
	W.2	Use grammatical conventions of English	1	1		1	1	1		1
	W.3	Write using appropriate vocabulary choice and variation	1				2			
	W.4	Construct sentences and develop paragraphs to organize writing supporting a central idea		1	1			2	4	
	W.6	Use various types of writing for specific purposes				1				4
		SUBTOTAL	3	3	2	2	5	5	5	5
Speaking	S.1	Use spoken language for daily activities within and beyond the school setting	1				1			
	S.2	Engage in conversations for personal expression and enjoyment	1		1		2		1	
	S.4	Use English to interact in the classroom				1				4
	S.5	Provide and obtain information; express and exchange opinions	1				2			
	S.6	Demonstrate comprehensible pronunciation and intonation for clarity in oral communication		1		1		1		1
	S.7	Present information, concepts, and ideas to an audience of listeners on a variety of topics		1				4		
	S.8	Use strategies to extend communicative competence			1				4	
		SUBTOTAL	3	2	2	2	5	5	5	5
		TOTAL*	16	15	14	14	20	20	20	20

DRAFT

* Some field test items were repeated across forms. The actual number of unique field test items by domain is: 18 Listening, 20 Reading, 10 Writing, and 9 Speaking.

A.8: Level III – Field Test Items

DOMAIN	STANDARD		NO. OF ITEMS F1	NO. OF ITEMS F2	NO. OF ITEMS F3	NO. OF POINTS F1	NO. OF POINTS F2	NO. OF POINTS F3
Listening	L.1	Follow simple and complex directions	1	1	1	1	1	1
	L.2	Understand spoken English to participate in social contexts	1	1	1	1	1	1
	L.3	Identify main ideas and supporting details from spoken English	1	1	1	1	1	1
	L.4	Identify meaning of vocabulary in the content areas	1			1		
	L.5	Identify speaker attitude and point of view	2	1	1	2	1	1
	L.6	Make inferences and predictions		2	1		2	1
		SUBTOTAL	6	6	5	6	6	5
Reading	R.3	Build vocabulary to develop concepts	2	1	1	2	1	1
	R.5	Read and demonstrate comprehension of main idea and supporting details	2	1	3	2	1	3
	R.7	Read for research purposes		1			1	
	R.8	Make inferences, predictions, and conclusions from reading	1	1	1	1	1	1
	R.10	Identify author's voice, attitude and point of view		1			1	
		SUBTOTAL	5	5	5	5	5	5
Writing	W.2	Use grammatical conventions of English	1	1	1	1	1	2
	W.3	Write using appropriate vocabulary choice and variation			1			4
	W.4	Construct sentences and develop paragraphs to organize writing supporting a central idea		1			4	
	W.6	Use various types of writing for specific purposes	1			4		
		SUBTOTAL	2	2	2	5	5	6
Speaking	S.1	Use spoken language for daily activities within and beyond the school setting	1			1		
	S.2	Engage in conversations for personal expression and enjoyment			1			1
	S.4	Use English to interact in the classroom			1			4
	S.5	Provide and obtain information; express and exchange opinions						
	S.6	Demonstrate comprehensible pronunciation and intonation for clarity in oral communication		1			1	
	S.7	Present information, concepts, and ideas to an audience of listeners on a variety of topics	1			4		
	S.8	Use strategies to extend communicative competence		1			4	
		SUBTOTAL	2	2	2	5	5	5
		TOTAL*	15	15	14	21	21	21

DRAFT

* Some field test items were repeated across forms. The actual number of unique field test items by domain is: 15 Listening, 15 Reading, 6 Writing, and 6 Speaking.

A.9: Level IV – Field Test Items

DOMAIN	STANDARD		NO. OF ITEMS F1	NO. OF ITEMS F2	NO. OF ITEMS F3	NO. OF POINTS F1	NO. OF POINTS F2	NO. OF POINTS F3
Listening	L.1	Follow simple and complex directions			1			1
	L.2	Understand spoken English to participate in social contexts			1			1
	L.3	Identify main ideas and supporting details from spoken English	2	2	1	2	2	1
	L.4	Identify meaning of vocabulary in the content areas	1			1		
	L.5	Identify speaker attitude and point of view	2	2	2	2	2	2
	L.6	Make inferences and predictions		1	1		1	1
		SUBTOTAL	5	5	6	5	5	6
Reading	R.3	Build vocabulary to develop concepts	1	1	1	1	1	1
	R.5	Read and demonstrate comprehension of main idea and supporting details	2	1	1	2	1	1
	R.7	Read for research purposes	1	1		1	1	
	R.8	Make inferences, predictions, and conclusions from reading	1	1	1	1	1	1
	R.10	Identify author's voice, attitude and point of view		1	2		1	2
		SUBTOTAL	5	5	5	5	5	5
Writing	W.2	Use grammatical conventions of English	1	1	2	2	2	2
	W.3	Write using appropriate vocabulary choice and variation			1			4
	W.4	Construct sentences and develop paragraphs to organize writing supporting a central idea	1			4		
	W.6	Use various types of writing for specific purposes		1			4	
		SUBTOTAL	2	2	3	6	6	6
Speaking	S.4	Use English to interact in the classroom			1			4
	S.5	Provide and obtain information; express and exchange opinions	1			1		
	S.6	Demonstrate comprehensible pronunciation and intonation for clarity in oral communication		1	1		1	1
	S.7	Present information, concepts, and ideas to an audience of listeners on a variety of topics	1			4		
	S.8	Use strategies to extend communicative competence		1			4	
		SUBTOTAL	2	2	2	5	5	5
		TOTAL*	14	14	16	21	21	22

DRAFT

* Some field test items were repeated across forms. The actual number of unique field test items by domain is: 15 Listening, 14 Reading, 7 Writing, and 6 Speaking.

A.10: Level V – Field Test Items

DOMAIN	STANDARD		NO. OF ITEMS F1	NO. OF ITEMS F2	NO. OF ITEMS F3	NO. OF POINTS F1	NO. OF POINTS F2	NO. OF POINTS F3
Listening	L.1	Follow simple and complex directions	2	1		2	1	
	L.2	Understand spoken English to participate in social contexts		1			1	
	L.3	Identify main ideas and supporting details from spoken English	2	1	2	2	1	2
	L.4	Identify meaning of vocabulary in the content areas	1		1	1		1
	L.5	Identify speaker attitude and point of view	1	1	1	1	1	1
	L.6	Make inferences and predictions		2	2		2	2
		SUBTOTAL	6	6	6	6	6	6
Reading	R.3	Build vocabulary to develop concepts	2	1	2	2	1	2
	R.5	Read and demonstrate comprehension of main idea and supporting details	1	2	1	1	2	1
	R.8	Make inferences, predictions, and conclusions from reading	1	1	1	1	1	1
	R.9	Analyze style and form of various genre	1		1	1		1
	R.10	Identify author's voice, attitude and point of view		1			1	
		SUBTOTAL	5	5	5	5	5	5
Writing	W.2	Use grammatical conventions of English	2	2	1	2	2	2
	W.3	Write using appropriate vocabulary choice and variation			1			4
	W.4	Construct sentences and develop paragraphs to organize writing supporting a central idea	1			4		
	W.6	Use various types of writing for specific purposes		1			4	
		SUBTOTAL	3	3	2	6	6	6
Speaking	S.2	Engage in conversations for personal expression and enjoyment			1			1
	S.4	Use English to interact in the classroom			1			4
	S.6	Demonstrate comprehensible pronunciation and intonation for clarity in oral communication	1	1		1	1	
	S.7	Present information, concepts, and ideas to an audience of listeners on a variety of topics	1			4		
	S.8	Use strategies to extend communicative competence		1			4	
		SUBTOTAL	2	2	2	5	5	5
		TOTAL*	16	16	15	22	22	22

DRAFT

* Some field test items were repeated across forms. The actual number of unique field test items by domain is: 17 Listening, 14 Reading, 8 Writing, and 6 Speaking.

APPENDIX B: READER AGREEMENT AND SCORE POINT DISTRIBUTION SUMMARY

Operational Items

Item ID	Item Name	Item Type	Reader Agreement %			Score Point Distribution %				
			Exact	Adjacent	Non-Adjacent	0	1	2	3	4
Level I										
3536736	Uppercase P	LW	93	7	0	31	69	NA	NA	NA
3537723	Fish	WW- V	93	6	0	53	25	21	NA	NA
3563044	Kids Drawing	ER	90	10	0	49	29	19	3	1
100000009583	"Run"	WW- A	92	8	0	45	21	34	NA	NA
100000009584	Lowercase g	LW	97	3	0	46	54	NA	NA	NA
100000017403	Kids in Library	SW- V	94	6	0	74	17	8	NA	NA
Level II										
3537800	"Path"	WW- A	94	6	0	18	15	67	NA	NA
3539399	"Table"	WW- A	92	8	0	18	42	41	NA	NA
3562388	Snack	SW- V	87	13	0	16	34	50	NA	NA
100000009658	"Close"	WW- A	94	6	0	18	38	44	NA	NA
100000055310	Kids in Class	ER	90	10	0	7	15	50	21	7
Level III										
3563477	Favorite Season	PW	88	12	0	3	12	53	26	6
100000009889	Wadlow	SW- SR	88	12	0	28	62	10	NA	NA
100000017435	Kids at Lunch	ER	88	12	0	3	18	46	24	9
100000056930	Hasna's Class	SW- SR	89	11	0	13	53	34	NA	NA
Level IV										
3539593	Continue Game	S-Co	92	8	0	21	25	54	NA	NA
3539646	Imelda and Luis	S-Co	94	6	0	18	21	61	NA	NA
3566110	Accomplishment	PW	90	10	0	3	10	43	32	12
100000055317	Pen Pals	ER	91	9	0	2	8	41	35	15
Level V										
3537943	Soccer Team	S-Co	89	11	0	12	26	61	NA	NA
3558502	Running	S-Co	88	12	0	14	30	56	NA	NA
100000017489	Cell Phone Driver	PW	90	10	0	3	13	44	30	9
100000055334	Label Drive	ER	90	10	0	4	13	41	31	11

Note: 1. LW = Letter Writing, WW = Word Writing, SW = Sentence Writing, S-Co = Sentence Combination, ER = Extended Response, PW = Paragraph Writing, V = Visual Stimulus, A = Aural Stimulus, SR = Short Reading
2. Due to rounding, the score point distribution may not add up to 100%.

DRAFT

Field Test Items

Item ID	Item Name	Item Type	Reader Agreement %			Score Point Distribution %				
			Exact	Adjacent	Non-Adjacent	0	1	2	3	4
Level I										
174275	Kids Reading	ER	81	18	0	36	34	25	4	1
174565	Kids Painting	ER	86	14	0	42	32	23	3	0
174625	Uppercase K	LW	94	6	0	28	72	NA	NA	NA
174628	Kid Sleeping	SW	84	16	0	68	16	17	NA	NA
174624	Lowercase h	LW	95	5	0	38	62	NA	NA	NA
174630	Girl Walking Dog	SW	90	9	1	77	12	11	NA	NA
174622	Lowercase f	LW	97	3	0	36	64	NA	NA	NA
174627	"Big"	WWA	95	5	0	39	14	46	NA	NA
Level II										
174263	"When"	WWA	93	7	0	15	32	53	NA	NA
174265	Kids Choosing Book	SW	80	19	1	21	32	48	NA	NA
174264	"Came"	WWA	95	4	0	10	21	69	NA	NA
174266	Crossing Guard	SW	77	23	1	26	32	42	NA	NA
177056	Tools to Learn	ER	80	19	1	7	18	36	27	13
187554	Playing Outside	ER	80	20	0	4	12	38	29	17
Level III										
177119	Game Directions	ER	75	24	0	4	18	48	23	7
17434	Special Event	ER	77	22	0	4	15	49	24	8
174333	Cesar Chavez	SWSR	82	18	0	24	64	13	NA	NA
177112	Learning in School	PW	74	25	1	4	18	46	23	8
Level IV										
174571	Juanita	S-Co	83	16	1	22	22	57	NA	NA
174574	Letter to Principal	ER	76	24	0	5	14	51	24	7
174576	George Washington	S-Co	84	15	0	23	32	45	NA	NA
174573	Volunteering	ER	79	21	0	2	10	50	31	7
174594	Leader You Admire	PW	73	27	1	3	11	48	29	9
Level V										
174601	Auto Mass Production	ER	77	23	0	2	13	44	31	10
176771	School Dress Code	ER	74	25	0	2	9	44	34	11
174606	Pat	S-Co	82	17	0	13	24	63	NA	NA
176753	Organization	PW	76	24	0	1	6	42	37	13

Note: 1. LW = Letter Writing, WW = Word Writing, SW = Sentence Writing, S-Co = Sentence Combination, ER = Extended Response, PW = Paragraph Writing, V = Visual Stimulus, A = Aural Stimulus, SR = Short Reading
2. Due to rounding, the score point distribution may not add up to 100%.

DRAFT

APPENDIX C: ELPA ITEM-LEVEL STATISTICS BY ASSESSMENT LEVEL

C.1: Level I – Operational Items

Item ID	Item Sequence	Domain	Type	Points	N	0	1/A	2/B	3/C	4/D	Omit/Invalid	P-Value	Item-Total Correlation
3540447	1	Listening	MC	1	9141	0.00	0.79	0.10	0.10	0.00	0.02	0.79	0.23
3467387	2	Listening	MC	1	9141	0.00	0.15	0.65	0.19	0.00	0.01	0.65	0.34
100000009514	3	Listening	MC	1	9141	0.00	0.31	0.55	0.11	0.00	0.02	0.55	0.37
100000009515	4	Listening	MC	1	9141	0.00	0.05	0.06	0.89	0.00	0.01	0.89	0.35
3536880	5	Listening	MC	1	9141	0.00	0.15	0.73	0.10	0.00	0.02	0.73	0.40
3484878	6	Listening	MC	1	9141	0.00	0.57	0.24	0.17	0.00	0.02	0.57	0.16
3477681	7	Listening	MC	1	9141	0.00	0.25	0.20	0.54	0.00	0.02	0.54	0.36
3540363	8	Listening	MC	1	9141	0.00	0.32	0.44	0.22	0.00	0.02	0.44	0.25
3540452	9	Listening	MC	1	9141	0.00	0.18	0.19	0.61	0.00	0.02	0.61	0.42
100000009509	10	Listening	MC	1	9141	0.00	0.49	0.25	0.25	0.00	0.02	0.49	0.25
100000009510	11	Listening	MC	1	9141	0.00	0.25	0.59	0.14	0.00	0.02	0.59	0.39
100000009511	12	Listening	MC	1	9141	0.00	0.17	0.20	0.60	0.00	0.02	0.60	0.36
3540453	16	Listening	MC	1	9141	0.00	0.33	0.21	0.44	0.00	0.02	0.44	0.31
3540454	17	Listening	MC	1	9141	0.00	0.37	0.20	0.40	0.00	0.02	0.37	0.06
3536612	18	Listening	MC	1	9141	0.00	0.16	0.29	0.53	0.00	0.02	0.53	0.25
3540455	19	Listening	MC	1	9141	0.00	0.20	0.19	0.58	0.00	0.03	0.58	0.33
3540775	1	Reading	MC	1	9141	0.00	0.96	0.01	0.02	0.00	0.01	0.96	0.20
3536628	2	Reading	MC	1	9141	0.00	0.08	0.15	0.76	0.00	0.02	0.76	0.46
3541699	3	Reading	MC	1	9141	0.00	0.75	0.21	0.03	0.00	0.01	0.75	0.33
3541548	5	Reading	MC	1	9141	0.00	0.83	0.06	0.09	0.00	0.02	0.83	0.32
3484912	6	Reading	MC	1	9141	0.00	0.37	0.49	0.12	0.00	0.03	0.49	0.38
3561062	8	Reading	MC	1	9141	0.00	0.18	0.77	0.03	0.00	0.01	0.77	0.45
3561060	9	Reading	MC	1	9141	0.00	0.28	0.61	0.10	0.00	0.01	0.61	0.43
3469887	10	Reading	MC	1	9141	0.00	0.64	0.17	0.16	0.00	0.02	0.64	0.16
100000009542	12	Reading	MC	1	9141	0.00	0.23	0.15	0.55	0.00	0.07	0.55	0.37
100000009526	13	Reading	MC	1	9141	0.00	0.32	0.20	0.41	0.00	0.07	0.41	0.36
100000009529	14	Reading	MC	1	9141	0.00	0.28	0.20	0.44	0.00	0.07	0.44	0.30
100000009545	15	Reading	MC	1	9141	0.00	0.33	0.44	0.16	0.00	0.07	0.44	0.34
100000009527	16	Reading	MC	1	9141	0.00	0.43	0.33	0.16	0.00	0.08	0.43	0.33
3526730	18	Reading	MC	1	9141	0.00	0.33	0.33	0.22	0.00	0.12	0.33	0.13
100000009538	19	Reading	MC	1	9141	0.00	0.35	0.22	0.30	0.00	0.13	0.30	0.36
100000009539	20	Reading	MC	1	9141	0.00	0.31	0.40	0.16	0.00	0.14	0.40	0.26
3484865	1	Speaking	CR	1	9141	0.12	0.88	0.00	0.00	0.00	0.00	0.88	0.42
100000017412	2	Speaking	CR	1	9141	0.11	0.89	0.00	0.00	0.00	0.00	0.89	0.40
3540678	3	Speaking	CR	1	9141	0.17	0.83	0.00	0.00	0.00	0.00	0.83	0.44
100000017413	4	Speaking	CR	1	9141	0.19	0.81	0.00	0.00	0.00	0.00	0.81	0.48
3536835	7	Speaking	CR	2	9141	0.24	0.45	0.31	0.00	0.00	0.00	0.54	0.50
100000012499	9	Speaking	CR	4	9141	0.12	0.15	0.25	0.31	0.17	0.00	0.57	0.60
100000017410	10	Speaking	CR	4	9141	0.35	0.26	0.22	0.13	0.05	0.00	0.32	0.58
3562622	11	Speaking	CR	4	9141	0.22	0.16	0.25	0.16	0.21	0.00	0.49	0.57

DRAFT

Item ID	Item Sequence	Domain	Type	Points	N	0	1/A	2/B	3/C	4/D	Omit/Invalid	P-Value	Item-Total Correlation
100000009565	1	Writing	MC	1	9141	0.00	0.34	0.57	0.07	0.00	0.02	0.57	0.25
100000009566	2	Writing	MC	1	9141	0.00	0.75	0.16	0.07	0.00	0.02	0.75	0.34
3467772	3	Writing	MC	1	9141	0.00	0.25	0.66	0.07	0.00	0.02	0.66	0.43
100000009569	4	Writing	MC	1	9141	0.00	0.27	0.27	0.39	0.00	0.06	0.39	0.45
100000009580	5	Writing	MC	1	9141	0.00	0.30	0.22	0.41	0.00	0.07	0.41	0.46
3561085	6	Writing	MC	1	9141	0.00	0.35	0.34	0.23	0.00	0.08	0.35	0.05
3536736	7	Writing	CR	1	9141	0.31	0.69	0.00	0.00	0.00	0.00	0.69	0.27
100000009584	8	Writing	CR	1	9141	0.46	0.54	0.00	0.00	0.00	0.00	0.54	0.36
3537723	9	Writing	CR	2	9141	0.54	0.25	0.21	0.00	0.00	0.00	0.34	0.55
100000009583	10	Writing	CR	2	9141	0.46	0.20	0.34	0.00	0.00	0.00	0.44	0.55
100000017403	11	Writing	CR	2	9141	0.76	0.16	0.08	0.00	0.00	0.00	0.16	0.50
3563044	12	Writing	CR	4	9141	0.55	0.26	0.17	0.02	0.01	0.00	0.17	0.56

DRAFT

C.2: Level II – Operational Items

Item ID	Item Sequence	Domain	Type	Points	N	0	1/A	2/B	3/C	4/D	Omit/Invalid	P-Value	Item-Total Correlation
100000009618	1	Listening	MC	1	16003	0.00	0.10	0.18	0.72	0.00	0.01	0.72	0.12
100000056914	2	Listening	MC	1	16003	0.00	0.66	0.24	0.09	0.00	0.01	0.66	0.35
100000056912	3	Listening	MC	1	16003	0.00	0.15	0.75	0.09	0.00	0.00	0.75	0.30
100000009619	4	Listening	MC	1	16003	0.00	0.04	0.63	0.32	0.00	0.01	0.63	0.25
3470853	8	Listening	MC	1	16003	0.00	0.09	0.10	0.81	0.00	0.00	0.81	0.42
100000056377	9	Listening	MC	1	16003	0.00	0.08	0.17	0.74	0.00	0.00	0.74	0.32
100000056913	10	Listening	MC	1	16003	0.00	0.32	0.47	0.20	0.00	0.01	0.47	0.21
3561090	13	Listening	MC	1	16003	0.00	0.11	0.06	0.82	0.00	0.00	0.82	0.40
3563425	14	Listening	MC	1	16003	0.00	0.19	0.16	0.65	0.00	0.01	0.65	0.37
3561094	15	Listening	MC	1	16003	0.00	0.71	0.12	0.16	0.00	0.01	0.71	0.32
3558375	16	Listening	MC	1	16003	0.00	0.81	0.05	0.14	0.00	0.01	0.81	0.35
3560392	17	Listening	MC	1	16003	0.00	0.10	0.55	0.35	0.00	0.01	0.55	0.32
3560391	18	Listening	MC	1	16003	0.00	0.24	0.05	0.71	0.00	0.01	0.71	0.43
100000055102	19	Listening	MC	1	16003	0.00	0.07	0.05	0.88	0.00	0.00	0.88	0.30
100000055103	20	Listening	MC	1	16003	0.00	0.08	0.67	0.24	0.00	0.00	0.67	0.32
100000055104	21	Listening	MC	1	16003	0.00	0.76	0.10	0.13	0.00	0.01	0.76	0.35
100000009609	22	Listening	MC	1	16003	0.00	0.18	0.60	0.21	0.00	0.01	0.60	0.33
100000009610	23	Listening	MC	1	16003	0.00	0.32	0.52	0.15	0.00	0.01	0.52	0.19
100000009611	24	Listening	MC	1	16003	0.00	0.55	0.06	0.38	0.00	0.01	0.55	0.29
100000009612	25	Listening	MC	1	16003	0.00	0.29	0.17	0.53	0.00	0.01	0.53	0.30
3540663	1	Reading	MC	1	16003	0.00	0.01	0.01	0.98	0.00	0.00	0.98	0.27
100000056382	2	Reading	MC	1	16003	0.00	0.24	0.74	0.02	0.00	0.00	0.74	0.45
3526731	5	Reading	MC	1	16003	0.00	0.65	0.26	0.08	0.00	0.01	0.65	0.42
100000009626	9	Reading	MC	1	16003	0.00	0.19	0.11	0.67	0.00	0.02	0.67	0.43
100000009625	10	Reading	MC	1	16003	0.00	0.19	0.64	0.15	0.00	0.02	0.64	0.38
100000009628	11	Reading	MC	1	16003	0.00	0.15	0.19	0.63	0.00	0.03	0.63	0.31
100000009627	12	Reading	MC	1	16003	0.00	0.57	0.15	0.24	0.00	0.03	0.57	0.33
3526419	13	Reading	MC	1	16003	0.00	0.25	0.29	0.43	0.00	0.03	0.43	0.43
3526418	14	Reading	MC	1	16003	0.00	0.14	0.54	0.29	0.00	0.03	0.54	0.29
3537258	15	Reading	MC	1	16003	0.00	0.08	0.79	0.10	0.00	0.03	0.79	0.48
3537247	16	Reading	MC	1	16003	0.00	0.74	0.13	0.10	0.00	0.03	0.74	0.51
100000055189	17	Reading	MC	1	16003	0.00	0.75	0.14	0.07	0.00	0.03	0.75	0.45
100000055190	18	Reading	MC	1	16003	0.00	0.20	0.19	0.58	0.00	0.03	0.58	0.48
100000055191	19	Reading	MC	1	16003	0.00	0.63	0.17	0.17	0.00	0.03	0.63	0.48
100000055187	20	Reading	MC	1	16003	0.00	0.26	0.21	0.49	0.00	0.04	0.49	0.33
100000055188	21	Reading	MC	1	16003	0.00	0.42	0.42	0.13	0.00	0.04	0.42	0.16
100000009634	22	Reading	MC	1	16003	0.00	0.30	0.35	0.31	0.00	0.04	0.35	0.11
100000009633	23	Reading	MC	1	16003	0.00	0.30	0.35	0.30	0.00	0.04	0.30	0.13
100000009632	24	Reading	MC	1	16003	0.00	0.48	0.32	0.16	0.00	0.04	0.48	0.25
100000009631	25	Reading	MC	1	16003	0.00	0.29	0.53	0.14	0.00	0.04	0.53	0.17
100000017416	1	Speaking	CR	1	16003	0.06	0.94	0.00	0.00	0.00	0.00	0.94	0.34
3540646	2	Speaking	CR	1	16003	0.14	0.86	0.00	0.00	0.00	0.00	0.86	0.45
3540649	3	Speaking	CR	1	16003	0.07	0.93	0.00	0.00	0.00	0.00	0.93	0.40
3540640	4	Speaking	CR	1	16003	0.13	0.87	0.00	0.00	0.00	0.00	0.87	0.46

DRAFT

Item ID	Item Sequence	Domain	Type	Points	N	0	1/A	2/B	3/C	4/D	Omit/Invalid	P-Value	Item-Total Correlation
100000055344	6	Speaking	CR	2	16003	0.17	0.37	0.47	0.00	0.00	0.00	0.65	0.52
3540643	7	Speaking	CR	2	16003	0.14	0.42	0.44	0.00	0.00	0.00	0.65	0.51
100000056531	10	Speaking	CR	4	16003	0.04	0.07	0.23	0.40	0.25	0.00	0.69	0.57
100000017420	11	Speaking	CR	4	16003	0.12	0.17	0.28	0.26	0.17	0.00	0.55	0.66
3562626	12	Speaking	CR	4	16003	0.05	0.09	0.27	0.20	0.39	0.00	0.70	0.52
3467778	1	Writing	MC	1	16003	0.00	0.86	0.08	0.05	0.00	0.01	0.86	0.44
100000055309	2	Writing	MC	1	16003	0.00	0.16	0.77	0.06	0.00	0.01	0.77	0.45
100000055308	3	Writing	MC	1	16003	0.00	0.15	0.78	0.06	0.00	0.01	0.78	0.48
100000009660	5	Writing	MC	1	16003	0.00	0.26	0.20	0.53	0.00	0.01	0.53	0.42
3559284	6	Writing	MC	1	16003	0.00	0.08	0.10	0.78	0.00	0.03	0.78	0.51
100000089936	7	Writing	MC	1	16003	0.00	0.14	0.73	0.12	0.00	0.01	0.73	0.47
100000009655	8	Writing	MC	1	16003	0.00	0.70	0.11	0.18	0.00	0.01	0.70	0.36
3537453	9	Writing	MC	1	16003	0.00	0.23	0.31	0.44	0.00	0.02	0.44	0.24
3537800	11	Writing	CR	2	16003	0.19	0.15	0.66	0.00	0.00	0.00	0.74	0.52
3539399	12	Writing	CR	2	16003	0.19	0.41	0.40	0.00	0.00	0.00	0.61	0.53
100000009658	13	Writing	CR	2	16003	0.19	0.37	0.43	0.00	0.00	0.00	0.62	0.55
3562388	14	Writing	CR	2	16003	0.17	0.33	0.50	0.00	0.00	0.00	0.67	0.61
100000055310	16	Writing	CR	4	16003	0.09	0.15	0.49	0.20	0.07	0.00	0.51	0.64

DRAFT

C.3: Level III – Operational Items

Item ID	Item Sequence	Domain	Type	Points	N	0	1/A	2/B	3/C	4/D	Omit/Invalid	P-Value	Item-Total Correlation
100000061235	1	Listening	MC	1	15081	0.00	0.07	0.12	0.24	0.55	0.01	0.55	0.21
100000017442	4	Listening	MC	1	15081	0.00	0.02	0.03	0.90	0.06	0.00	0.90	0.37
100000017437	5	Listening	MC	1	15081	0.00	0.11	0.74	0.11	0.04	0.01	0.74	0.34
100000017441	6	Listening	MC	1	15081	0.00	0.75	0.09	0.11	0.04	0.00	0.75	0.35
100000017440	7	Listening	MC	1	15081	0.00	0.05	0.06	0.80	0.07	0.01	0.80	0.39
100000056916	8	Listening	MC	1	15081	0.00	0.16	0.63	0.14	0.07	0.00	0.63	0.34
100000056917	9	Listening	MC	1	15081	0.00	0.68	0.05	0.12	0.16	0.01	0.68	0.29
3561098	10	Listening	MC	1	15081	0.00	0.09	0.14	0.60	0.16	0.01	0.60	0.32
3470725	11	Listening	MC	1	15081	0.00	0.05	0.56	0.12	0.27	0.01	0.56	0.32
100000009812	13	Listening	MC	1	15081	0.00	0.80	0.04	0.10	0.06	0.00	0.80	0.33
100000009817	14	Listening	MC	1	15081	0.00	0.16	0.14	0.60	0.10	0.00	0.60	0.37
100000055139	18	Listening	MC	1	15081	0.00	0.05	0.02	0.25	0.68	0.00	0.68	0.44
100000055140	19	Listening	MC	1	15081	0.00	0.07	0.65	0.09	0.18	0.00	0.65	0.21
100000055135	20	Listening	MC	1	15081	0.00	0.13	0.10	0.20	0.56	0.01	0.56	0.33
100000055136	21	Listening	MC	1	15081	0.00	0.50	0.08	0.24	0.19	0.01	0.50	0.37
100000018786	22	Listening	MC	1	15081	0.00	0.03	0.76	0.07	0.13	0.00	0.76	0.27
100000018785	23	Listening	MC	1	15081	0.00	0.06	0.11	0.03	0.79	0.00	0.79	0.35
100000018784	24	Listening	MC	1	15081	0.00	0.45	0.12	0.05	0.38	0.01	0.45	0.34
100000018783	25	Listening	MC	1	15081	0.00	0.12	0.23	0.47	0.17	0.01	0.47	0.16
3533550	1	Reading	MC	1	15081	0.00	0.04	0.06	0.08	0.82	0.01	0.82	0.35
3534017	2	Reading	MC	1	15081	0.00	0.07	0.07	0.72	0.14	0.00	0.72	0.43
100000056927	3	Reading	MC	1	15081	0.00	0.13	0.12	0.21	0.53	0.01	0.53	0.33
100000056926	4	Reading	MC	1	15081	0.00	0.23	0.53	0.18	0.06	0.01	0.53	0.39
100000009853	6	Reading	MC	1	15081	0.00	0.10	0.36	0.09	0.44	0.00	0.44	0.34
100000009854	7	Reading	MC	1	15081	0.00	0.16	0.12	0.11	0.60	0.01	0.60	0.39
100000009855	8	Reading	MC	1	15081	0.00	0.14	0.62	0.08	0.16	0.01	0.62	0.36
100000009858	9	Reading	MC	1	15081	0.00	0.23	0.09	0.09	0.58	0.01	0.58	0.41
100000055204	14	Reading	MC	1	15081	0.00	0.22	0.37	0.28	0.13	0.01	0.37	0.37
100000055200	15	Reading	MC	1	15081	0.00	0.09	0.13	0.10	0.68	0.01	0.68	0.46
100000055203	16	Reading	MC	1	15081	0.00	0.09	0.21	0.57	0.12	0.01	0.57	0.38
100000055199	17	Reading	MC	1	15081	0.00	0.23	0.23	0.10	0.44	0.01	0.44	0.26
100000055202	18	Reading	MC	1	15081	0.00	0.18	0.30	0.27	0.24	0.01	0.30	0.05
3537588	19	Reading	MC	1	15081	0.00	0.15	0.12	0.61	0.12	0.01	0.61	0.46
3526459	20	Reading	MC	1	15081	0.00	0.26	0.15	0.23	0.35	0.01	0.35	0.26
3526458	21	Reading	MC	1	15081	0.00	0.08	0.48	0.11	0.32	0.01	0.48	0.23
3526461	22	Reading	MC	1	15081	0.00	0.14	0.50	0.24	0.11	0.01	0.50	0.35
100000085124	23	Reading	MC	1	15081	0.00	0.43	0.32	0.13	0.10	0.01	0.32	0.24
100000085122	24	Reading	MC	1	15081	0.00	0.12	0.30	0.43	0.13	0.01	0.43	0.26
100000085123	25	Reading	MC	1	15081	0.00	0.43	0.19	0.15	0.20	0.01	0.43	0.23
3529874	26	Reading	MC	1	15081	0.00	0.29	0.29	0.17	0.23	0.01	0.29	0.16
3529880	27	Reading	MC	1	15081	0.00	0.37	0.19	0.21	0.21	0.01	0.37	0.13
3540626	1	Speaking	CR	1	15081	0.07	0.93	0.00	0.00	0.00	0.00	0.93	0.41
3540561	2	Speaking	CR	1	15081	0.05	0.95	0.00	0.00	0.00	0.00	0.95	0.34
3540625	3	Speaking	CR	1	15081	0.11	0.89	0.00	0.00	0.00	0.00	0.89	0.43

DRAFT

Item ID	Item Sequence	Domain	Type	Points	N	0	1/A	2/B	3/C	4/D	Omit/Invalid	P-Value	Item-Total Correlation
3563708	4	Speaking	CR	1	15081	0.14	0.86	0.00	0.00	0.00	0.00	0.86	0.48
3540565	6	Speaking	CR	2	15081	0.18	0.41	0.41	0.00	0.00	0.00	0.62	0.50
100000056534	7	Speaking	CR	2	15081	0.14	0.34	0.52	0.00	0.00	0.00	0.69	0.53
100000012506	8	Speaking	CR	4	15081	0.02	0.04	0.16	0.39	0.39	0.00	0.77	0.58
100000017448	10	Speaking	CR	4	15081	0.04	0.08	0.22	0.35	0.31	0.00	0.71	0.66
100000017446	11	Speaking	CR	4	15081	0.04	0.05	0.17	0.18	0.56	0.00	0.80	0.58
3526783	1	Writing	MC	1	15081	0.00	0.79	0.08	0.04	0.08	0.01	0.79	0.40
3537520	2	Writing	MC	1	15081	0.00	0.16	0.07	0.07	0.69	0.01	0.69	0.42
100000057816	3	Writing	MC	1	15081	0.00	0.08	0.55	0.08	0.28	0.01	0.55	0.28
100000057815	4	Writing	MC	1	15081	0.00	0.31	0.12	0.40	0.16	0.01	0.40	0.22
3526794	5	Writing	MC	1	15081	0.00	0.84	0.04	0.02	0.09	0.01	0.84	0.35
3526795	6	Writing	MC	1	15081	0.00	0.12	0.03	0.03	0.81	0.01	0.81	0.46
100000057810	7	Writing	MC	1	15081	0.00	0.11	0.12	0.09	0.68	0.01	0.68	0.42
3526811	8	Writing	MC	1	15081	0.00	0.76	0.08	0.09	0.07	0.01	0.76	0.45
100000009889	10	Writing	CR	2	15081	0.29	0.61	0.10	0.00	0.00	0.00	0.41	0.50
100000056930	11	Writing	CR	2	15081	0.14	0.52	0.34	0.00	0.00	0.00	0.60	0.49
3563477	12	Writing	CR	4	15081	0.05	0.11	0.52	0.25	0.06	0.00	0.54	0.59
100000017435	13	Writing	CR	4	15081	0.05	0.17	0.45	0.24	0.09	0.00	0.53	0.60

