

# **Comparability of Student Scores Obtained from Paper and Computer Administrations**

**Office of Assessment and Information Services  
Oregon Department of Education**

**April 2007**

## **Executive Summary**

This study explored the comparability in scores across two administration modes of the state Math, Reading and Science tests; 1) the computer-based progressive test administration and 2) the paper and pencil administration. The standard administration mode is the TESA, the computer-based progressive test, with over 90 percent of all students accessing the state tests through the system. However, an increasingly small number of schools request waivers due to insufficient technology for testing and administer the state tests via paper and pencil. While the tests are built to the same grade-level test specifications, monitoring for administration mode effects ensures unanticipated effects of test mode can be identified and mitigated to have minimal impact on test scores.

Results suggest that average scores and standard errors are quite similar across TESA and paper tests. Although the differences were still quite small (less than half a scale score point), third graders tended to show slightly larger differences. None of the differences were statistically significant, and all TESA and paper scores correlated highly with each other.

This study provides evidence that scores are comparable across TESA and paper delivery modes.

# Comparability of Student Scores Obtained from Paper and Computer Administrations

## Introduction

Since the advent of Oregon's online testing system in 2001, the Oregon Department of Education (ODE) has offered districts a choice of using the Technology-Enhanced Student Assessment (TESA) or conventional paper and pencil assessments. At the outset, schools had varying capabilities for testing via computer. Over the last six years, TESA administration of multiple choice tests has grown to account for over 90 percent of the students tested. In the 2007-08 school year paper tests will be made available as an accommodation only.

Offering this choice carries with it the responsibility for ensuring comparability of scores, such that a student should be indifferent—at least from the perspective of score level obtained—in choosing between TESA and paper tests. Prior to TESA launch, the Department researched the comparability of paper and pencil tests and found that computer-adaptive tests have been shown to provide equivalent results to paper and pencil tests, but require only about half the testing time (Wainer, Dorans, Green, Mislevy, Steinberg, and Thissen, 2000). When TESA was launched, the Department conducted several equating studies to ensure that TESA scores would be comparable to the existing paper reporting scale (Choi & Tinkler, 2002). This study suggested that scores are comparable when the administration mode of the test changes, although the youngest test users who have less experience with computers may experience more difficulty with computer-based tests than with paper-and-pencil tests. They also identified slight “mode effects” at both the item and scale level, which led to the decision to maintain separate item parameters for computer administered and paper test items. This ensured that scores derived from the two systems would have a common meaning.

The purpose of the present study is to confirm the link between TESA and paper test scales, allowing districts the flexibility to choose between test administration mode. To this end, the study is intended to yield a set of paper and pencil test forms that may be used interchangeably with TESA. The same test forms may be used as “linking blocks” of items for creating future paper and pencil test forms that are equated to the TESA scale.

## Procedure

### Equating Design

The “single group with counterbalancing” design (Kolen & Brennan, 2004) was chosen to equate the TESA and paper tests. Students were asked to take tests under both modes, counterbalanced for order of administration by classroom unit. The assumption underlying this design is that order effects, such as practice or fatigue, cancel out. Scale scores obtained via the TESA test were used to anchor the estimation of paper item difficulty parameters and score table using the Rasch calibration program Winsteps.

## Subjects and Test Administration Procedure

A sample of 180 schools was drawn to participate in the TESA-paper comparability study. Schools eligible to be selected were those not already identified in the NAEP sample (N = 257) and those not on the list of schools that were planning to administer only paper assessments (N = 56). The remaining 2824 school/grade combinations were rank ordered on the basis of SES indicators. Every 40th school/grade was selected to receive a letter requesting participation (see Appendix A), with the stipulation that a school would only be requested to supply up to 100 students at one grade and subject area.<sup>1</sup> A \$2.00 per student stipend was offered as an incentive to participate in the study. Approximately 60 schools drawn in the original sample declined to participate and in most cases they were not replaced. A small number of grade or subject substitutions were allowed for administrative convenience. The expected number of student participants by subject and grade is given in Table 1.<sup>2</sup>

Schools were instructed to counterbalance the order of testing mode by classroom (see instructions to participants in Appendix B). That is, if four classes were being tested, two were instructed to take the paper test first, and the other two classes were to take the TESA exam first.

Test administration took place during the three-week period February 12 through March 7, 2007.

Table 1. Count of Selected Sample Schools and Students

Subject	Grade	Schools	Students
Mathematics	3	6	355
	4	6	327
	5	8	541
	6	9	485
	7	6	407
	8	7	462
	10	5	272
Reading	3	6	309
	4	10	520
	5	9	339
	6	8	420
	7	5	407
	8	4	316
	10	6	368
Science	5	8	562
	8	8	551
	10	7	418

<sup>1</sup> The limit of 100 students was used to: (1) increase the number of participating schools (avoiding one or two schools dominating a grade's sample) and (2) make it more feasible to administer TESA exams during a short testing window.

<sup>2</sup> Counts in Table 1 are based on faxed agreements to participate, received by December 14, 2006.

## Test Form Construction

The TESA tests consisted of operational computer-administered tests developed for 2006-07, which adhere to the blueprints described in the ODE Test Specifications documents.<sup>3</sup> The tests adapt to the student's estimated ability level, maximizing test information (and minimizing standard errors of measurement), while adhering to content balance constraints.<sup>4</sup> Each test is drawn from a pool of items that are written (and independently judged) to measure grade level academic content.