DRAFT

C.4: Level IV – Operational Items

Item ID	Item Sequence	Domain	Type	Points	N	0	1/A	2/B	3/C	4/D	Omit/Invalid	P-Value	Item-Total Correlation
100000056491	1	Listening	MC	1	10947	0.00	0.78	0.05	0.13	0.02	0.01	0.78	0.26
100000017454	2	Listening	MC	1	10947	0.00	0.75	0.08	0.08	0.08	0.00	0.75	0.32
100000017455	3	Listening	MC	1	10947	0.00	0.14	0.07	0.71	0.07	0.00	0.71	0.30
100000017456	4	Listening	MC	1	10947	0.00	0.08	0.53	0.35	0.03	0.00	0.53	0.22
100000017457	5	Listening	MC	1	10947	0.00	0.72	0.07	0.04	0.17	0.00	0.72	0.36
100000056493	6	Listening	MC	1	10947	0.00	0.05	0.10	0.08	0.76	0.00	0.76	0.34
3526745	7	Listening	MC	1	10947	0.00	0.64	0.09	0.07	0.19	0.00	0.64	0.38
100000008945	8	Listening	MC	1	10947	0.00	0.19	0.04	0.08	0.70	0.00	0.70	0.27
100000056492	9	Listening	MC	1	10947	0.00	0.10	0.55	0.21	0.14	0.00	0.55	0.25
100000008946	10	Listening	MC	1	10947	0.00	0.18	0.09	0.56	0.16	0.00	0.56	0.25
100000056490	11	Listening	MC	1	10947	0.00	0.50	0.18	0.09	0.23	0.00	0.50	0.31
100000008952	17	Listening	MC	1	10947	0.00	0.74	0.10	0.07	0.09	0.00	0.74	0.32
100000008949	18	Listening	MC	1	10947	0.00	0.11	0.15	0.54	0.19	0.00	0.54	0.27
100000008953	19	Listening	MC	1	10947	0.00	0.15	0.21	0.50	0.13	0.00	0.50	0.15
100000008950	20	Listening	MC	1	10947	0.00	0.20	0.56	0.06	0.18	0.00	0.56	0.41
100000008954	21	Listening	MC	1	10947	0.00	0.09	0.09	0.19	0.62	0.01	0.62	0.27
100000055211	22	Listening	MC	1	10947	0.00	0.15	0.66	0.08	0.11	0.00	0.66	0.41
100000055212	23	Listening	MC	1	10947	0.00	0.48	0.14	0.07	0.30	0.00	0.30	0.05
100000055214	24	Listening	MC	1	10947	0.00	0.63	0.12	0.10	0.15	0.00	0.63	0.35
100000055215	25	Listening	MC	1	10947	0.00	0.16	0.41	0.24	0.19	0.00	0.16	0.01
100000055216	26	Listening	MC	1	10947	0.00	0.15	0.22	0.10	0.53	0.01	0.53	0.29
3535829	1	Reading	MC	1	10947	0.00	0.14	0.06	0.72	0.07	0.00	0.72	0.33
100000056501	2	Reading	MC	1	10947	0.00	0.22	0.67	0.07	0.03	0.00	0.67	0.43
100000056503	3	Reading	MC	1	10947	0.00	0.12	0.06	0.12	0.70	0.01	0.70	0.32
3535677	4	Reading	MC	1	10947	0.00	0.68	0.14	0.10	0.07	0.00	0.68	0.48
100000056500	5	Reading	MC	1	10947	0.00	0.13	0.15	0.42	0.29	0.01	0.42	0.12
100000055152	11	Reading	MC	1	10947	0.00	0.10	0.62	0.14	0.14	0.00	0.62	0.32
100000055154	12	Reading	MC	1	10947	0.00	0.66	0.09	0.10	0.15	0.00	0.66	0.39
100000055155	13	Reading	MC	1	10947	0.00	0.29	0.12	0.08	0.50	0.01	0.50	0.35
100000055156	14	Reading	MC	1	10947	0.00	0.40	0.16	0.16	0.27	0.01	0.40	0.17
100000008958	15	Reading	MC	1	10947	0.00	0.13	0.15	0.54	0.17	0.00	0.54	0.38
100000008959	16	Reading	MC	1	10947	0.00	0.10	0.72	0.08	0.11	0.00	0.72	0.40
100000008960	17	Reading	MC	1	10947	0.00	0.09	0.08	0.76	0.08	0.01	0.76	0.49
100000008961	18	Reading	MC	1	10947	0.00	0.30	0.16	0.06	0.47	0.01	0.47	0.23
100000008962	19	Reading	MC	1	10947	0.00	0.51	0.15	0.16	0.18	0.01	0.51	0.35
100000008963	20	Reading	MC	1	10947	0.00	0.24	0.58	0.10	0.07	0.01	0.58	0.34
3558745	21	Reading	MC	1	10947	0.00	0.59	0.13	0.13	0.14	0.01	0.59	0.27
3558748	22	Reading	MC	1	10947	0.00	0.20	0.51	0.10	0.18	0.01	0.51	0.27
3558753	23	Reading	MC	1	10947	0.00	0.16	0.11	0.50	0.22	0.01	0.50	0.23
3558750	24	Reading	MC	1	10947	0.00	0.06	0.27	0.48	0.17	0.01	0.48	0.28
3558739	25	Reading	MC	1	10947	0.00	0.21	0.61	0.09	0.08	0.01	0.61	0.25
3558738	26	Reading	MC	1	10947	0.00	0.31	0.14	0.40	0.14	0.01	0.40	0.32
3566131	1	Speaking	CR	1	10947	0.12	0.88	0.00	0.00	0.00	0.00	0.88	0.51
3566132	2	Speaking	CR	1	10947	0.13	0.87	0.00	0.00	0.00	0.00	0.87	0.55

DRAFT

Item ID	Item Sequence	Domain	Type	Points	N	0	1/A	2/B	3/C	4/D	Omit/Invalid	P-Value	Item-Total Correlation
3540821	3	Speaking	CR	1	10947	0.20	0.80	0.00	0.00	0.00	0.00	0.80	0.56
3540819	4	Speaking	CR	1	10947	0.07	0.93	0.00	0.00	0.00	0.00	0.93	0.49
3540820	5	Speaking	CR	1	10947	0.07	0.93	0.00	0.00	0.00	0.00	0.93	0.47
100000056536	7	Speaking	CR	2	10947	0.13	0.34	0.53	0.00	0.00	0.00	0.70	0.59
3538699	8	Speaking	CR	2	10947	0.15	0.34	0.51	0.00	0.00	0.00	0.68	0.59
100000012509	9	Speaking	CR	4	10947	0.03	0.05	0.16	0.39	0.37	0.00	0.76	0.63
100000017467	11	Speaking	CR	4	10947	0.05	0.07	0.18	0.34	0.37	0.00	0.73	0.68
100000055348	12	Speaking	CR	4	10947	0.04	0.03	0.15	0.14	0.64	0.00	0.82	0.60
3526775	1	Writing	MC	1	10947	0.00	0.17	0.03	0.78	0.02	0.00	0.78	0.28
3471438	2	Writing	MC	1	10947	0.00	0.02	0.20	0.73	0.04	0.00	0.73	0.39
100000056517	3	Writing	MC	1	10947	0.00	0.05	0.67	0.06	0.21	0.00	0.67	0.26
3558613	4	Writing	MC	1	10947	0.00	0.05	0.25	0.03	0.65	0.00	0.65	0.30
3537533	5	Writing	MC	1	10947	0.00	0.37	0.09	0.13	0.40	0.01	0.40	0.26
100000056522	6	Writing	MC	1	10947	0.00	0.04	0.05	0.84	0.07	0.00	0.84	0.37
3526803	7	Writing	MC	1	10947	0.00	0.78	0.09	0.07	0.06	0.00	0.78	0.40
100000056520	8	Writing	MC	1	10947	0.00	0.03	0.75	0.05	0.16	0.00	0.75	0.37
3558575	9	Writing	MC	1	10947	0.00	0.04	0.70	0.03	0.22	0.00	0.70	0.29
3558574	10	Writing	MC	1	10947	0.00	0.62	0.06	0.09	0.23	0.00	0.62	0.33
3539646	11	Writing	CR	2	10947	0.18	0.21	0.60	0.00	0.00	0.00	0.71	0.48
3539593	12	Writing	CR	2	10947	0.22	0.24	0.54	0.00	0.00	0.00	0.66	0.47
3566110	14	Writing	CR	4	10947	0.05	0.10	0.42	0.30	0.12	0.00	0.58	0.65
100000055317	15	Writing	CR	4	10947	0.04	0.08	0.40	0.33	0.14	0.00	0.61	0.65

DRAFT

C.5: Level V – Operational Items

Item ID	Item Sequence	Domain	Type	Points	N	0	1/A	2/B	3/C	4/D	Omit/Invalid	P-Value	Item-Total Correlation
100000061241	1	Listening	MC	1	10908	0.00	0.03	0.08	0.87	0.01	0.01	0.87	0.41
100000017475	2	Listening	MC	1	10908	0.00	0.29	0.19	0.07	0.44	0.00	0.44	0.26
100000017476	3	Listening	MC	1	10908	0.00	0.06	0.74	0.11	0.07	0.01	0.74	0.40
100000055163	4	Listening	MC	1	10908	0.00	0.06	0.09	0.81	0.03	0.00	0.81	0.44
100000055164	5	Listening	MC	1	10908	0.00	0.06	0.50	0.07	0.36	0.01	0.50	0.32
3566062	8	Listening	MC	1	10908	0.00	0.03	0.05	0.84	0.07	0.00	0.84	0.42
100000056495	9	Listening	MC	1	10908	0.00	0.04	0.04	0.74	0.17	0.00	0.74	0.35
3566089	10	Listening	MC	1	10908	0.00	0.07	0.13	0.56	0.24	0.00	0.56	0.26
3526756	11	Listening	MC	1	10908	0.00	0.05	0.16	0.09	0.69	0.00	0.69	0.45
100000009039	12	Listening	MC	1	10908	0.00	0.32	0.48	0.13	0.06	0.00	0.48	0.28
100000056480	13	Listening	MC	1	10908	0.00	0.10	0.37	0.31	0.21	0.00	0.37	0.24
100000056469	18	Listening	MC	1	10908	0.00	0.72	0.05	0.20	0.03	0.00	0.72	0.38
100000056470	19	Listening	MC	1	10908	0.00	0.06	0.75	0.07	0.11	0.00	0.75	0.35
100000056471	20	Listening	MC	1	10908	0.00	0.09	0.11	0.09	0.71	0.00	0.71	0.38
100000056933	21	Listening	MC	1	10908	0.00	0.10	0.66	0.12	0.12	0.00	0.66	0.27
100000056472	22	Listening	MC	1	10908	0.00	0.07	0.60	0.05	0.28	0.01	0.60	0.38
3526516	23	Listening	MC	1	10908	0.00	0.51	0.12	0.32	0.04	0.01	0.51	0.18
3536855	24	Listening	MC	1	10908	0.00	0.53	0.22	0.20	0.05	0.01	0.53	0.30
3536878	25	Listening	MC	1	10908	0.00	0.27	0.19	0.10	0.44	0.01	0.27	0.15
3536738	26	Listening	MC	1	10908	0.00	0.20	0.08	0.51	0.21	0.01	0.51	0.34
3558509	1	Reading	MC	1	10908	0.00	0.15	0.04	0.77	0.03	0.00	0.77	0.51
100000056505	2	Reading	MC	1	10908	0.00	0.68	0.12	0.16	0.04	0.00	0.68	0.34
100000056506	3	Reading	MC	1	10908	0.00	0.27	0.49	0.12	0.11	0.01	0.49	0.37
100000009007	5	Reading	MC	1	10908	0.00	0.12	0.06	0.77	0.04	0.00	0.77	0.51
100000009009	6	Reading	MC	1	10908	0.00	0.05	0.07	0.09	0.79	0.00	0.79	0.60
100000009008	7	Reading	MC	1	10908	0.00	0.12	0.75	0.05	0.08	0.00	0.75	0.52
100000009006	8	Reading	MC	1	10908	0.00	0.06	0.06	0.80	0.08	0.00	0.80	0.58
100000009010	9	Reading	MC	1	10908	0.00	0.08	0.74	0.12	0.06	0.00	0.74	0.52
100000009005	10	Reading	MC	1	10908	0.00	0.43	0.12	0.14	0.31	0.00	0.43	0.21
100000057368	15	Reading	MC	1	10908	0.00	0.21	0.55	0.20	0.04	0.01	0.55	0.33
100000057370	16	Reading	MC	1	10908	0.00	0.09	0.27	0.50	0.13	0.01	0.50	0.32
100000057369	17	Reading	MC	1	10908	0.00	0.18	0.15	0.18	0.49	0.00	0.49	0.33
100000057373	18	Reading	MC	1	10908	0.00	0.23	0.13	0.30	0.34	0.01	0.34	0.13
100000057371	19	Reading	MC	1	10908	0.00	0.47	0.14	0.22	0.16	0.01	0.47	0.34
100000057372	20	Reading	MC	1	10908	0.00	0.07	0.62	0.17	0.13	0.01	0.62	0.34
3562741	21	Reading	MC	1	10908	0.00	0.22	0.63	0.08	0.06	0.02	0.63	0.31
3562743	22	Reading	MC	1	10908	0.00	0.19	0.15	0.22	0.42	0.02	0.42	0.37
3562744	23	Reading	MC	1	10908	0.00	0.51	0.22	0.18	0.07	0.02	0.51	0.29
3562737	24	Reading	MC	1	10908	0.00	0.11	0.16	0.14	0.56	0.02	0.56	0.44
3562746	25	Reading	MC	1	10908	0.00	0.30	0.39	0.15	0.14	0.02	0.39	0.11
3562742	26	Reading	MC	1	10908	0.00	0.12	0.17	0.15	0.54	0.02	0.54	0.39
3566126	1	Speaking	CR	1	10908	0.11	0.89	0.00	0.00	0.00	0.00	0.89	0.51
3541261	2	Speaking	CR	1	10908	0.12	0.88	0.00	0.00	0.00	0.00	0.88	0.56
3541263	3	Speaking	CR	1	10908	0.17	0.83	0.00	0.00	0.00	0.00	0.83	0.62
3541258	4	Speaking	CR	1	10908	0.16	0.84	0.00	0.00	0.00	0.00	0.84	0.59

DRAFT

Item ID	Item Sequence	Domain	Type	Points	N	0	1/A	2/B	3/C	4/D	Omit/Invalid	P-Value	Item-Total Correlation
3541259	5	Speaking	CR	1	10908	0.12	0.88	0.00	0.00	0.00	0.00	0.88	0.55
3562634	7	Speaking	CR	2	10908	0.10	0.21	0.69	0.00	0.00	0.00	0.79	0.67
100000055355	8	Speaking	CR	2	10908	0.07	0.19	0.73	0.00	0.00	0.00	0.83	0.67
100000012511	9	Speaking	CR	4	10908	0.03	0.05	0.12	0.28	0.53	0.00	0.81	0.69
100000017492	11	Speaking	CR	4	10908	0.06	0.06	0.15	0.28	0.45	0.00	0.75	0.74
100000055353	12	Speaking	CR	4	10908	0.03	0.02	0.11	0.08	0.76	0.00	0.88	0.65
3526786	1	Writing	MC	1	10908	0.00	0.03	0.06	0.04	0.87	0.00	0.87	0.50
100000009014	2	Writing	MC	1	10908	0.00	0.56	0.16	0.17	0.11	0.01	0.56	0.47
100000056526	3	Writing	MC	1	10908	0.00	0.41	0.48	0.06	0.04	0.01	0.48	0.27
3470663	4	Writing	MC	1	10908	0.00	0.10	0.24	0.38	0.27	0.00	0.38	0.19
100000056527	5	Writing	MC	1	10908	0.00	0.21	0.21	0.47	0.11	0.01	0.47	0.12
3526809	7	Writing	MC	1	10908	0.00	0.05	0.84	0.04	0.06	0.01	0.84	0.47
3526810	8	Writing	MC	1	10908	0.00	0.06	0.08	0.77	0.08	0.00	0.77	0.51
100000056529	9	Writing	MC	1	10908	0.00	0.10	0.75	0.09	0.06	0.01	0.75	0.51
3558631	10	Writing	MC	1	10908	0.00	0.08	0.04	0.72	0.16	0.01	0.72	0.49
3558630	11	Writing	MC	1	10908	0.00	0.21	0.68	0.05	0.06	0.01	0.68	0.57
3537943	13	Writing	CR	2	10908	0.14	0.26	0.60	0.00	0.00	0.00	0.73	0.53
3558502	14	Writing	CR	2	10908	0.16	0.29	0.55	0.00	0.00	0.00	0.70	0.44
100000017489	15	Writing	CR	4	10908	0.05	0.13	0.43	0.29	0.09	0.00	0.56	0.66
100000055334	16	Writing	CR	4	10908	0.10	0.12	0.39	0.29	0.11	0.00	0.55	0.68

DRAFT

C.6: Level I – Field Test Items

Item ID	Item Sequence	Domain	Type	Points	N	0	1/A	2/B	3/C	4/D	Omit/Invalid	P-Value	Item-Total Correlation
100000133352	8	Speaking	CR	4	1755	0.24	0.24	0.25	0.18	0.09	0.00	0.41	0.63
100000166540	13	Listening	MC	1	1877	0.00	0.29	0.16	0.54	0.00	0.01	0.54	0.33
100000166541	15	Listening	MC	1	1877	0.00	0.19	0.63	0.16	0.00	0.02	0.63	0.35
100000166543	14	Listening	MC	1	1877	0.00	0.32	0.21	0.46	0.00	0.02	0.46	0.30
100000166548	14	Listening	MC	1	1766	0.00	0.23	0.17	0.55	0.00	0.04	0.55	0.36
100000166550	13	Listening	MC	1	1766	0.00	0.64	0.19	0.13	0.00	0.04	0.64	0.31
100000166552	15	Listening	MC	1	1766	0.00	0.26	0.41	0.28	0.00	0.05	0.41	0.12
100000166555	4	Reading	MC	1	1867	0.00	0.02	0.02	0.95	0.00	0.01	0.95	0.24
100000166565	13	Listening	MC	1	1867	0.00	0.21	0.22	0.56	0.00	0.01	0.56	0.43
100000166566	14	Listening	MC	1	1867	0.00	0.66	0.18	0.15	0.00	0.01	0.66	0.34
100000166567	15	Listening	MC	1	1867	0.00	0.19	0.57	0.23	0.00	0.01	0.57	0.34
100000166595	4	Reading	MC	1	1877	0.00	0.08	0.76	0.14	0.00	0.02	0.76	0.33
100000166598	16	Reading	MC	1	1867	0.00	0.13	0.31	0.52	0.00	0.04	0.52	0.37
100000166600	10	Reading	MC	1	1766	0.00	0.33	0.52	0.12	0.00	0.03	0.52	0.42
100000166602	16	Reading	MC	1	1766	0.00	0.26	0.21	0.44	0.00	0.09	0.44	0.45
100000168722	6	Reading	MC	1	1755	0.00	0.90	0.05	0.04	0.00	0.01	0.90	0.31
100000168726	7	Reading	MC	1	1867	0.00	0.22	0.16	0.60	0.00	0.01	0.60	0.43
100000169217	5	Listening	MC	1	1755	0.00	0.03	0.96	0.01	0.00	0.01	0.96	0.28
100000169218	6	Listening	MC	1	1755	0.00	0.02	0.95	0.02	0.00	0.01	0.95	0.20
100000169219	7	Listening	MC	1	1755	0.00	0.93	0.03	0.03	0.00	0.01	0.93	0.21
100000169221	9	Listening	MC	1	1876	0.00	0.33	0.47	0.18	0.00	0.02	0.47	0.22
100000169222	8	Listening	MC	1	1876	0.00	0.54	0.20	0.24	0.00	0.02	0.54	0.24
100000169223	10	Listening	MC	1	1876	0.00	0.77	0.09	0.12	0.00	0.02	0.77	0.35
100000169249	9	Reading	MC	1	1876	0.00	0.31	0.14	0.52	0.00	0.02	0.52	0.52
100000169309	17	Reading	MC	1	1877	0.00	0.53	0.27	0.08	0.00	0.11	0.53	0.20
100000169310	15	Reading	MC	1	1755	0.00	0.32	0.23	0.35	0.00	0.10	0.35	0.38
100000169417	17	Reading	MC	1	1755	0.00	0.39	0.27	0.21	0.00	0.13	0.21	0.27
100000169418	18	Reading	MC	1	1755	0.00	0.38	0.25	0.22	0.00	0.15	0.38	0.18
100000169419	18	Reading	MC	1	1876	0.00	0.22	0.53	0.12	0.00	0.13	0.53	0.37
100000169953	5	Speaking	CR	1	1877	0.11	0.89	0.00	0.00	0.00	0.00	0.89	0.41
100000169954	6	Speaking	CR	1	1877	0.09	0.91	0.00	0.00	0.00	0.00	0.91	0.34
100000169957	5	Speaking	CR	1	1867	0.23	0.77	0.00	0.00	0.00	0.00	0.77	0.51
100000169958	8	Speaking	CR	2	1877	0.19	0.42	0.39	0.00	0.00	0.00	0.60	0.55
100000169960	8	Speaking	CR	2	1867	0.26	0.40	0.35	0.00	0.00	0.00	0.55	0.62
100000169961	9	Speaking	CR	4	1876	0.18	0.15	0.26	0.15	0.26	0.00	0.54	0.61
100000170522	7	Reading	MC	1	1877	0.00	0.13	0.21	0.65	0.00	0.01	0.65	0.45
100000174275	13	Writing	CR	4	1877	0.54	0.24	0.18	0.03	0.01	0.00	0.18	0.57
100000174565	13	Writing	CR	4	1867	0.59	0.22	0.17	0.02	0.00	0.00	0.16	0.57
100000174618	4	Writing	MC	1	1766	0.00	0.25	0.18	0.55	0.00	0.02	0.55	0.51
100000174620	4	Writing	MC	1	1755	0.00	0.29	0.14	0.55	0.00	0.02	0.55	0.44
100000174621	4	Writing	MC	1	1876	0.00	0.16	0.77	0.04	0.00	0.02	0.77	0.45
100000174624	10	Writing	CR	1	1755	0.39	0.61	0.00	0.00	0.00	0.00	0.61	0.36
100000174625	10	Writing	CR	1	1766	0.30	0.70	0.00	0.00	0.00	0.00	0.70	0.30

DRAFT

Item ID	Item Sequence	Domain	Type	Points	N	0	1/A	2/B	3/C	4/D	Omit/Invalid	P-Value	Item-Total Correlation
100000174628	14	Writing	CR	2	1766	0.74	0.13	0.13	0.00	0.00	0.00	0.20	0.58
100000174630	14	Writing	CR	2	1755	0.81	0.09	0.10	0.00	0.00	0.00	0.14	0.49
100000176048	6	Speaking	CR	1	1867	0.11	0.89	0.00	0.00	0.00	0.00	0.89	0.43
100000177033	4	Reading	MC	1	1766	0.00	0.20	0.74	0.03	0.00	0.02	0.74	0.36
100000187659	7	Speaking	CR	4	1766	0.13	0.17	0.26	0.27	0.16	0.00	0.54	0.65
1000000166558	11	Reading	MC	1	1877	0.00	0.38	0.48	0.13	0.00	0.01	0.48	0.41
1000000174622	10	Writing	CR	1	1876	0.37	0.63	0.00	0.00	0.00	0.00	0.63	0.36
1000000174627	12	Writing	CR	2	1876	0.42	0.14	0.44	0.00	0.00	0.00	0.51	0.57

DRAFT

C.7: Level II – Field Test Items

Item ID	Item Sequence	Domain	Type	Points	N	0	1/A	2/B	3/C	4/D	Omit/Invalid	P-Value	Item-Total Correlation
100000168560	5	Listening	MC	1	4042	0.00	0.13	0.72	0.14	0.00	0.01	0.72	0.29
100000169724	5	Listening	MC	1	4008	0.00	0.73	0.15	0.12	0.00	0.01	0.73	0.27
100000169725	5	Listening	MC	1	3986	0.00	0.82	0.12	0.05	0.00	0.01	0.82	0.37
100000168540	6	Listening	MC	1	4042	0.00	0.06	0.07	0.86	0.00	0.00	0.86	0.37
100000168563	7	Listening	MC	1	3967	0.00	0.80	0.09	0.10	0.00	0.01	0.80	0.38
100000177102	7	Listening	MC	1	4042	0.00	0.06	0.91	0.02	0.00	0.01	0.91	0.39
100000177662	9	Listening	MC	1	4008	0.00	0.07	0.57	0.36	0.00	0.01	0.57	0.23
100000179508	9	Listening	MC	1	3986	0.00	0.12	0.12	0.76	0.00	0.00	0.76	0.12
100000169717	10	Listening	MC	1	4008	0.00	0.61	0.10	0.29	0.00	0.01	0.29	0.04
100000169756	10	Listening	MC	1	3986	0.00	0.16	0.20	0.63	0.00	0.01	0.63	0.21
100000168530	11	Listening	MC	1	3967	0.00	0.33	0.45	0.21	0.00	0.01	0.45	0.18
100000168531	11	Listening	MC	1	4042	0.00	0.05	0.49	0.45	0.00	0.01	0.45	0.28
100000169721	11	Listening	MC	1	4008	0.00	0.45	0.19	0.35	0.00	0.01	0.45	0.34
100000169759	11	Listening	MC	1	3986	0.00	0.10	0.83	0.06	0.00	0.01	0.83	0.44
100000168553	12	Listening	MC	1	3967	0.00	0.08	0.32	0.59	0.00	0.01	0.59	0.25
100000169718	12	Listening	MC	1	4008	0.00	0.54	0.12	0.34	0.00	0.01	0.34	0.06
100000169758	12	Listening	MC	1	3986	0.00	0.30	0.24	0.45	0.00	0.01	0.30	0.13
1000000168352	12	Listening	MC	1	4042	0.00	0.49	0.22	0.28	0.00	0.01	0.49	0.28
100000166557	3	Reading	MC	1	3967	0.00	0.01	0.99	0.00	0.00	0.00	0.99	0.18
100000166596	3	Reading	MC	1	4042	0.00	0.98	0.01	0.01	0.00	0.00	0.98	0.18
100000168467	4	Reading	MC	1	3986	0.00	0.22	0.18	0.58	0.00	0.02	0.58	0.46
100000168498	4	Reading	MC	1	4008	0.00	0.26	0.46	0.26	0.00	0.02	0.46	0.27
100000168725	4	Reading	MC	1	4042	0.00	0.04	0.91	0.04	0.00	0.00	0.91	0.38
100000170521	4	Reading	MC	1	3967	0.00	0.04	0.80	0.16	0.00	0.01	0.80	0.34
100000168466	5	Reading	MC	1	3986	0.00	0.55	0.26	0.17	0.00	0.02	0.55	0.47
100000168501	5	Reading	MC	1	4008	0.00	0.39	0.34	0.24	0.00	0.03	0.39	0.24
100000168468	6	Reading	MC	1	3986	0.00	0.10	0.16	0.72	0.00	0.02	0.72	0.44
100000168716	6	Reading	MC	1	3967	0.00	0.22	0.13	0.64	0.00	0.02	0.64	0.54
100000168775	6	Reading	MC	1	4042	0.00	0.07	0.57	0.34	0.00	0.02	0.57	0.47
100000177642	6	Reading	MC	1	4008	0.00	0.25	0.25	0.46	0.00	0.03	0.46	0.27
100000168470	7	Reading	MC	1	3986	0.00	0.21	0.19	0.58	0.00	0.02	0.58	0.42
100000168502	7	Reading	MC	1	4008	0.00	0.20	0.18	0.60	0.00	0.02	0.60	0.47
100000168715	7	Reading	MC	1	3967	0.00	0.13	0.09	0.76	0.00	0.02	0.76	0.51
100000170532	7	Reading	MC	1	4042	0.00	0.79	0.10	0.09	0.00	0.02	0.79	0.42
100000168471	8	Reading	MC	1	3986	0.00	0.60	0.17	0.21	0.00	0.02	0.60	0.39
100000168718	8	Reading	MC	1	3967	0.00	0.47	0.31	0.20	0.00	0.03	0.47	0.34
100000175031	8	Reading	MC	1	4042	0.00	0.73	0.12	0.13	0.00	0.03	0.73	0.45
100000177643	8	Reading	MC	1	4008	0.00	0.16	0.25	0.57	0.00	0.03	0.57	0.45
100000169039	5	Speaking	CR	1	4042	0.07	0.93	0.00	0.00	0.00	0.00	0.93	0.41
100000174687	5	Speaking	CR	1	4008	0.04	0.96	0.00	0.00	0.00	0.00	0.96	0.32
100000174688	5	Speaking	CR	1	3967	0.05	0.95	0.00	0.00	0.00	0.00	0.95	0.37
100000187606	5	Speaking	CR	1	3986	0.04	0.96	0.00	0.00	0.00	0.00	0.96	0.34
100000169041	8	Speaking	CR	2	4042	0.08	0.34	0.58	0.00	0.00	0.00	0.75	0.52
100000169043	9	Speaking	CR	2	4042	0.12	0.31	0.57	0.00	0.00	0.00	0.73	0.55

DRAFT

Item ID	Item Sequence	Domain	Type	Points	N	0	1/A	2/B	3/C	4/D	Omit/Invalid	P-Value	Item-Total Correlation
100000187662	9	Speaking	CR	4	3967	0.04	0.10	0.26	0.35	0.24	0.00	0.66	0.61
100000138889	10	Speaking	CR	4	4008	0.11	0.17	0.27	0.26	0.18	0.00	0.56	0.69
100000177166	11	Speaking	CR	4	3986	0.05	0.06	0.25	0.20	0.44	0.00	0.73	0.60
100000169017	4	Writing	MC	1	4008	0.00	0.09	0.14	0.73	0.00	0.04	0.73	0.47
100000169018	4	Writing	MC	1	4042	0.00	0.10	0.81	0.08	0.00	0.01	0.81	0.50
100000169021	4	Writing	MC	1	3967	0.00	0.10	0.10	0.78	0.00	0.01	0.78	0.53
100000169022	4	Writing	MC	1	3986	0.00	0.46	0.33	0.19	0.00	0.03	0.46	0.21
100000174263	10	Writing	CR	2	4042	0.18	0.30	0.52	0.00	0.00	0.00	0.67	0.53
100000174264	10	Writing	CR	2	3967	0.12	0.20	0.68	0.00	0.00	0.00	0.78	0.55
100000177056	14	Writing	CR	4	4008	0.11	0.17	0.35	0.25	0.12	0.00	0.52	0.65
100000187554	14	Writing	CR	4	3986	0.07	0.12	0.36	0.28	0.17	0.00	0.59	0.62
100000174265	15	Writing	CR	2	4042	0.22	0.31	0.47	0.00	0.00	0.00	0.63	0.57
100000174266	15	Writing	CR	2	3967	0.28	0.31	0.41	0.00	0.00	0.00	0.57	0.53

DRAFT

C.8: Level III – Field Test Items

Item ID	Item Sequence	Domain	Type	Points	N	0	1/A	2/B	3/C	4/D	Omit/Invalid	P-Value	Item-Total Correlation
3526862	2	Listening	MC	1	4715	0.00	0.60	0.20	0.10	0.11	0.01	0.60	0.29
3526859	3	Listening	MC	1	4715	0.00	0.23	0.17	0.13	0.46	0.01	0.46	0.21
100000174047	6	Listening	MC	1	5546	0.00	0.08	0.06	0.82	0.03	0.00	0.82	0.34
100000174049	7	Listening	MC	1	5546	0.00	0.59	0.09	0.10	0.21	0.01	0.59	0.33
100000169776	12	Listening	MC	1	4715	0.00	0.83	0.11	0.02	0.04	0.01	0.83	0.29
100000176868	12	Listening	MC	1	4820	0.00	0.15	0.65	0.11	0.08	0.00	0.65	0.26
100000174195	14	Listening	MC	1	5546	0.00	0.13	0.13	0.20	0.53	0.00	0.53	0.22
100000174193	15	Listening	MC	1	5546	0.00	0.21	0.52	0.18	0.08	0.00	0.52	0.24
100000176490	15	Listening	MC	1	4820	0.00	0.15	0.69	0.10	0.06	0.00	0.69	0.33
100000200928	15	Listening	MC	1	4715	0.00	0.37	0.44	0.09	0.10	0.00	0.44	0.14
100000174192	16	Listening	MC	1	5546	0.00	0.12	0.65	0.08	0.14	0.01	0.65	0.37
100000200919	16	Listening	MC	1	4820	0.00	0.06	0.03	0.09	0.82	0.00	0.82	0.37
100000200927	16	Listening	MC	1	4715	0.00	0.02	0.32	0.55	0.10	0.00	0.55	0.37
100000169766	17	Listening	MC	1	4715	0.00	0.09	0.32	0.51	0.07	0.01	0.32	0.31
100000169874	17	Listening	MC	1	4820	0.00	0.15	0.04	0.34	0.47	0.00	0.47	0.12
100000167861	5	Reading	MC	1	4715	0.00	0.21	0.13	0.59	0.05	0.01	0.59	0.42
100000169863	5	Reading	MC	1	4820	0.00	0.18	0.46	0.27	0.09	0.01	0.46	0.31
100000200835	9	Reading	MC	1	5546	0.00	0.39	0.18	0.27	0.15	0.00	0.39	0.16
100000169860	10	Reading	MC	1	4715	0.00	0.73	0.05	0.07	0.13	0.01	0.73	0.42
100000169869	10	Reading	MC	1	4820	0.00	0.48	0.21	0.10	0.21	0.00	0.48	0.24
100000200839	10	Reading	MC	1	5546	0.00	0.69	0.07	0.18	0.05	0.00	0.69	0.40
100000169856	11	Reading	MC	1	4820	0.00	0.06	0.12	0.66	0.16	0.01	0.66	0.44
100000169857	11	Reading	MC	1	4715	0.00	0.61	0.17	0.12	0.09	0.01	0.61	0.35
100000200837	11	Reading	MC	1	5546	0.00	0.17	0.49	0.19	0.14	0.01	0.49	0.35
100000169858	12	Reading	MC	1	4715	0.00	0.09	0.28	0.05	0.57	0.01	0.57	0.37
100000176337	12	Reading	MC	1	4820	0.00	0.33	0.17	0.09	0.41	0.01	0.41	0.30
100000200838	12	Reading	MC	1	5546	0.00	0.15	0.22	0.54	0.09	0.01	0.54	0.41
100000169867	13	Reading	MC	1	4820	0.00	0.37	0.14	0.29	0.18	0.01	0.29	0.09
100000169870	13	Reading	MC	1	4715	0.00	0.30	0.13	0.40	0.17	0.01	0.40	0.19
100000200832	13	Reading	MC	1	5546	0.00	0.07	0.06	0.76	0.10	0.01	0.76	0.51
100000169497	5	Speaking	CR	1	4715	0.08	0.92	0.00	0.00	0.00	0.00	0.92	0.43
100000169499	5	Speaking	CR	1	4820	0.16	0.84	0.00	0.00	0.00	0.00	0.84	0.44
100000169910	5	Speaking	CR	1	5546	0.10	0.90	0.00	0.00	0.00	0.00	0.90	0.44
100000187668	9	Speaking	CR	4	4715	0.02	0.05	0.19	0.41	0.33	0.00	0.75	0.62
100000200744	9	Speaking	CR	4	4820	0.05	0.12	0.33	0.32	0.18	0.00	0.62	0.65
100000169911	11	Speaking	CR	4	5546	0.04	0.04	0.18	0.16	0.58	0.00	0.80	0.60
100000174325	9	Writing	MC	1	4820	0.00	0.06	0.76	0.10	0.07	0.01	0.76	0.35
100000187597	9	Writing	MC	1	4715	0.00	0.69	0.16	0.04	0.11	0.01	0.69	0.35
100000174333	11	Writing	CR	2	5546	0.25	0.62	0.13	0.00	0.00	0.00	0.44	0.49
100000177112	13	Writing	CR	4	5546	0.06	0.17	0.45	0.23	0.08	0.00	0.53	0.61
100000017434	14	Writing	CR	4	4820	0.08	0.15	0.47	0.23	0.08	0.00	0.52	0.61
100000177119	14	Writing	CR	4	4715	0.08	0.17	0.46	0.21	0.07	0.00	0.50	0.58