The paper tests were constructed specifically for the comparability study to match the test blueprints in terms of content and item difficulty and avoid item overlap with the TESA pools. The operational TESA administration provided the computer-based tests used in the study, however, the operational paper form is secure and in use during 2006-07, so it could not be used. The paper tests developed for the study were built following the same process as for operational administered paper forms. See the tables in Appendix C for summaries of the number of items by strand. Both the TESA and paper tests are designed to align to the test specifications, in terms of both emphasis and coverage.

---

<sup>3</sup> See Documents 3.1, 3.2 and 3.3 for detailed test specifications by grade.

<sup>4</sup> See Document 7 for an analysis of the performance of the CAT algorithm in meeting content specifications.

## Results

### Participation in the Study

Final counts of students with pairs of tests are shown in Table 2. Each TESA and paper test record was matched on the basis of Oregon's unique student identifier. For a variety of practical reasons, it was anticipated that a lower number of matched student records would be obtained than the potential number listed in Table 1.

Table 2. Count of Participating Sample Schools and Students

Subject	Grade	Schools	Students
Mathematics	3	6	251
	4	5	266
	5	7	396
	6	8	292
	7	4	322
	8	6	304
	10	3	156
Reading	3	6	234
	4	10	392
	5	6	187
	6	5	152
	7	4	384
	8	3	135
	10	3	207
Science	5	7	460
	8	5	388
	10	7	229

### Quality Control

Prior to scoring, paper test responses were reviewed to identify any faulty or mis-keyed items. Three mathematics items had low or negative point-biserial correlations. These items were deemed to contain misleading wording that may have confused students and were not counted in the final scores. Two item key errors were corrected. No item problems were found in the operational TESA items.

The pairing of TESA and paper scored records was checked by matching on additional fields (last name, first name, date of birth). Oregon's SSID system is mature and well-implemented by districts, so few records needed manual correction in order to ensure a correct match. Typically this occurred with hand-bubbled SSIDs, and was not an issue with precoded answer sheets.

### Common Person Calibration and Scoring

After item responses were scored as correct/incorrect/omit, response files and "person anchor" files were prepared for Winsteps input. The student's ability estimate on

TESA was used to anchor the estimation of paper item difficulty parameters. A sample Winsteps control file is included in Appendix D. A sample set of Winsteps output is provided in Appendix E. The scoring table produced by Winsteps was used to look up scale scores and standard errors of measurement corresponding to each student's raw score (number of correct answers). The complete set of raw to scale score tables is provided in Appendix F.

### Score Comparisons between TESA and Paper Tests

Descriptive statistics for TESA and paper scale scores are shown in Table 3. The means for students taking both the computer-progressive and paper administrations of the test were nearly identical. On average scores differed by between .01 and .40 of a RIT scale score point, with an average difference of .15. Standard deviations were similar across the two tests as well.

Table 3. Scale Score Means and Standard Deviations by Subject, Grade and Mode of Assessment

Subject	Grade	N	Mean		Standard Deviation	
			TESA	Paper	TESA	Paper
Mathematics	3	251	203.01	202.99	10.36	9.72
	4	266	212.02	212.08	10.38	11.04
	5	396	220.85	221.29	10.50	12.17
	6	292	222.08	222.25	10.27	11.65
	7	322	230.37	230.52	9.49	10.86
	8	304	232.00	232.14	9.43	10.83
	10	156	230.73	230.63	8.68	9.01
	All	1987	221.62	221.78	13.82	14.60
Reading	3	234	211.59	211.05	13.49	12.39
	4	392	215.79	215.77	12.17	12.18
	5	187	220.01	220.16	10.15	11.10
	6	152	223.13	223.38	10.51	12.10
	7	384	236.13	236.19	12.03	10.90
	8	135	229.27	229.12	10.14	9.56
	10	207	238.29	238.26	8.70	8.83
	All	1691	224.78	224.74	15.15	15.06
Science	5	460	227.72	227.71	10.96	10.23
	8	388	234.86	234.78	11.40	10.52
	10	229	240.35	240.21	12.42	11.08
	All	1077	232.98	232.92	12.46	11.60

Paired sample t-tests were used to evaluate the differences between TESA and equated paper scale scores. Table 4 provides the mean difference, standard error of the difference, t value, degrees of freedom and two-tailed significance level for each grade/subject combination and subject totals across grades. None of the average scores for each grade/subject differed significantly across TSEA or paper administered tests.

Table 4. Paired Samples T-test of Mean Difference between TESA and Paper Scale Scores

Subject	Grade	Mean difference (TESA – paper)	Std. Error of mean difference	t value	Df	Sig. (2 tailed)
Mathematics	3	0.02	0.46	0.043	250	0.966
	4	-0.06	0.38	-0.168	265	0.867
	5	-0.44	0.40	-1.076	395	0.283
	6	-0.18	0.48	-0.374	291	0.709
	7	-0.15	0.39	-0.390	321	0.697
	8	-0.15	0.42	-0.354	303	0.724
	10	0.10	0.55	0.183	155	0.855
	All	-0.16	0.16	-0.960	1986	0.337
Reading	3	0.54	0.57	0.948	233	0.344
	4	0.02	0.40	0.058	391	0.954
	5	-0.15	0.50	-0.303	186	0.763
	6	-0.26	0.59	-0.436	151	0.664
	7	-0.06	0.39	-0.147	383	0.883
	8	0.15	0.66	0.225	134	0.823
	10	0.03	0.51	0.057	206	0.955
	All	0.04	0.19	0.228	1690	0.820
Science	5	0.01	0.33	0.027	459	0.978
	8	0.08	0.37	0.228	387	0.820
	10	0.14	0.48	0.297	228	0.766
	All	0.06	0.22	0.295	1076	0.768