DRAFT

C.9: Level IV – Field Test Items

Item ID	Item Sequence	Domain	Type	Points	N	0	1/A	2/B	3/C	4/D	Omit/Invalid	P-Value	Item-Total Correlation
100000167865	6	Listening	MC	1	3796	0.00	0.81	0.07	0.03	0.09	0.00	0.81	0.34
100000167866	7	Listening	MC	1	3796	0.00	0.07	0.14	0.02	0.77	0.00	0.77	0.23
100000176840	12	Listening	MC	1	3573	0.00	0.34	0.12	0.34	0.20	0.00	0.12	0.00
100000167868	13	Listening	MC	1	3573	0.00	0.72	0.09	0.08	0.10	0.01	0.72	0.34
100000176841	13	Listening	MC	1	3578	0.00	0.17	0.53	0.12	0.17	0.00	0.53	0.29
100000167970	14	Listening	MC	1	3578	0.00	0.05	0.88	0.05	0.02	0.00	0.88	0.44
100000167979	14	Listening	MC	1	3573	0.00	0.07	0.80	0.09	0.03	0.00	0.80	0.42
100000204385	14	Listening	MC	1	3796	0.00	0.13	0.07	0.77	0.02	0.00	0.77	0.39
100000167973	15	Listening	MC	1	3573	0.00	0.31	0.10	0.32	0.27	0.00	0.31	0.18
100000167982	15	Listening	MC	1	3578	0.00	0.70	0.25	0.03	0.02	0.00	0.70	0.38
100000204387	15	Listening	MC	1	3796	0.00	0.71	0.09	0.04	0.16	0.00	0.71	0.45
100000167981	16	Listening	MC	1	3573	0.00	0.67	0.12	0.05	0.16	0.01	0.67	0.35
100000167983	16	Listening	MC	1	3578	0.00	0.08	0.20	0.65	0.07	0.00	0.65	0.28
100000204386	16	Listening	MC	1	3796	0.00	0.26	0.18	0.51	0.05	0.00	0.51	0.31
100000204388	17	Listening	MC	1	3796	0.00	0.05	0.06	0.62	0.26	0.00	0.62	0.34
100000167959	6	Reading	MC	1	3796	0.00	0.70	0.06	0.04	0.20	0.00	0.70	0.51
100000168003	6	Reading	MC	1	3573	0.00	0.73	0.09	0.09	0.08	0.01	0.73	0.43
100000168004	6	Reading	MC	1	3578	0.00	0.08	0.04	0.13	0.74	0.01	0.74	0.45
100000167961	7	Reading	MC	1	3796	0.00	0.02	0.06	0.87	0.05	0.00	0.87	0.42
100000167996	7	Reading	MC	1	3573	0.00	0.61	0.14	0.16	0.08	0.00	0.61	0.42
100000167962	8	Reading	MC	1	3796	0.00	0.29	0.08	0.12	0.50	0.00	0.50	0.32
100000167986	8	Reading	MC	1	3573	0.00	0.16	0.06	0.64	0.14	0.00	0.64	0.43
100000167995	8	Reading	MC	1	3578	0.00	0.08	0.67	0.15	0.10	0.01	0.67	0.47
100000167964	9	Reading	MC	1	3796	0.00	0.10	0.21	0.15	0.53	0.00	0.53	0.36
100000167994	9	Reading	MC	1	3573	0.00	0.11	0.16	0.61	0.12	0.00	0.61	0.46
100000167997	9	Reading	MC	1	3578	0.00	0.37	0.15	0.38	0.09	0.01	0.38	0.30
100000167965	10	Reading	MC	1	3796	0.00	0.58	0.15	0.08	0.18	0.01	0.58	0.38
100000167993	10	Reading	MC	1	3578	0.00	0.06	0.69	0.09	0.15	0.01	0.69	0.48
100000167998	10	Reading	MC	1	3573	0.00	0.29	0.13	0.11	0.47	0.01	0.47	0.24
3540817	6	Speaking	CR	1	3573	0.16	0.84	0.00	0.00	0.00	0.00	0.84	0.58
100000176694	6	Speaking	CR	1	3578	0.09	0.91	0.00	0.00	0.00	0.00	0.91	0.52
1000000174670	6	Speaking	CR	1	3796	0.05	0.95	0.00	0.00	0.00	0.00	0.95	0.43
100000134324	10	Speaking	CR	4	3578	0.03	0.07	0.23	0.34	0.32	0.00	0.71	0.66
100000187670	10	Speaking	CR	4	3573	0.03	0.04	0.16	0.34	0.42	0.00	0.77	0.65
100000176695	12	Speaking	CR	4	3796	0.05	0.02	0.16	0.13	0.65	0.00	0.82	0.59
100000174591	6	Writing	MC	1	3796	0.00	0.41	0.17	0.30	0.12	0.01	0.41	0.23
100000174593	12	Writing	MC	1	3796	0.00	0.05	0.41	0.36	0.18	0.00	0.36	0.31
100000174571	13	Writing	CR	2	3573	0.22	0.21	0.57	0.00	0.00	0.00	0.68	0.49
100000174576	13	Writing	CR	2	3578	0.24	0.31	0.45	0.00	0.00	0.00	0.60	0.42
100000174573	16	Writing	CR	4	3578	0.07	0.10	0.47	0.29	0.07	0.00	0.54	0.63
100000174574	16	Writing	CR	4	3573	0.10	0.13	0.47	0.23	0.07	0.00	0.51	0.63
100000174594	16	Writing	CR	4	3796	0.05	0.11	0.46	0.28	0.09	0.00	0.56	0.66

DRAFT

C.10: Level V – Field Test Items

Item D	Item Sequence	Domain	Type	Points	N	0	1/A	2/B	3/C	4/D	Omit/Invalid	P-Value	Item-Total Correlation
100000009038	14	Listening	MC	1	3667	0.00	0.41	0.18	0.13	0.27	0.01	0.41	0.28
100000134326	10	Speaking	CR	4	3247	0.06	0.08	0.16	0.28	0.42	0.00	0.73	0.79
100000168051	10	Reading	MC	1	3994	0.00	0.74	0.15	0.06	0.04	0.00	0.74	0.44
100000168052	11	Reading	MC	1	3994	0.00	0.14	0.12	0.40	0.33	0.00	0.40	0.26
100000168053	14	Reading	MC	1	3994	0.00	0.55	0.11	0.18	0.14	0.01	0.55	0.36
100000168054	13	Reading	MC	1	3994	0.00	0.06	0.69	0.05	0.19	0.01	0.69	0.36
100000168055	12	Reading	MC	1	3994	0.00	0.07	0.35	0.53	0.05	0.01	0.53	0.39
100000168068	4	Reading	MC	1	3667	0.00	0.10	0.04	0.06	0.79	0.01	0.79	0.46
100000168069	4	Reading	MC	1	3247	0.00	0.64	0.09	0.20	0.07	0.01	0.64	0.40
100000168102	6	Listening	MC	1	3247	0.00	0.62	0.07	0.24	0.07	0.00	0.62	0.39
100000168104	7	Listening	MC	1	3247	0.00	0.19	0.42	0.22	0.16	0.01	0.42	0.34
100000168110	16	Listening	MC	1	3247	0.00	0.05	0.77	0.07	0.11	0.00	0.77	0.58
100000168111	17	Listening	MC	1	3247	0.00	0.14	0.05	0.07	0.73	0.00	0.73	0.47
100000168224	11	Reading	MC	1	3667	0.00	0.55	0.07	0.18	0.19	0.00	0.55	0.28
100000168225	11	Reading	MC	1	3247	0.00	0.22	0.17	0.57	0.04	0.00	0.57	0.30
100000168226	12	Reading	MC	1	3667	0.00	0.07	0.22	0.25	0.46	0.00	0.46	0.33
100000168227	13	Reading	MC	1	3247	0.00	0.15	0.44	0.26	0.14	0.00	0.44	0.23
100000168228	14	Reading	MC	1	3667	0.00	0.62	0.13	0.16	0.09	0.00	0.62	0.53
100000168229	13	Reading	MC	1	3667	0.00	0.17	0.10	0.45	0.26	0.00	0.45	0.27
100000168230	14	Reading	MC	1	3247	0.00	0.66	0.07	0.18	0.09	0.00	0.66	0.39
100000168273	15	Listening	MC	1	3667	0.00	0.19	0.09	0.05	0.67	0.00	0.67	0.43
100000168274	15	Listening	MC	1	3247	0.00	0.12	0.06	0.72	0.10	0.00	0.72	0.53
100000168276	16	Listening	MC	1	3667	0.00	0.05	0.61	0.05	0.29	0.00	0.29	0.17
100000168277	17	Listening	MC	1	3667	0.00	0.06	0.11	0.07	0.75	0.00	0.75	0.51
100000168292	12	Speaking	CR	4	3994	0.04	0.02	0.12	0.09	0.72	0.00	0.86	0.68
100000174583	6	Writing	MC	1	3667	0.00	0.69	0.13	0.10	0.06	0.01	0.69	0.55
100000174584	6	Writing	MC	1	3247	0.00	0.08	0.03	0.78	0.09	0.01	0.78	0.50
100000174585	12	Writing	MC	1	3667	0.00	0.11	0.08	0.73	0.07	0.00	0.73	0.49
100000174599	12	Writing	MC	1	3247	0.00	0.10	0.12	0.63	0.14	0.01	0.63	0.44
100000174601	17	Writing	CR	4	3667	0.09	0.11	0.41	0.29	0.09	0.00	0.54	0.64
100000174606	13	Writing	CR	2	3994	0.13	0.24	0.62	0.00	0.00	0.00	0.75	0.52
100000174718	6	Speaking	CR	1	3247	0.20	0.80	0.00	0.00	0.00	0.00	0.80	0.63
100000174719	6	Speaking	CR	1	3994	0.11	0.89	0.00	0.00	0.00	0.00	0.89	0.55
100000174728	6	Speaking	CR	1	3667	0.21	0.79	0.00	0.00	0.00	0.00	0.79	0.60
100000176753	15	Writing	CR	4	3994	0.05	0.06	0.41	0.36	0.13	0.00	0.62	0.64
100000176771	17	Writing	CR	4	3247	0.11	0.08	0.41	0.31	0.09	0.00	0.55	0.66
100000187673	10	Speaking	CR	4	3667	0.04	0.06	0.12	0.28	0.50	0.00	0.78	0.72
100000197314	14	Listening	MC	1	3247	0.00	0.68	0.11	0.09	0.12	0.00	0.68	0.42
100000198197	15	Listening	MC	1	3994	0.00	0.15	0.08	0.21	0.55	0.01	0.55	0.27
100000198198	16	Listening	MC	1	3994	0.00	0.17	0.10	0.13	0.59	0.01	0.59	0.29
100000198199	17	Listening	MC	1	3994	0.00	0.15	0.15	0.42	0.27	0.01	0.42	0.24
100000200737	13	Listening	MC	1	3994	0.00	0.06	0.09	0.20	0.65	0.00	0.65	0.27
100000200738	14	Listening	MC	1	3994	0.00	0.12	0.70	0.08	0.09	0.01	0.70	0.32
100000202225	6	Listening	MC	1	3667	0.00	0.32	0.16	0.26	0.25	0.00	0.26	0.16
100000202226	7	Listening	MC	1	3667	0.00	0.15	0.60	0.12	0.12	0.01	0.60	0.25

DRAFT

C.11.a: DIF Analysis for Level I Ethnicity Comparison (Asians and Whites)

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000166540	2.84	-0.26	0.15	-1.75	A
100000166541	0.35	-0.11	0.16	-0.67	A
100000166543	0.25	-0.09	0.15	-0.58	A
100000166548	2.07	0.25	0.16	1.53	A
100000166550	2.62	0.30	0.17	1.71	A
100000166552	2.18	-0.24	0.15	-1.55	A
100000166565	4.14	0.35	0.17	2.08	A
100000166566	0.70	0.16	0.17	0.93	A
100000166567	2.68	-0.27	0.16	-1.71	A
100000169217	3.93	-0.79	0.37	-2.11	B
100000169218	1.13	0.44	0.34	1.26	A
100000169219	0.07	0.10	0.26	0.40	A
100000169221	0.01	0.02	0.14	0.17	A
100000169222	10.28	0.47	0.14	3.27	B
100000169223	0.00	0.02	0.18	0.10	A
1000000166558	9.12	-0.49	0.16	-3.08	B
100000166555	0.00	0.09	0.46	0.19	A
100000166595	1.95	0.26	0.17	1.48	A
100000166598	0.15	0.08	0.16	0.48	A
100000166600	19.77	-0.80	0.18	-4.48	C
100000166602	4.30	-0.38	0.17	-2.20	A
100000168722	3.35	-0.57	0.29	-1.96	A
100000168726	0.23	-0.10	0.17	-0.57	A
100000169249	10.50	-0.54	0.16	-3.31	B
100000169309	3.16	0.15	0.08	1.81	A
100000169310	13.13	-0.39	0.11	-3.66	A
100000169417	0.90	-0.12	0.12	-1.01	A
100000169418	0.14	-0.06	0.14	-0.44	A
100000169419	0.02	-0.03	0.15	-0.22	A
100000170522	17.68	-0.75	0.18	-4.25	B
100000177033	1.50	-0.29	0.22	-1.32	A
1000000174622	0.00	-0.01	0.16	-0.09	A
1000000174627	42.10	-1.02	0.16	-6.48	C
100000174275	9.01	-0.44	0.14	-3.04	B
100000174565	0.71	-0.13	0.15	-0.84	A
100000174618	0.86	0.18	0.18	1.01	A
100000174620	1.47	-0.20	0.15	-1.29	A
100000174621	0.92	-0.22	0.20	-1.06	A
100000174624	9.08	-0.48	0.16	-3.08	B
100000174625	8.20	-0.56	0.19	-2.90	B
100000174628	0.67	-0.15	0.18	-0.82	A
100000174630	1.15	-0.20	0.18	-1.08	A
100000133352	15.60	0.51	0.13	3.92	B
100000169953	0.00	0.02	0.27	0.08	A
100000169954	0.84	0.28	0.27	1.04	A
100000169957	0.21	0.13	0.24	0.57	A
100000169958	2.49	0.22	0.14	1.57	A
100000169960	5.35	0.35	0.15	2.31	A
100000169961	11.08	0.44	0.13	3.31	B
100000176048	3.50	-0.69	0.35	-1.97	A
100000187659	7.33	0.39	0.14	2.75	A

DRAFT

C.11.b: DIF Analysis for Level I Ethnicity Comparison (African Americans and Whites)

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000166540	0.63	-0.23	0.26	-0.90	A
100000166541	0.61	-0.25	0.28	-0.92	A
100000166543	1.48	-0.32	0.25	-1.32	A
100000166548	1.93	0.42	0.27	1.55	A
100000166550	0.14	0.14	0.28	0.52	A
100000166552	3.19	-0.45	0.25	-1.85	A
100000166565	0.00	-0.05	0.36	-0.15	A
100000166566	3.06	-0.76	0.40	-1.88	A
100000166567	0.03	0.01	0.37	0.01	A
100000169217	2.57	-2.26	1.27	-1.78	A
100000169218	0.04	-0.23	0.98	-0.23	A
100000169219	0.19	-0.62	0.80	-0.77	A
100000169221	0.30	0.20	0.30	0.68	A
100000169222	0.42	0.25	0.31	0.81	A
100000169223	0.90	0.50	0.41	1.21	A
1000000166558	0.06	0.10	0.27	0.38	A
100000166555	.	.	.	.	.
100000166595	1.68	0.41	0.29	1.42	A
100000166598	1.11	0.47	0.38	1.25	A
100000166600	0.36	0.21	0.28	0.76	A
100000166602	1.15	0.35	0.29	1.22	A
100000168722	1.08	0.70	0.54	1.31	A
100000168726	2.68	0.69	0.38	1.82	A
100000169249	0.01	-0.02	0.34	-0.06	A
100000169309	7.08	0.41	0.15	2.72	A
100000169310	0.57	-0.20	0.24	-0.86	A
100000169417	1.09	0.35	0.29	1.19	A
100000169418	0.04	0.12	0.33	0.36	A
100000169419	1.33	0.41	0.32	1.30	A
100000170522	0.85	-0.31	0.29	-1.08	A
100000177033	1.25	0.36	0.30	1.22	A
1000000174622	1.97	0.52	0.33	1.57	A
1000000174627	0.67	-0.28	0.34	-0.82	A
100000174275	0.17	0.11	0.26	0.43	A
100000174565	3.52	0.82	0.40	2.05	B
100000174618	0.97	0.34	0.30	1.14	A
100000174620	0.19	0.26	0.40	0.66	A
100000174621	0.07	0.22	0.44	0.50	A
100000174624	0.03	0.15	0.39	0.38	A
100000174625	0.31	-0.22	0.31	-0.71	A
100000174628	0.00	0.00	0.32	-0.01	A
100000174630	0.27	-0.23	0.45	-0.52	A
100000133352	0.89	-0.25	0.31	-0.81	A
100000169953	0.00	-0.11	0.50	-0.23	A
100000169954	0.18	-0.34	0.51	-0.65	A
100000169957	0.04	-0.02	0.46	-0.05	A
100000169958	3.47	-0.44	0.24	-1.86	A
100000169960	3.07	-0.55	0.34	-1.65	A
100000169961	3.74	-0.61	0.31	-1.97	A
100000176048	7.16	-1.82	0.88	-2.07	B
100000187659	2.02	-0.36	0.24	-1.50	A

DRAFT

C.11.c: DIF Analysis for Level I Ethnicity Comparison (Hispanics and Whites)

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000166540	0.14	0.05	0.12	0.43	A
100000166541	0.12	-0.05	0.13	-0.41	A
100000166543	0.00	-0.01	0.12	-0.08	A
100000166548	1.22	-0.14	0.12	-1.17	A
100000166550	3.02	0.22	0.12	1.79	A
100000166552	1.25	-0.14	0.12	-1.18	A
100000166565	0.02	-0.02	0.12	-0.21	A
100000166566	0.58	0.10	0.12	0.82	A
100000166567	2.71	-0.19	0.11	-1.70	A
100000169217	18.91	-1.37	0.33	-4.20	C
100000169218	0.01	-0.01	0.30	-0.05	A
100000169219	5.38	-0.59	0.24	-2.42	B
100000169221	0.23	-0.06	0.11	-0.54	A
100000169222	5.57	0.27	0.11	2.42	A
100000169223	1.27	0.17	0.14	1.20	A
1000000166558	1.62	0.17	0.13	1.35	A
100000166555	6.87	0.70	0.26	2.71	B
100000166595	0.20	-0.07	0.15	-0.51	A
100000166598	3.92	-0.23	0.11	-2.04	A
100000166600	0.14	-0.05	0.12	-0.44	A
100000166602	4.23	-0.27	0.13	-2.11	A
100000168722	0.16	-0.10	0.20	-0.51	A
100000168726	1.65	0.16	0.12	1.34	A
100000169249	0.91	0.13	0.13	1.02	A
100000169309	11.48	0.22	0.06	3.42	A
100000169310	1.27	0.11	0.09	1.17	A
100000169417	0.20	0.05	0.10	0.50	A
100000169418	10.01	0.38	0.12	3.18	A
100000169419	0.12	0.05	0.12	0.40	A
100000170522	4.71	-0.29	0.13	-2.23	A
100000177033	8.09	0.40	0.14	2.90	A
1000000174622	4.15	0.26	0.12	2.10	A
1000000174627	14.93	-0.47	0.12	-3.95	B
100000174275	1.56	0.15	0.12	1.26	A
100000174565	1.70	0.15	0.12	1.29	A
100000174618	1.07	-0.15	0.13	-1.10	A
100000174620	1.70	-0.17	0.13	-1.37	A
100000174621	0.05	0.04	0.14	0.29	A
100000174624	5.88	0.30	0.12	2.49	A
100000174625	0.77	0.12	0.13	0.94	A
100000174628	0.54	-0.12	0.17	-0.73	A
100000174630	0.16	-0.07	0.17	-0.40	A
100000133352	0.62	-0.08	0.11	-0.77	A
100000169953	0.01	0.04	0.22	0.19	A
100000169954	0.22	0.13	0.22	0.57	A
100000169957	4.38	0.32	0.15	2.16	A
100000169958	4.79	-0.25	0.12	-2.16	A
100000169960	0.81	0.10	0.11	0.90	A
100000169961	0.34	0.06	0.11	0.57	A
100000176048	6.12	-0.51	0.20	-2.53	B
100000187659	0.00	-0.01	0.11	-0.06	A

DRAFT

C.11.d: DIF Analysis for Level I Gender Comparison

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000166540	2.41	0.16	0.10	1.60	A
100000166541	2.78	-0.18	0.10	-1.73	A
100000166543	0.00	0.01	0.10	0.07	A
100000166548	0.07	0.03	0.10	0.31	A
100000166550	0.04	0.03	0.11	0.24	A
100000166552	2.76	-0.17	0.10	-1.71	A
100000166565	0.28	0.06	0.11	0.58	A
100000166566	1.01	0.11	0.11	1.06	A
100000166567	4.51	0.22	0.10	2.17	A
100000169217	1.07	0.28	0.25	1.13	A
100000169218	0.32	0.17	0.24	0.69	A
100000169219	0.43	-0.15	0.20	-0.75	A
100000169221	1.98	-0.14	0.10	-1.46	A
100000169222	4.05	0.20	0.10	2.06	A
100000169223	2.45	-0.19	0.12	-1.62	A
1000000166558	0.04	0.02	0.10	0.24	A
100000166555	0.00	-0.03	0.24	-0.11	A
100000166595	2.66	-0.20	0.12	-1.68	A
100000166598	0.16	0.05	0.10	0.45	A
100000166600	4.75	-0.24	0.11	-2.25	A
100000166602	3.01	0.20	0.11	1.80	A
100000168722	0.40	-0.13	0.18	-0.72	A
100000168726	0.95	0.11	0.11	1.03	A
100000169249	0.66	-0.10	0.11	-0.87	A
100000169309	4.28	0.12	0.06	2.09	A
100000169310	0.07	0.02	0.08	0.30	A
100000169417	6.81	0.23	0.09	2.65	A
100000169418	4.95	0.24	0.10	2.28	A
100000169419	2.94	-0.18	0.10	-1.77	A
100000170522	2.24	0.17	0.11	1.54	A
100000177033	0.28	0.07	0.12	0.59	A
1000000174622	0.04	-0.03	0.10	-0.26	A
1000000174627	3.75	-0.20	0.10	-1.93	A
100000174275	3.50	-0.19	0.10	-1.90	A
100000174565	4.27	-0.21	0.10	-2.08	A
100000174618	1.16	-0.13	0.11	-1.14	A
100000174620	2.10	0.16	0.11	1.51	A
100000174621	0.51	-0.10	0.13	-0.79	A
100000174624	0.00	-0.01	0.11	-0.10	A
100000174625	0.05	0.03	0.11	0.28	A
100000174628	0.01	-0.01	0.13	-0.08	A
100000174630	9.51	-0.44	0.15	-3.03	B
100000133352	3.18	0.17	0.09	1.80	A
100000169953	1.45	-0.22	0.17	-1.30	A
100000169954	0.70	0.17	0.18	0.93	A
100000169957	0.00	0.01	0.14	0.07	A
100000169958	2.15	-0.14	0.09	-1.47	A
100000169960	0.09	0.03	0.10	0.30	A
100000169961	3.72	0.17	0.09	1.93	A
100000176048	0.12	0.08	0.18	0.44	A
100000187659	0.62	-0.07	0.09	-0.78	A

DRAFT

C.12.a: DIF Analysis for Level II Ethnicity Comparison (Asians and Whites)

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
1000000168352	1.39	0.13	0.11	1.24	A
100000168530	0.00	0.00	0.10	-0.05	A
100000168531	2.45	0.17	0.10	1.62	A
100000168540	0.72	0.10	0.11	0.90	A
100000168553	1.74	-0.14	0.10	-1.37	A
100000168560	3.70	-0.16	0.08	-1.95	A
100000168563	0.48	-0.11	0.14	-0.77	A
100000169717	1.13	-0.13	0.12	-1.12	A
100000169718	0.55	-0.09	0.11	-0.80	A
100000169721	0.43	-0.08	0.11	-0.71	A
100000169724	0.50	0.09	0.12	0.77	A
100000169725	2.12	-0.22	0.14	-1.51	A
100000169756	0.04	-0.03	0.10	-0.26	A
100000169758	3.62	0.20	0.10	1.95	A
100000169759	4.80	-0.34	0.15	-2.26	A
100000177102	1.78	-0.30	0.21	-1.44	A
100000177662	0.87	-0.10	0.11	-0.98	A
100000179508	1.99	-0.16	0.11	-1.46	A
100000166557	1.37	-1.10	0.78	-1.41	A
100000166596	0.93	-0.50	0.44	-1.13	A
100000168466	13.46	-0.42	0.11	-3.71	A
100000168467	0.08	-0.04	0.11	-0.34	A
100000168468	2.18	-0.18	0.12	-1.53	A
100000168470	0.67	-0.09	0.11	-0.86	A
100000168471	6.22	-0.28	0.11	-2.57	A
100000168498	3.61	0.22	0.11	1.96	A
100000168501	0.17	0.05	0.11	0.46	A
100000168502	0.08	-0.04	0.12	-0.34	A
100000168715	1.93	-0.21	0.15	-1.45	A
100000168716	0.68	-0.11	0.13	-0.89	A
100000168718	0.05	-0.03	0.10	-0.28	A
100000168725	0.12	0.09	0.21	0.45	A
100000168775	0.22	0.06	0.12	0.52	A
100000170521	0.15	-0.06	0.14	-0.45	A
100000170532	5.43	-0.33	0.14	-2.40	A
100000175031	4.75	-0.29	0.13	-2.25	A
100000177642	0.39	-0.07	0.11	-0.68	A
100000177643	4.20	-0.26	0.12	-2.10	A
100000169017	3.65	-0.27	0.14	-1.96	A
100000169018	3.25	-0.31	0.17	-1.89	A
100000169021	4.80	-0.36	0.16	-2.23	A
100000169022	30.45	0.53	0.10	5.48	B
100000174263	21.11	-0.52	0.11	-4.60	B
100000174264	3.11	0.21	0.13	1.71	A
100000174265	3.25	-0.19	0.10	-1.78	A
100000174266	4.78	0.22	0.10	2.23	A

DRAFT

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000177056	5.33	-0.22	0.10	-2.28	A
100000187554	6.88	0.24	0.09	2.69	A
100000138889	10.23	0.32	0.10	3.19	A
100000169039	0.54	0.20	0.23	0.85	A
100000169041	0.99	-0.10	0.11	-0.98	A
100000169043	0.00	-0.01	0.11	-0.07	A
100000174687	0.00	-0.07	0.39	-0.17	A
100000174688	0.13	-0.14	0.28	-0.49	A
100000177166	10.74	0.30	0.09	3.25	A
100000187606	2.54	0.48	0.28	1.75	A
100000187662	14.49	-0.36	0.09	-3.88	A

DRAFT

C.12.b: DIF Analysis for Level II Ethnicity Comparison (African Americans and Whites)

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
1000000168352	0.15	-0.09	0.19	-0.48	A
100000168530	0.50	-0.15	0.19	-0.81	A
100000168531	2.85	0.34	0.19	1.78	A
100000168540	2.36	0.30	0.18	1.63	A
100000168553	2.50	-0.35	0.21	-1.72	A
100000168560	0.00	-0.01	0.14	-0.06	A
100000168563	0.04	-0.08	0.25	-0.32	A
100000169717	0.01	0.02	0.31	0.07	A
100000169718	0.75	-0.27	0.27	-1.01	A
100000169721	0.03	-0.08	0.28	-0.30	A
100000169724	0.00	-0.05	0.30	-0.18	A
100000169725	1.88	0.38	0.26	1.45	A
100000169756	0.00	-0.03	0.21	-0.15	A
100000169758	1.36	-0.25	0.20	-1.24	A
100000169759	0.60	0.27	0.29	0.93	A
100000177102	0.61	0.28	0.29	0.96	A
100000177662	0.02	0.07	0.27	0.28	A
100000179508	0.08	0.09	0.23	0.39	A
100000166557	0.00	0.34	0.78	0.44	A
100000166596	0.00	0.12	0.48	0.24	A
100000168466	0.00	-0.03	0.23	-0.14	A
100000168467	0.47	0.18	0.23	0.80	A
100000168468	0.59	-0.21	0.24	-0.86	A
100000168470	0.67	-0.20	0.22	-0.91	A
100000168471	0.31	0.15	0.22	0.68	A
100000168498	0.00	0.05	0.28	0.20	A
100000168501	0.89	0.30	0.28	1.07	A
100000168502	0.01	0.02	0.30	0.08	A
100000168715	0.62	0.21	0.23	0.89	A
100000168716	0.18	0.12	0.22	0.54	A
100000168718	0.00	-0.02	0.20	-0.10	A
100000168725	0.00	0.06	0.32	0.18	A
100000168775	6.28	0.54	0.21	2.59	B
100000170521	0.01	0.05	0.24	0.22	A
100000170532	0.00	-0.04	0.22	-0.16	A
100000175031	0.10	0.09	0.21	0.43	A
100000177642	0.00	0.06	0.28	0.20	A
100000177643	0.30	-0.20	0.29	-0.70	A
100000169017	0.20	0.21	0.34	0.64	A
100000169018	3.33	0.44	0.23	1.90	A
100000169021	0.00	-0.05	0.26	-0.18	A
100000169022	0.1316	0.0965	0.2058	0.4689	A
100000174263	0.087	-0.057	0.192	-0.297	A
100000174264	9.314	0.656	0.211	3.109	B
100000174265	0.59	0.141	0.185	0.762	A

DRAFT

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000174266	3.339	0.341	0.184	1.853	A
100000177056	6.551	0.629	0.253	2.486	B
100000187554	2.832	0.324	0.188	1.723	A
100000138889	0.151	-0.085	0.241	-0.353	A
100000169039	1.4447	-0.5137	0.3892	-1.3199	A
100000169041	4.523	-0.403	0.187	-2.155	A
100000169043	1.611	-0.24	0.191	-1.257	A
100000174687	0.0898	0.664	0.8407	0.7898	A
100000174688	2.1202	-1.4278	0.8816	-1.6196	A
100000177166	6.737	-0.523	0.204	-2.564	B
100000187606	0.4735	0.5385	0.556	0.9685	A
100000187662	14.774	-0.732	0.187	-3.914	B

DRAFT

C.12.c: DIF Analysis for Level II Ethnicity Comparison (Hispanics and Whites)

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
1000000168352	9.31	0.24	0.08	3.09	A
100000168530	0.00	0.00	0.07	-0.05	A
100000168531	4.24	0.16	0.08	2.09	A
100000168540	0.00	0.01	0.08	0.06	A
100000168553	1.61	0.10	0.08	1.31	A
100000168560	0.00	0.00	0.06	0.03	A
100000168563	2.12	-0.15	0.10	-1.51	A
100000169717	25.22	-0.42	0.08	-5.04	A
100000169718	17.14	-0.33	0.08	-4.16	A
100000169721	0.00	0.00	0.08	-0.03	A
100000169724	3.39	-0.16	0.09	-1.88	A
100000169725	1.44	0.13	0.10	1.25	A
100000169756	0.21	0.04	0.08	0.50	A
100000169758	1.89	0.11	0.08	1.42	A
100000169759	1.54	-0.14	0.11	-1.29	A
100000177102	2.74	-0.26	0.15	-1.73	A
100000177662	9.53	-0.24	0.08	-3.12	A
100000179508	7.27	-0.24	0.09	-2.74	A
100000166557	0.00	-0.03	0.31	-0.09	A
100000166596	1.32	-0.38	0.29	-1.29	A
100000168466	0.32	0.05	0.08	0.61	A
100000168467	0.68	0.07	0.08	0.86	A
100000168468	6.51	-0.23	0.09	-2.60	A
100000168470	1.93	0.12	0.08	1.43	A
100000168471	1.22	0.09	0.08	1.15	A
100000168498	0.00	0.01	0.08	0.08	A
100000168501	3.65	0.15	0.08	1.94	A
100000168502	4.29	-0.18	0.08	-2.11	A
100000168715	0.82	0.09	0.10	0.95	A
100000168716	0.35	-0.06	0.09	-0.64	A
100000168718	8.24	0.23	0.08	2.91	A
100000168725	1.86	0.23	0.16	1.44	A
100000168775	7.84	0.25	0.09	2.84	A
100000170521	1.99	0.14	0.09	1.45	A
100000170532	4.45	-0.21	0.10	-2.15	A
100000175031	4.18	-0.20	0.09	-2.09	A
100000177642	0.00	0.00	0.08	-0.01	A
100000177643	1.53	-0.11	0.08	-1.28	A
100000169017	0.14	0.04	0.09	0.42	A
100000169018	0.03	0.02	0.11	0.22	A
100000169021	4.00	0.20	0.10	2.03	A
100000169022	17.5825	0.3187	0.0752	4.238	A
100000174263	2.386	0.123	0.08	1.537	A
100000174264	44.936	0.574	0.086	6.674	B
100000174265	0.997	0.077	0.077	1	A
100000174266	29.396	0.388	0.072	5.389	A

DRAFT

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000177056	1.598	0.087	0.069	1.261	A
100000187554	78.211	0.623	0.07	8.9	B
100000138889	6.201	0.174	0.07	2.486	A
100000169039	0.1977	-0.0953	0.1779	-0.5357	A
100000169041	19.66	-0.356	0.081	-4.395	A
100000169043	2.22	-0.12	0.081	-1.481	A
100000174687	8.9941	0.7792	0.2563	3.0402	B
100000174688	0.0027	-0.0068	0.1794	-0.0379	A
100000177166	30.4	-0.401	0.073	-5.493	A
100000187606	4.5892	0.4711	0.2105	2.238	B
100000187662	27.983	-0.365	0.069	-5.29	A

DRAFT

C.12.d: DIF Analysis for Level II Gender Comparison

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
1000000168352	1.46	0.08	0.07	1.24	A
100000168530	0.86	0.06	0.07	0.96	A
100000168531	31.89	0.39	0.07	5.65	A
100000168540	1.59	-0.10	0.07	-1.30	A
100000168553	17.55	0.29	0.07	4.22	A
100000168560	0.13	0.02	0.05	0.39	A
100000168563	0.00	0.00	0.09	-0.04	A
100000169717	8.90	0.22	0.07	3.02	A
100000169718	0.01	0.01	0.07	0.15	A
100000169721	10.53	0.23	0.07	3.27	A
100000169724	0.00	0.00	0.08	0.01	A
100000169725	1.68	0.12	0.09	1.34	A
100000169756	0.36	-0.04	0.07	-0.63	A
100000169758	2.99	0.13	0.07	1.76	A
100000169759	5.05	0.22	0.10	2.30	A
100000177102	0.23	0.07	0.13	0.54	A
100000177662	8.12	-0.19	0.07	-2.88	A
100000179508	1.40	0.09	0.08	1.22	A
100000166557	0.22	0.20	0.31	0.63	A
100000166596	0.97	0.28	0.25	1.12	A
100000168466	0.00	0.01	0.08	0.09	A
100000168467	0.03	-0.02	0.07	-0.21	A
100000168468	0.50	-0.06	0.08	-0.74	A
100000168470	10.34	-0.23	0.07	-3.25	A
100000168471	9.58	0.23	0.07	3.12	A
100000168498	1.86	-0.09	0.07	-1.40	A
100000168501	13.98	0.26	0.07	3.77	A
100000168502	0.54	-0.06	0.07	-0.77	A
100000168715	1.06	-0.09	0.09	-1.07	A
100000168716	0.01	-0.01	0.08	-0.15	A
100000168718	0.01	-0.01	0.07	-0.13	A
100000168725	5.62	-0.33	0.13	-2.43	A
100000168775	0.47	-0.05	0.07	-0.73	A
100000170521	4.29	0.18	0.09	2.12	A
100000170532	1.39	-0.11	0.09	-1.22	A
100000175031	9.38	-0.25	0.08	-3.09	A
100000177642	0.43	-0.05	0.07	-0.69	A
100000177643	0.00	0.00	0.07	0.04	A
100000169017	3.29	-0.16	0.08	-1.86	A
100000169018	0.81	0.09	0.10	0.95	A
100000169021	0.82	0.09	0.09	0.95	A
100000169022	0.34	-0.04	0.07	-0.62	A
100000174263	0.22	-0.03	0.07	-0.47	A
100000174264	0.55	-0.06	0.08	-0.74	A
100000174265	23.70	-0.32	0.07	-4.89	A
100000174266	11.13	-0.22	0.06	-3.36	A

DRAFT

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000177056	40.27	-0.39	0.06	-6.31	A
100000187554	54.66	-0.46	0.06	-7.40	B
100000138889	0.01	0.01	0.06	0.08	A
100000169039	0.38	-0.10	0.15	-0.69	A
100000169041	1.42	-0.08	0.07	-1.19	A
100000169043	1.27	-0.08	0.07	-1.13	A
100000174687	1.36	0.24	0.19	1.26	A
100000174688	2.55	0.28	0.17	1.68	A
100000177166	0.12	0.02	0.06	0.36	A
100000187606	0.00	0.03	0.18	0.15	A
100000187662	6.74	0.16	0.06	2.61	A