Correlations between TESA and paper test scale scores by subject and grade are provided in Table 5. The Pearson correlation coefficients, with and without correction for attenuation<sup>5</sup>, are shown for all cases (Column A). Because participation in the study was a relatively “low stakes” event for many students, some may have not put forth their best effort. By using the Rasch “person fit” statistic it is possible to identify tests for which student item responses are unexpected (e.g., answering easy items incorrectly). Such patterns indicate tests in which the student was careless, rushed, or otherwise unengaged. Under the assumption that such tests include construct-irrelevant variance and provide misleading information, researchers have developed methods for identifying them (Wise, et al., 2007; Kingsbury and Houser, 2007). The second set of correlations below (Column B) include only the cases in which the “outfit” mean square statistic for the paper test falls between 0.5 and 1.3.

<sup>5</sup> Correction for attenuation provides a theoretical estimate of the correlation if the tests were free of measurement error (Nunnally & Bernstein, 1994)



Table 5. Correlation between TESA and Paper Scale Scores by Grade

Subject	Grade	A) All Cases		B) Cases Meeting Model Fit Criterion <sup>6</sup>	
		Observed Correlation	Corrected for Attenuation	Observed Correlation	Corrected for Attenuation
Mathematics	3	0.737	0.873	0.784	0.929
	4	0.831	0.961	0.832	0.962
	5	0.756	0.854	0.788	0.890
	6	0.732	0.841	0.792	0.910
	7	0.771	0.907	0.803	0.945
	8	0.751	0.884	0.750	0.882
	10	0.700	0.870	0.694	0.863
	Median	0.751	0.873	0.788	0.910
	Range	0.700 – 0.831	0.841 – 0.961	0.784 – 0.832	0.863 – 0.962
Reading	3	0.778	0.865	0.813	0.904
	4	0.792	0.870	0.848	0.932
	5	0.801	0.910	0.797	0.906
	6	0.803	0.902	0.847	0.952
	7	0.784	0.881	0.820	0.922
	8	0.700	0.824	0.807	0.950
	10	0.649	0.792	0.748	0.912
	Median	0.784	0.870	0.813	0.922
	Range	0.649 – 0.843	0.792 – 0.910	0.748 – 0.838	0.904 – 0.952
Science	5	0.779	0.881	0.808	0.914
	8	0.801	0.896	0.834	0.933
	10	0.814	0.910	0.851	0.945
	Median	0.801	0.896	0.834	0.933
	Range	0.779 – 0.814	0.881 – 0.910	0.808 – 0.851	0.914 – 0.945

<sup>6</sup> Students are included if their “person fit” statistic meets the criterion for individual validity

## Discussion

Given the assumptions of the equating design, the study provided item difficulty parameters for the paper test items that should yield comparable scores to the computer administered tests, removing “mode effects” that might otherwise contribute construct irrelevant variance to the measures. The TESA scale was used as the basis for scaling the paper test, since TESA is, and will remain the dominant mode of administration under Oregon’s Statewide Assessment System. Students who need to take a paper test are able to do so without concern that the score is biased in either direction.

While not addressed in the current study, additional information about the consistency of student classification across TESA and paper forms is described Oregon’s Statewide Assessment Technical Report, Volume 4: Reliability and Validity. See specifically Sections 4.1-4.4 describing this work (ODE, 2007; McCall, 2005; and Doran & Cohen, 2006). This work provides evidence that students are classified consistently into performance level classifications across TESA and paper test administration modes.

Looking forward, the comparability study also offers a means to check the continuity of the TESA scale when a new TESA delivery platform is launched in 2007-08. The paper linking blocks created in the present study can be re-administered in 2007-08 along with the new version of TESA to verify that the TESA scale is unaffected.

## **Appendix A**

### **Letter of Invitation to Participate in Study**



OREGON DEPARTMENT OF EDUCATION  
Public Service Building, 255 Capitol Street NE, Salem, Oregon 97310  
Phone (503) 947-5600 • Fax (503) 378-5156 • www.ode.state.or.us

November 8, 2006

To: «Principal», Principal  
«School\_Name»

From: Doug Kosty, Assistant Superintendent  
Office of Assessment & Information Services

Re: Paper/Pencil and TESA Comparability Study

This letter is to inform you that your school has been selected to participate in a study of the Oregon Statewide Assessment System required by the U. S. Department of Education. The purpose of this study is to demonstrate that students who use the TESA testing system receive a comparable score when they use the paper/pencil version of the state test.

Schools were randomly selected for this study which will require students in selected classrooms to take a short form of the paper/pencil test in addition to their regular participation in TESA. As an incentive to schools to participate, we are offering a \$2.00 per student stipend for completing both portions of the test.

The study will take place between February 12-28, 2007. Students in the study would need to complete both the regular TESA test and the short paper/pencil test in that time frame.

Your school has been selected to participate in the following study:

- **Subject to be Assessed:** «Subject»
- **Grade level to be Assessed:** «Grade»

To confirm your participation in the study, please return the enclosed Paper/Pencil and TESA Comparability Study form by FAX to Peggy Kouf at 503-378-5156.

As always, your willingness to assist us in this important work is greatly appreciated.

If you have questions about the study, please contact Barbara Wolfe at 503-947-5823 ([barbara.wolfe@state.or.us](mailto:barbara.wolfe@state.or.us)) or Peggy Kouf at 503-947-5829 ([peggy.kouf@state.or.us](mailto:peggy.kouf@state.or.us).)

Enclosure: Paper/Pencil and TESA Comparability Study Confirmation Form  
cc: «DTC», District Testing Coordinator

## **Appendix B**

### **Instructions to Participating Schools**

To: Schools Involved in Comparability Study  
Re: Materials and Procedures

You will soon be receiving a shipment of testing materials for the Paper/Pencil and TESA Comparability Study. All the materials you need for the paper/pencil portion of the assessment are contained in the shipment. If you ordered pre-coded answer sheets, those will be included.

If you have any questions about the shipment, please contact Katie Cowan at [Katie.cowan@state.or.us](mailto:Katie.cowan@state.or.us) or 503-947-5842.

For the purposes of the study, we want approximately half of the students to take the paper/pencil test first and half to take the TESA portion of the test first. You can handle this in one of two ways:

- 1) If it is feasible in your school, simply give about half the students the paper/pencil test first and half the students the TESA test first. When you return the answer sheets, include a separate page saying who was in each group.
- 2) If it is difficult to split up the students in your school for testing, simply e-mail Steve Slater at [steve.slater@state.or.us](mailto:steve.slater@state.or.us) telling him what grade and subject you are testing and which you plan to do first – paper/pencil or TESA. That way he can keep track of the order in which different schools will be conducting the tests for the study.

I want to stress the importance of having as many students as possible complete both portions of the test. Remember, these are only students who are normally tested under standard test administration. Having both tests completed is critical to the validity of the study.

When your tests have been returned and the TESA portions verified, we will calculate your per student stipend based on completing both portions of the test.

If you have questions or concerns, please do not hesitate to contact me.

Barbara Wolfe  
Director for Assessment Communications  
[Barbara.wolfe@state.or.us](mailto:Barbara.wolfe@state.or.us)  
503-947-5823

## **Appendix C**

### **Number of Items by Strand**

Appendix C-1. Mathematics Paper Test Number of Items by Strand

Grade	Strand	Number of Items	Percent of Test	Target Percent of Test
3	Calculation and Estimation	7	23%	25%
3	Measurement	6	20%	20%
3	Statistics and Probability	4	13%	15%
3	Algebraic Relationships	6	20%	20%
3	Geometry	7	23%	20%
4	Calculation and Estimation	6	20%	20%
4	Measurement	6	20%	20%
4	Statistics and Probability	6	20%	20%
4	Algebraic Relationships	6	20%	20%
4	Geometry	6	20%	20%
5	Calculation and Estimation	6	20%	20%
5	Measurement	6	20%	20%
5	Statistics and Probability	6	20%	20%
5	Algebraic Relationships	6	20%	20%
5	Geometry	6	20%	20%
6	Calculation and Estimation	5	17%	15%
6	Measurement	6	20%	20%
6	Statistics and Probability	6	20%	20%
6	Algebraic Relationships	7	23%	25%
6	Geometry	6	20%	20%
7	Calculation and Estimation	4	13%	15%
7	Measurement	5	17%	15%
7	Statistics and Probability	6	20%	20%
7	Algebraic Relationships	9	30%	30%
7	Geometry	6	20%	20%
8	Calculation and Estimation	4	13%	15%
8	Measurement	5	17%	15%
8	Statistics and Probability	6	20%	20%
8	Algebraic Relationships	9	30%	30%
8	Geometry	6	20%	20%
10	Calculation and Estimation	3	10%	10%
10	Measurement	3	10%	10%
10	Statistics and Probability	6	20%	20%
10	Algebraic Relationships	10	33%	35%
10	Geometry	8	27%	25%



Appendix C-2. Reading Paper Test Number of Items by Strand

Grade	Strand	Number of Items	Percent of Test	Target Percent of Test
3	Vocabulary	6	24%	28%
3	Read to Perform a Task	4	16%	16%
3	Demonstrate General Understanding	9	36%	28%
3	Develop an Interpretation	6	24%	28%
4	Vocabulary	8	27%	25%
4	Read to Perform a Task	4	13%	13%
4	Demonstrate General Understanding	7	23%	25%
4	Develop an Interpretation	8	27%	25%
4	Examine Content and Structure: Informative Text	3	10%	12%
5	Vocabulary	7	23%	21%
5	Read to Perform a Task	4	13%	13%
5	Demonstrate General Understanding	7	23%	21%
5	Develop an Interpretation	6	20%	21%
5	Examine Content and Structure: Informative Text	3	10%	12%
5	Examine Content and Structure: Literary Text	3	10%	12%
6	Vocabulary	6	20%	20%
6	Read to Perform a Task	3	10%	12%
6	Demonstrate General Understanding	6	20%	20%
6	Develop an Interpretation	7	23%	22%
6	Examine Content and Structure: Informative Text	4	13%	14%
6	Examine Content and Structure: Literary Text	4	13%	14%
7	Vocabulary	5	17%	20%
7	Read to Perform a Task	4	13%	12%
7	Demonstrate General Understanding	6	20%	20%
7	Develop an Interpretation	5	17%	20%
7	Examine Content and Structure: Informative Text	5	17%	14%
7	Examine Content and Structure: Literary Text	5	17%	14%
8	Vocabulary	7	23%	20%
8	Read to Perform a Task	3	10%	12%
8	Demonstrate General Understanding	6	20%	18%
8	Develop an Interpretation	6	20%	20%
8	Examine Content and Structure: Informative Text	4	13%	15%
8	Examine Content and Structure: Literary Text	4	13%	15%
10	Vocabulary	5	17%	20%
10	Read to Perform a Task	3	10%	12%
10	Demonstrate General Understanding	6	20%	16%
10	Develop an Interpretation	6	20%	20%
10	Examine Content and Structure: Informative Text	5	17%	16%
10	Examine Content and Structure: Literary Text	5	17%	16%