DRAFT

C.13.a: DIF Analysis for Level III Ethnicity Comparison (Asians and Whites)

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000169766	1.87	-0.16	0.11	-1.41	A
100000169776	0.61	0.11	0.13	0.84	A
100000169874	0.10	-0.03	0.10	-0.36	A
100000174047	0.01	0.02	0.11	0.14	A
100000174049	1.33	-0.11	0.09	-1.20	A
100000174192	0.20	-0.05	0.09	-0.50	A
100000174193	3.79	-0.17	0.08	-1.99	A
100000174195	0.62	0.07	0.08	0.83	A
100000176490	3.69	0.22	0.11	1.97	A
100000176868	5.50	0.25	0.11	2.40	A
100000200919	1.17	-0.16	0.14	-1.15	A
100000200927	0.01	-0.01	0.11	-0.13	A
100000200928	0.52	0.08	0.10	0.77	A
3526859	20.84	-0.32	0.07	-4.58	A
3526862	4.86	0.16	0.07	2.24	A
100000167861	0.31	0.07	0.11	0.61	A
100000169856	1.61	-0.15	0.12	-1.33	A
100000169857	0.19	-0.05	0.11	-0.49	A
100000169858	0.46	0.08	0.11	0.73	A
100000169860	0.59	-0.11	0.13	-0.84	A
100000169863	0.43	0.07	0.10	0.71	A
100000169867	0.58	-0.09	0.11	-0.81	A
100000169869	0.53	0.08	0.10	0.78	A
100000169870	2.94	0.19	0.11	1.77	A
100000176337	0.87	0.10	0.10	0.98	A
100000200832	3.34	-0.22	0.12	-1.88	A
100000200835	11.74	-0.29	0.08	-3.46	A
100000200837	0.25	-0.05	0.09	-0.55	A
100000200838	12.42	-0.33	0.09	-3.57	A
100000200839	12.51	-0.37	0.10	-3.58	A
100000017434	7.24	-0.25	0.09	-2.70	A
100000174325	0.93	-0.13	0.13	-1.02	A
100000174333	9.77	-0.27	0.09	-3.11	A
100000177112	7.67	0.22	0.08	2.76	A
100000177119	0.20	0.04	0.10	0.44	A
100000187597	5.95	0.28	0.11	2.46	A
100000169497	0.45	0.14	0.19	0.74	A
100000169499	4.20	0.35	0.16	2.12	A
100000169910	13.15	0.52	0.14	3.63	B
100000169911	64.61	0.67	0.09	7.92	C
100000187668	4.64	0.20	0.10	2.12	A
100000200744	2.73	0.15	0.09	1.69	A

DRAFT

C.13.b: DIF Analysis for Level III Ethnicity Comparison (African Americans and Whites)

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000169766	0.90	-0.23	0.21	-1.07	A
100000169776	1.28	-0.29	0.25	-1.20	A
100000169874	0.22	-0.10	0.17	-0.55	A
100000174047	5.69	0.51	0.21	2.47	B
100000174049	1.38	-0.23	0.18	-1.27	A
100000174192	1.04	-0.21	0.19	-1.11	A
100000174193	0.32	0.11	0.17	0.66	A
100000174195	0.19	-0.09	0.17	-0.51	A
100000176490	0.45	0.15	0.20	0.78	A
100000176868	0.17	0.10	0.20	0.52	A
100000200919	2.98	-0.50	0.27	-1.84	A
100000200927	0.01	0.00	0.20	-0.02	A
100000200928	0.10	-0.08	0.19	-0.41	A
3526859	1.85	0.18	0.13	1.42	A
3526862	2.04	0.20	0.13	1.49	A
100000167861	4.20	0.46	0.21	2.18	B
100000169856	0.00	-0.02	0.20	-0.11	A
100000169857	0.28	0.13	0.20	0.63	A
100000169858	1.47	0.25	0.20	1.30	A
100000169860	0.85	0.24	0.22	1.08	A
100000169863	0.21	-0.10	0.18	-0.55	A
100000169867	0.50	0.16	0.20	0.81	A
100000169869	0.51	0.14	0.18	0.80	A
100000169870	0.10	-0.08	0.19	-0.42	A
100000176337	2.95	-0.34	0.19	-1.82	A
100000200832	0.23	0.14	0.23	0.60	A
100000200835	0.32	-0.11	0.17	-0.66	A
100000200837	1.10	0.20	0.18	1.14	A
100000200838	0.15	-0.09	0.19	-0.48	A
100000200839	1.24	0.23	0.19	1.21	A
100000017434	1.79	0.23	0.17	1.38	A
100000174325	3.82	-0.50	0.24	-2.06	B
100000174333	0.59	-0.14	0.18	-0.77	A
100000177112	0.36	0.10	0.16	0.62	A
100000177119	3.55	0.32	0.18	1.77	A
100000187597	0.51	-0.16	0.21	-0.79	A
100000169497	3.37	-0.85	0.43	-1.97	A
100000169499	0.18	-0.18	0.31	-0.57	A
100000169910	1.84	-0.67	0.43	-1.56	A
100000169911	11.93	-0.68	0.20	-3.44	B
100000187668	0.36	0.10	0.17	0.60	A
100000200744	0.17	-0.07	0.17	-0.40	A

DRAFT

C.13.c: DIF Analysis for Level III Ethnicity Comparison (Hispanics and Whites)

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000169766	12.90	-0.28	0.08	-3.62	A
100000169776	17.99	-0.40	0.09	-4.29	A
100000169874	5.63	-0.16	0.07	-2.40	A
100000174047	3.73	-0.17	0.09	-1.98	A
100000174049	2.49	-0.11	0.07	-1.61	A
100000174192	0.42	-0.05	0.07	-0.68	A
100000174193	2.20	-0.10	0.06	-1.51	A
100000174195	2.41	-0.10	0.06	-1.58	A
100000176490	2.07	-0.11	0.08	-1.47	A
100000176868	21.18	0.33	0.07	4.62	A
100000200919	9.36	-0.28	0.09	-3.09	A
100000200927	0.63	0.06	0.07	0.83	A
100000200928	3.03	-0.12	0.07	-1.77	A
3526859	5.78	0.12	0.05	2.42	A
3526862	1.59	0.06	0.05	1.28	A
100000167861	2.35	0.12	0.07	1.57	A
100000169856	2.68	-0.13	0.08	-1.67	A
100000169857	0.88	-0.07	0.07	-0.98	A
100000169858	3.04	-0.13	0.07	-1.77	A
100000169860	0.34	0.05	0.08	0.63	A
100000169863	7.71	0.20	0.07	2.81	A
100000169867	1.35	-0.09	0.07	-1.20	A
100000169869	1.64	-0.09	0.07	-1.31	A
100000169870	0.01	-0.01	0.07	-0.15	A
100000176337	3.62	-0.14	0.07	-1.94	A
100000200832	0.08	0.03	0.08	0.32	A
100000200835	2.57	-0.10	0.06	-1.63	A
100000200837	6.61	0.17	0.07	2.59	A
100000200838	0.06	-0.02	0.07	-0.27	A
100000200839	0.01	0.01	0.07	0.12	A
100000017434	3.92	-0.13	0.07	-1.97	A
100000174325	1.71	0.11	0.08	1.35	A
100000174333	0.11	0.02	0.07	0.32	A
100000177112	25.69	0.30	0.06	5.05	A
100000177119	3.05	0.11	0.06	1.73	A
100000187597	7.74	0.22	0.08	2.82	A
100000169497	3.96	-0.30	0.15	-2.03	A
100000169499	13.86	0.40	0.11	3.76	A
100000169910	5.77	-0.32	0.13	-2.46	A
100000169911	19.09	-0.29	0.07	-4.34	A
100000187668	1.50	-0.08	0.07	-1.22	A
100000200744	0.15	0.03	0.06	0.40	A

DRAFT

C.13.d: DIF Analysis for Level III Gender Comparison

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000169766	0.55	-0.05	0.07	-0.78	A
100000169776	4.56	-0.18	0.08	-2.17	A
100000169874	18.89	0.26	0.06	4.37	A
100000174047	0.42	-0.05	0.08	-0.69	A
100000174049	8.09	-0.17	0.06	-2.87	A
100000174192	0.00	0.01	0.06	0.09	A
100000174193	0.48	0.04	0.06	0.72	A
100000174195	0.31	0.03	0.06	0.59	A
100000176490	0.21	0.03	0.07	0.50	A
100000176868	0.69	-0.05	0.06	-0.86	A
100000200919	0.00	0.00	0.08	-0.04	A
100000200927	3.27	-0.12	0.06	-1.84	A
100000200928	12.77	0.22	0.06	3.60	A
3526859	20.44	0.19	0.04	4.54	A
3526862	0.34	0.03	0.04	0.61	A
100000167861	14.11	0.25	0.07	3.78	A
100000169856	1.43	-0.08	0.07	-1.23	A
100000169857	0.65	0.05	0.06	0.84	A
100000169858	0.54	-0.05	0.06	-0.76	A
100000169860	3.68	0.14	0.07	1.96	A
100000169863	14.17	0.24	0.06	3.79	A
100000169867	1.49	-0.08	0.06	-1.26	A
100000169869	1.48	-0.08	0.06	-1.25	A
100000169870	0.65	-0.05	0.06	-0.83	A
100000176337	8.43	-0.18	0.06	-2.94	A
100000200832	26.78	-0.39	0.07	-5.19	A
100000200835	2.90	-0.10	0.06	-1.73	A
100000200837	0.44	-0.04	0.06	-0.69	A
100000200838	40.31	-0.38	0.06	-6.37	A
100000200839	0.73	-0.06	0.06	-0.89	A
100000017434	145.89	-0.71	0.06	-11.98	C
100000174325	6.97	-0.20	0.07	-2.67	A
100000174333	5.63	-0.14	0.06	-2.38	A
100000177112	117.68	-0.57	0.05	-10.79	B
100000177119	147.99	-0.72	0.06	-12.12	C
100000187597	0.03	-0.01	0.07	-0.20	A
100000169497	0.39	-0.09	0.13	-0.69	A
100000169499	5.82	0.22	0.09	2.46	A
100000169910	0.32	-0.07	0.11	-0.62	A
100000169911	0.68	-0.05	0.06	-0.83	A
100000187668	0.02	0.01	0.06	0.14	A
100000200744	56.51	0.42	0.06	7.55	A

DRAFT

C.14.a: DIF Analysis for Level IV Ethnicity Comparison (Asians and Whites)

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000167865	0.23	-0.07	0.13	-0.55	A
100000167866	2.09	-0.18	0.12	-1.50	A
100000167868	5.19	-0.24	0.10	-2.33	A
100000167970	3.01	-0.38	0.21	-1.82	A
100000167973	1.39	-0.16	0.13	-1.24	A
100000167979	3.42	0.33	0.17	1.95	A
100000167981	6.80	0.37	0.14	2.65	A
100000167982	2.13	-0.23	0.15	-1.53	A
100000167983	0.56	0.11	0.13	0.82	A
100000176840	0.06	-0.06	0.18	-0.33	A
100000176841	0.37	-0.09	0.13	-0.67	A
100000204385	0.24	0.07	0.13	0.55	A
100000204386	4.28	0.22	0.10	2.12	A
100000204387	0.75	-0.11	0.12	-0.92	A
100000204388	6.54	0.29	0.11	2.60	A
100000167959	0.63	-0.11	0.13	-0.86	A
100000167961	1.28	-0.20	0.16	-1.20	A
100000167962	4.57	-0.23	0.10	-2.19	A
100000167964	2.51	0.17	0.11	1.63	A
100000167965	0.01	0.01	0.11	0.14	A
100000167986	5.64	-0.36	0.15	-2.42	A
100000167993	0.00	-0.02	0.15	-0.10	A
100000167994	4.29	-0.32	0.15	-2.12	A
100000167995	2.14	-0.24	0.16	-1.53	A
100000167996	0.06	-0.03	0.10	-0.29	A
100000167997	3.56	0.26	0.14	1.94	A
100000167998	8.98	-0.40	0.13	-3.09	A
100000168003	1.62	0.21	0.16	1.32	A
100000168004	0.00	-0.01	0.16	-0.07	A
100000174571	3.13	-0.25	0.14	-1.75	A
100000174573	0.12	-0.04	0.13	-0.35	A
100000174574	2.82	0.21	0.12	1.69	A
100000174576	7.22	-0.34	0.13	-2.69	A
100000174591	10.15	-0.33	0.10	-3.24	A
100000174593	14.38	0.43	0.11	3.84	A
100000174594	0.89	0.09	0.09	0.97	A
100000134324	1.39	0.15	0.12	1.21	A
1000000174670	7.56	0.66	0.24	2.79	B
100000176694	4.39	0.55	0.25	2.20	B
100000176695	45.28	0.75	0.11	6.72	C
100000187670	3.00	-0.22	0.13	-1.72	A
3540817	9.58	0.69	0.22	3.18	B

DRAFT

C.14.b: DIF Analysis for Level IV Ethnicity Comparison (African Americans and Whites)

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000167865	0.72	-0.25	0.26	-0.96	A
100000167866	0.00	-0.04	0.23	-0.18	A
100000167868	0.25	0.08	0.14	0.58	A
100000167970	0.86	-0.30	0.28	-1.08	A
100000167973	9.18	0.84	0.27	3.14	B
100000167979	1.99	0.37	0.24	1.52	A
100000167981	1.23	0.24	0.21	1.17	A
100000167982	0.26	0.12	0.20	0.61	A
100000167983	0.07	0.07	0.19	0.35	A
100000176840	1.56	-0.37	0.27	-1.40	A
100000176841	3.42	0.36	0.19	1.94	A
100000204385	1.70	-0.41	0.28	-1.45	A
100000204386	0.65	-0.19	0.21	-0.92	A
100000204387	0.72	0.23	0.23	0.97	A
100000204388	2.97	0.39	0.21	1.82	A
100000167959	0.94	-0.28	0.26	-1.07	A
100000167961	0.01	-0.03	0.31	-0.08	A
100000167962	0.06	-0.07	0.21	-0.35	A
100000167964	0.18	-0.12	0.22	-0.54	A
100000167965	0.12	0.09	0.21	0.44	A
100000167986	0.84	-0.24	0.23	-1.03	A
100000167993	0.03	0.06	0.21	0.28	A
100000167994	1.49	0.30	0.23	1.33	A
100000167995	5.88	0.51	0.20	2.48	B
100000167996	0.05	0.04	0.15	0.30	A
100000167997	0.15	-0.10	0.20	-0.49	A
100000167998	1.21	-0.25	0.20	-1.22	A
100000168003	0.00	0.02	0.24	0.08	A
100000168004	0.60	0.18	0.21	0.86	A
100000174571	1.32	0.23	0.21	1.14	A
100000174573	5.35	0.42	0.18	2.38	A
100000174574	3.31	0.35	0.19	1.82	A
100000174576	2.52	0.28	0.17	1.63	A
100000174591	1.96	0.32	0.21	1.51	A
100000174593	2.15	0.33	0.22	1.55	A
100000174594	0.00	0.01	0.18	0.03	A
100000134324	3.93	-0.35	0.17	-2.07	A
1000000174670	0.00	-0.15	0.53	-0.28	A
100000176694	2.97	-0.76	0.40	-1.88	A
100000176695	0.09	-0.07	0.24	-0.29	A
100000187670	8.29	-0.54	0.20	-2.71	B
3540817	1.37	0.41	0.31	1.29	A

DRAFT

C.14.c: DIF Analysis for Level IV Ethnicity Comparison (Hispanics and Whites)

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000167865	29.33	-0.55	0.10	-5.41	B
100000167866	0.60	-0.08	0.09	-0.82	A
100000167868	42.24	-0.42	0.06	-6.49	A
100000167970	5.96	-0.33	0.13	-2.50	A
100000167973	9.80	0.26	0.08	3.16	A
100000167979	0.00	-0.01	0.10	-0.06	A
100000167981	3.04	0.15	0.08	1.78	A
100000167982	0.13	0.04	0.09	0.41	A
100000167983	0.09	-0.03	0.09	-0.34	A
100000176840	1.51	0.15	0.12	1.29	A
100000176841	0.99	0.09	0.08	1.03	A
100000204385	0.56	0.08	0.10	0.79	A
100000204386	0.00	0.00	0.08	0.00	A
100000204387	1.56	-0.12	0.09	-1.29	A
100000204388	2.70	0.14	0.08	1.69	A
100000167959	2.16	0.15	0.10	1.52	A
100000167961	0.38	-0.09	0.13	-0.69	A
100000167962	6.87	0.21	0.08	2.64	A
100000167964	9.92	0.26	0.08	3.18	A
100000167965	0.38	0.06	0.08	0.66	A
100000167986	0.32	-0.05	0.08	-0.61	A
100000167993	5.03	-0.22	0.10	-2.28	A
100000167994	5.76	0.21	0.08	2.43	A
100000167995	6.78	0.26	0.10	2.65	A
100000167996	3.45	0.12	0.06	1.89	A
100000167997	0.08	0.03	0.09	0.33	A
100000167998	5.03	-0.17	0.08	-2.28	A
100000168003	5.26	0.22	0.09	2.34	A
100000168004	0.35	0.07	0.10	0.64	A
100000174571	11.38	0.26	0.08	3.38	A
100000174573	2.56	-0.13	0.08	-1.57	A
100000174574	63.57	0.59	0.07	7.95	B
100000174576	5.90	0.19	0.08	2.41	A
100000174591	6.06	0.20	0.08	2.49	A
100000174593	2.17	0.12	0.08	1.51	A
100000174594	0.21	0.03	0.07	0.46	A
100000134324	4.69	-0.17	0.08	-2.15	A
1000000174670	2.37	-0.37	0.23	-1.62	A
100000176694	20.23	-0.83	0.18	-4.58	C
100000176695	2.19	0.13	0.09	1.48	A
100000187670	54.29	-0.54	0.07	-7.35	B
3540817	15.98	0.53	0.13	4.01	B

DRAFT

C.14.d: DIF Analysis for Level IV Gender Comparison

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000167865	2.15	-0.14	0.09	-1.52	A
100000167866	0.43	0.06	0.08	0.70	A
100000167868	1.10	-0.06	0.06	-1.08	A
100000167970	1.13	0.13	0.12	1.12	A
100000167973	0.04	-0.02	0.08	-0.23	A
100000167979	19.66	0.43	0.10	4.46	A
100000167981	0.00	0.00	0.08	0.06	A
100000167982	0.09	0.03	0.08	0.34	A
100000167983	21.01	0.34	0.07	4.61	A
100000176840	13.56	0.40	0.11	3.70	A
100000176841	9.61	0.22	0.07	3.13	A
100000204385	78.43	0.77	0.09	8.76	C
100000204386	0.82	0.07	0.07	0.94	A
100000204387	0.18	0.04	0.08	0.47	A
100000204388	0.12	0.03	0.07	0.38	A
100000167959	8.12	0.25	0.08	2.89	A
100000167961	2.07	-0.17	0.11	-1.51	A
100000167962	1.73	-0.09	0.07	-1.35	A
100000167964	0.02	-0.01	0.07	-0.17	A
100000167965	0.13	0.03	0.07	0.39	A
100000167986	7.22	0.22	0.08	2.72	A
100000167993	0.01	0.01	0.09	0.15	A
100000167994	0.33	0.05	0.08	0.61	A
100000167995	15.12	0.33	0.08	3.92	A
100000167996	39.40	0.35	0.06	6.27	A
100000167997	46.77	0.51	0.08	6.81	B
100000167998	21.81	0.34	0.07	4.69	A
100000168003	2.28	0.13	0.09	1.55	A
100000168004	1.48	-0.11	0.09	-1.26	A
100000174571	19.67	-0.33	0.07	-4.42	A
100000174573	168.24	-0.93	0.07	-12.86	C
100000174574	90.42	-0.66	0.07	-9.54	C
100000174576	0.12	-0.02	0.07	-0.34	A
100000174591	2.44	-0.11	0.07	-1.59	A
100000174593	3.44	0.14	0.07	1.89	A
100000174594	92.41	-0.63	0.07	-9.61	B
100000134324	45.32	0.45	0.07	6.67	B
1000000174670	0.51	-0.15	0.18	-0.80	A
100000176694	0.15	0.07	0.16	0.46	A
100000176695	0.00	0.00	0.08	0.04	A
100000187670	6.06	0.17	0.07	2.49	A
3540817	3.93	0.25	0.12	2.05	A

DRAFT

C.15.a: DIF Analysis for Level V Ethnicity Comparison (Asians and Whites)

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000009038	0.11	0.03	0.08	0.37	A
100000168102	7.29	-0.36	0.13	-2.76	A
100000168104	4.44	-0.26	0.12	-2.16	A
100000168110	1.34	-0.22	0.18	-1.26	A
100000168111	0.79	-0.14	0.15	-0.97	A
100000168273	0.07	-0.04	0.13	-0.33	A
100000168274	0.99	-0.17	0.16	-1.08	A
100000168276	4.45	0.26	0.12	2.17	A
100000168277	4.45	-0.32	0.15	-2.14	A
100000197314	0.00	0.00	0.13	-0.03	A
100000198197	5.68	0.23	0.10	2.43	A
100000198198	13.74	0.37	0.10	3.72	A
100000198199	10.60	-0.32	0.10	-3.30	A
100000200737	0.00	0.01	0.10	0.09	A
100000200738	7.80	0.29	0.10	2.83	A
100000202225	4.63	-0.28	0.12	-2.22	A
100000202226	18.66	-0.51	0.12	-4.35	B
100000168051	4.80	0.26	0.12	2.24	A
100000168052	3.42	0.19	0.10	1.89	A
100000168053	20.17	0.46	0.10	4.53	B
100000168054	4.32	-0.22	0.11	-2.11	A
100000168055	4.65	-0.23	0.10	-2.20	A
100000168068	3.19	-0.28	0.15	-1.85	A
100000168069	8.70	0.39	0.13	2.97	A
100000168224	21.12	0.53	0.11	4.65	B
100000168225	3.69	-0.24	0.12	-1.97	A
100000168226	9.21	-0.26	0.08	-3.08	A
100000168227	4.43	-0.25	0.11	-2.15	A
100000168228	8.32	-0.40	0.14	-2.95	A
100000168229	0.10	0.04	0.11	0.36	A
100000168230	0.04	-0.03	0.13	-0.26	A
100000174583	8.82	-0.44	0.15	-3.00	B
100000174584	1.30	0.18	0.15	1.18	A
100000174585	15.24	0.49	0.13	3.81	B
100000174599	2.90	0.23	0.13	1.76	A
100000174601	86.72	0.99	0.11	9.38	C
100000174606	0.66	-0.09	0.11	-0.81	A
100000176753	5.54	-0.22	0.09	-2.35	A
100000176771	0.41	-0.07	0.11	-0.64	A
100000134326	6.55	0.29	0.11	2.60	A
100000168292	16.82	0.49	0.12	4.03	B
100000174718	8.11	0.47	0.17	2.77	B
100000174719	7.99	0.50	0.17	2.85	B
100000174728	90.36	1.41	0.16	8.75	C
100000187673	21.03	0.51	0.11	4.73	B

DRAFT

C.15.b: DIF Analysis for Level V Ethnicity Comparison (African Americans and Whites)

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000009038	0.00	0.01	0.13	0.07	A
100000168102	0.04	0.07	0.22	0.31	A
100000168104	0.48	-0.17	0.22	-0.80	A
100000168110	2.22	0.41	0.26	1.57	A
100000168111	2.71	-0.50	0.27	-1.81	A
100000168273	5.05	0.46	0.20	2.31	B
100000168274	0.07	0.09	0.25	0.38	A
100000168276	10.63	-0.67	0.20	-3.36	B
100000168277	0.18	-0.12	0.22	-0.53	A
100000197314	0.00	-0.02	0.22	-0.11	A
100000198197	0.82	-0.17	0.17	-0.99	A
100000198198	2.04	0.25	0.17	1.49	A
100000198199	2.44	-0.28	0.17	-1.63	A
100000200737	0.93	0.17	0.17	1.04	A
100000200738	2.57	-0.31	0.18	-1.67	A
100000202225	0.15	-0.11	0.22	-0.49	A
100000202226	2.34	-0.29	0.18	-1.59	A
100000168051	0.00	-0.01	0.20	-0.03	A
100000168052	1.18	-0.21	0.18	-1.19	A
100000168053	0.54	-0.14	0.18	-0.82	A
100000168054	17.97	-0.86	0.20	-4.20	C
100000168055	0.13	-0.08	0.18	-0.46	A
100000168068	0.57	-0.20	0.22	-0.87	A
100000168069	0.00	0.04	0.22	0.17	A
100000168224	0.07	0.07	0.19	0.35	A
100000168225	2.22	0.32	0.20	1.57	A
100000168226	1.28	-0.17	0.14	-1.20	A
100000168227	0.65	-0.19	0.21	-0.93	A
100000168228	1.20	-0.26	0.21	-1.20	A
100000168229	2.42	0.34	0.20	1.68	A
100000168230	6.26	-0.58	0.23	-2.56	B
100000174583	0.21	-0.12	0.21	-0.56	A
100000174584	21.56	1.10	0.25	4.41	C
100000174585	0.33	0.14	0.21	0.65	A
100000174599	5.92	0.59	0.23	2.55	B
100000174601	89.79	1.78	0.18	9.91	C
100000174606	6.96	0.44	0.17	2.67	B
100000176753	0.03	0.03	0.16	0.18	A
100000176771	9.60	0.57	0.18	3.09	B
100000134326	0.03	-0.04	0.20	-0.18	A
100000168292	2.30	-0.31	0.21	-1.52	A
100000174718	0.69	-0.30	0.32	-0.96	A
100000174719	3.11	-0.59	0.31	-1.87	A
100000174728	10.26	0.76	0.24	3.14	B
100000187673	0.03	-0.03	0.17	-0.17	A

DRAFT

C.15.c: DIF Analysis for Level V Ethnicity Comparison (Hispanics and Whites)

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000009038	0.17	-0.02	0.06	-0.44	A
100000168102	4.62	-0.20	0.09	-2.20	A
100000168104	7.73	-0.25	0.09	-2.82	A
100000168110	2.45	-0.20	0.12	-1.63	A
100000168111	7.33	-0.28	0.10	-2.75	A
100000168273	0.25	-0.05	0.09	-0.54	A
100000168274	0.56	-0.09	0.11	-0.81	A
100000168276	23.97	0.42	0.09	4.92	A
100000168277	2.39	-0.16	0.10	-1.59	A
100000197314	0.58	0.08	0.10	0.81	A
100000198197	2.91	-0.14	0.08	-1.74	A
100000198198	0.02	-0.01	0.08	-0.17	A
100000198199	0.40	0.05	0.08	0.67	A
100000200737	0.11	-0.03	0.08	-0.37	A
100000200738	19.20	-0.39	0.09	-4.40	A
100000202225	16.81	-0.36	0.09	-4.13	A
100000202226	1.05	-0.08	0.08	-1.06	A
100000168051	6.55	-0.25	0.10	-2.60	A
100000168052	3.46	0.15	0.08	1.90	A
100000168053	3.87	-0.16	0.08	-2.01	A
100000168054	1.40	-0.11	0.09	-1.22	A
100000168055	3.53	-0.16	0.08	-1.92	A
100000168068	9.56	-0.34	0.11	-3.14	A
100000168069	3.49	0.18	0.10	1.92	A
100000168224	5.38	0.19	0.08	2.36	A
100000168225	1.02	0.09	0.09	1.05	A
100000168226	9.04	-0.18	0.06	-3.04	A
100000168227	0.16	0.04	0.08	0.44	A
100000168228	16.45	-0.38	0.09	-4.09	A
100000168229	13.91	0.30	0.08	3.77	A
100000168230	50.75	-0.66	0.09	-7.09	C
100000174583	6.82	-0.26	0.10	-2.65	A
100000174584	0.19	0.06	0.12	0.49	A
100000174585	3.43	-0.20	0.10	-1.91	A
100000174599	5.30	0.23	0.10	2.35	A
100000174601	309.53	1.37	0.08	17.35	C
100000174606	1.60	0.10	0.08	1.27	A
100000176753	2.58	-0.12	0.07	-1.58	A
100000176771	3.47	0.15	0.08	1.89	A
100000134326	5.47	-0.20	0.08	-2.35	A
100000168292	4.31	0.21	0.10	2.09	A
100000174718	3.30	-0.26	0.14	-1.85	A
100000174719	2.55	-0.27	0.16	-1.69	A
100000174728	0.00	0.02	0.13	0.12	A
100000187673	8.67	-0.24	0.08	-2.96	A

DRAFT

C.15.d: DIF Analysis for Level V Gender Comparison

Item ID	Mantel Chi-Square	Mantel-Haenszel/ Liu-Agresti Log-Odds Ratio	Log-Odds Ratio SE	Standardized Log-Odds Ratio	ETS Criteria
100000009038	1.22	0.06	0.05	1.13	A
100000168102	0.99	0.08	0.08	1.04	A
100000168104	31.35	0.44	0.08	5.61	B
100000168110	0.97	0.11	0.11	1.03	A
100000168111	2.57	-0.15	0.09	-1.64	A
100000168273	1.03	-0.08	0.08	-1.05	A
100000168274	5.22	0.22	0.10	2.33	A
100000168276	9.36	-0.23	0.08	-3.10	A
100000168277	0.35	0.06	0.09	0.64	A
100000197314	2.96	0.15	0.08	1.76	A
100000198197	6.36	0.17	0.07	2.55	A
100000198198	12.21	0.24	0.07	3.53	A
100000198199	0.62	0.06	0.07	0.82	A
100000200737	3.10	0.13	0.07	1.80	A
100000200738	0.26	0.04	0.07	0.55	A
100000202225	6.69	0.20	0.08	2.62	A
100000202226	17.83	-0.30	0.07	-4.25	A
100000168051	6.48	-0.21	0.08	-2.59	A
100000168052	1.84	0.10	0.07	1.39	A
100000168053	4.01	-0.14	0.07	-2.04	A
100000168054	15.70	-0.30	0.08	-4.01	A
100000168055	11.30	-0.24	0.07	-3.38	A
100000168068	1.88	0.13	0.09	1.41	A
100000168069	9.03	0.25	0.08	3.05	A
100000168224	0.34	-0.04	0.07	-0.62	A
100000168225	0.00	0.01	0.08	0.07	A
100000168226	36.13	0.32	0.05	5.98	A
100000168227	0.08	-0.02	0.07	-0.32	A
100000168228	7.38	0.23	0.08	2.75	A
100000168229	3.13	-0.13	0.07	-1.80	A
100000168230	3.36	0.15	0.08	1.87	A
100000174583	9.67	-0.28	0.09	-3.15	A
100000174584	3.42	0.19	0.10	1.89	A
100000174585	0.02	0.02	0.09	0.19	A
100000174599	0.14	-0.03	0.08	-0.41	A
100000174601	15.87	-0.26	0.07	-3.99	A
100000174606	7.62	-0.20	0.07	-2.78	A
100000176753	135.12	-0.75	0.07	-11.57	C
100000176771	93.78	-0.70	0.07	-9.74	C
100000134326	2.96	0.13	0.07	1.74	A
100000168292	1.44	0.10	0.08	1.21	A
100000174718	6.19	0.30	0.12	2.54	A
100000174719	0.63	-0.11	0.13	-0.85	A
100000174728	1.66	0.14	0.10	1.33	A
100000187673	6.45	0.18	0.07	2.54	A

DRAFT

APPENDIX D: RAW SCORE TO SCALE SCORE CONVERSION TABLES FOR THE ELPA

D.1.a: Level I

Total ELPA

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-7.1366	1.8455	327	57.7
1	-5.8831	1.0339	366	32.3
2	-5.1235	0.7512	390	23.5
3	-4.6558	0.6274	405	19.6
4	-4.3095	0.5540	415	17.3
5	-4.0310	0.5040	424	15.8
6	-3.7959	0.4670	431	14.6
7	-3.5916	0.4382	438	13.7
8	-3.4098	0.4150	443	13.0
9	-3.2457	0.3957	449	12.4
10	-3.0957	0.3794	453	11.9
11	-2.9571	0.3654	458	11.4
12	-2.8281	0.3532	462	11.0
13	-2.7072	0.3424	465	10.7
14	-2.5932	0.3330	469	10.4
15	-2.4852	0.3245	472	10.1
16	-2.3824	0.3169	476	9.9
17	-2.2841	0.3102	479	9.7
18	-2.1898	0.3041	482	9.5
19	-2.0990	0.2986	484	9.3
20	-2.0114	0.2936	487	9.2
21	-1.9265	0.2892	490	9.0
22	-1.8440	0.2851	492	8.9
23	-1.7638	0.2815	495	8.8
24	-1.6854	0.2783	497	8.7
25	-1.6088	0.2754	500	8.6
26	-1.5337	0.2728	502	8.5
27	-1.4599	0.2705	504	8.5
28	-1.3873	0.2685	507	8.4
29	-1.3157	0.2668	509	8.3
30	-1.2448	0.2654	511	8.3
31	-1.1747	0.2642	513	8.3
32	-1.1052	0.2632	515	8.2

DRAFT

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
33	-1.0361	0.2625	518	8.2
34	-0.9674	0.2620	520	8.2
35	-0.8988	0.2617	522	8.2
36	-0.8303	0.2617	524	8.2
37	-0.7618	0.2619	526	8.2
38	-0.6931	0.2624	528	8.2
39	-0.6240	0.2631	531	8.2
40	-0.5546	0.2640	533	8.3
41	-0.4845	0.2653	535	8.3
42	-0.4138	0.2668	537	8.3
43	-0.3421	0.2686	539	8.4
44	-0.2695	0.2707	542	8.5
45	-0.1956	0.2731	544	8.5
46	-0.1202	0.2759	546	8.6
47	-0.0432	0.2791	549	8.7
48	0.0357	0.2828	551	8.8
49	0.1168	0.2870	554	9.0
50	0.2005	0.2917	556	9.1
51	0.2871	0.2970	559	9.3
52	0.3771	0.3031	562	9.5
53	0.4711	0.3100	565	9.7
54	0.5695	0.3179	568	9.9
55	0.6734	0.3269	571	10.2
56	0.7836	0.3372	574	10.5
57	0.9012	0.3491	578	10.9
58	1.0279	0.3630	582	11.3
59	1.1655	0.3792	586	11.9
60	1.3163	0.3982	591	12.4
61	1.4839	0.4209	596	13.2
62	1.6724	0.4484	602	14.0
63	1.8887	0.4826	609	15.1
64	2.1427	0.5276	617	16.5
65	2.4544	0.5931	627	18.5
66	2.8698	0.7068	640	22.1
67	3.5470	0.9823	661	30.7
68	4.7125	1.8040	697	56.4

DRAFT

D.1.b: Level I ELPA by Domain- Listening

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-5.6318	1.8647	35	7.2
1	-4.3294	1.0666	40	4.1
2	-3.5023	0.7942	44	3.1
3	-2.9682	0.6784	46	2.6
4	-2.5543	0.6133	47	2.4
5	-2.2042	0.5729	49	2.2
6	-1.8915	0.5475	50	2.1
7	-1.6006	0.5326	51	2.1
8	-1.3210	0.5264	52	2.0
9	-1.0436	0.5282	53	2.0
10	-0.7599	0.5386	54	2.1
11	-0.4595	0.5595	55	2.2
12	-0.1277	0.5954	57	2.3
13	0.2609	0.6563	58	2.5
14	0.7608	0.7691	60	3.0
15	1.5428	1.0423	63	4.0
16	2.8063	1.8485	68	7.1

DRAFT

D.1.c: Level I ELPA by Domain- Reading

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-6.0639	1.9137	32	7.2
1	-4.6418	1.1363	38	4.3
2	-3.6878	0.8564	41	3.2
3	-3.0681	0.7294	43	2.7
4	-2.5912	0.6573	45	2.5
5	-2.1899	0.6128	47	2.3
6	-1.8328	0.5845	48	2.2
7	-1.5021	0.5672	49	2.1
8	-1.1860	0.5585	51	2.1
9	-0.8754	0.5575	52	2.1
10	-0.5613	0.5649	53	2.1
11	-0.2331	0.5827	54	2.2
12	0.1239	0.6152	55	2.3
13	0.5353	0.6726	57	2.5
14	1.0556	0.7816	59	2.9
15	1.8560	1.0506	62	3.9
16	3.1309	1.8527	67	6.9