Appendix C-3. Science Paper Test Number of Items by Strand

Grade	Strand	Number of Items	Percent of Test	Target Percent of Test
5	Physical Science	10	33%	33%
5	Life Science	10	33%	33%
5	Earth and Space Science	10	33%	33%
8	Physical Science	11	37%	33%
8	Life Science	10	33%	33%
8	Earth and Space Science	9	30%	33%
10	Physical Science	10	33%	33%
10	Life Science	10	33%	33%
10	Earth and Space Science	10	33%	33%

## **Appendix D**

### **Winsteps Control File (Sample)**

```
&INST
TITLE='March 2007 OSAT, Reading, GRADE: 4 all, FORM: Paper Comparability Study'
NI=30
NAME1=1
NAMELEN=30
ITEM1=37
XWIDE=1
DATA=RL Gr4 all.DAT
PAFILE=RL Gr4 all.ANC
CODES=01
MPROX=30
MUCON=50
CONVERGE=L
LCONV=0.0001
UMEAN=200
USCALE=10
PTBIS=N
MRANGE=40
TABLES=00100000000000000000100
ASCII=Y
MAXPAG=70
&END
R0500270 01
R0500290 02
R0500300 03
R0500320 04
R0500330 05
R0500340 06
R0500360 07
R0465020 08
R0465050 09
R0465060 10
R0465070 11
R0465210 12
R0465240 13
R0500380 14
R0500390 15
R0500420 16
R0500440 17
R0500450 18
R0500460 19
R0406150 20
R0406160 21
R0406190 22
R0406200 23
R0224870 24
R0224880 25
R0237350 26
R0237420 27
R0237370 28
R0237410 29
R0237380 30
END NAMES
```

## **Appendix E**

### **Winsteps Output (Sample)**

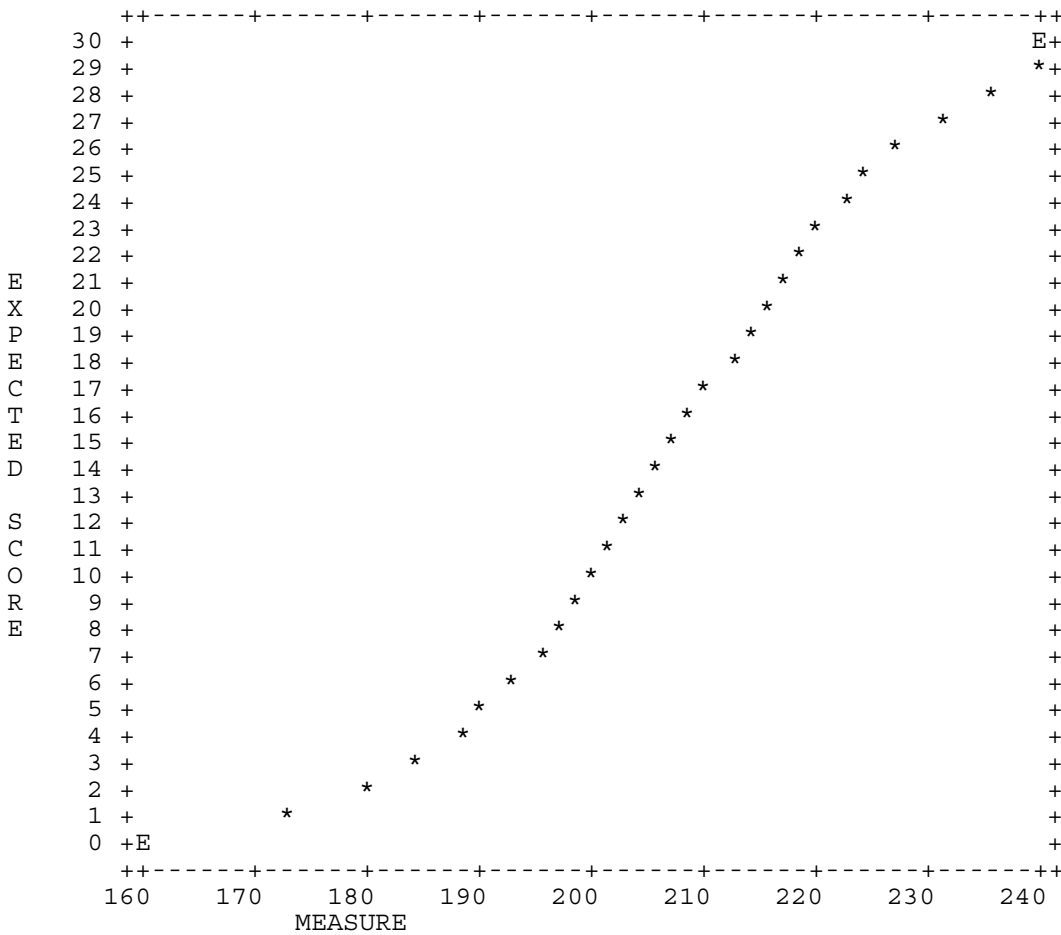
TABLE 20.1 March 2007 OSAT, Reading, GRA RL Gr4 All Output.txt Apr 6 7:34 2007  
 INPUT: 392 PERSONS 30 ITEMS MEASURED: 392 PERSONS 30 ITEMS 2 CATS 3.61.1

TABLE OF MEASURES ON COMPLETE TEST

SCORE	MEASURE	S.E.	SCORE	MEASURE	S.E.	SCORE	MEASURE	S.E.
0	160.30E	18.38	11	201.94	3.92	22	218.68	4.26
1	172.66	10.23	12	203.45	3.86	23	220.57	4.44
2	180.05	7.39	13	204.93	3.82	24	222.65	4.68
3	184.58	6.17	14	206.38	3.80	25	224.99	5.01
4	187.94	5.47	15	207.82	3.79	26	227.72	5.47
5	190.67	5.01	16	209.26	3.80	27	231.09	6.18
6	193.01	4.68	17	210.71	3.83	28	235.62	7.40
7	195.08	4.44	18	212.19	3.87	29	243.01	10.23
8	196.97	4.25	19	213.71	3.93	30	255.37E	18.38
9	198.72	4.11	20	215.28	4.01			
10	200.36	4.01	21	216.93	4.12			

CURRENT VALUES, UMEAN=200.000 USCALE=10.000  
 TO SET MEASURE RANGE AS 0-100, UMEAN=41.760 USCALE=10.519  
 TO SET MEASURE RANGE TO MATCH RAW SCORE RANGE, UMEAN=12.528 USCALE=3.156  
 TEST SLOPE=.95 INTERCEPT=207.83

RAW SCORE-MEASURE OGIVE FOR COMPLETE TEST



PERSON 1 111121121312 2 1 1 1  
 1 212243271976877033726022484815684311339  
 T S M S T

ITEMS 1 1212233333 24

T S M S T

TABLE 20.2 March 2007 OSAT, Reading, GRA RL Gr4 All Output.txt Apr 6 7:34 2007  
 INPUT: 392 PERSONS 30 ITEMS MEASURED: 392 PERSONS 30 ITEMS 2 CATS 3.61.1

TABLE OF SAMPLE NORMS (500/100) AND FREQUENCIES CORRESPONDING TO COMPLETE TEST

SCORE	MEASURE	S.E.	NORMED	S.E.	FREQUENCY	%	CUM.FREQ.	%	PERCENTILE
0	160.30E	18.38	43	151	0	.0	0	.0	0
1	172.66	10.23	145	84	0	.0	0	.0	0
2	180.05	7.39	206	61	0	.0	0	.0	0
3	184.58	6.17	243	51	3	.8	3	.8	1
4	187.94	5.47	271	45	3	.8	6	1.5	1
5	190.67	5.01	293	41	5	1.3	11	2.8	2
6	193.01	4.68	312	38	6	1.5	17	4.3	4
7	195.08	4.44	329	36	7	1.8	24	6.1	5
8	196.97	4.25	345	35	1	.3	25	6.4	6
9	198.72	4.11	359	34	9	2.3	34	8.7	8
10	200.36	4.01	373	33	15	3.8	49	12.5	11
11	201.94	3.92	386	32	8	2.0	57	14.5	14
12	203.45	3.86	398	32	15	3.8	72	18.4	16
13	204.93	3.82	410	31	4	1.0	76	19.4	19
14	206.38	3.80	422	31	23	5.9	99	25.3	22
15	207.82	3.79	434	31	8	2.0	107	27.3	26
16	209.26	3.80	446	31	5	1.3	112	28.6	28
17	210.71	3.83	458	31	25	6.4	137	34.9	32
18	212.19	3.87	470	32	15	3.8	152	38.8	37
19	213.71	3.93	483	32	28	7.1	180	45.9	42
20	215.28	4.01	496	33	20	5.1	200	51.0	48
21	216.93	4.12	509	34	12	3.1	212	54.1	53
22	218.68	4.26	523	35	32	8.2	244	62.2	58
23	220.57	4.44	539	37	20	5.1	264	67.3	65
24	222.65	4.68	556	39	26	6.6	290	74.0	71
25	224.99	5.01	575	41	29	7.4	319	81.4	78
26	227.72	5.47	598	45	21	5.4	340	86.7	84
27	231.09	6.18	625	51	25	6.4	365	93.1	90
28	235.62	7.40	663	61	18	4.6	383	97.7	95
29	243.01	10.23	724	84	7	1.8	390	99.5	99
30	255.37E	18.38	825	151	2	.5	392	100.0	99

THE NORMED SCALE IS EQUIVALENT TO UIMEAN= -1274.91 USCALE= 8.22



```

*****
*
*           * * * * W I N S T E P S * * * *
*           -----
*
*           - RASCH ANALYSIS FOR TWO-FACET MODELS -
*
*           PERSON, ITEM & RESPONSE STRUCTURE MEASUREMENT AND FIT ANALYSIS
*
* INQUIRE: WINSTEPS
*           PO BOX 811322, CHICAGO ILLINOIS 60681-1322
*           Tel. & FAX (312) 264-2352
*           www.winsteps.com
*
*           COPYRIGHT (C) JOHN MICHAEL LINACRE, 1991-2006
*
*                               MAY 22, 2006           VERSION 3.61.*
*****

```

TABLE 0.1 March 2007 OSAT, Reading, GRAD RL Gr4 All Output.txt Apr 6 7:34 2007

---

TITLE= March 2007 OSAT, Reading, GRADE: 4 all, FORM: Paper Comparab  
CONTROL FILE: J:\ADAT\Slater\2007 Assessment\Comparability studies\PP-TESA comparability\W  
OUTPUT FILE: J:\ADAT\Slater\2007 Assessment\Comparability studies\PP-TESA comparability\W  
DATE: Apr 6 7:34 2007  
392 PERSON Records Input.