DRAFT

D.1.d: Level I ELPA by Domain- Writing

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-4.7025	1.8607	33	7.4
1	-3.4083	1.0624	38	4.2
2	-2.5874	0.7915	42	3.2
3	-2.0570	0.6758	44	2.7
4	-1.6475	0.6085	45	2.4
5	-1.3051	0.5642	47	2.3
6	-1.0047	0.5338	48	2.1
7	-0.7312	0.5135	49	2.1
8	-0.4743	0.5015	50	2.0
9	-0.2256	0.4971	51	2.0
10	0.0224	0.5002	52	2.0
11	0.2777	0.5117	53	2.0
12	0.5495	0.5329	54	2.1
13	0.8503	0.5661	55	2.3
14	1.1971	0.6139	57	2.5
15	1.6126	0.6775	58	2.7
16	2.1297	0.7672	61	3.1
17	2.8623	0.9853	63	3.9
18	3.9782	1.7499	68	7.0

DRAFT

D.1.e: Level I ELPA by Domain- Speaking

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-5.9405	1.8800	32	7.9
1	-4.5938	1.0980	38	4.6
2	-3.6963	0.8370	41	3.5
3	-3.0987	0.7180	44	3.0
4	-2.6404	0.6395	46	2.7
5	-2.2685	0.5828	47	2.4
6	-1.9532	0.5423	49	2.3
7	-1.6750	0.5145	50	2.2
8	-1.4200	0.4969	51	2.1
9	-1.1783	0.4876	52	2.0
10	-0.9419	0.4861	53	2.0
11	-0.7033	0.4923	54	2.1
12	-0.4543	0.5071	55	2.1
13	-0.1851	0.5327	56	2.2
14	0.1195	0.5739	58	2.4
15	0.4851	0.6406	59	2.7
16	0.9673	0.7595	61	3.2
17	1.7380	1.0384	64	4.4
18	2.9971	1.8472	70	7.8

DRAFT

D.1.f: Level I ELPA by Domain- Comprehension

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-6.1940	1.8953	34	6.7
1	-4.8219	1.1048	39	3.9
2	-3.9341	0.8194	42	2.9
3	-3.3739	0.6888	44	2.4
4	-2.9542	0.6121	46	2.2
5	-2.6112	0.5623	47	2.0
6	-2.3147	0.5285	48	1.9
7	-2.0484	0.5049	49	1.8
8	-1.8021	0.4887	50	1.7
9	-1.5689	0.4781	50	1.7
10	-1.3435	0.4722	51	1.7
11	-1.1216	0.4705	52	1.7
12	-0.8994	0.4729	53	1.7
13	-0.6729	0.4798	54	1.7
14	-0.4374	0.4917	54	1.7
15	-0.1871	0.5102	55	1.8
16	0.0865	0.5376	56	1.9
17	0.3966	0.5788	57	2.1
18	0.7675	0.6441	59	2.3
19	1.2529	0.7606	60	2.7
20	2.0232	1.0372	63	3.7
21	3.2796	1.8459	68	6.6

DRAFT

D.2.a: Level II

Total ELPA

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-6.1387	1.8406	358	57.5
1	-4.8979	1.0251	397	32.0
2	-4.1566	0.7387	420	23.1
3	-3.7075	0.6124	434	19.1
4	-3.3799	0.5370	444	16.8
5	-3.1198	0.4854	453	15.2
6	-2.9030	0.4473	459	14.0
7	-2.7165	0.4176	465	13.1
8	-2.5522	0.3937	470	12.3
9	-2.4051	0.3739	475	11.7
10	-2.2716	0.3572	479	11.2
11	-2.1492	0.3430	483	10.7
12	-2.0358	0.3307	486	10.3
13	-1.9301	0.3199	490	10.0
14	-1.8308	0.3104	493	9.7
15	-1.7371	0.3020	496	9.4
16	-1.6481	0.2945	498	9.2
17	-1.5634	0.2878	501	9.0
18	-1.4823	0.2819	504	8.8
19	-1.4044	0.2765	506	8.6
20	-1.3293	0.2716	508	8.5
21	-1.2568	0.2672	511	8.4
22	-1.1864	0.2633	513	8.2
23	-1.1180	0.2597	515	8.1
24	-1.0514	0.2565	517	8.0
25	-0.9864	0.2536	519	7.9
26	-0.9227	0.2510	521	7.8
27	-0.8603	0.2486	523	7.8
28	-0.7991	0.2465	525	7.7
29	-0.7387	0.2446	527	7.6
30	-0.6793	0.2430	529	7.6
31	-0.6206	0.2415	531	7.5
32	-0.5626	0.2402	532	7.5
33	-0.5052	0.2391	534	7.5
34	-0.4482	0.2382	536	7.4
35	-0.3917	0.2374	538	7.4
36	-0.3355	0.2368	540	7.4
37	-0.2796	0.2363	541	7.4
38	-0.2238	0.2360	543	7.4
39	-0.1681	0.2359	545	7.4
40	-0.1125	0.2359	546	7.4
41	-0.0568	0.2360	548	7.4
42	-0.0011	0.2363	550	7.4
43	0.0549	0.2367	552	7.4

DRAFT

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
44	0.1111	0.2373	553	7.4
45	0.1675	0.2380	555	7.4
46	0.2244	0.2389	557	7.5
47	0.2817	0.2400	559	7.5
48	0.3396	0.2412	561	7.5
49	0.3981	0.2426	562	7.6
50	0.4574	0.2442	564	7.6
51	0.5174	0.2460	566	7.7
52	0.5784	0.2479	568	7.7
53	0.6404	0.2501	570	7.8
54	0.7035	0.2526	572	7.9
55	0.7680	0.2552	574	8.0
56	0.8339	0.2582	576	8.1
57	0.9014	0.2614	578	8.2
58	0.9706	0.2649	580	8.3
59	1.0418	0.2688	583	8.4
60	1.1152	0.2731	585	8.5
61	1.1911	0.2778	587	8.7
62	1.2696	0.2829	590	8.8
63	1.3513	0.2887	592	9.0
64	1.4364	0.2950	595	9.2
65	1.5255	0.3020	598	9.4
66	1.6191	0.3099	601	9.7
67	1.7178	0.3187	604	10.0
68	1.8225	0.3287	607	10.3
69	1.9343	0.3401	610	10.6
70	2.0544	0.3534	614	11.0
71	2.1846	0.3688	618	11.5
72	2.3274	0.3872	623	12.1
73	2.4859	0.4096	628	12.8
74	2.6649	0.4376	633	13.7
75	2.8720	0.4740	640	14.8
76	3.1197	0.5239	647	16.4
77	3.4316	0.5977	657	18.7
78	3.8607	0.7234	671	22.6
79	4.5767	1.0111	693	31.6
80	5.7959	1.8315	731	57.2

DRAFT

D.2.b: Level II ELPA by Domain- Listening

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-4.6763	1.8474	39	7.1
1	-3.4162	1.0397	44	4.0
2	-2.6411	0.7635	47	2.9
3	-2.1515	0.6471	49	2.5
4	-1.7770	0.5816	50	2.2
5	-1.4639	0.5402	51	2.1
6	-1.1878	0.5124	52	2.0
7	-0.9353	0.4937	53	1.9
8	-0.6981	0.4815	54	1.9
9	-0.4700	0.4744	55	1.8
10	-0.2465	0.4719	56	1.8
11	-0.0233	0.4736	57	1.8
12	0.2036	0.4799	58	1.8
13	0.4391	0.4914	59	1.9
14	0.6888	0.5094	60	2.0
15	0.9614	0.5365	61	2.1
16	1.2700	0.5773	62	2.2
17	1.6390	0.6424	63	2.5
18	2.1220	0.7588	65	2.9
19	2.8892	1.0356	68	4.0
20	4.1432	1.8449	73	7.1

DRAFT

D.2.c: Level II ELPA by Domain- Reading

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-4.9047	1.9211	37	7.2
1	-3.4694	1.1394	42	4.3
2	-2.5256	0.8420	46	3.2
3	-1.9393	0.7014	48	2.6
4	-1.5066	0.6201	49	2.3
5	-1.1552	0.5687	51	2.1
6	-0.8519	0.5348	52	2.0
7	-0.5787	0.5121	53	1.9
8	-0.3245	0.4973	54	1.9
9	-0.0820	0.4886	55	1.8
10	0.1546	0.4850	56	1.8
11	0.3900	0.4861	56	1.8
12	0.6287	0.4919	57	1.8
13	0.8757	0.5030	58	1.9
14	1.1370	0.5207	59	2.0
15	1.4213	0.5473	60	2.1
16	1.7420	0.5877	62	2.2
17	2.1232	0.6520	63	2.4
18	2.6189	0.7673	65	2.9
19	3.3993	1.0421	68	3.9
20	4.6628	1.8487	72	6.9

DRAFT

D.2.d: Level II ELPA by Domain- Writing

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-4.2118	1.8419	35	7.4
1	-2.9664	1.0290	40	4.1
2	-2.2146	0.7475	43	3.0
3	-1.7499	0.6271	45	2.5
4	-1.4011	0.5589	46	2.2
5	-1.1140	0.5157	48	2.1
6	-0.8635	0.4872	49	1.9
7	-0.6357	0.4687	49	1.9
8	-0.4216	0.4578	50	1.8
9	-0.2146	0.4532	51	1.8
10	-0.0091	0.4544	52	1.8
11	0.2001	0.4613	53	1.8
12	0.4184	0.4741	54	1.9
13	0.6519	0.4935	55	2.0
14	0.9084	0.5206	56	2.1
15	1.1980	0.5574	57	2.2
16	1.5359	0.6075	58	2.4
17	1.9467	0.6783	60	2.7
18	2.4805	0.7927	62	3.2
19	3.2972	1.0562	65	4.2
20	4.5748	1.8509	70	7.4

DRAFT

D.2.e: Level II ELPA by Domain- Speaking

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-4.9126	1.8653	36	7.8
1	-3.6093	1.0664	42	4.5
2	-2.7864	0.7887	45	3.3
3	-2.2660	0.6645	47	2.8
4	-1.8752	0.5907	49	2.5
5	-1.5556	0.5430	50	2.3
6	-1.2785	0.5117	52	2.1
7	-1.0276	0.4917	53	2.1
8	-0.7922	0.4797	54	2.0
9	-0.5654	0.4736	55	2.0
10	-0.3422	0.4719	56	2.0
11	-0.1189	0.4740	57	2.0
12	0.1083	0.4799	57	2.0
13	0.3432	0.4905	58	2.1
14	0.5917	0.5077	59	2.1
15	0.8622	0.5344	61	2.2
16	1.1688	0.5758	62	2.4
17	1.5366	0.6422	63	2.7
18	2.0204	0.7600	65	3.2
19	2.7905	1.0375	69	4.4
20	4.0477	1.8463	74	7.8

DRAFT

D.2.f: Level II ELPA by Domain- Comprehension

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-4.9117	1.8380	39	6.5
1	-3.6760	1.0226	43	3.6
2	-2.9368	0.7392	46	2.6
3	-2.4844	0.6170	47	2.2
4	-2.1488	0.5462	48	1.9
5	-1.8768	0.4995	49	1.8
6	-1.6444	0.4663	50	1.7
7	-1.4388	0.4416	51	1.6
8	-1.2523	0.4226	52	1.5
9	-1.0801	0.4079	52	1.4
10	-0.9186	0.3964	53	1.4
11	-0.7652	0.3873	53	1.4
12	-0.6180	0.3803	54	1.4
13	-0.4754	0.3751	54	1.3
14	-0.3362	0.3713	55	1.3
15	-0.1993	0.3690	55	1.3
16	-0.0636	0.3679	56	1.3
17	0.0717	0.3680	56	1.3
18	0.2075	0.3694	57	1.3
19	0.3449	0.3721	57	1.3
20	0.4848	0.3761	58	1.3
21	0.6282	0.3816	58	1.4
22	0.7765	0.3889	59	1.4
23	0.9313	0.3982	59	1.4
24	1.0944	0.4100	60	1.5
25	1.2684	0.4249	61	1.5
26	1.4569	0.4440	61	1.6
27	1.6648	0.4689	62	1.7
28	1.8999	0.5023	63	1.8
29	2.1749	0.5491	64	1.9
30	2.5137	0.6198	65	2.2
31	2.9697	0.7418	67	2.6
32	3.7129	1.0246	69	3.6
33	4.9518	1.8393	74	6.5

DRAFT

D.3.a: Level III

Total ELPA

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-4.5202	1.8333	409	57.3
1	-3.2973	1.0132	447	31.7
2	-2.5785	0.7245	469	22.6
3	-2.1488	0.5975	483	18.7
4	-1.8380	0.5221	493	16.3
5	-1.5927	0.4709	500	14.7
6	-1.3889	0.4334	507	13.5
7	-1.2138	0.4045	512	12.6
8	-1.0596	0.3815	517	11.9
9	-0.9213	0.3627	521	11.3
10	-0.7955	0.3471	525	10.8
11	-0.6797	0.3339	529	10.4
12	-0.5720	0.3227	532	10.1
13	-0.4709	0.3131	535	9.8
14	-0.3755	0.3048	538	9.5
15	-0.2849	0.2976	541	9.3
16	-0.1982	0.2912	544	9.1
17	-0.1151	0.2856	546	8.9
18	-0.0349	0.2807	549	8.8
19	0.0427	0.2764	551	8.6
20	0.1180	0.2725	554	8.5
21	0.1913	0.2691	556	8.4
22	0.2629	0.2661	558	8.3
23	0.3329	0.2633	560	8.2
24	0.4017	0.2609	563	8.2
25	0.4691	0.2587	565	8.1
26	0.5356	0.2568	567	8.0
27	0.6011	0.2551	569	8.0
28	0.6658	0.2536	571	7.9
29	0.7297	0.2522	573	7.9
30	0.7930	0.2510	575	7.8
31	0.8557	0.2500	577	7.8
32	0.9180	0.2491	579	7.8
33	0.9799	0.2483	581	7.8
34	1.0414	0.2477	583	7.7
35	1.1026	0.2472	584	7.7
36	1.1637	0.2469	586	7.7
37	1.2245	0.2466	588	7.7
38	1.2853	0.2465	590	7.7
39	1.3461	0.2465	592	7.7
40	1.4068	0.2466	594	7.7
41	1.4677	0.2468	596	7.7
42	1.5287	0.2472	598	7.7
43	1.5899	0.2477	600	7.7

DRAFT

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
44	1.6514	0.2483	602	7.8
45	1.7133	0.2491	604	7.8
46	1.7756	0.2500	605	7.8
47	1.8383	0.2510	607	7.8
48	1.9015	0.2522	609	7.9
49	1.9655	0.2535	611	7.9
50	2.0301	0.2550	613	8.0
51	2.0955	0.2567	615	8.0
52	2.1619	0.2585	618	8.1
53	2.2292	0.2605	620	8.1
54	2.2977	0.2627	622	8.2
55	2.3673	0.2652	624	8.3
56	2.4383	0.2678	626	8.4
57	2.5108	0.2707	628	8.5
58	2.5849	0.2739	631	8.6
59	2.6609	0.2774	633	8.7
60	2.7389	0.2812	636	8.8
61	2.8192	0.2854	638	8.9
62	2.9019	0.2900	641	9.1
63	2.9874	0.2950	643	9.2
64	3.0761	0.3006	646	9.4
65	3.1683	0.3068	649	9.6
66	3.2646	0.3137	652	9.8
67	3.3654	0.3215	655	10.0
68	3.4715	0.3302	658	10.3
69	3.5838	0.3401	662	10.6
70	3.7032	0.3514	666	11.0
71	3.8313	0.3646	670	11.4
72	3.9698	0.3800	674	11.9
73	4.1211	0.3984	679	12.5
74	4.2887	0.4209	684	13.2
75	4.4774	0.4490	690	14.0
76	4.6949	0.4854	697	15.2
77	4.9541	0.5352	705	16.7
78	5.2787	0.6089	715	19.0
79	5.7222	0.7339	729	22.9
80	6.4546	1.0198	752	31.9
81	7.6868	1.8368	790	57.4

DRAFT

D.3.b: Level III ELPA by Domain- Listening

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-3.0442	1.8495	45	7.1
1	-1.7790	1.0430	50	4.0
2	-0.9973	0.7676	53	3.0
3	-0.5016	0.6517	55	2.5
4	-0.1211	0.5868	57	2.3
5	0.1983	0.5461	58	2.1
6	0.4811	0.5192	59	2.0
7	0.7410	0.5016	60	1.9
8	0.9866	0.4907	61	1.9
9	1.2244	0.4853	62	1.9
10	1.4593	0.4850	63	1.9
11	1.6963	0.4896	64	1.9
12	1.9405	0.4997	64	1.9
13	2.1981	0.5166	65	2.0
14	2.4778	0.5427	67	2.1
15	2.7929	0.5828	68	2.2
16	3.1681	0.6471	69	2.5
17	3.6571	0.7626	71	2.9
18	4.4301	1.0384	74	4.0
19	5.6882	1.8465	79	7.1

DRAFT

D.3.c: Level III ELPA by Domain- Reading

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-2.4626	1.8499	46	6.9
1	-1.1970	1.0428	51	3.9
2	-0.4172	0.7655	53	2.9
3	0.0738	0.6471	55	2.4
4	0.4470	0.5794	57	2.2
5	0.7563	0.5354	58	2.0
6	1.0261	0.5052	59	1.9
7	1.2700	0.4837	60	1.8
8	1.4963	0.4686	61	1.8
9	1.7108	0.4583	61	1.7
10	1.9177	0.4521	62	1.7
11	2.1208	0.4495	63	1.7
12	2.3229	0.4504	64	1.7
13	2.5276	0.4549	64	1.7
14	2.7382	0.4635	65	1.7
15	2.9588	0.4769	66	1.8
16	3.1952	0.4966	67	1.9
17	3.4554	0.5253	68	2.0
18	3.7526	0.5677	69	2.1
19	4.1109	0.6344	70	2.4
20	4.5840	0.7524	72	2.8
21	5.3420	1.0313	75	3.9
22	6.5898	1.8426	80	6.9

DRAFT

D.3.d: Level III ELPA by Domain- Writing

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-2.9746	1.8427	40	7.4
1	-1.7261	1.0322	45	4.1
2	-0.9645	0.7562	48	3.0
3	-0.4830	0.6432	50	2.6
4	-0.1098	0.5839	52	2.3
5	0.2102	0.5505	53	2.2
6	0.5022	0.5323	54	2.1
7	0.7806	0.5241	55	2.1
8	1.0540	0.5227	56	2.1
9	1.3288	0.5264	57	2.1
10	1.6096	0.5341	58	2.1
11	1.9006	0.5452	60	2.2
12	2.2052	0.5591	61	2.2
13	2.5270	0.5759	62	2.3
14	2.8700	0.5961	63	2.4
15	3.2402	0.6217	65	2.5
16	3.6479	0.6574	67	2.6
17	4.1150	0.7139	68	2.9
18	4.6934	0.8183	71	3.3
19	5.5529	1.0785	74	4.3
20	6.8695	1.8687	79	7.5

DRAFT

D.3.e: Level III ELPA by Domain- Speaking

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-3.6340	1.8536	42	7.8
1	-2.3627	1.0441	47	4.4
2	-1.5870	0.7587	50	3.2
3	-1.1107	0.6330	52	2.7
4	-0.7571	0.5616	54	2.4
5	-0.4676	0.5176	55	2.2
6	-0.2149	0.4900	56	2.1
7	0.0164	0.4733	57	2.0
8	0.2357	0.4642	58	1.9
9	0.4491	0.4605	59	1.9
10	0.6611	0.4611	60	1.9
11	0.8755	0.4653	61	2.0
12	1.0953	0.4731	62	2.0
13	1.3246	0.4852	63	2.0
14	1.5683	0.5035	64	2.1
15	1.8350	0.5310	65	2.2
16	2.1381	0.5731	66	2.4
17	2.5031	0.6403	68	2.7
18	2.9851	0.7594	70	3.2
19	3.7557	1.0384	73	4.4
20	5.0150	1.8475	78	7.8

DRAFT

D.3.f: Level III ELPA by Domain- Comprehension

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-3.1908	1.8439	45	6.5
1	-1.9405	1.0324	49	3.7
2	-1.1818	0.7518	52	2.7
3	-0.7112	0.6312	53	2.2
4	-0.3583	0.5615	55	2.0
5	-0.0696	0.5156	56	1.8
6	0.1790	0.4831	57	1.7
7	0.4004	0.4590	57	1.6
8	0.6024	0.4408	58	1.6
9	0.7904	0.4268	59	1.5
10	0.9678	0.4161	59	1.5
11	1.1374	0.4080	60	1.4
12	1.3013	0.4022	61	1.4
13	1.4614	0.3983	61	1.4
14	1.6191	0.3962	62	1.4
15	1.7758	0.3958	62	1.4
16	1.9329	0.3972	63	1.4
17	2.0918	0.4003	63	1.4
18	2.2539	0.4054	64	1.4
19	2.4211	0.4128	65	1.5
20	2.5955	0.4228	65	1.5
21	2.7796	0.4360	66	1.5
22	2.9771	0.4536	67	1.6
23	3.1931	0.4770	67	1.7
24	3.4354	0.5089	68	1.8
25	3.7167	0.5544	69	2.0
26	4.0610	0.6238	70	2.2
27	4.5216	0.7446	72	2.6
28	5.2686	1.0263	75	3.6
29	6.5095	1.8400	79	6.5

DRAFT

D.4.a: Level IV

Total ELPA

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-4.3428	1.8320	414	57.3
1	-3.1231	1.0108	452	31.6
2	-2.4089	0.7214	475	22.5
3	-1.9836	0.5938	488	18.6
4	-1.6772	0.5180	498	16.2
5	-1.4361	0.4665	505	14.6
6	-1.2364	0.4286	511	13.4
7	-1.0654	0.3993	517	12.5
8	-0.9155	0.3759	521	11.7
9	-0.7815	0.3567	526	11.1
10	-0.6601	0.3406	529	10.6
11	-0.5488	0.3269	533	10.2
12	-0.4458	0.3152	536	9.9
13	-0.3497	0.3050	539	9.5
14	-0.2594	0.2961	542	9.3
15	-0.1741	0.2883	545	9.0
16	-0.0931	0.2813	547	8.8
17	-0.0156	0.2752	550	8.6
18	0.0585	0.2697	552	8.4
19	0.1299	0.2648	554	8.3
20	0.1989	0.2604	556	8.1
21	0.2656	0.2564	558	8.0
22	0.3305	0.2529	560	7.9
23	0.3936	0.2497	562	7.8
24	0.4552	0.2469	564	7.7
25	0.5155	0.2443	566	7.6
26	0.5746	0.2420	568	7.6
27	0.6327	0.2399	570	7.5
28	0.6898	0.2381	572	7.4
29	0.7461	0.2365	573	7.4
30	0.8016	0.2350	575	7.3
31	0.8566	0.2338	577	7.3
32	0.9110	0.2327	578	7.3
33	0.9649	0.2317	580	7.2
34	1.0184	0.2310	582	7.2
35	1.0716	0.2303	583	7.2
36	1.1245	0.2298	585	7.2
37	1.1772	0.2294	587	7.2
38	1.2298	0.2292	588	7.2
39	1.2823	0.2291	590	7.2
40	1.3348	0.2291	592	7.2
41	1.3872	0.2292	593	7.2
42	1.4398	0.2294	595	7.2
43	1.4925	0.2298	597	7.2

DRAFT

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
44	1.5454	0.2302	598	7.2
45	1.5986	0.2308	600	7.2
46	1.6520	0.2316	602	7.2
47	1.7058	0.2324	603	7.3
48	1.7601	0.2334	605	7.3
49	1.8148	0.2345	607	7.3
50	1.8700	0.2357	608	7.4
51	1.9259	0.2370	610	7.4
52	1.9824	0.2385	612	7.5
53	2.0397	0.2402	614	7.5
54	2.0978	0.2420	616	7.6
55	2.1569	0.2440	617	7.6
56	2.2169	0.2461	619	7.7
57	2.2781	0.2485	621	7.8
58	2.3404	0.2510	623	7.8
59	2.4041	0.2538	625	7.9
60	2.4693	0.2568	627	8.0
61	2.5360	0.2600	629	8.1
62	2.6046	0.2636	631	8.2
63	2.6751	0.2674	634	8.4
64	2.7477	0.2716	636	8.5
65	2.8227	0.2762	638	8.6
66	2.9004	0.2813	641	8.8
67	2.9810	0.2868	643	9.0
68	3.0650	0.2929	646	9.2
69	3.1528	0.2997	649	9.4
70	3.2449	0.3072	651	9.6
71	3.3418	0.3157	654	9.9
72	3.4445	0.3252	658	10.2
73	3.5537	0.3360	661	10.5
74	3.6708	0.3484	665	10.9
75	3.7971	0.3628	669	11.3
76	3.9348	0.3797	673	11.9
77	4.0865	0.3998	678	12.5
78	4.2560	0.4243	683	13.3
79	4.4487	0.4548	689	14.2
80	4.6731	0.4941	696	15.4
81	4.9429	0.5473	704	17.1
82	5.2835	0.6247	715	19.5
83	5.7508	0.7531	730	23.5
84	6.5179	1.0399	754	32.5
85	7.7832	1.8512	793	57.9

DRAFT

D.4.b: Level IV ELPA by Domain- Listening

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-2.6424	1.8436	47	7.1
1	-1.3920	1.0332	52	4.0
2	-0.6299	0.7553	55	2.9
3	-0.1523	0.6382	56	2.5
4	0.2112	0.5725	58	2.2
5	0.5141	0.5310	59	2.0
6	0.7808	0.5035	60	1.9
7	1.0245	0.4849	61	1.9
8	1.2533	0.4728	62	1.8
9	1.4733	0.4659	63	1.8
10	1.6888	0.4633	64	1.8
11	1.9038	0.4647	64	1.8
12	2.1220	0.4703	65	1.8
13	2.3475	0.4803	66	1.8
14	2.5852	0.4957	67	1.9
15	2.8414	0.5179	68	2.0
16	3.1254	0.5496	69	2.1
17	3.4516	0.5954	70	2.3
18	3.8462	0.6656	72	2.6
19	4.3654	0.7864	74	3.0
20	5.1828	1.0633	77	4.1
21	6.4819	1.8640	82	7.2

DRAFT

D.4.c: Level IV ELPA by Domain- Reading

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-2.4163	1.8425	46	6.9
1	-1.1687	1.0311	51	3.9
2	-0.4111	0.7522	53	2.8
3	0.0617	0.6342	55	2.4
4	0.4199	0.5676	57	2.1
5	0.7170	0.5253	58	2.0
6	0.9773	0.4967	59	1.9
7	1.2138	0.4772	60	1.8
8	1.4350	0.4641	60	1.7
9	1.6462	0.4560	61	1.7
10	1.8521	0.4521	62	1.7
11	2.0562	0.4521	63	1.7
12	2.2620	0.4560	63	1.7
13	2.4733	0.4641	64	1.7
14	2.6944	0.4772	65	1.8
15	2.9309	0.4967	66	1.9
16	3.1912	0.5253	67	2.0
17	3.4883	0.5676	68	2.1
18	3.8466	0.6342	69	2.4
19	4.3195	0.7523	71	2.8
20	5.0773	1.0312	74	3.9
21	6.3250	1.8426	79	6.9

DRAFT

D.4.d: Level IV ELPA by Domain- Writing

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-2.9510	1.8352	40	7.3
1	-1.7218	1.0187	45	4.1
2	-0.9882	0.7371	48	2.9
3	-0.5364	0.6187	50	2.5
4	-0.1959	0.5533	51	2.2
5	0.0868	0.5128	52	2.1
6	0.3356	0.4865	53	1.9
7	0.5634	0.4692	54	1.9
8	0.7780	0.4581	55	1.8
9	0.9846	0.4519	56	1.8
10	1.1876	0.4498	57	1.8
11	1.3904	0.4517	58	1.8
12	1.5969	0.4578	58	1.8
13	1.8110	0.4684	59	1.9
14	2.0375	0.4843	60	1.9
15	2.2825	0.5066	61	2.0
16	2.5540	0.5368	62	2.1
17	2.8636	0.5780	63	2.3
18	3.2303	0.6361	65	2.5
19	3.6888	0.7240	67	2.9
20	4.3167	0.8732	69	3.5
21	5.3328	1.1792	73	4.7
22	6.8374	1.9465	79	7.8

DRAFT

D.4.e: Level IV ELPA by Domain- Speaking

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-3.4482	1.8523	43	7.8
1	-2.1797	1.0425	48	4.4
2	-1.4067	0.7572	51	3.2
3	-0.9330	0.6303	53	2.6
4	-0.5843	0.5557	55	2.3
5	-0.3035	0.5070	56	2.1
6	-0.0638	0.4740	57	2.0
7	0.1497	0.4515	58	1.9
8	0.3464	0.4365	58	1.8
9	0.5327	0.4276	59	1.8
10	0.7136	0.4237	60	1.8
11	0.8930	0.4244	61	1.8
12	1.0749	0.4294	62	1.8
13	1.2633	0.4392	62	1.8
14	1.4625	0.4545	63	1.9
15	1.6788	0.4767	64	2.0
16	1.9205	0.5084	65	2.1
17	2.2014	0.5542	66	2.3
18	2.5459	0.6244	68	2.6
19	3.0078	0.7459	70	3.1
20	3.7573	1.0277	73	4.3
21	5.0003	1.8409	78	7.7

DRAFT

D.4.f: Level IV ELPA by Domain- Comprehension

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-2.8043	1.8393	46	6.5
1	-1.5654	1.0249	50	3.6
2	-0.8212	0.7429	53	2.6
3	-0.3631	0.6219	55	2.2
4	-0.0211	0.5524	56	2.0
5	0.2581	0.5070	57	1.8
6	0.4986	0.4752	58	1.7
7	0.7130	0.4520	59	1.6
8	0.9092	0.4346	59	1.5
9	1.0922	0.4217	60	1.5
10	1.2658	0.4121	60	1.5
11	1.4326	0.4052	61	1.4
12	1.5948	0.4007	62	1.4
13	1.7542	0.3982	62	1.4
14	1.9125	0.3978	63	1.4
15	2.0711	0.3992	63	1.4
16	2.2318	0.4027	64	1.4
17	2.3961	0.4083	65	1.4
18	2.5659	0.4164	65	1.5
19	2.7437	0.4273	66	1.5
20	2.9322	0.4417	66	1.6
21	3.1354	0.4606	67	1.6
22	3.3586	0.4856	68	1.7
23	3.6103	0.5193	69	1.8
24	3.9038	0.5667	70	2.0
25	4.2639	0.6381	71	2.3
26	4.7452	0.7603	73	2.7
27	5.5192	1.0412	76	3.7
28	6.7839	1.8501	80	6.6

DRAFT

D.5.a: Level V

Total ELPA

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-3.9290	1.8296	427	57.2
1	-2.7151	1.0070	465	31.5
2	-2.0080	0.7168	487	22.4
3	-1.5889	0.5890	500	18.4
4	-1.2878	0.5132	510	16.0
5	-1.0512	0.4619	517	14.4
6	-0.8556	0.4243	523	13.3
7	-0.6881	0.3953	528	12.4
8	-0.5411	0.3722	533	11.6
9	-0.4097	0.3533	537	11.0
10	-0.2906	0.3375	541	10.5
11	-0.1812	0.3242	544	10.1
12	-0.0799	0.3127	548	9.8
13	0.0148	0.3028	550	9.5
14	0.1038	0.2942	553	9.2
15	0.1882	0.2867	556	9.0
16	0.2685	0.2801	558	8.8
17	0.3452	0.2742	561	8.6
18	0.4190	0.2690	563	8.4
19	0.4900	0.2644	565	8.3
20	0.5588	0.2603	567	8.1
21	0.6256	0.2566	570	8.0
22	0.6906	0.2534	572	7.9
23	0.7541	0.2505	574	7.8
24	0.8162	0.2479	576	7.7
25	0.8771	0.2457	577	7.7
26	0.9369	0.2437	579	7.6
27	0.9959	0.2420	581	7.6
28	1.0541	0.2405	583	7.5
29	1.1116	0.2392	585	7.5
30	1.1686	0.2382	587	7.4
31	1.2251	0.2373	588	7.4
32	1.2812	0.2366	590	7.4
33	1.3370	0.2360	592	7.4
34	1.3927	0.2356	594	7.4
35	1.4481	0.2354	595	7.4
36	1.5035	0.2353	597	7.4
37	1.5588	0.2353	599	7.4
38	1.6142	0.2355	600	7.4
39	1.6697	0.2357	602	7.4
40	1.7254	0.2361	604	7.4
41	1.7812	0.2366	606	7.4
42	1.8374	0.2372	607	7.4
43	1.8938	0.2379	609	7.4

DRAFT

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
44	1.9505	0.2387	611	7.5
45	2.0077	0.2396	613	7.5
46	2.0653	0.2406	615	7.5
47	2.1235	0.2417	616	7.6
48	2.1822	0.2429	618	7.6
49	2.2415	0.2442	620	7.6
50	2.3014	0.2456	622	7.7
51	2.3621	0.2471	624	7.7
52	2.4236	0.2487	626	7.8
53	2.4859	0.2505	628	7.8
54	2.5491	0.2524	630	7.9
55	2.6133	0.2545	632	8.0
56	2.6787	0.2567	634	8.0
57	2.7452	0.2591	636	8.1
58	2.8129	0.2617	638	8.2
59	2.8821	0.2644	640	8.3
60	2.9529	0.2675	642	8.4
61	3.0253	0.2707	645	8.5
62	3.0995	0.2743	647	8.6
63	3.1758	0.2781	649	8.7
64	3.2543	0.2824	652	8.8
65	3.3353	0.2870	654	9.0
66	3.4191	0.2921	657	9.1
67	3.5061	0.2977	660	9.3
68	3.5965	0.3039	662	9.5
69	3.6910	0.3109	665	9.7
70	3.7901	0.3187	668	10.0
71	3.8944	0.3275	672	10.2
72	4.0050	0.3376	675	10.6
73	4.1228	0.3492	679	10.9
74	4.2493	0.3626	683	11.3
75	4.3865	0.3784	687	11.8
76	4.5367	0.3974	692	12.4
77	4.7037	0.4205	697	13.1
78	4.8924	0.4495	703	14.0
79	5.1110	0.4871	710	15.2
80	5.3726	0.5383	718	16.8
81	5.7016	0.6137	728	19.2
82	6.1527	0.7403	742	23.1
83	6.8970	1.0270	766	32.1
84	8.1412	1.8422	804	57.6

DRAFT

D.5.b: Level V ELPA by Domain- Listening

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-2.3399	1.8493	48	7.1
1	-1.0746	1.0433	53	4.0
2	-0.2917	0.7690	56	3.0
3	0.2067	0.6541	58	2.5
4	0.5906	0.5900	59	2.3
5	0.9138	0.5497	61	2.1
6	1.2007	0.5231	62	2.0
7	1.4644	0.5053	63	1.9
8	1.7136	0.4939	64	1.9
9	1.9540	0.4876	65	1.9
10	2.1905	0.4857	65	1.9
11	2.4271	0.4880	66	1.9
12	2.6681	0.4947	67	1.9
13	2.9182	0.5064	68	1.9
14	3.1833	0.5246	69	2.0
15	3.4720	0.5516	70	2.1
16	3.7977	0.5922	72	2.3
17	4.1846	0.6566	73	2.5
18	4.6867	0.7717	75	3.0
19	5.4745	1.0458	78	4.0
20	6.7436	1.8510	83	7.1

DRAFT

D.5.c: Level V ELPA by Domain- Reading

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-2.1039	1.8456	47	6.9
1	-0.8481	1.0368	52	3.9
2	-0.0786	0.7602	55	2.9
3	0.4063	0.6436	57	2.4
4	0.7765	0.5782	58	2.2
5	1.0858	0.5366	59	2.0
6	1.3580	0.5086	60	1.9
7	1.6064	0.4893	61	1.8
8	1.8390	0.4762	62	1.8
9	2.0615	0.4679	63	1.8
10	2.2782	0.4637	64	1.7
11	2.4927	0.4632	64	1.7
12	2.7085	0.4666	65	1.7
13	2.9294	0.4741	66	1.8
14	3.1596	0.4865	67	1.8
15	3.4049	0.5053	68	1.9
16	3.6735	0.5330	69	2.0
17	3.9787	0.5745	70	2.2
18	4.3446	0.6402	71	2.4
19	4.8249	0.7572	73	2.8
20	5.5900	1.0347	76	3.9
21	6.8428	1.8445	81	6.9

DRAFT

D.5.d: Level V ELPA by Domain- Writing

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-2.6215	1.8449	42	7.4
1	-1.3676	1.0355	47	4.1
2	-0.6010	0.7581	50	3.0
3	-0.1193	0.6412	52	2.6
4	0.2478	0.5755	53	2.3
5	0.5542	0.5343	54	2.1
6	0.8246	0.5073	55	2.0
7	1.0725	0.4898	56	2.0
8	1.3067	0.4792	57	1.9
9	1.5335	0.4740	58	1.9
10	1.7576	0.4734	59	1.9
11	1.9829	0.4766	60	1.9
12	2.2129	0.4833	61	1.9
13	2.4512	0.4935	62	2.0
14	2.7014	0.5075	63	2.0
15	2.9679	0.5259	64	2.1
16	3.2569	0.5503	65	2.2
17	3.5772	0.5834	66	2.3
18	3.9441	0.6307	68	2.5
19	4.3858	0.7036	70	2.8
20	4.9647	0.8294	72	3.3
21	5.8666	1.1102	75	4.4
22	7.2460	1.8976	81	7.6

DRAFT

D.5.e: Level V ELPA by Domain- Speaking

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-2.8822	1.8328	45	7.7
1	-1.6621	1.0105	50	4.2
2	-0.9492	0.7204	53	3.0
3	-0.5243	0.5947	55	2.5
4	-0.2148	0.5231	56	2.2
5	0.0340	0.4774	57	2.0
6	0.2468	0.4469	58	1.9
7	0.4368	0.4262	59	1.8
8	0.6123	0.4125	60	1.7
9	0.7787	0.4042	60	1.7
10	0.9403	0.4006	61	1.7
11	1.1009	0.4013	62	1.7
12	1.2637	0.4065	62	1.7
13	1.4326	0.4164	63	1.7
14	1.6122	0.4320	64	1.8
15	1.8082	0.4547	65	1.9
16	2.0291	0.4871	66	2.0
17	2.2884	0.5338	67	2.2
18	2.6100	0.6050	68	2.5
19	3.0466	0.7275	70	3.1
20	3.7661	1.0114	73	4.2
21	4.9836	1.8298	78	7.7

DRAFT

D.5.f: Level V ELPA by Domain- Comprehension

Raw Score	Theta	SE(Theta)	Scale Score	SE(Scale Score)
0	-2.6641	1.8415	47	6.5
1	-1.4193	1.0289	51	3.7
2	-0.6669	0.7483	54	2.7
3	-0.2008	0.6282	55	2.2
4	0.1490	0.5593	57	2.0
5	0.4358	0.5144	58	1.8
6	0.6836	0.4828	58	1.7
7	0.9052	0.4598	59	1.6
8	1.1084	0.4426	60	1.6
9	1.2983	0.4296	61	1.5
10	1.4786	0.4200	61	1.5
11	1.6519	0.4130	62	1.5
12	1.8204	0.4084	62	1.4
13	1.9860	0.4057	63	1.4
14	2.1501	0.4050	64	1.4
15	2.3145	0.4061	64	1.4
16	2.4804	0.4090	65	1.5
17	2.6497	0.4141	65	1.5
18	2.8240	0.4214	66	1.5
19	3.0056	0.4314	67	1.5
20	3.1973	0.4448	67	1.6
21	3.4027	0.4624	68	1.6
22	3.6270	0.4859	69	1.7
23	3.8781	0.5178	70	1.8
24	4.1688	0.5631	71	2.0
25	4.5232	0.6322	72	2.2
26	4.9948	0.7523	74	2.7
27	5.7538	1.0323	76	3.7
28	7.0037	1.8437	81	6.5

DRAFT

APPENDIX E: IRT STATISTICS

E.1: Level I Items

Item ID	Type	N-Count	Rasch Difficulty	SE of Rasch Difficulty	INFIT	OUTFIT	Step 1 Value	Step 2 Value	Step 3 Value	Step 4 Value
3540447	OP	9141	-2.47	0.03	1.08	1.28				
3467387	OP	9141	-1.83	0.02	1.10	1.10				
100000009514	OP	9141	-1.21	0.02	1.01	1.01				
100000009515	OP	9141	-3.33	0.04	0.96	0.82				
3536880	OP	9141	-1.98	0.03	0.92	0.86				
3484878	OP	9141	-1.27	0.02	1.20	1.27				
3477681	OP	9141	-1.05	0.02	1.01	1.01				
3540363	OP	9141	-0.77	0.02	1.10	1.14				
3540452	OP	9141	-1.34	0.02	0.94	0.92				
100000009509	OP	9141	-0.81	0.02	1.10	1.16				
100000009510	OP	9141	-1.18	0.02	0.96	0.95				
100000009511	OP	9141	-1.31	0.02	0.99	0.98				
3540453	OP	9141	-0.73	0.02	1.05	1.07				
3540454	OP	9141	-0.22	0.02	1.27	1.48				
3536612	OP	9141	-0.86	0.02	1.11	1.17				
3540455	OP	9141	-1.25	0.02	1.02	1.06				
3540775	OP	9141	-4.38	0.05	1.01	0.89				
3536628	OP	9141	-2.00	0.03	0.81	0.73				
3541699	OP	9141	-2.24	0.03	1.04	0.99				
3541548	OP	9141	-2.50	0.03	0.87	0.76				
3484912	OP	9141	-1.08	0.02	1.01	1.01				
3561062	OP	9141	-2.33	0.03	0.89	0.76				
3561060	OP	9141	-1.12	0.02	0.91	0.88				
3469887	OP	9141	-1.66	0.02	1.21	1.28				
100000009542	OP	9141	-0.95	0.02	0.99	0.96				
100000009526	OP	9141	-0.18	0.02	1.06	1.09				
100000009529	OP	9141	-0.49	0.02	1.07	1.09				
100000009545	OP	9141	-0.40	0.02	1.04	1.06				
100000009527	OP	9141	-0.58	0.02	1.01	1.05				
3526730	OP	9141	-0.08	0.02	1.19	1.28				
100000009538	OP	9141	0.14	0.02	0.97	1.00				
100000009539	OP	9141	-0.42	0.02	1.08	1.10				
100000009565	OP	9141	-1.08	0.02	1.09	1.13				
100000009566	OP	9141	-2.00	0.03	0.94	0.85				
3467772	OP	9141	-1.41	0.02	0.89	0.85				
100000009569	OP	9141	-0.34	0.02	0.90	0.92				
100000009580	OP	9141	-0.47	0.02	0.90	0.89				
3561085	OP	9141	-0.20	0.02	1.26	1.43				
3536736	OP	9141	-2.05	0.03	1.15	1.26				
100000009584	OP	9141	-1.06	0.02	1.00	1.01				
3537723	OP	9141	-0.06	0.02	0.93	0.92	0.17	-0.17		

DRAFT

Item ID	Type	N-Count	Rasch Difficulty	SE of Rasch Difficulty	INFIT	OUTFIT	Step 1 Value	Step 2 Value	Step 3 Value	Step 4 Value
100000009583	OP	9141	-0.55	0.02	1.09	1.03	-0.28	0.28		
100000017403	OP	9141	0.65	0.02	0.81	0.66	0.04	-0.04		
3563044	OP	9141	1.06	0.01	0.90	0.87	-1.30	-1.05	1.41	0.95
3484865	OP	9141	-3.58	0.04	1.18	0.85				
100000017412	OP	9141	-3.53	0.04	1.07	0.79				
3540678	OP	9141	-2.89	0.03	0.94	0.77				
100000017413	OP	9141	-2.55	0.03	0.84	0.69				
3536835	OP	9141	-1.09	0.02	0.96	0.97	-0.85	0.85		
100000012499	OP	9141	-1.10	0.01	1.00	1.01	-0.90	-0.73	0.13	1.50
100000017410	OP	9141	0.00	0.01	1.06	1.00	-0.93	-0.47	0.24	1.15
3562622	OP	9141	-0.96	0.01	1.12	1.21	-0.28	-0.62	0.58	0.32
100000166540	FT	1877	-1.04	0.05	1.03	1.04				
100000166541	FT	1877	-1.45	0.05	1.01	1.03				
100000166543	FT	1877	-0.62	0.05	1.06	1.08				
100000166548	FT	1766	-1.11	0.05	1.02	1.00				
100000166550	FT	1766	-1.54	0.05	1.06	1.02				
100000166552	FT	1766	-0.42	0.05	1.25	1.38				
100000166565	FT	1867	-1.28	0.05	0.94	0.92				
100000166566	FT	1867	-1.77	0.05	1.02	1.00				
100000166567	FT	1867	-1.31	0.05	1.03	1.04				
100000169217	FT	1755	-4.26	0.12	0.94	0.57				
100000169218	FT	1755	-4.24	0.12	0.99	0.81				
100000169219	FT	1755	-3.71	0.10	1.00	0.99				
100000169221	FT	1876	-0.73	0.05	1.14	1.20				
100000169222	FT	1876	-1.08	0.05	1.12	1.13				
100000169223	FT	1876	-2.27	0.06	0.98	0.92				
1000000166558	FT	1877	-0.73	0.05	0.96	0.94				
100000166555	FT	1867	-4.39	0.11	0.96	0.84				
100000166595	FT	1877	-2.18	0.06	1.00	1.00				
100000166598	FT	1867	-1.09	0.05	1.01	0.99				
100000166600	FT	1766	-0.96	0.05	0.96	0.93				
100000166602	FT	1766	-0.54	0.05	0.93	0.90				
100000168722	FT	1755	-3.34	0.08	0.93	0.79				
100000168726	FT	1867	-1.48	0.05	0.94	0.89				
100000169249	FT	1876	-0.99	0.05	0.86	0.82				
100000169309	FT	5510	-1.01	0.03	1.18	1.22				
100000169310	FT	3631	-0.11	0.04	0.92	0.93				
100000169417	FT	3631	0.71	0.04	1.00	1.09				
100000169418	FT	1755	-0.26	0.05	1.12	1.23				
100000169419	FT	1876	-1.01	0.05	0.99	0.97				
100000170522	FT	1877	-1.56	0.05	0.92	0.88				
100000177033	FT	1766	-2.08	0.06	0.99	0.89				
1000000174622	FT	1876	-1.49	0.05	0.99	1.00				
1000000174627	FT	1876	-0.93	0.03	0.95	0.94	0.85	-0.85		
100000174275	FT	1877	0.97	0.03	0.91	0.86	-1.10	-1.08	1.15	1.03

DRAFT

Item ID	Type	N-Count	Rasch Difficulty	SE of Rasch Difficulty	INFIT	OUTFIT	Step 1 Value	Step 2 Value	Step 3 Value	Step 4 Value
100000174565	FT	1867	1.21	0.03	0.91	0.91	-1.30	-1.40	0.97	1.72
100000174618	FT	1766	-1.08	0.05	0.88	0.83				
100000174620	FT	1755	-1.07	0.05	0.92	0.89				
100000174621	FT	1876	-2.28	0.06	0.88	0.75				
100000174624	FT	1755	-1.32	0.05	0.98	0.98				
100000174625	FT	1766	-1.88	0.06	1.04	1.08				
100000174628	FT	1766	0.44	0.04	0.82	0.64	0.54	-0.54		
100000174630	FT	1755	0.66	0.04	0.89	0.66	0.80	-0.80		
100000133352	FT	1755	-0.43	0.02	0.98	0.96	-0.93	-0.48	0.29	1.12
100000169953	FT	1877	-3.21	0.08	0.87	0.66				
100000169954	FT	1877	-3.44	0.08	0.91	0.76				
100000169957	FT	1867	-2.41	0.06	0.81	0.74				
100000169958	FT	1877	-1.33	0.04	0.92	0.92	-0.77	0.77		
100000169960	FT	1867	-1.20	0.04	0.83	0.81	-0.61	0.61		
100000169961	FT	1876	-1.05	0.02	1.16	1.20	-0.39	-0.67	0.82	0.24
100000176048	FT	1867	-3.40	0.08	0.82	0.58				
100000187659	FT	1766	-0.98	0.03	0.97	0.98	-0.91	-0.67	0.23	1.34

DRAFT

E.2: Level II Items

Item ID	Type	N-Count	Rasch Difficulty	SE of Rasch Difficulty	INFIT	OUTFIT	Step 1 Value	Step 2 Value	Step 3 Value	Step 4 Value
100000009618	OP	16003	-0.47	0.02	1.24	1.45				
100000056914	OP	16003	-0.06	0.02	1.01	1.01				
100000056912	OP	16003	-0.62	0.02	1.05	1.06				
100000009619	OP	16003	0.06	0.02	1.13	1.17				
3470853	OP	16003	-1.06	0.02	0.96	0.85				
100000056377	OP	16003	-0.58	0.02	1.04	1.02				
100000056913	OP	16003	0.77	0.02	1.15	1.23				
3561090	OP	16003	-1.47	0.02	1.23	1.10				
3563425	OP	16003	-0.02	0.02	1.00	0.99				
3561094	OP	16003	-0.33	0.02	1.04	1.06				
3558375	OP	16003	-0.96	0.02	0.98	0.94				
3560392	OP	16003	0.58	0.02	1.05	1.08				
3560391	OP	16003	-0.48	0.02	0.99	0.93				
100000055102	OP	16003	-1.67	0.03	0.99	1.10				
100000055103	OP	16003	-0.13	0.02	1.04	1.08				
100000055104	OP	16003	-0.68	0.02	0.99	1.00				
100000009609	OP	16003	0.22	0.02	1.05	1.06				
100000009610	OP	16003	0.64	0.02	1.18	1.25				
100000009611	OP	16003	0.62	0.02	1.08	1.13				
100000009612	OP	16003	0.57	0.02	1.07	1.11				
3540663	OP	16003	-3.29	0.05	0.76	0.42				
100000056382	OP	16003	-0.56	0.02	0.93	0.84				
3526731	OP	16003	-0.10	0.02	0.97	0.97				
100000009626	OP	16003	-0.46	0.02	1.07	1.01				
100000009625	OP	16003	-0.10	0.02	1.02	1.02				
100000009628	OP	16003	-0.11	0.02	1.11	1.23				
100000009627	OP	16003	0.04	0.02	1.12	1.13				
3526419	OP	16003	1.07	0.02	0.93	0.95				
3526418	OP	16003	0.47	0.02	1.08	1.09				
3537258	OP	16003	-0.81	0.02	0.87	0.71				
3537247	OP	16003	-0.51	0.02	0.85	0.72				
100000055189	OP	16003	-0.59	0.02	0.92	0.79				
100000055190	OP	16003	0.32	0.02	0.89	0.86				
100000055191	OP	16003	0.02	0.02	0.91	0.85				
100000055187	OP	16003	0.71	0.02	1.03	1.05				
100000055188	OP	16003	1.10	0.02	1.18	1.29				
100000009634	OP	16003	1.37	0.02	1.20	1.37				
100000009633	OP	16003	1.70	0.02	1.18	1.39				
100000009632	OP	16003	1.02	0.02	1.13	1.20				
100000009631	OP	16003	0.64	0.02	1.19	1.25				
3467778	OP	16003	-1.06	0.02	0.72	0.57				
100000055309	OP	16003	-0.80	0.02	0.94	0.84				

DRAFT

Item ID	Type	N-Count	Rasch Difficulty	SE of Rasch Difficulty	INFIT	OUTFIT	Step 1 Value	Step 2 Value	Step 3 Value	Step 4 Value
100000055308	OP	16003	-0.85	0.02	0.91	0.77				
100000009660	OP	16003	0.54	0.02	0.95	0.95				
3559284	OP	16003	-0.86	0.02	0.87	0.74				
100000089936	OP	16003	-0.45	0.02	0.89	0.81				
100000009655	OP	16003	-0.35	0.02	1.01	0.99				
3537453	OP	16003	0.90	0.02	1.10	1.16				
3537800	OP	16003	-0.15	0.01	0.94	1.03	0.92	-0.92		
3539399	OP	16003	0.38	0.01	0.82	0.84	-0.31	0.31		
100000009658	OP	16003	0.10	0.01	0.94	0.94	-0.59	0.59		
3562388	OP	16003	0.15	0.01	0.88	0.84	-0.72	0.72		
100000055310	OP	16003	0.80	0.01	0.85	0.85	-1.50	-1.51	1.04	1.97
100000017416	OP	16003	-2.51	0.04	1.02	0.77				
3540646	OP	16003	-1.54	0.03	0.97	0.78				
3540649	OP	16003	-2.25	0.03	0.84	0.56				
3540640	OP	16003	-1.52	0.02	0.87	0.70				
100000055344	OP	16003	0.05	0.01	0.96	0.94	-0.61	0.61		
3540643	OP	16003	0.04	0.01	0.93	0.92	-0.83	0.83		
100000056531	OP	16003	-0.11	0.01	0.98	0.99	-0.92	-1.00	0.14	1.78
100000017420	OP	16003	0.50	0.01	0.92	0.89	-1.11	-0.66	0.46	1.30
3562626	OP	16003	-0.29	0.01	1.24	1.31	-0.80	-0.83	1.12	0.52
1000000168352	FT	4042	0.73	0.03	1.08	1.19				
100000168530	FT	3967	0.96	0.03	1.17	1.29				
100000168531	FT	4042	0.92	0.03	1.09	1.13				
100000168540	FT	8009	-1.58	0.04	0.95	0.86				
100000168553	FT	3967	0.25	0.04	1.12	1.18				
100000168560	FT	8009	-0.43	0.03	1.08	1.11				
100000168563	FT	3967	-0.96	0.04	0.97	0.88				
100000169717	FT	4008	1.78	0.04	1.24	1.58				
100000169718	FT	4008	1.52	0.04	1.26	1.50				
100000169721	FT	4008	0.97	0.03	1.00	1.04				
100000169724	FT	4008	-0.45	0.04	1.07	1.19				
100000169725	FT	3986	-1.08	0.04	0.96	0.89				
100000169756	FT	3986	0.07	0.04	1.15	1.22				
100000169758	FT	3986	1.74	0.04	1.17	1.42				
100000169759	FT	3986	-1.12	0.04	0.90	0.77				
100000177102	FT	4042	-2.09	0.06	0.91	0.65				
100000177662	FT	4008	0.37	0.03	1.13	1.19				
100000179508	FT	3986	-0.66	0.04	1.17	1.56				
100000166557	FT	3967	-4.11	0.14	0.96	0.96				
100000166596	FT	4042	-3.83	0.12	1.01	0.89				
100000168466	FT	3986	0.47	0.03	0.89	0.87				
100000168467	FT	3986	0.32	0.04	0.91	0.88				
100000168468	FT	3986	-0.40	0.04	0.93	0.85				
100000168470	FT	3986	0.36	0.04	0.95	0.95				
100000168471	FT	3986	0.21	0.04	0.98	0.94				

DRAFT

Item ID	Type	N-Count	Rasch Difficulty	SE of Rasch Difficulty	INFIT	OUTFIT	Step 1 Value	Step 2 Value	Step 3 Value	Step 4 Value
100000168498	FT	4008	0.90	0.03	1.07	1.13				
100000168501	FT	4008	1.24	0.04	1.08	1.22				
100000168502	FT	4008	0.23	0.04	0.90	0.86				
100000168715	FT	3967	-0.65	0.04	0.87	0.73				
100000168716	FT	3967	0.05	0.04	0.84	0.76				
100000168718	FT	3967	0.87	0.03	1.02	1.04				
100000168725	FT	4042	-2.07	0.06	0.90	0.75				
100000168775	FT	4042	0.34	0.03	0.91	0.89				
100000170521	FT	3967	-0.90	0.04	1.02	0.91				
100000170532	FT	4042	-0.86	0.04	0.95	0.85				
100000175031	FT	4042	-0.50	0.04	0.93	0.84				
100000177642	FT	4008	0.89	0.03	1.08	1.14				
100000177643	FT	4008	0.40	0.03	0.92	0.88				
100000169017	FT	4008	-0.49	0.04	0.90	0.80				
100000169018	FT	4042	-1.03	0.04	0.86	0.70				
100000169021	FT	3967	-0.80	0.04	0.84	0.68				
100000169022	FT	3986	0.93	0.03	1.15	1.21				
100000174263	FT	4042	-0.07	0.02	0.98	0.99	-0.31	0.31		
100000174264	FT	3967	-0.56	0.03	0.91	0.93	-0.01	0.01		
100000174265	FT	4042	0.12	0.02	0.92	0.92	-0.30	0.30		
100000174266	FT	3967	0.42	0.02	1.00	1.01	-0.25	0.25		
100000177056	FT	4008	0.63	0.02	0.91	0.90	-1.22	-0.94	0.62	1.53
100000187554	FT	3986	0.30	0.02	0.95	0.96	-1.10	-1.19	0.72	1.57
100000138889	FT	4008	0.46	0.02	0.88	0.86	-1.07	-0.58	0.40	1.25
100000169039	FT	4042	-2.31	0.06	0.86	0.62				
100000169041	FT	4042	-0.67	0.03	0.94	0.91	-0.83	0.83		
100000169043	FT	4042	-0.45	0.03	0.92	0.91	-0.55	0.55		
100000174687	FT	4008	-2.90	0.08	0.91	0.56				
100000174688	FT	3967	-2.62	0.07	0.87	0.59				
100000177166	FT	3986	-0.30	0.02	1.06	1.11	-0.38	-1.16	1.01	0.52
100000187606	FT	3986	-2.75	0.08	0.89	0.50				
100000187662	FT	3967	-0.10	0.02	1.02	1.01	-1.26	-0.82	0.43	1.65

DRAFT

E.3: Level III Items

Item ID	Type	N-Count	Rasch Difficulty	SE of Rasch Difficulty	INFIT	OUTFIT	Step 1 Value	Step 2 Value	Step 3 Value	Step 4 Value
100000061235	OP	15081	1.88	0.02	1.14	1.20				
100000017442	OP	15081	-0.26	0.03	0.85	0.76				
100000017437	OP	15081	0.73	0.02	1.11	1.17				
100000017441	OP	15081	0.81	0.02	1.03	1.01				
100000017440	OP	15081	0.49	0.02	0.96	0.86				
100000056916	OP	15081	1.52	0.02	1.02	1.04				
100000056917	OP	15081	1.25	0.02	1.07	1.12				
3561098	OP	15081	1.57	0.02	1.05	1.08				
3470725	OP	15081	1.89	0.02	1.04	1.06				
100000009812	OP	15081	0.60	0.02	0.96	0.98				
100000009817	OP	15081	1.73	0.02	0.98	0.98				
100000055139	OP	15081	1.30	0.02	0.92	0.88				
100000055140	OP	15081	1.37	0.02	1.14	1.25				
100000055135	OP	15081	1.87	0.02	1.02	1.04				
100000055136	OP	15081	2.16	0.02	0.98	1.00				
100000018786	OP	15081	0.88	0.02	1.02	1.05				
100000018785	OP	15081	0.69	0.02	0.95	0.91				
100000018784	OP	15081	2.54	0.02	1.02	1.08				
100000018783	OP	15081	2.35	0.02	1.17	1.30				
3533550	OP	15081	0.25	0.02	1.06	1.12				
3534017	OP	15081	0.80	0.02	1.04	1.00				
100000056927	OP	15081	1.96	0.02	1.02	1.02				
100000056926	OP	15081	2.05	0.02	0.96	1.01				
100000009853	OP	15081	2.41	0.02	0.99	1.05				
100000009854	OP	15081	1.63	0.02	0.97	0.96				
100000009855	OP	15081	1.59	0.02	1.00	1.03				
100000009858	OP	15081	1.74	0.02	0.95	0.94				
100000055204	OP	15081	2.73	0.02	0.95	0.98				
100000055200	OP	15081	1.28	0.02	0.90	0.84				
100000055203	OP	15081	1.80	0.02	0.98	0.98				
100000055199	OP	15081	2.40	0.02	1.06	1.14				
100000055202	OP	15081	3.13	0.02	1.20	1.54				
3537588	OP	15081	1.63	0.02	0.91	0.87				
3526459	OP	15081	2.82	0.02	1.02	1.16				
3526458	OP	15081	2.18	0.02	1.11	1.17				
3526461	OP	15081	2.25	0.02	1.00	1.03				
100000085124	OP	15081	2.73	0.02	1.01	1.11				
100000085122	OP	15081	2.53	0.02	1.07	1.17				
100000085123	OP	15081	2.46	0.02	1.09	1.16				
3529874	OP	15081	3.13	0.02	1.08	1.39				
3529880	OP	15081	2.74	0.02	1.17	1.33				
3526783	OP	15081	0.80	0.02	0.87	0.82				
3537520	OP	15081	1.49	0.02	0.89	0.86				

DRAFT

Item ID	Type	N-Count	Rasch Difficulty	SE of Rasch Difficulty	INFIT	OUTFIT	Step 1 Value	Step 2 Value	Step 3 Value	Step 4 Value
100000057816	OP	15081	1.87	0.02	1.07	1.10				
100000057815	OP	15081	2.58	0.02	1.11	1.21				
3526794	OP	15081	0.33	0.02	0.93	0.92				
3526795	OP	15081	0.42	0.02	0.89	0.76				
100000057810	OP	15081	1.21	0.02	0.95	0.95				
3526811	OP	15081	0.82	0.02	0.91	0.84				
100000009889	OP	15081	2.73	0.02	0.95	0.94	-1.69	1.69		
100000056930	OP	15081	1.54	0.01	0.95	0.95	-1.19	1.19		
3563477	OP	15081	1.97	0.01	0.89	0.89	-1.71	-1.70	2.26	
100000017435	OP	15081	2.25	0.01	1.01	1.01	-2.37	-0.81	0.82	2.37
3540626	OP	15081	-1.00	0.04	1.04	0.75				
3540561	OP	15081	-1.00	0.04	0.80	0.70				
3540625	OP	15081	-0.42	0.03	0.92	0.77				
3563708	OP	15081	-0.07	0.03	0.88	0.74				
3540565	OP	15081	1.51	0.01	0.98	0.99	-0.72	0.72		
100000056534	OP	15081	1.34	0.01	0.91	0.89	-0.61	0.61		
100000012506	OP	15081	0.66	0.01	1.02	1.04	-0.67	-1.15	0.13	1.69
100000017448	OP	15081	1.04	0.01	0.91	0.91	-1.14	-0.80	0.37	1.57
100000017446	OP	15081	0.91	0.01	0.96	0.99	-0.48	-0.81	0.93	0.37
100000169766	FT	4715	2.95	0.03	0.98	1.15				
100000169776	FT	4715	0.26	0.04	1.02	1.14				
100000169874	FT	4820	2.31	0.03	1.21	1.32				
100000174047	FT	5546	0.39	0.04	1.00	0.97				
100000174049	FT	5546	1.74	0.03	1.03	1.03				
100000174192	FT	5546	1.45	0.03	0.99	1.00				
100000174193	FT	5546	2.11	0.03	1.11	1.20				
100000174195	FT	5546	2.03	0.03	1.14	1.22				
100000176490	FT	4820	1.24	0.03	1.01	1.04				
100000176868	FT	4820	1.45	0.03	1.08	1.15				
100000200919	FT	4820	0.38	0.04	0.96	0.90				
100000200927	FT	4715	1.84	0.03	0.99	1.01				
100000200928	FT	4715	2.36	0.03	1.19	1.32				
3526859	FT	9535	2.31	0.02	1.10	1.19				
3526862	FT	9535	1.62	0.02	1.07	1.10				
100000167861	FT	4715	1.64	0.03	0.94	0.92				
100000169856	FT	4820	1.41	0.03	0.92	0.88				
100000169857	FT	4715	1.54	0.03	1.01	1.01				
100000169858	FT	4715	1.77	0.03	0.98	0.98				
100000169860	FT	4715	0.92	0.04	0.95	0.89				
100000169863	FT	4820	2.35	0.03	1.02	1.07				
100000169867	FT	4820	3.21	0.03	1.18	1.44				
100000169869	FT	4820	2.29	0.03	1.09	1.17				
100000169870	FT	4715	2.57	0.03	1.14	1.23				
100000176337	FT	4820	2.61	0.03	1.02	1.09				
100000200832	FT	5546	0.86	0.03	0.85	0.74				

DRAFT

Item ID	Type	N-Count	Rasch Difficulty	SE of Rasch Difficulty	INFIT	OUTFIT	Step 1 Value	Step 2 Value	Step 3 Value	Step 4 Value
100000200835	FT	5546	2.72	0.03	1.16	1.28				
100000200837	FT	5546	2.25	0.03	0.99	1.02				
100000200838	FT	5546	2.02	0.03	0.95	0.95				
100000200839	FT	5546	1.26	0.03	0.97	0.94				
100000017434	FT	4820	2.09	0.02	0.90	0.90	-1.49	-1.37	1.02	1.84
100000174325	FT	4820	0.84	0.04	0.99	1.01				
100000174333	FT	5546	2.61	0.03	0.92	0.91	-1.60	1.60		
100000177112	FT	5546	2.04	0.02	0.92	0.92	-1.84	-1.11	1.04	1.91
100000177119	FT	4715	2.12	0.02	0.96	0.96	-1.63	-1.26	1.00	1.89
100000187597	FT	4715	1.16	0.03	1.01	1.00				
100000169497	FT	4715	-0.76	0.06	0.84	0.63				
100000169499	FT	4820	0.24	0.04	0.88	0.80				
100000169910	FT	5546	-0.39	0.05	0.84	0.72				
100000169911	FT	5546	0.88	0.02	1.02	1.20	-0.17	-1.03	1.02	0.18
100000187668	FT	4715	0.83	0.02	0.91	0.94	-0.99	-0.99	0.23	1.76
100000200744	FT	4820	1.60	0.02	0.88	0.88	-1.32	-0.95	0.61	1.66

DRAFT

E.4: Level IV Items

Item ID	Type	N-Count	Rasch Difficulty	SE of Rasch Difficulty	INFIT	OUTFIT	Step 1 Value	Step 2 Value	Step 3 Value	Step 4 Value
100000056491	OP	10947	0.77	0.03	1.09	1.13				
100000017454	OP	10947	0.96	0.02	1.04	1.12				
100000017455	OP	10947	1.36	0.02	1.01	1.03				
100000017456	OP	10947	2.15	0.02	1.13	1.18				
100000017457	OP	10947	1.20	0.02	0.99	0.99				
100000056493	OP	10947	0.91	0.02	1.02	1.08				
3526745	OP	10947	1.13	0.02	1.19	1.24				
100000008945	OP	10947	1.23	0.02	1.10	1.14				
100000056492	OP	10947	2.06	0.02	1.10	1.14				
100000008946	OP	10947	2.10	0.02	1.10	1.13				
100000056490	OP	10947	2.26	0.02	1.03	1.06				
100000008952	OP	10947	1.18	0.02	1.00	0.98				
100000008949	OP	10947	2.22	0.02	1.06	1.12				
100000008953	OP	10947	2.25	0.02	1.19	1.29				
100000008950	OP	10947	2.27	0.02	0.94	0.96				
100000008954	OP	10947	1.67	0.02	1.09	1.13				
100000055211	OP	10947	1.50	0.02	0.96	0.92				
100000055212	OP	10947	3.28	0.02	1.24	1.53				
100000055214	OP	10947	1.60	0.02	1.02	1.00				
100000055215	OP	10947	4.22	0.03	1.17	1.84				
100000055216	OP	10947	2.16	0.02	1.05	1.09				
3535829	OP	10947	1.12	0.02	1.05	1.06				
100000056501	OP	10947	1.44	0.02	0.94	0.91				
100000056503	OP	10947	1.36	0.02	1.02	1.02				
3535677	OP	10947	1.59	0.02	0.85	0.80				
100000056500	OP	10947	2.66	0.02	1.19	1.34				
100000055152	OP	10947	1.71	0.02	1.03	1.03				
100000055154	OP	10947	1.50	0.02	0.97	0.94				
100000055155	OP	10947	2.33	0.02	0.99	1.02				
100000055156	OP	10947	2.83	0.02	1.14	1.32				
100000008958	OP	10947	2.21	0.02	0.96	0.98				
100000008959	OP	10947	1.22	0.02	0.96	0.92				
100000008960	OP	10947	1.08	0.02	0.84	0.73				
100000008961	OP	10947	2.43	0.02	1.11	1.18				
100000008962	OP	10947	2.21	0.02	0.99	1.02				
100000008963	OP	10947	1.91	0.02	1.01	1.03				
3558745	OP	10947	1.91	0.02	1.07	1.11				
3558748	OP	10947	2.23	0.02	1.08	1.13				
3558753	OP	10947	2.31	0.02	1.10	1.17				
3558750	OP	10947	2.35	0.02	1.05	1.10				
3558739	OP	10947	1.73	0.02	1.11	1.17				
3558738	OP	10947	2.89	0.02	1.00	1.10				
3526775	OP	10947	0.62	0.03	1.14	1.24				

DRAFT

Item ID	Type	N-Count	Rasch Difficulty	SE of Rasch Difficulty	INFIT	OUTFIT	Step 1 Value	Step 2 Value	Step 3 Value	Step 4 Value
3471438	OP	10947	1.17	0.02	0.96	0.91				
100000056517	OP	10947	1.38	0.02	1.11	1.15				
3558613	OP	10947	1.54	0.02	1.05	1.10				
3537533	OP	10947	2.65	0.02	1.03	1.11				
100000056522	OP	10947	0.34	0.03	0.98	0.98				
3526803	OP	10947	0.54	0.03	1.09	1.08				
100000056520	OP	10947	1.02	0.02	0.99	0.98				
3558575	OP	10947	1.31	0.02	1.07	1.09				
3558574	OP	10947	1.54	0.02	1.08	1.07				
3539646	OP	10947	1.31	0.02	1.12	1.15	0.06	-0.06		
3539593	OP	10947	1.59	0.01	1.07	1.12	0.04	-0.04		
3566110	OP	10947	2.25	0.01	1.17	1.21	-2.36	-1.40	0.57	3.20
100000055317	OP	10947	1.67	0.01	0.81	0.82	-1.31	-1.45	0.78	1.98
3566131	OP	10947	-0.11	0.03	0.91	0.71				
3566132	OP	10947	0.02	0.03	0.83	0.62				
3540821	OP	10947	0.70	0.03	0.80	0.69				
3540819	OP	10947	-0.75	0.04	0.83	0.52				
3540820	OP	10947	-0.78	0.04	0.83	0.52				
100000056536	OP	10947	1.34	0.02	0.84	0.81	-0.59	0.59		
3538699	OP	10947	1.43	0.02	0.88	0.85	-0.66	0.66		
100000012509	OP	10947	0.91	0.01	0.93	0.94	-0.89	-0.96	0.32	1.54
100000017467	OP	10947	1.21	0.01	0.87	0.87	-0.66	-0.85	0.27	1.23
100000055348	OP	10947	0.93	0.01	0.98	1.03	0.13	-1.08	1.03	-0.08
100000167865	FT	3796	0.64	0.04	0.98	0.98				
100000167866	FT	3796	0.93	0.04	1.10	1.15				
100000167868	FT	7151	0.97	0.03	1.02	1.04				
100000167970	FT	3578	0.11	0.05	0.89	0.75				
100000167973	FT	3573	3.12	0.04	1.11	1.28				
100000167979	FT	3573	0.52	0.05	0.94	0.88				
100000167981	FT	3573	1.32	0.04	1.02	1.02				
100000167982	FT	3578	1.41	0.04	0.98	0.97				
100000167983	FT	3578	1.68	0.04	1.07	1.10				
100000176840	FT	3573	4.43	0.05	1.10	2.28*				
100000176841	FT	3578	2.24	0.04	1.06	1.07				
100000204385	FT	3796	0.87	0.04	0.96	0.93				
100000204386	FT	3796	2.25	0.04	1.03	1.06				
100000204387	FT	3796	1.28	0.04	0.92	0.88				
100000204388	FT	3796	1.72	0.04	1.01	1.02				
100000167959	FT	3796	1.33	0.04	0.86	0.78				
100000167961	FT	3796	0.11	0.05	0.90	0.76				
100000167962	FT	3796	2.29	0.04	1.02	1.05				
100000167964	FT	3796	2.17	0.04	0.98	1.00				
100000167965	FT	3796	1.91	0.04	0.98	0.98				
100000167986	FT	3573	1.48	0.04	0.94	0.89				
100000167993	FT	3578	1.44	0.04	0.89	0.82				

DRAFT

Item ID	Type	N-Count	Rasch Difficulty	SE of Rasch Difficulty	INFIT	OUTFIT	Step 1 Value	Step 2 Value	Step 3 Value	Step 4 Value
100000167994	FT	3573	1.65	0.04	0.91	0.88				
100000167995	FT	3578	1.56	0.04	0.89	0.82				
100000167996	FT	7151	1.64	0.03	0.93	0.90				
100000167997	FT	3578	2.95	0.04	1.00	1.09				
100000167998	FT	3573	2.34	0.04	1.08	1.18				
100000168003	FT	3573	0.97	0.04	0.94	0.89				
100000168004	FT	3578	1.16	0.04	0.91	0.84				
100000174571	FT	3573	1.49	0.02	1.06	1.10	0.20	-0.20		
100000174573	FT	3578	2.29	0.02	0.87	0.89	-1.17	-1.84	0.76	2.25
100000174574	FT	3573	2.23	0.02	0.90	0.90	-1.20	-1.61	0.94	1.86
100000174576	FT	3578	1.95	0.02	1.12	1.14	-0.25	0.25		
100000174591	FT	3796	2.74	0.04	1.10	1.20				
100000174593	FT	3796	2.96	0.04	1.02	1.07				
100000174594	FT	3796	2.00	0.02	0.82	0.82	-1.54	-1.48	0.95	2.07
100000134324	FT	3578	1.34	0.02	0.88	0.91	-1.03	-0.89	0.51	1.41
1000000174670	FT	3796	-1.01	0.08	0.83	0.45				
100000176694	FT	3578	-0.35	0.06	0.80	0.54				
100000176695	FT	3796	0.99	0.02	1.03	1.12	0.58	-1.49	1.14	-0.23
100000187670	FT	3573	0.88	0.02	0.91	0.91	-0.61	-0.96	0.26	1.30
3540817	FT	3573	0.25	0.05	0.77	0.64				

Note: * indicates that an item had an INFIT or OUTFIT value greater than 2.0 and so distorted or degraded the measurement system.