TABLE 0.2 March 2007 OSAT, Reading, GRAD RL Gr4 All Output.txt Apr 6 7:34 2007  
 INPUT: 392 PERSONS 30 ITEMS WINSTEPS 3.61.1

-----  
 CONVERGENCE TABLE  
 -----

PROX ITERATION	ACTIVE COUNT			EXTREME 5 RANGE		MAX LOGIT CHANGE	
	PERSONS	ITEMS	CATS	PERSONS	ITEMS	MEASURES	STRUCTURE
1	392	30	2	6.28	1.71	-1.7681	
2	392	30	2	6.28	1.71	.8750	
JMLE ITERATION	MAX SCORE RESIDUAL*	MAX LOGIT CHANGE	LEAST CONVERGED		CATEGORY	STRUCTURE	
			PERSON	ITEM	CAT	RESIDUAL	CHANGE
1	1.73	-.0202	392	10*			
2	.50	-.0056	392	10*			
3	.15	-.0025	392	10*			
4	.04	-.0012	392	10*			
5	.01	-.0005	392	10*			
6	.00	-.0002	392	21*			
7	.00	-.0001	392	24*			

-----  
 Standardized Residuals N(0,1) Mean: -.01 S.D.: 1.10  
 -----

## **Appendix F**

### **Raw to Scale Score Tables for Paper Tests**

Appendix F-1. Mathematics Raw to Scale Score Conversion

Subject	Grade	Strand	Raw Score	Scale Score	SEM
MA	GR3	T	0	149	14
MA	GR3	T	1	156	10
MA	GR3	T	2	164	8
MA	GR3	T	3	169	6
MA	GR3	T	4	173	6
MA	GR3	T	5	176	5
MA	GR3	T	6	178	5
MA	GR3	T	7	181	5
MA	GR3	T	8	183	5
MA	GR3	T	9	185	5
MA	GR3	T	10	187	4
MA	GR3	T	11	189	4
MA	GR3	T	12	191	4
MA	GR3	T	13	193	4
MA	GR3	T	14	195	4
MA	GR3	T	15	196	4
MA	GR3	T	16	198	4
MA	GR3	T	17	200	4
MA	GR3	T	18	202	4
MA	GR3	T	19	204	4
MA	GR3	T	20	206	4
MA	GR3	T	21	208	5
MA	GR3	T	22	210	5
MA	GR3	T	23	212	5
MA	GR3	T	24	215	5
MA	GR3	T	25	217	5
MA	GR3	T	26	220	6
MA	GR3	T	27	224	6
MA	GR3	T	28	229	8
MA	GR3	T	29	236	10
MA	GR3	T	30	244	14

Appendix F-1. Mathematics Raw to Scale Score Conversion

Subject	Grade	Strand	Raw Score	Scale Score	SEM
MA	GR4	T	0	158	15
MA	GR4	T	1	166	11
MA	GR4	T	2	174	8
MA	GR4	T	3	179	7
MA	GR4	T	4	183	6
MA	GR4	T	5	186	6
MA	GR4	T	6	189	5
MA	GR4	T	7	192	5
MA	GR4	T	8	194	5
MA	GR4	T	9	197	5
MA	GR4	T	10	199	5
MA	GR4	T	11	201	4
MA	GR4	T	12	203	4
MA	GR4	T	13	205	4
MA	GR4	T	14	206	4
MA	GR4	T	15	208	4
MA	GR4	T	16	210	4
MA	GR4	T	17	212	4
MA	GR4	T	18	214	4
MA	GR4	T	19	215	4
MA	GR4	T	20	217	4
MA	GR4	T	21	219	4
MA	GR4	T	22	221	5
MA	GR4	T	23	224	5
MA	GR4	T	24	226	5
MA	GR4	T	25	229	5
MA	GR4	T	26	231	6
MA	GR4	T	27	235	6
MA	GR4	T	28	240	8
MA	GR4	T	29	247	10
MA	GR4	T	30	255	14

Appendix F-1. Mathematics Raw to Scale Score Conversion

Subject	Grade	Strand	Raw Score	Scale Score	SEM
MA	GR5	T	0	164	14
MA	GR5	T	1	172	10
MA	GR5	T	2	179	8
MA	GR5	T	3	184	6
MA	GR5	T	4	188	6
MA	GR5	T	5	191	5
MA	GR5	T	6	194	5
MA	GR5	T	7	196	5
MA	GR5	T	8	198	5
MA	GR5	T	9	200	5
MA	GR5	T	10	202	4
MA	GR5	T	11	204	4
MA	GR5	T	12	206	4
MA	GR5	T	13	208	4
MA	GR5	T	14	210	4
MA	GR5	T	15	211	4
MA	GR5	T	16	213	4
MA	GR5	T	17	215	4
MA	GR5	T	18	217	4
MA	GR5	T	19	218	4
MA	GR5	T	20	220	4
MA	GR5	T	21	222	4
MA	GR5	T	22	224	5
MA	GR5	T	23	226	5
MA	GR5	T	24	229	5
MA	GR5	T	25	231	5
MA	GR5	T	26	234	6
MA	GR5	T	27	238	6
MA	GR5	T	28	242	8
MA	GR5	T	29	250	10
MA	GR5	T	30	257	14