DRAFT

E.5: Level V Items

Item ID	Type	N-Count	Rasch Difficulty	SE of Rasch Difficulty	INFIT	OUTFIT	Step 1 Value	Step 2 Value	Step 3 Value	Step 4 Value
100000061241	OP	10908	0.73	0.03	0.85	0.76				
100000017475	OP	10908	3.24	0.02	1.14	1.28				
100000017476	OP	10908	1.56	0.02	0.98	0.95				
100000055163	OP	10908	1.01	0.03	0.95	0.87				
100000055164	OP	10908	2.84	0.02	1.07	1.13				
3566062	OP	10908	0.74	0.03	0.95	0.97				
100000056495	OP	10908	1.53	0.02	1.06	1.10				
3566089	OP	10908	2.40	0.02	1.17	1.24				
3526756	OP	10908	1.69	0.02	0.99	0.99				
100000009039	OP	10908	3.10	0.02	1.10	1.28				
100000056480	OP	10908	3.50	0.02	1.12	1.31				
100000056469	OP	10908	1.78	0.02	1.00	1.03				
100000056470	OP	10908	1.56	0.02	1.01	1.05				
100000056471	OP	10908	1.74	0.02	1.02	1.05				
100000056933	OP	10908	2.11	0.02	1.13	1.23				
100000056472	OP	10908	2.29	0.02	1.03	1.04				
3526516	OP	10908	2.80	0.02	1.21	1.35				
3536855	OP	10908	2.64	0.02	1.10	1.14				
3536878	OP	10908	3.94	0.02	1.12	1.57				
3536738	OP	10908	2.69	0.02	1.05	1.08				
3558509	OP	10908	1.35	0.02	0.87	0.77				
100000056505	OP	10908	1.97	0.02	1.06	1.09				
100000056506	OP	10908	2.85	0.02	1.00	1.05				
100000009007	OP	10908	1.46	0.02	0.83	0.72				
100000009009	OP	10908	1.32	0.03	0.75	0.60				
100000009008	OP	10908	1.53	0.02	0.87	0.79				
100000009006	OP	10908	1.20	0.03	0.76	0.63				
100000009010	OP	10908	1.60	0.02	0.85	0.76				
100000009005	OP	10908	3.17	0.02	1.15	1.35				
100000057368	OP	10908	2.69	0.02	1.05	1.14				
100000057370	OP	10908	2.87	0.02	1.06	1.16				
100000057369	OP	10908	2.90	0.02	1.05	1.12				
100000057373	OP	10908	3.66	0.02	1.23	1.54				
100000057371	OP	10908	3.03	0.02	1.04	1.13				
100000057372	OP	10908	2.25	0.02	1.05	1.06				
3562741	OP	10908	2.17	0.02	1.11	1.14				
3562743	OP	10908	3.05	0.02	0.98	1.05				
3562744	OP	10908	2.69	0.02	1.10	1.18				
3562737	OP	10908	2.46	0.02	0.96	0.94				
3562746	OP	10908	3.26	0.02	1.25	1.48				
3562742	OP	10908	2.49	0.02	1.01	1.02				
3526786	OP	10908	0.82	0.03	0.73	0.59				
100000009014	OP	10908	2.42	0.02	0.93	0.92				

DRAFT

Item ID	Type	N-Count	Rasch Difficulty	SE of Rasch Difficulty	INFIT	OUTFIT	Step 1 Value	Step 2 Value	Step 3 Value	Step 4 Value
100000056526	OP	10908	2.85	0.02	1.13	1.20				
3470663	OP	10908	3.17	0.02	1.16	1.29				
100000056527	OP	10908	3.02	0.02	1.28	1.50				
3526809	OP	10908	0.86	0.03	0.86	0.76				
3526810	OP	10908	1.37	0.02	0.87	0.79				
100000056529	OP	10908	1.51	0.02	0.88	0.82				
3558631	OP	10908	1.80	0.02	0.88	0.82				
3558630	OP	10908	1.86	0.02	0.85	0.77				
3537943	OP	10908	1.32	0.02	1.27	1.26	-1.07	1.07		
3558502	OP	10908	1.86	0.02	1.21	1.26	-0.53	0.53		
100000017489	OP	10908	2.40	0.01	1.05	1.07	-2.38	-1.34	0.58	3.14
100000055334	OP	10908	2.56	0.01	0.88	0.88	-1.32	-1.33	0.78	1.86
3566126	OP	10908	0.13	0.03	0.93	0.73				
3541261	OP	10908	0.54	0.03	0.72	0.58				
3541263	OP	10908	0.67	0.03	0.86	0.65				
3541258	OP	10908	0.76	0.03	0.78	0.62				
3541259	OP	10908	0.17	0.03	0.91	0.70				
3562634	OP	10908	1.34	0.02	0.76	0.72	-0.21	0.21		
100000055355	OP	10908	0.88	0.02	0.82	0.79	-0.31	0.31		
100000012511	OP	10908	1.23	0.01	0.85	0.86	-0.67	-0.89	0.25	1.31
100000017492	OP	10908	1.67	0.01	0.78	0.78	-0.56	-0.74	0.33	0.97
100000055353	OP	10908	0.86	0.01	0.97	1.24	0.27	-1.16	1.54	-0.65
100000009038	FT	7661	3.29	0.03	1.03	1.22				
100000168102	FT	3247	2.23	0.04	1.04	1.03				
100000168104	FT	3247	3.21	0.04	1.05	1.17				
100000168110	FT	3247	1.29	0.05	0.82	0.67				
100000168111	FT	3247	1.59	0.04	0.95	0.98				
100000168273	FT	3667	1.96	0.04	0.97	0.97				
100000168274	FT	3247	1.61	0.04	0.88	0.79				
100000168276	FT	3667	3.90	0.04	1.14	1.48				
100000168277	FT	3667	1.49	0.04	0.88	0.81				
100000197314	FT	3247	1.89	0.04	1.02	0.98				
100000198197	FT	3994	2.56	0.04	1.12	1.18				
100000198198	FT	3994	2.36	0.04	1.10	1.14				
100000198199	FT	3994	3.18	0.03	1.12	1.23				
100000200737	FT	3994	2.04	0.04	1.12	1.20				
100000200738	FT	3994	1.76	0.04	1.07	1.11				
100000202225	FT	3667	4.06	0.04	1.12	1.51				
100000202226	FT	3667	2.33	0.04	1.16	1.25				
100000168051	FT	3994	1.56	0.04	0.96	0.90				
100000168052	FT	3994	3.29	0.04	1.06	1.25				
100000168053	FT	3994	2.56	0.03	1.02	1.05				
100000168054	FT	3994	1.83	0.04	1.03	1.03				
100000168055	FT	3994	2.67	0.03	0.97	1.00				
100000168068	FT	3667	1.24	0.04	0.94	0.86				

DRAFT

Item ID	Type	N-Count	Rasch Difficulty	SE of Rasch Difficulty	INFIT	OUTFIT	Step 1 Value	Step 2 Value	Step 3 Value	Step 4 Value
100000168069	FT	3247	2.12	0.04	1.03	0.99				
100000168224	FT	3667	2.59	0.04	1.11	1.20				
100000168225	FT	3247	2.48	0.04	1.14	1.20				
100000168226	FT	6914	2.92	0.03	1.00	1.12				
100000168227	FT	3247	3.13	0.04	1.20	1.36				
100000168228	FT	3667	2.23	0.04	0.86	0.80				
100000168229	FT	3667	3.05	0.04	1.10	1.20				
100000168230	FT	3247	2.01	0.04	1.05	1.07				
100000174583	FT	3667	1.86	0.04	0.85	0.75				
100000174584	FT	3247	1.25	0.05	0.90	0.83				
100000174585	FT	3667	1.60	0.04	0.91	0.88				
100000174599	FT	3247	2.13	0.04	0.99	0.99				
100000174601	FT	3667	2.69	0.02	0.95	0.95	-1.25	-1.54	0.71	2.08
100000174606	FT	3994	1.66	0.03	1.02	1.09	-0.21	0.21		
100000176753	FT	3994	2.16	0.02	0.89	0.90	-1.09	-1.95	0.81	2.23
100000176771	FT	3247	2.66	0.02	0.99	1.01	-0.90	-1.97	0.68	2.19
100000134326	FT	3247	1.68	0.02	0.75	0.75	-0.81	-0.59	0.28	1.12
100000168292	FT	3994	1.13	0.02	0.88	0.94	0.39	-1.37	1.34	-0.37
100000174718	FT	3247	1.11	0.05	0.76	0.66				
100000174719	FT	3994	0.34	0.05	0.79	0.58				
100000174728	FT	3667	1.22	0.04	0.79	0.70				
100000187673	FT	3667	1.45	0.02	0.86	0.89	-0.65	-0.52	0.15	1.01

DRAFT

APPENDIX F: ITEM INFORMATION AT DIFFERENT CUT SCORES BY GRADE

F.1: Grade K

Total ELPA

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
3540447	-2.47	0.23	0.16	0.12	0.07
3467387	-1.83	0.25	0.22	0.18	0.11
100000009514	-1.21	0.23	0.25	0.23	0.17
100000009515	-3.33	0.15	0.09	0.06	0.03
3536880	-1.98	0.25	0.20	0.16	0.10
3484878	-1.27	0.23	0.25	0.22	0.17
3477681	-1.05	0.22	0.25	0.24	0.19
3540363	-0.77	0.19	0.24	0.25	0.21
3540452	-1.34	0.24	0.25	0.22	0.16
100000009509	-0.81	0.20	0.25	0.25	0.21
100000009510	-1.18	0.23	0.25	0.23	0.18
100000009511	-1.31	0.23	0.25	0.22	0.16
3540453	-0.73	0.19	0.24	0.25	0.22
3540454	-0.22	0.14	0.21	0.24	0.25
3536612	-0.86	0.20	0.25	0.25	0.20
3540455	-1.25	0.23	0.25	0.23	0.17
3540775	-4.38	0.07	0.03	0.02	0.01
3536628	-2.00	0.25	0.20	0.16	0.10
3541699	-2.24	0.24	0.18	0.14	0.08
3541548	-2.50	0.22	0.16	0.11	0.07
3484912	-1.08	0.22	0.25	0.24	0.18
3561062	-2.33	0.23	0.17	0.13	0.08
3561060	-1.12	0.22	0.25	0.23	0.18
3469887	-1.66	0.25	0.23	0.19	0.13
100000009542	-0.95	0.21	0.25	0.24	0.20
100000009526	-0.18	0.14	0.21	0.24	0.25
100000009529	-0.49	0.17	0.23	0.25	0.23
100000009545	-0.40	0.16	0.22	0.25	0.24
100000009527	-0.58	0.17	0.24	0.25	0.23
3526730	-0.08	0.13	0.20	0.23	0.25
100000009538	0.14	0.11	0.18	0.22	0.25
100000009539	-0.42	0.16	0.23	0.25	0.24
100000009565	-1.08	0.22	0.25	0.24	0.18
100000009566	-2.00	0.25	0.20	0.16	0.10
3467772	-1.41	0.24	0.24	0.21	0.15
100000009569	-0.34	0.15	0.22	0.25	0.24
100000009580	-0.47	0.16	0.23	0.25	0.23
3561085	-0.20	0.14	0.21	0.24	0.25
3536736	-2.05	0.25	0.20	0.15	0.10

DRAFT

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000009584	-1.06	0.22	0.25	0.24	0.19
3537723	-0.06	0.18	0.40	0.58	0.70
100000009583	-0.55	0.21	0.40	0.52	0.60
100000017403	0.65	0.19	0.40	0.56	0.67
3563044	1.06	0.57	0.75	0.76	0.76
3484865	-3.58	0.13	0.07	0.05	0.03
100000017412	-3.53	0.13	0.07	0.05	0.03
3540678	-2.89	0.19	0.12	0.08	0.05
100000017413	-2.55	0.22	0.15	0.11	0.06
3536835	-1.09	0.25	0.37	0.43	0.46
100000012499	-1.10	0.50	0.93	1.10	1.03
100000017410	0.00	0.44	0.84	1.08	1.16
3562622	-0.96	0.34	0.81	1.21	1.49

DRAFT

F.2: Grade 1

Total ELPA

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000009618	-0.47	0.23	0.25	0.17	0.11
100000056914	-0.06	0.20	0.25	0.21	0.14
100000056912	-0.62	0.24	0.24	0.16	0.09
100000009619	0.06	0.19	0.24	0.22	0.15
3470853	-1.06	0.25	0.22	0.12	0.07
100000056377	-0.58	0.24	0.24	0.16	0.10
100000056913	0.77	0.12	0.19	0.25	0.22
3561090	-1.47	0.24	0.18	0.08	0.05
3563425	-0.02	0.20	0.25	0.21	0.15
3561094	-0.33	0.22	0.25	0.18	0.12
3558375	-0.96	0.25	0.22	0.13	0.07
3560392	0.58	0.14	0.21	0.25	0.20
3560391	-0.48	0.23	0.25	0.17	0.11
100000055102	-1.67	0.22	0.16	0.07	0.04
100000055103	-0.13	0.21	0.25	0.20	0.14
100000055104	-0.68	0.24	0.24	0.15	0.09
100000009609	0.22	0.18	0.24	0.23	0.17
100000009610	0.64	0.14	0.20	0.25	0.21
100000009611	0.62	0.14	0.21	0.25	0.21
100000009612	0.57	0.14	0.21	0.25	0.20
3540663	-3.29	0.08	0.04	0.02	0.01
100000056382	-0.56	0.24	0.25	0.16	0.10
3526731	-0.10	0.20	0.25	0.21	0.14
100000009626	-0.46	0.23	0.25	0.17	0.11
100000009625	-0.10	0.20	0.25	0.21	0.14
100000009628	-0.11	0.21	0.25	0.20	0.14
100000009627	0.04	0.19	0.24	0.22	0.15
3526419	1.07	0.10	0.16	0.25	0.24
3526418	0.47	0.15	0.22	0.24	0.19
3537258	-0.81	0.25	0.23	0.14	0.08
3537247	-0.51	0.24	0.25	0.17	0.10
100000055189	-0.59	0.24	0.24	0.16	0.10
100000055190	0.32	0.17	0.23	0.24	0.18
100000055191	0.02	0.19	0.24	0.22	0.15
100000055187	0.71	0.13	0.20	0.25	0.21
100000055188	1.10	0.10	0.16	0.24	0.24
100000009634	1.37	0.08	0.14	0.23	0.25
100000009633	1.70	0.06	0.11	0.21	0.25
100000009632	1.02	0.10	0.17	0.25	0.24
100000009631	0.64	0.14	0.20	0.25	0.21
3467778	-1.06	0.25	0.22	0.12	0.07
100000055309	-0.80	0.25	0.23	0.14	0.08
100000055308	-0.85	0.25	0.23	0.14	0.08
100000009660	0.54	0.14	0.21	0.25	0.20

DRAFT

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
3559284	-0.86	0.25	0.23	0.13	0.08
100000089936	-0.45	0.23	0.25	0.17	0.11
100000009655	-0.35	0.22	0.25	0.18	0.12
3537453	0.90	0.11	0.18	0.25	0.23
3537800	-0.15	0.43	0.79	0.54	0.22
3539399	0.38	0.41	0.58	0.47	0.28
100000009658	0.10	0.40	0.51	0.44	0.29
3562388	0.15	0.39	0.48	0.42	0.30
100000055310	0.80	0.74	0.60	0.60	0.61
100000017416	-2.51	0.15	0.09	0.03	0.02
3540646	-1.54	0.23	0.17	0.08	0.04
3540649	-2.25	0.17	0.11	0.04	0.02
3540640	-1.52	0.23	0.17	0.08	0.04
100000055344	0.05	0.40	0.51	0.44	0.29
3540643	0.04	0.38	0.46	0.41	0.30
100000056531	-0.11	1.00	1.00	0.66	0.50
100000017420	0.50	0.86	1.02	0.84	0.58
3562626	-0.29	0.83	1.04	1.03	0.65

DRAFT

F.3: Grade 2

Total ELPA

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000009618	-0.47	0.25	0.23	0.11	0.07
100000056914	-0.06	0.22	0.25	0.15	0.09
100000056912	-0.62	0.25	0.22	0.10	0.06
100000009619	0.06	0.21	0.25	0.16	0.10
3470853	-1.06	0.24	0.18	0.07	0.04
100000056377	-0.58	0.25	0.22	0.10	0.06
100000056913	0.77	0.15	0.23	0.22	0.16
3561090	-1.47	0.22	0.14	0.05	0.03
3563425	-0.02	0.22	0.25	0.15	0.09
3561094	-0.33	0.24	0.24	0.13	0.07
3558375	-0.96	0.25	0.19	0.08	0.04
3560392	0.58	0.17	0.24	0.21	0.15
3560391	-0.48	0.25	0.23	0.11	0.06
100000055102	-1.67	0.20	0.12	0.04	0.02
100000055103	-0.13	0.23	0.25	0.14	0.09
100000055104	-0.68	0.25	0.21	0.10	0.05
100000009609	0.22	0.20	0.25	0.18	0.11
100000009610	0.64	0.16	0.24	0.21	0.15
100000009611	0.62	0.16	0.24	0.21	0.15
100000009612	0.57	0.17	0.24	0.21	0.15
3540663	-3.29	0.07	0.03	0.01	0.00
100000056382	-0.56	0.25	0.22	0.11	0.06
3526731	-0.10	0.23	0.25	0.15	0.09
100000009626	-0.46	0.25	0.23	0.11	0.07
100000009625	-0.10	0.23	0.25	0.15	0.09
100000009628	-0.11	0.23	0.25	0.15	0.09
100000009627	0.04	0.22	0.25	0.16	0.10
3526419	1.07	0.12	0.20	0.24	0.19
3526418	0.47	0.18	0.24	0.20	0.14
3537258	-0.81	0.25	0.20	0.09	0.05
3537247	-0.51	0.25	0.22	0.11	0.06
100000055189	-0.59	0.25	0.22	0.10	0.06
100000055190	0.32	0.19	0.25	0.19	0.12
100000055191	0.02	0.22	0.25	0.16	0.10
100000055187	0.71	0.15	0.23	0.22	0.16
100000055188	1.10	0.12	0.20	0.24	0.20
100000009634	1.37	0.10	0.18	0.25	0.22
100000009633	1.70	0.07	0.14	0.25	0.24
100000009632	1.02	0.13	0.21	0.24	0.19
100000009631	0.64	0.16	0.24	0.21	0.15
3467778	-1.06	0.24	0.18	0.07	0.04
100000055309	-0.80	0.25	0.20	0.09	0.05
100000055308	-0.85	0.25	0.20	0.08	0.05
100000009660	0.54	0.17	0.24	0.21	0.14

DRAFT

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
3559284	-0.86	0.25	0.20	0.08	0.05
100000089936	-0.45	0.24	0.23	0.11	0.07
100000009655	-0.35	0.24	0.24	0.12	0.07
3537453	0.90	0.14	0.22	0.23	0.18
3537800	-0.15	0.57	0.82	0.25	0.10
3539399	0.38	0.48	0.59	0.30	0.16
100000009658	0.10	0.45	0.52	0.31	0.18
3562388	0.15	0.43	0.49	0.31	0.20
100000055310	0.80	0.69	0.57	0.61	0.51
100000017416	-2.51	0.12	0.06	0.02	0.01
3540646	-1.54	0.21	0.13	0.05	0.02
3540649	-2.25	0.15	0.08	0.02	0.01
3540640	-1.52	0.22	0.13	0.05	0.03
100000055344	0.05	0.45	0.52	0.31	0.19
3540643	0.04	0.42	0.46	0.31	0.20
100000056531	-0.11	1.06	0.86	0.51	0.38
100000017420	0.50	0.95	0.99	0.61	0.38
3562626	-0.29	0.93	1.09	0.70	0.34

DRAFT

F.4: Grade 3
Total ELPA

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000061235	1.88	0.07	0.18	0.24	0.19
100000017442	-0.26	0.24	0.20	0.07	0.04
100000017437	0.73	0.17	0.25	0.15	0.09
100000017441	0.81	0.16	0.25	0.16	0.10
100000017440	0.49	0.19	0.25	0.13	0.08
100000056916	1.52	0.10	0.21	0.22	0.16
100000056917	1.25	0.12	0.23	0.20	0.14
3561098	1.57	0.09	0.21	0.23	0.17
3470725	1.89	0.07	0.18	0.24	0.20
100000009812	0.60	0.18	0.25	0.14	0.08
100000009817	1.73	0.08	0.19	0.24	0.18
100000055139	1.30	0.11	0.23	0.20	0.14
100000055140	1.37	0.11	0.22	0.21	0.15
100000055135	1.87	0.07	0.18	0.24	0.19
100000055136	2.16	0.06	0.15	0.25	0.22
100000018786	0.88	0.15	0.25	0.16	0.10
100000018785	0.69	0.17	0.25	0.15	0.09
100000018784	2.54	0.04	0.12	0.24	0.24
100000018783	2.35	0.05	0.14	0.25	0.23
3533550	0.25	0.21	0.24	0.11	0.06
3534017	0.80	0.16	0.25	0.16	0.10
100000056927	1.96	0.07	0.17	0.25	0.20
100000056926	2.05	0.06	0.16	0.25	0.21
100000009853	2.41	0.04	0.13	0.25	0.24
100000009854	1.63	0.09	0.20	0.23	0.17
100000009855	1.59	0.09	0.21	0.23	0.17
100000009858	1.74	0.08	0.19	0.24	0.18
100000055204	2.73	0.03	0.10	0.23	0.25
100000055200	1.28	0.11	0.23	0.20	0.14
100000055203	1.80	0.08	0.19	0.24	0.19
100000055199	2.40	0.05	0.13	0.25	0.24
100000055202	3.13	0.02	0.07	0.20	0.25
3537588	1.63	0.09	0.20	0.23	0.17
3526459	2.82	0.03	0.10	0.23	0.25
3526458	2.18	0.06	0.15	0.25	0.22
3526461	2.25	0.05	0.15	0.25	0.23
100000085124	2.73	0.03	0.10	0.23	0.25
100000085122	2.53	0.04	0.12	0.24	0.24
100000085123	2.46	0.04	0.13	0.25	0.24
3529874	3.13	0.02	0.07	0.20	0.25
3529880	2.74	0.03	0.10	0.23	0.25
3526783	0.80	0.16	0.25	0.16	0.10
3537520	1.49	0.10	0.22	0.22	0.16
100000057816	1.87	0.07	0.18	0.24	0.19

DRAFT

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000057815	2.58	0.04	0.12	0.24	0.24
3526794	0.33	0.20	0.24	0.11	0.07
3526795	0.42	0.19	0.24	0.12	0.07
100000057810	1.21	0.12	0.24	0.20	0.13
3526811	0.82	0.16	0.25	0.16	0.10
100000009889	2.73	0.28	0.28	0.25	0.18
100000056930	1.54	0.37	0.36	0.22	0.14
3563477	1.97	0.56	0.51	0.51	0.36
100000017435	2.25	0.61	0.61	0.47	0.34
3540626	-1.00	0.24	0.13	0.04	0.02
3540561	-1.00	0.24	0.13	0.04	0.02
3540625	-0.42	0.25	0.18	0.06	0.03
3563708	-0.07	0.23	0.22	0.08	0.05
3540565	1.51	0.45	0.44	0.18	0.10
100000056534	1.34	0.47	0.45	0.17	0.09
100000012506	0.66	1.12	0.68	0.34	0.22
100000017448	1.04	0.98	0.77	0.36	0.22
100000017446	0.91	1.10	1.09	0.25	0.11

DRAFT

F.5: Grade 4

Total ELPA

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000061235	1.88	0.08	0.20	0.23	0.18
100000017442	-0.26	0.25	0.18	0.06	0.03
100000017437	0.73	0.18	0.25	0.13	0.08
100000017441	0.81	0.17	0.25	0.14	0.09
100000017440	0.49	0.20	0.24	0.11	0.07
100000056916	1.52	0.11	0.23	0.21	0.14
100000056917	1.25	0.13	0.24	0.18	0.12
3561098	1.57	0.10	0.22	0.21	0.15
3470725	1.89	0.08	0.20	0.23	0.18
100000009812	0.60	0.19	0.24	0.12	0.07
100000009817	1.73	0.09	0.21	0.22	0.16
100000055139	1.30	0.13	0.24	0.19	0.12
100000055140	1.37	0.12	0.24	0.19	0.13
100000055135	1.87	0.08	0.20	0.23	0.18
100000055136	2.16	0.06	0.17	0.25	0.20
100000018786	0.88	0.16	0.25	0.15	0.09
100000018785	0.69	0.18	0.25	0.13	0.08
100000018784	2.54	0.05	0.14	0.25	0.23
100000018783	2.35	0.05	0.15	0.25	0.22
3533550	0.25	0.22	0.22	0.09	0.05
3534017	0.80	0.17	0.25	0.14	0.08
100000056927	1.96	0.07	0.19	0.24	0.19
100000056926	2.05	0.07	0.18	0.24	0.19
100000009853	2.41	0.05	0.15	0.25	0.22
100000009854	1.63	0.10	0.22	0.21	0.15
100000009855	1.59	0.10	0.22	0.21	0.15
100000009858	1.74	0.09	0.21	0.22	0.17
100000055204	2.73	0.04	0.12	0.24	0.24
100000055200	1.28	0.13	0.24	0.18	0.12
100000055203	1.80	0.09	0.21	0.23	0.17
100000055199	2.40	0.05	0.15	0.25	0.22
100000055202	3.13	0.03	0.09	0.22	0.25
3537588	1.63	0.10	0.22	0.21	0.15
3526459	2.82	0.03	0.11	0.24	0.25
3526458	2.18	0.06	0.17	0.25	0.21
3526461	2.25	0.06	0.16	0.25	0.21
100000085124	2.73	0.04	0.12	0.24	0.24
100000085122	2.53	0.05	0.14	0.25	0.23
100000085123	2.46	0.05	0.14	0.25	0.23
3529874	3.13	0.03	0.09	0.22	0.25
3529880	2.74	0.04	0.12	0.24	0.24
3526783	0.80	0.17	0.25	0.14	0.08
3537520	1.49	0.11	0.23	0.20	0.14
100000057816	1.87	0.08	0.20	0.23	0.18

DRAFT

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000057815	2.58	0.04	0.13	0.25	0.24
3526794	0.33	0.21	0.23	0.10	0.06
3526795	0.42	0.21	0.24	0.10	0.06
100000057810	1.21	0.13	0.24	0.18	0.12
3526811	0.82	0.17	0.25	0.14	0.09
100000009889	2.73	0.27	0.28	0.23	0.17
100000056930	1.54	0.37	0.35	0.19	0.12
3563477	1.97	0.53	0.53	0.47	0.32
100000017435	2.25	0.61	0.60	0.43	0.31
3540626	-1.00	0.23	0.11	0.03	0.02
3540561	-1.00	0.23	0.11	0.03	0.02
3540625	-0.42	0.25	0.17	0.05	0.03
3563708	-0.07	0.24	0.20	0.07	0.04
3540565	1.51	0.47	0.41	0.15	0.09
100000056534	1.34	0.49	0.42	0.14	0.08
100000012506	0.66	1.10	0.63	0.31	0.19
100000017448	1.04	0.98	0.72	0.32	0.19
100000017446	0.91	1.16	0.98	0.19	0.09

DRAFT

F.6: Grade 5

Total ELPA

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000061235	1.88	0.10	0.22	0.22	0.15
100000017442	-0.26	0.25	0.16	0.05	0.03
100000017437	0.73	0.20	0.24	0.11	0.06
100000017441	0.81	0.19	0.24	0.12	0.07
100000017440	0.49	0.22	0.23	0.09	0.05
100000056916	1.52	0.13	0.24	0.18	0.12
100000056917	1.25	0.15	0.25	0.16	0.10
3561098	1.57	0.12	0.24	0.19	0.12
3470725	1.89	0.10	0.22	0.22	0.15
100000009812	0.60	0.21	0.23	0.10	0.06
100000009817	1.73	0.11	0.23	0.20	0.14
100000055139	1.30	0.15	0.25	0.16	0.10
100000055140	1.37	0.14	0.25	0.17	0.11
100000055135	1.87	0.10	0.22	0.22	0.15
100000055136	2.16	0.08	0.19	0.24	0.18
100000018786	0.88	0.19	0.25	0.12	0.07
100000018785	0.69	0.21	0.24	0.11	0.06
100000018784	2.54	0.06	0.16	0.25	0.21
100000018783	2.35	0.07	0.18	0.24	0.20
3533550	0.25	0.24	0.21	0.08	0.04
3534017	0.80	0.20	0.24	0.12	0.07
100000056927	1.96	0.09	0.21	0.22	0.16
100000056926	2.05	0.09	0.20	0.23	0.17
100000009853	2.41	0.06	0.17	0.25	0.20
100000009854	1.63	0.12	0.24	0.19	0.13
100000009855	1.59	0.12	0.24	0.19	0.13
100000009858	1.74	0.11	0.23	0.20	0.14
100000055204	2.73	0.05	0.14	0.25	0.23
100000055200	1.28	0.15	0.25	0.16	0.10
100000055203	1.80	0.10	0.22	0.21	0.14
100000055199	2.40	0.06	0.17	0.25	0.20
100000055202	3.13	0.03	0.10	0.24	0.25
3537588	1.63	0.12	0.24	0.19	0.13
3526459	2.82	0.04	0.13	0.25	0.23
3526458	2.18	0.08	0.19	0.24	0.18
3526461	2.25	0.07	0.19	0.24	0.19
100000085124	2.73	0.05	0.14	0.25	0.23
100000085122	2.53	0.06	0.16	0.25	0.21
100000085123	2.46	0.06	0.17	0.25	0.21
3529874	3.13	0.03	0.10	0.24	0.25
3529880	2.74	0.05	0.14	0.25	0.23
3526783	0.80	0.20	0.24	0.12	0.07
3537520	1.49	0.13	0.24	0.18	0.12
100000057816	1.87	0.10	0.22	0.22	0.15

DRAFT

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000057815	2.58	0.05	0.15	0.25	0.22
3526794	0.33	0.23	0.21	0.08	0.04
3526795	0.42	0.23	0.22	0.09	0.05
100000057810	1.21	0.16	0.25	0.15	0.09
3526811	0.82	0.19	0.24	0.12	0.07
100000009889	2.73	0.27	0.29	0.21	0.14
100000056930	1.54	0.38	0.34	0.17	0.10
3563477	1.97	0.50	0.55	0.42	0.26
100000017435	2.25	0.62	0.59	0.39	0.26
3540626	-1.00	0.21	0.09	0.02	0.01
3540561	-1.00	0.21	0.09	0.02	0.01
3540625	-0.42	0.25	0.14	0.04	0.02
3563708	-0.07	0.25	0.18	0.06	0.03
3540565	1.51	0.49	0.37	0.13	0.07
100000056534	1.34	0.51	0.37	0.12	0.06
100000012506	0.66	1.02	0.57	0.26	0.16
100000017448	1.04	0.97	0.66	0.27	0.15
100000017446	0.91	1.24	0.83	0.15	0.06

DRAFT

F.7: Grade 6

Total ELPA

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000056491	0.77	0.22	0.22	0.11	0.06
100000017454	0.96	0.21	0.23	0.13	0.07
100000017455	1.36	0.17	0.25	0.16	0.10
100000017456	2.15	0.10	0.23	0.23	0.17
100000017457	1.20	0.19	0.24	0.15	0.09
100000056493	0.91	0.21	0.23	0.12	0.07
3526745	1.13	0.19	0.24	0.14	0.08
100000008945	1.23	0.19	0.24	0.15	0.09
100000056492	2.06	0.11	0.23	0.23	0.16
100000008946	2.10	0.11	0.23	0.23	0.17
100000056490	2.26	0.09	0.22	0.24	0.18
100000008952	1.18	0.19	0.24	0.15	0.09
100000008949	2.22	0.10	0.22	0.24	0.18
100000008953	2.25	0.09	0.22	0.24	0.18
100000008950	2.27	0.09	0.22	0.24	0.18
100000008954	1.67	0.14	0.25	0.19	0.13
100000055211	1.50	0.16	0.25	0.18	0.11
100000055212	3.28	0.04	0.13	0.23	0.25
100000055214	1.60	0.15	0.25	0.19	0.12
100000055215	4.22	0.02	0.06	0.15	0.21
100000055216	2.16	0.10	0.23	0.23	0.17
3535829	1.12	0.20	0.24	0.14	0.08
100000056501	1.44	0.17	0.25	0.17	0.11
100000056503	1.36	0.17	0.25	0.16	0.10
3535677	1.59	0.15	0.25	0.19	0.12
100000056500	2.66	0.07	0.19	0.25	0.22
100000055152	1.71	0.14	0.25	0.20	0.13
100000055154	1.50	0.16	0.25	0.18	0.11
100000055155	2.33	0.09	0.21	0.24	0.19
100000055156	2.83	0.06	0.17	0.25	0.23
100000008958	2.21	0.10	0.22	0.24	0.18
100000008959	1.22	0.19	0.24	0.15	0.09
100000008960	1.08	0.20	0.24	0.14	0.08
100000008961	2.43	0.08	0.21	0.25	0.20
100000008962	2.21	0.10	0.22	0.24	0.18
100000008963	1.91	0.12	0.24	0.21	0.15
3558745	1.91	0.12	0.24	0.21	0.15
3558748	2.23	0.10	0.22	0.24	0.18
3558753	2.31	0.09	0.22	0.24	0.19
3558750	2.35	0.09	0.21	0.24	0.19
3558739	1.73	0.14	0.25	0.20	0.13
3558738	2.89	0.06	0.16	0.25	0.23
3526775	0.62	0.23	0.20	0.10	0.05
3471438	1.17	0.19	0.24	0.15	0.09

DRAFT

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000056517	1.38	0.17	0.25	0.17	0.10
3558613	1.54	0.16	0.25	0.18	0.11
3537533	2.65	0.07	0.19	0.25	0.22
100000056522	0.34	0.25	0.18	0.08	0.04
3526803	0.54	0.24	0.20	0.09	0.05
100000056520	1.02	0.20	0.23	0.13	0.08
3558575	1.31	0.18	0.25	0.16	0.10
3558574	1.54	0.16	0.25	0.18	0.11
3539646	1.31	0.68	0.25	0.07	0.03
3539593	1.59	0.67	0.25	0.07	0.03
3566110	2.25	0.50	0.39	0.35	0.30
100000055317	1.67	0.63	0.57	033	0.20
3566131	-0.11	0.25	0.14	0.05	0.03
3566132	0.02	0.25	0.15	0.06	0.03
3540821	0.70	0.23	0.21	0.10	0.06
3540819	-0.75	0.21	0.08	0.03	0.01
3540820	-0.78	0.21	0.08	0.03	0.01
100000056536	1.34	0.52	0.29	0.11	0.06
3538699	1.43	0.51	0.29	0.11	0.06
100000012509	0.91	0.92	0.53	0.25	0.14
100000017467	1.21	1.07	0.52	0.20	0.11
100000055348	0.93	1.52	0.47	0.10	0.04

DRAFT

F.8: Grade 7

Total ELPA

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000056491	0.77	0.23	0.19	0.09	0.05
100000017454	0.96	0.22	0.21	0.11	0.06
100000017455	1.36	0.18	0.24	0.14	0.09
100000017456	2.15	0.11	0.24	0.22	0.15
100000017457	1.20	0.20	0.23	0.13	0.07
100000056493	0.91	0.22	0.21	0.10	0.06
3526745	1.13	0.21	0.22	0.12	0.07
100000008945	1.23	0.20	0.23	0.13	0.08
100000056492	2.06	0.12	0.25	0.21	0.14
100000008946	2.10	0.12	0.24	0.21	0.15
100000056490	2.26	0.10	0.24	0.23	0.16
100000008952	1.18	0.20	0.23	0.13	0.07
100000008949	2.22	0.11	0.24	0.22	0.16
100000008953	2.25	0.10	0.24	0.22	0.16
100000008950	2.27	0.10	0.24	0.23	0.16
100000008954	1.67	0.15	0.25	0.17	0.11
100000055211	1.50	0.17	0.24	0.16	0.10
100000055212	3.28	0.04	0.15	0.24	0.24
100000055214	1.60	0.16	0.25	0.17	0.10
100000055215	4.22	0.02	0.08	0.17	0.23
100000055216	2.16	0.11	0.24	0.22	0.15
3535829	1.12	0.21	0.22	0.12	0.07
100000056501	1.44	0.18	0.24	0.15	0.09
100000056503	1.36	0.18	0.24	0.14	0.09
3535677	1.59	0.16	0.25	0.17	0.10
100000056500	2.66	0.07	0.21	0.25	0.20
100000055152	1.71	0.15	0.25	0.18	0.11
100000055154	1.50	0.17	0.24	0.16	0.10
100000055155	2.33	0.10	0.23	0.23	0.17
100000055156	2.83	0.06	0.19	0.25	0.21
100000008958	2.21	0.11	0.24	0.22	0.16
100000008959	1.22	0.20	0.23	0.13	0.08
100000008960	1.08	0.21	0.22	0.12	0.07
100000008961	2.43	0.09	0.23	0.24	0.18
100000008962	2.21	0.11	0.24	0.22	0.16
100000008963	1.91	0.13	0.25	0.20	0.13
3558745	1.91	0.13	0.25	0.20	0.13
3558748	2.23	0.10	0.24	0.22	0.16
3558753	2.31	0.10	0.24	0.23	0.17
3558750	2.35	0.10	0.23	0.23	0.17
3558739	1.73	0.15	0.25	0.18	0.11
3558738	2.89	0.06	0.19	0.25	0.22
3526775	0.62	0.24	0.18	0.08	0.05
3471438	1.17	0.20	0.23	0.13	0.07

DRAFT

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000056517	1.38	0.18	0.24	0.15	0.09
3558613	1.54	0.17	0.25	0.16	0.10
3537533	2.65	0.07	0.21	0.25	0.20
100000056522	0.34	0.25	0.15	0.07	0.03
3526803	0.54	0.24	0.17	0.08	0.04
100000056520	1.02	0.21	0.21	0.11	0.06
3558575	1.31	0.19	0.24	0.14	0.08
3558574	1.54	0.17	0.25	0.16	0.10
3539646	1.31	0.66	0.19	0.06	0.03
3539593	1.59	0.66	0.19	0.06	0.03
3566110	2.25	0.49	0.38	0.34	0.28
100000055317	1.67	0.62	0.53	0.29	0.17
3566131	-0.11	0.24	0.11	0.04	0.02
3566132	0.02	0.25	0.12	0.05	0.03
3540821	0.70	0.24	0.19	0.09	0.05
3540819	-0.75	0.20	0.07	0.02	0.01
3540820	-0.78	0.20	0.06	0.02	0.01
100000056536	1.34	0.52	0.23	0.09	0.05
3538699	1.43	0.50	0.24	0.10	0.05
100000012509	0.91	0.89	0.46	0.21	0.11
100000017467	1.21	1.03	0.43	0.17	0.09
100000055348	0.93	1.48	0.33	0.07	0.03

DRAFT

F.9: Grade 8

Total ELPA

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000056491	0.77	0.24	0.19	0.08	0.04
100000017454	0.96	0.24	0.20	0.09	0.05
100000017455	1.36	0.21	0.23	0.12	0.07
100000017456	2.15	0.13	0.25	0.20	0.13
100000017457	1.20	0.22	0.22	0.11	0.06
100000056493	0.91	0.24	0.20	0.09	0.05
3526745	1.13	0.22	0.22	0.10	0.06
100000008945	1.23	0.22	0.23	0.11	0.06
100000056492	2.06	0.14	0.25	0.19	0.12
100000008946	2.10	0.14	0.25	0.19	0.13
100000056490	2.26	0.12	0.24	0.21	0.14
100000008952	1.18	0.22	0.22	0.11	0.06
100000008949	2.22	0.13	0.24	0.20	0.14
100000008953	2.25	0.12	0.24	0.21	0.14
100000008950	2.27	0.12	0.24	0.21	0.14
100000008954	1.67	0.18	0.25	0.15	0.09
100000055211	1.50	0.19	0.24	0.14	0.08
100000055212	3.28	0.05	0.16	0.25	0.23
100000055214	1.60	0.18	0.25	0.15	0.09
100000055215	4.22	0.02	0.08	0.19	0.24
100000055216	2.16	0.13	0.25	0.20	0.13
3535829	1.12	0.22	0.22	0.10	0.06
100000056501	1.44	0.20	0.24	0.13	0.08
100000056503	1.36	0.21	0.23	0.12	0.07
3535677	1.59	0.18	0.24	0.14	0.09
100000056500	2.66	0.09	0.22	0.24	0.18
100000055152	1.71	0.17	0.25	0.16	0.09
100000055154	1.50	0.19	0.24	0.14	0.08
100000055155	2.33	0.12	0.24	0.21	0.15
100000055156	2.83	0.08	0.20	0.24	0.20
100000008958	2.21	0.13	0.24	0.20	0.14
100000008959	1.22	0.22	0.22	0.11	0.06
100000008960	1.08	0.23	0.21	0.10	0.06
100000008961	2.43	0.11	0.23	0.22	0.16
100000008962	2.21	0.13	0.24	0.20	0.14
100000008963	1.91	0.15	0.25	0.18	0.11
3558745	1.91	0.15	0.25	0.18	0.11
3558748	2.23	0.12	0.24	0.20	0.14
3558753	2.31	0.12	0.24	0.21	0.15
3558750	2.35	0.11	0.24	0.21	0.15
3558739	1.73	0.17	0.25	0.16	0.10
3558738	2.89	0.07	0.20	0.25	0.20
3526775	0.62	0.25	0.17	0.07	0.04
3471438	1.17	0.22	0.22	0.11	0.06

DRAFT

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000056517	1.38	0.20	0.24	0.13	0.07
3558613	1.54	0.19	0.24	0.14	0.08
3537533	2.65	0.09	0.22	0.24	0.18
100000056522	0.34	0.25	0.15	0.05	0.03
3526803	0.54	0.25	0.16	0.06	0.03
100000056520	1.02	0.23	0.21	0.10	0.05
3558575	1.31	0.21	0.23	0.12	0.07
3558574	1.54	0.19	0.24	0.14	0.08
3539646	1.31	0.61	0.18	0.04	0.02
3539593	1.59	0.61	0.18	0.05	0.02
3566110	2.25	0.41	0.19	0.07	0.04
100000055317	1.67	0.61	0.51	0.25	0.14
3566131	-0.11	0.23	0.11	0.04	0.02
3566132	0.02	0.24	0.12	0.04	0.02
3540821	0.70	0.25	0.18	0.07	0.04
3540819	-0.75	0.18	0.06	0.02	0.01
3540820	-0.78	0.17	0.06	0.02	0.01
100000056536	1.34	0.49	0.22	0.07	0.04
3538699	1.43	0.48	0.23	0.08	0.04
100000012509	0.91	0.82	0.44	0.17	0.09
100000017467	1.21	0.94	0.41	0.14	0.07
100000055348	0.93	1.35	0.30	0.06	0.02

DRAFT

F.10: Grade 9
Total ELPA

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000061241	0.73	0.24	0.15	0.06	0.03
100000017475	3.24	0.10	0.19	0.25	0.20
100000017476	1.56	0.24	0.22	0.11	0.06
100000055163	1.01	0.25	0.18	0.07	0.04
100000055164	2.84	0.13	0.23	0.23	0.17
3566062	0.74	0.24	0.15	0.06	0.03
100000056495	1.53	0.24	0.22	0.11	0.06
3566089	2.40	0.17	0.25	0.19	0.12
3526756	1.69	0.23	0.23	0.13	0.07
100000009039	3.10	0.11	0.21	0.24	0.19
100000056480	3.50	0.08	0.17	0.25	0.22
100000056469	1.78	0.22	0.24	0.13	0.08
100000056470	1.56	0.24	0.22	0.11	0.06
100000056471	1.74	0.23	0.24	0.13	0.07
100000056933	2.11	0.20	0.25	0.17	0.10
100000056472	2.29	0.18	0.25	0.18	0.11
3526516	2.80	0.13	0.23	0.23	0.16
3536855	2.64	0.15	0.24	0.21	0.15
3536878	3.94	0.05	0.13	0.24	0.25
3536738	2.69	0.14	0.24	0.22	0.15
3558509	1.35	0.25	0.21	0.10	0.05
100000056505	1.97	0.21	0.25	0.15	0.09
100000056506	2.85	0.13	0.23	0.23	0.17
100000009007	1.46	0.24	0.22	0.11	0.06
100000009009	1.32	0.25	0.21	0.10	0.05
100000009008	1.53	0.24	0.22	0.11	0.06
100000009006	1.20	0.25	0.19	0.09	0.05
100000009010	1.60	0.24	0.23	0.12	0.07
100000009005	3.17	0.10	0.20	0.25	0.20
100000057368	2.69	0.14	0.24	0.22	0.15
100000057370	2.87	0.13	0.23	0.23	0.17
100000057369	2.90	0.12	0.22	0.23	0.17
100000057373	3.66	0.07	0.15	0.25	0.23
100000057371	3.03	0.11	0.21	0.24	0.18
100000057372	2.25	0.18	0.25	0.18	0.11
3562741	2.17	0.19	0.25	0.17	0.10
3562743	3.05	0.11	0.21	0.24	0.19
3562744	2.69	0.14	0.24	0.22	0.15
3562737	2.46	0.16	0.25	0.20	0.13
3562746	3.26	0.09	0.19	0.25	0.20
3562742	2.49	0.16	0.25	0.20	0.13
3526786	0.82	0.24	0.16	0.06	0.03
100000009014	2.42	0.17	0.25	0.19	0.13
100000056526	2.85	0.13	0.23	0.23	0.17

DRAFT

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
3470663	3.17	0.10	0.20	0.25	0.20
100000056527	3.02	0.11	0.21	0.24	0.18
3526809	0.86	0.25	0.16	0.07	0.03
3526810	1.37	0.25	0.21	0.10	0.05
100000056529	1.51	0.24	0.22	0.11	0.06
3558631	1.80	0.22	0.24	0.14	0.08
3558630	1.86	0.22	0.24	0.14	0.08
3537943	1.32	0.35	0.21	0.08	0.04
3558502	1.86	0.38	0.16	0.05	0.03
100000017489	2.40	0.43	0.37	0.30	0.21
100000055334	2.56	0.63	0.44	0.18	0.09
3566126	0.13	0.20	0.10	0.03	0.02
3541261	0.54	0.23	0.13	0.05	0.03
3541263	0.67	0.24	0.14	0.06	0.03
3541258	0.76	0.24	0.15	0.06	0.03
3541259	0.17	0.20	0.10	0.04	0.02
3562634	1.34	0.38	0.14	0.04	0.02
100000055355	0.88	0.38	0.15	0.04	0.02
100000012511	1.23	0.66	0.32	0.11	0.06
100000017492	1.67	0.71	0.28	0.08	0.04
100000055353	0.86	0.80	0.16	0.03	0.01

DRAFT

F.11: Grade 10
Total ELPA

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000061241	0.73	0.24	0.14	0.05	0.03
100000017475	3.24	0.10	0.20	0.24	0.19
100000017476	1.56	0.24	0.22	0.11	0.06
100000055163	1.01	0.25	0.17	0.07	0.03
100000055164	2.84	0.14	0.23	0.22	0.15
3566062	0.74	0.24	0.15	0.05	0.03
100000056495	1.53	0.24	0.22	0.10	0.06
3566089	2.40	0.18	0.25	0.18	0.11
3526756	1.69	0.24	0.23	0.12	0.06
100000009039	3.10	0.11	0.21	0.24	0.18
100000056480	3.50	0.08	0.18	0.25	0.21
100000056469	1.78	0.23	0.24	0.13	0.07
100000056470	1.56	0.24	0.22	0.11	0.06
100000056471	1.74	0.23	0.23	0.12	0.07
100000056933	2.11	0.21	0.25	0.16	0.09
100000056472	2.29	0.19	0.25	0.17	0.10
3526516	2.80	0.14	0.23	0.22	0.15
3536855	2.64	0.16	0.24	0.21	0.13
3536878	3.94	0.06	0.13	0.24	0.24
3536738	2.69	0.15	0.24	0.21	0.14
3558509	1.35	0.25	0.20	0.09	0.05
100000056505	1.97	0.22	0.24	0.14	0.08
100000056506	2.85	0.14	0.23	0.22	0.15
100000009007	1.46	0.25	0.21	0.10	0.05
100000009009	1.32	0.25	0.20	0.09	0.05
100000009008	1.53	0.24	0.22	0.10	0.06
100000009006	1.20	0.25	0.19	0.08	0.04
100000009010	1.60	0.24	0.22	0.11	0.06
100000009005	3.17	0.11	0.21	0.24	0.18
100000057368	2.69	0.15	0.24	0.21	0.14
100000057370	2.87	0.14	0.23	0.22	0.16
100000057369	2.90	0.13	0.23	0.23	0.16
100000057373	3.66	0.07	0.16	0.25	0.23
100000057371	3.03	0.12	0.22	0.23	0.17
100000057372	2.25	0.19	0.25	0.17	0.10
3562741	2.17	0.20	0.25	0.16	0.09
3562743	3.05	0.12	0.22	0.24	0.17
3562744	2.69	0.15	0.24	0.21	0.14
3562737	2.46	0.17	0.25	0.19	0.12
3562746	3.26	0.10	0.20	0.25	0.19
3562742	2.49	0.17	0.25	0.19	0.12
3526786	0.82	0.24	0.15	0.06	0.03
100000009014	2.42	0.18	0.25	0.19	0.11
100000056526	2.85	0.14	0.23	0.22	0.15

DRAFT

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
3470663	3.17	0.11	0.21	0.24	0.18
100000056527	3.02	0.12	0.22	0.23	0.17
3526809	0.86	0.24	0.16	0.06	0.03
3526810	1.37	0.25	0.20	0.09	0.05
100000056529	1.51	0.24	0.22	0.10	0.05
3558631	1.80	0.23	0.24	0.13	0.07
3558630	1.86	0.23	0.24	0.13	0.07
3537943	1.32	0.34	0.20	0.07	0.04
3558502	1.86	0.35	0.16	0.05	0.02
100000017489	2.40	0.42	0.37	0.29	0.20
100000055334	2.56	0.62	0.43	0.16	0.08
3566126	0.13	0.19	0.09	0.03	0.01
3541261	0.54	0.22	0.13	0.05	0.02
3541263	0.67	0.23	0.14	0.05	0.03
3541258	0.76	0.24	0.15	0.06	0.03
3541259	0.17	0.19	0.10	0.03	0.02
3562634	1.34	0.35	0.13	0.04	0.02
100000055355	0.88	0.35	0.14	0.04	0.02
100000012511	1.23	062	0.30	0.10	0.05
100000017492	1.67	0.66	0.26	0.08	0.04
100000055353	0.86	0.70	0.14	0.02	0.01

DRAFT

F.12: Grade 11
Total ELPA

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000061241	0.73	0.23	0.11	0.05	0.03
100000017475	3.24	0.11	0.23	0.24	0.19
100000017476	1.56	0.24	0.19	0.10	0.06
100000055163	1.01	0.25	0.14	0.06	0.03
100000055164	2.84	0.14	0.25	0.21	0.15
3566062	0.74	0.23	0.12	0.05	0.03
100000056495	1.53	0.25	0.19	0.09	0.06
3566089	2.40	0.18	0.25	0.17	0.11
3526756	1.69	0.24	0.20	0.11	0.06
100000009039	3.10	0.12	0.24	0.23	0.18
100000056480	3.50	0.09	0.21	0.25	0.21
100000056469	1.78	0.23	0.21	0.11	0.07
100000056470	1.56	0.24	0.19	0.10	0.06
100000056471	1.74	0.24	0.21	0.11	0.07
100000056933	2.11	0.21	0.24	0.14	0.09
100000056472	2.29	0.19	0.24	0.16	0.10
3526516	2.80	0.15	0.25	0.21	0.15
3536855	2.64	0.16	0.25	0.19	0.13
3536878	3.94	0.06	0.17	0.25	0.24
3536738	2.69	0.16	0.25	0.20	0.14
3558509	1.35	0.25	0.17	0.08	0.05
100000056505	1.97	0.22	0.23	0.13	0.08
100000056506	2.85	0.14	0.25	0.21	0.15
100000009007	1.46	0.25	0.18	0.09	0.05
100000009009	1.32	0.25	0.17	0.08	0.05
100000009008	1.53	0.25	0.19	0.09	0.06
100000009006	1.20	0.25	0.16	0.07	0.04
100000009010	1.60	0.24	0.20	0.10	0.06
100000009005	3.17	0.11	0.23	0.24	0.18
100000057368	2.69	0.16	0.25	0.20	0.14
100000057370	2.87	0.14	0.25	0.21	0.16
100000057369	2.90	0.14	0.24	0.22	0.16
100000057373	3.66	0.08	0.19	0.25	0.23
100000057371	3.03	0.13	0.24	0.23	0.17
100000057372	2.25	0.20	0.24	0.16	0.10
3562741	2.17	0.21	0.24	0.15	0.09
3562743	3.05	0.12	0.24	0.23	0.17
3562744	2.69	0.16	0.25	0.20	0.14
3562737	2.46	0.18	0.25	0.18	0.12
3562746	3.26	0.11	0.23	0.24	0.19
3562742	2.49	0.18	0.25	0.18	0.12
3526786	0.82	0.24	0.12	0.05	0.03
100000009014	2.42	0.18	0.25	0.17	0.11
100000056526	2.85	0.14	0.25	0.21	0.15

DRAFT

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
3470663	3.17	0.11	0.23	0.24	0.18
100000056527	3.02	0.13	0.24	0.23	0.17
3526809	0.86	0.24	0.13	0.05	0.03
3526810	1.37	0.25	0.17	0.08	0.05
100000056529	1.51	0.25	0.19	0.09	0.05
3558631	1.80	0.23	0.21	0.12	0.07
3558630	1.86	0.23	0.22	0.12	0.07
3537943	1.32	0.33	0.16	0.07	0.04
3558502	1.86	0.34	0.12	0.04	0.02
100000017489	2.40	0.42	0.36	0.27	0.20
100000055334	2.56	0.62	0.34	0.15	0.08
3566126	0.13	0.18	0.07	0.03	0.01
3541261	0.54	0.22	0.10	0.04	0.02
3541263	0.67	0.23	0.11	0.05	0.03
3541258	0.76	0.23	0.12	0.05	0.03
3541259	0.17	0.19	0.07	0.03	0.02
3562634	1.34	0.34	0.09	0.03	0.02
100000055355	0.88	0.34	0.10	0.04	0.02
100000012511	1.23	0.61	0.23	0.09	0.05
100000017492	1.67	0.64	0.19	0.07	0.04
100000055353	0.86	0.66	0.09	0.02	0.01

DRAFT

F.13: Grade 12
Total ELPA

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
100000061241	0.73	0.23	0.10	0.04	0.02
100000017475	3.24	0.12	0.24	0.22	0.16
100000017476	1.56	0.25	0.17	0.08	0.04
100000055163	1.01	0.24	0.12	0.05	0.03
100000055164	2.84	0.15	0.25	0.19	0.13
3566062	0.74	0.23	0.10	0.04	0.02
100000056495	1.53	0.25	0.17	0.08	0.04
3566089	2.40	0.19	0.24	0.15	0.09
3526756	1.69	0.24	0.18	0.09	0.05
100000009039	3.10	0.13	0.25	0.21	0.15
100000056480	3.50	0.10	0.22	0.24	0.19
100000056469	1.78	0.24	0.19	0.10	0.05
100000056470	1.56	0.25	0.17	0.08	0.04
100000056471	1.74	0.24	0.19	0.09	0.05
100000056933	2.11	0.22	0.22	0.12	0.07
100000056472	2.29	0.20	0.23	0.14	0.08
3526516	2.80	0.16	0.25	0.19	0.12
3536855	2.64	0.17	0.25	0.17	0.11
3536878	3.94	0.07	0.19	0.25	0.23
3536738	2.69	0.17	0.25	0.18	0.11
3558509	1.35	0.25	0.15	0.07	0.04
100000056505	1.97	0.23	0.21	0.11	0.06
100000056506	2.85	0.15	0.25	0.19	0.13
100000009007	1.46	0.25	0.16	0.07	0.04
100000009009	1.32	0.25	0.15	0.07	0.04
100000009008	1.53	0.25	0.17	0.08	0.04
100000009006	1.20	0.25	0.14	0.06	0.03
100000009010	1.60	0.25	0.18	0.08	0.05
100000009005	3.17	0.12	0.24	0.22	0.16
100000057368	2.69	0.17	0.25	0.18	0.11
100000057370	2.87	0.15	0.25	0.19	0.13
100000057369	2.90	0.15	0.25	0.20	0.13
100000057373	3.66	0.08	0.21	0.25	0.20
100000057371	3.03	0.14	0.25	0.21	0.14
100000057372	2.25	0.21	0.23	0.14	0.08
3562741	2.17	0.21	0.22	0.13	0.07
3562743	3.05	0.13	0.25	0.21	0.15
3562744	2.69	0.17	0.25	0.18	0.11
3562737	2.46	0.19	0.24	0.15	0.09
3562746	3.26	0.11	0.24	0.23	0.17
3562742	2.49	0.19	0.24	0.16	0.10
3526786	0.82	0.23	0.10	0.04	0.02
100000009014	2.42	0.19	0.24	0.15	0.09
100000056526	2.85	0.15	0.25	0.19	0.13

DRAFT

Item ID	Rasch Difficulty	Item Information at Low Intermediate Cut	Item Information at High Intermediate Cut	Item Information at Proficient Cut	Item Information at Advanced Proficient Cut
3470663	3.17	0.12	0.24	0.22	0.16
100000056527	3.02	0.14	0.25	0.21	0.14
3526809	0.86	0.23	0.11	0.04	0.02
3526810	1.37	0.25	0.15	0.07	0.04
100000056529	1.51	0.25	0.17	0.08	0.04
3558631	1.80	0.24	0.19	0.10	0.05
3558630	1.86	0.24	0.20	0.10	0.06
3537943	1.32	0.32	0.14	0.05	0.03
3558502	1.86	0.32	0.09	0.03	0.02
100000017489	2.40	0.41	0.34	0.25	0.16
100000055334	2.56	0.61	0.29	0.12	0.06
3566126	0.13	0.17	0.06	0.02	0.01
3541261	0.54	0.21	0.08	0.03	0.02
3541263	0.67	0.22	0.09	0.04	0.02
3541258	0.76	0.23	0.10	0.04	0.02
3541259	0.17	0.18	0.06	0.02	0.01
3562634	1.34	0.31	0.08	0.03	0.01
100000055355	0.88	0.31	0.08	0.03	0.01
100000012511	1.23	0.57	0.19	0.07	0.04
100000017492	1.67	0.59	0.16	0.05	0.03
100000055353	0.86	0.56	0.06	0.01	0.01

DRAFT

APPENDIX G: CLASSIFICATION DISTRIBUTION ACROSS PERFORMANCE LEVELS FOR THE ELPA, MEAP, AND MME

G.1: Grade 3- Percentage of ELLs Scoring at Each MEAP ELA and Math Performance Level by Each ELPA Performance Level

ELA			ELPA Performance Levels				
			1 Advanced Proficient (N=491)	2 Proficient (N=1260)	3 High Intermediate (N=3191)	4 Low Intermediate (N=277)	5 Basic (N=6)
MEAP Performance Levels	1 Exceeded Standards	%	28.90	5.20	0.50	0.00	0.00
	2 Met Standards	%	69.70	88.90	58.00	15.50	0.00
	3 Basic	%	1.40	5.90	40.20	71.50	83.30
	4 Apprentice	%	0.00	0.10	1.30	13.00	16.70

Math			ELPA Performance Levels				
			1 Advanced Proficient (N=491)	2 Proficient (N=1260)	3 High Intermediate (N=3191)	4 Low Intermediate (N=277)	5 Basic (N=6)
MEAP Performance Levels	1 Exceeded Standards	%	85.50	62.00	24.30	4.70	0.00
	2 Met Standards	%	13.80	36.40	57.90	44.00	16.70
	3 Basic	%	0.60	1.60	17.60	50.50	83.30
	4 Apprentice	%	0.00	0.00	0.20	0.70	0.00

DRAFT

G.2: Grade 4- Percentage of ELLs Scoring at Each MEAP ELA and Math Performance Level by Each ELPA Performance Level

ELA			ELPA Performance Levels				
			1 Advanced Proficient (N=812)	2 Proficient (N=1274)	3 High Intermediate (N=2220)	4 Low Intermediate (N=188)	5 Basic (N=4)
MEAP Performance Levels	1 Exceeded Standards	%	5.00	0.50	0.20	0.00	0.00
	2 Met Standards	%	89.00	68.10	28.30	3.70	0.00
	3 Basic	%	5.70	30.90	65.80	76.10	75.00
	4 Apprentice	%	0.20	0.50	5.70	20.20	25.00

Math			ELPA Performance Levels				
			1 Advanced Proficient (N=812)	2 Proficient (N=1274)	3 High Intermediate (N=2220)	4 Low Intermediate (N=188)	5 Basic (N=4)
MEAP Performance Levels	1 Exceeded Standards	%	58.90	27.40	9.10	6.40	0.00
	2 Met Standards	%	38.90	64.20	56.40	23.90	25.00
	3 Basic	%	2.20	8.20	31.50	57.40	25.00
	4 Apprentice	%	0.00	0.20	3.00	12.20	50.00

DRAFT

G.3: Grade 5- Percentage of ELLs Scoring at Each MEAP ELA, Math, and Science Performance Level by Each ELPA Performance Level

ELA			ELPA Performance Levels				
			1 Advanced Proficient (N=778)	2 Proficient (N=1176)	3 High Intermediate (N=1670)	4 Low Intermediate (N=141)	5 Basic (N=6)
MEAP Performance Levels	1 Exceeded Standards	%	11.30	0.90	0.20	0.00	0.00
	2 Met Standards	%	80.50	63.20	22.80	2.80	0.00
	3 Basic	%	7.10	34.40	65.70	50.40	33.30
	4 Apprentice	%	1.20	1.50	11.20	46.80	66.70

Math			ELPA Performance Levels				
			1 Advanced Proficient (N=778)	2 Proficient (N=1176)	3 High Intermediate (N=1670)	4 Low Intermediate (N=141)	5 Basic (N=6)
MEAP Performance Levels	1 Exceeded Standards	%	58.00	25.50	9.70	2.80	0.00
	2 Met Standards	%	33.70	46.60	32.90	8.50	16.70
	3 Basic	%	7.80	25.20	45.60	52.50	66.70
	4 Apprentice	%	0.50	2.70	11.90	36.20	16.70

DRAFT

(G.3 continued)

Science			ELPA Performance Levels				
			1 Advanced Proficient (N=778)	2 Proficient (N=1176)	3 High Intermediate (N=1670)	4 Low Intermediate (N=141)	5 Basic (N=6)
MEAP Performance Levels	1 Exceeded Standards	%	44.00	10.60	3.00	1.40	0.00
	2 Met Standards	%	51.50	64.60	35.20	5.70	16.70
	3 Basic	%	4.20	22.00	44.10	39.00	16.70
	4 Apprentice	%	0.30	2.70	17.70	53.90	66.70

DRAFT

G.4: Grade 6- Percentage of ELLs Scoring at Each MEAP ELA, Math, and Social Studies Performance Level by Each ELPA Performance Level

ELA			ELPA Performance Levels				
			1 Advanced Proficient (N=138)	2 Proficient (N=562)	3 High Intermediate (N=2010)	4 Low Intermediate (N=561)	5 Basic (N=20)
MEAP Performance Levels	1 Exceeded Standards	%	24.60	5.30	0.70	0.00	0.00
	2 Met Standards	%	71.70	84.30	47.10	9.30	0.00
	3 Basic	%	3.60	10.30	49.50	76.30	65.00
	4 Apprentice	%	0.00	0.00	2.70	14.40	35.00

Math			ELPA Performance Levels				
			1 Advanced Proficient (N=138)	2 Proficient (N=562)	3 High Intermediate (N=2010)	4 Low Intermediate (N=561)	5 Basic (N=20)
MEAP Performance Levels	1 Exceeded Standards	%	82.60	52.70	22.40	6.20	0.00
	2 Met Standards	%	16.70	35.40	41.30	19.60	5.00
	3 Basic	%	0.70	10.70	32.20	54.00	65.00
	4 Apprentice	%	0.00	1.20	4.10	20.10	30.00

DRAFT

(G.4 continued)

Social Studies			ELPA Performance Levels				
			1 Advanced Proficient (N=138)	2 Proficient (N=562)	3 High Intermediate (N=2010)	4 Low Intermediate (N=561)	5 Basic (N=20)
MEAP Performance Levels	1 Exceeded Standards	%	70.30	25.40	6.60	0.70	0.00
	2 Met Standards	%	23.20	54.80	32.80	7.50	5.00
	3 Basic	%	6.50	14.90	30.90	18.90	10.00
	4 Apprentice	%	0.00	4.80	29.60	72.90	85.00

DRAFT

G.5: Grade 7- Percentage of ELLs Scoring at Each MEAP ELA and Math Performance Level by Each ELPA Performance Level

ELA			ELPA Performance Levels				
			1 Advanced Proficient (N=238)	2 Proficient (N=618)	3 High Intermediate (N=1641)	4 Low Intermediate (N=595)	5 Basic (N=21)
MEAP Performance Levels	1 Exceeded Standards	%	19.30	1.80	0.50	0.00	0.00
	2 Met Standards	%	79.80	85.30	51.20	10.40	4.80
	3 Basic	%	0.80	12.10	40.50	52.40	33.30
	4 Apprentice	%	0.00	0.80	7.70	37.10	61.90

Math			ELPA Performance Levels				
			1 Advanced Proficient (N=238)	2 Proficient (N=618)	3 High Intermediate (N=1641)	4 Low Intermediate (N=595)	5 Basic (N=21)
MEAP Performance Levels	1 Exceeded Standards	%	79.80	47.60	19.10	6.20	4.80
	2 Met Standards	%	19.30	42.90	47.80	26.20	14.30
	3 Basic	%	0.80	9.10	30.20	52.90	71.40
	4 Apprentice	%	0.00	0.50	2.90	14.60	9.50

DRAFT

G.6: Grade 8- Percentage of ELLs Scoring at Each MEAP ELA, Math, and Science Performance Level by Each ELPA Performance Level

ELA			ELPA Performance Levels				
			1 Advanced Proficient (N=279)	2 Proficient (N=612)	3 High Intermediate (N=1771)	4 Low Intermediate (N=522)	5 Basic (N=30)
MEAP Performance Levels	1 Exceeded Standards	%	23.30	4.20	0.90	0.00	0.00
	2 Met Standards	%	73.10	77.50	42.30	10.50	3.30
	3 Basic	%	3.60	17.30	47.10	42.30	20.00
	4 Apprentice	%	0.00	1.00	9.70	47.10	76.70

Math			ELPA Performance Levels				
			1 Advanced Proficient (N=279)	2 Proficient (N=612)	3 High Intermediate (N=1771)	4 Low Intermediate (N=522)	5 Basic (N=30)
MEAP Performance Levels	1 Exceeded Standards	%	59.50	30.90	14.90	5.20	0.00
	2 Met Standards	%	31.90	47.90	38.70	24.10	16.70
	3 Basic	%	6.80	18.00	33.90	42.90	53.30
	4 Apprentice	%	1.80	3.30	12.50	27.80	30.00

DRAFT

(G.6 continued)

Science			ELPA Performance Levels				
			1 Advanced Proficient (N=279)	2 Proficient (N=612)	3 High Intermediate (N=1771)	4 Low Intermediate (N=522)	5 Basic (N=30)
MEAP Performance Level	1 Exceeded Standards	%	41.90	12.70	2.60	0.60	0.00
	2 Met Standards	%	53.40	71.70	39.70	13.40	0.00
	3 Basic	%	4.30	13.60	44.60	46.00	30.00
	4 Apprentice	%	0.40	2.00	13.10	40.00	70.00

DRAFT

G.7: Grade 11- Percentage of ELLs Scoring at Each MME ELA, Math, Science, and Social Studies Performance Level by Each ELPA Performance Level

ELA			ELPA Performance Levels				
			1 Advanced Proficient (N=159)	2 Proficient (N=408)	3 High Intermediate (N=824)	4 Low Intermediate (N=493)	5 Basic (N=60)
MME Performance Levels	1 Advanced	%	0.00	0.00	0.00	0.00	0.00
	2 Proficient	%	59.10	25.70	4.00	2.40	0.00
	3 Partially Proficient	%	39.00	68.60	60.70	19.50	3.30
	4 Not Proficient	%	1.90	5.60	35.30	78.10	96.70

Math			ELPA Performance Levels				
			1 Advanced Proficient (N=159)	2 Proficient (N=408)	3 High Intermediate (N=824)	4 Low Intermediate (N=493)	5 Basic (N=60)
MME Performance Levels	1 Advanced	%	10.10	5.10	1.30	0.80	0.00
	2 Proficient	%	49.70	30.40	13.80	5.90	1.70
	3 Partially Proficient	%	20.80	21.10	15.20	9.70	3.30
	4 Not Proficient	%	19.50	43.40	69.70	83.60	95.00

DRAFT

(G.7 continued)

Science			ELPA Performance Levels				
			1 Advanced Proficient (N=159)	2 Proficient (N=408)	3 High Intermediate (N=824)	4 Low Intermediate (N=493)	5 Basic (N=60)
MME Performance Levels	1 Advanced	%	3.80	1.00	0.00	0.00	0.00
	2 Proficient	%	61.00	38.00	11.40	3.00	0.00
	3 Partially Proficient	%	23.30	23.80	14.90	5.90	1.70
	4 Not Proficient	%	11.90	37.30	73.70	91.10	98.30

Social Studies			ELPA Performance Levels				
			1 Advanced Proficient (N=159)	2 Proficient (N=408)	3 High Intermediate (N=824)	4 Low Intermediate (N=493)	5 Basic (N=60)
MME Performance Levels	1 Advanced	%	45.30	17.40	4.60	2.60	0.00
	2 Proficient	%	45.30	60.50	45.90	23.30	26.70
	3 Partially Proficient	%	8.20	15.40	27.90	31.80	20.00
	4 Not Proficient	%	1.30	6.60	21.60	42.20	53.30

DRAFT