Appendix F-1. Mathematics Raw to Scale Score Conversion

Subject	Grade	Strand	Raw Score	Scale Score	SEM
MA	GR6	T	0	167	14
MA	GR6	T	1	175	10
MA	GR6	T	2	182	8
MA	GR6	T	3	187	7
MA	GR6	T	4	191	6
MA	GR6	T	5	194	5
MA	GR6	T	6	197	5
MA	GR6	T	7	200	5
MA	GR6	T	8	202	5
MA	GR6	T	9	204	5
MA	GR6	T	10	206	5
MA	GR6	T	11	208	4
MA	GR6	T	12	210	4
MA	GR6	T	13	212	4
MA	GR6	T	14	214	4
MA	GR6	T	15	216	4
MA	GR6	T	16	218	4
MA	GR6	T	17	220	4
MA	GR6	T	18	222	4
MA	GR6	T	19	224	5
MA	GR6	T	20	226	5
MA	GR6	T	21	228	5
MA	GR6	T	22	230	5
MA	GR6	T	23	233	5
MA	GR6	T	24	236	5
MA	GR6	T	25	239	6
MA	GR6	T	26	243	7
MA	GR6	T	27	248	8
MA	GR6	T	28	256	11
MA	GR6	T	29	264	15



Appendix F-1. Mathematics Raw to Scale Score Conversion

Subject	Grade	Strand	Raw Score	Scale Score	SEM
MA	GR7	T	0	174	14
MA	GR7	T	1	182	10
MA	GR7	T	2	190	8
MA	GR7	T	3	195	7
MA	GR7	T	4	199	6
MA	GR7	T	5	202	6
MA	GR7	T	6	205	5
MA	GR7	T	7	207	5
MA	GR7	T	8	210	5
MA	GR7	T	9	212	5
MA	GR7	T	10	214	5
MA	GR7	T	11	216	5
MA	GR7	T	12	218	4
MA	GR7	T	13	220	4
MA	GR7	T	14	222	4
MA	GR7	T	15	224	4
MA	GR7	T	16	226	4
MA	GR7	T	17	228	4
MA	GR7	T	18	230	4
MA	GR7	T	19	232	4
MA	GR7	T	20	234	4
MA	GR7	T	21	236	5
MA	GR7	T	22	238	5
MA	GR7	T	23	240	5
MA	GR7	T	24	243	5
MA	GR7	T	25	245	5
MA	GR7	T	26	248	6
MA	GR7	T	27	252	6
MA	GR7	T	28	257	8
MA	GR7	T	29	264	10
MA	GR7	T	30	272	14

Appendix F-1. Mathematics Raw to Scale Score Conversion

Subject	Grade	Strand	Raw Score	Scale Score	SEM
MA	GR8	T	0	177	14
MA	GR8	T	1	184	10
MA	GR8	T	2	192	8
MA	GR8	T	3	197	7
MA	GR8	T	4	201	6
MA	GR8	T	5	204	6
MA	GR8	T	6	207	5
MA	GR8	T	7	210	5
MA	GR8	T	8	212	5
MA	GR8	T	9	214	5
MA	GR8	T	10	216	5
MA	GR8	T	11	218	4
MA	GR8	T	12	220	4
MA	GR8	T	13	222	4
MA	GR8	T	14	224	4
MA	GR8	T	15	226	4
MA	GR8	T	16	228	4
MA	GR8	T	17	230	4
MA	GR8	T	18	232	4
MA	GR8	T	19	234	4
MA	GR8	T	20	236	5
MA	GR8	T	21	238	5
MA	GR8	T	22	240	5
MA	GR8	T	23	242	5
MA	GR8	T	24	245	5
MA	GR8	T	25	248	5
MA	GR8	T	26	251	6
MA	GR8	T	27	255	6
MA	GR8	T	28	260	8
MA	GR8	T	29	267	10
MA	GR8	T	30	275	14

Appendix F-1. Mathematics Raw to Scale Score Conversion

Subject	Grade	Strand	Raw Score	Scale Score	SEM
MA	GR10	T	0	177	15
MA	GR10	T	1	184	11
MA	GR10	T	2	192	8
MA	GR10	T	3	198	7
MA	GR10	T	4	202	6
MA	GR10	T	5	205	5
MA	GR10	T	6	208	5
MA	GR10	T	7	210	5
MA	GR10	T	8	212	5
MA	GR10	T	9	215	5
MA	GR10	T	10	217	5
MA	GR10	T	11	219	4
MA	GR10	T	12	221	4
MA	GR10	T	13	222	4
MA	GR10	T	14	224	4
MA	GR10	T	15	226	4
MA	GR10	T	16	228	4
MA	GR10	T	17	230	4
MA	GR10	T	18	232	5
MA	GR10	T	19	234	5
MA	GR10	T	20	236	5
MA	GR10	T	21	239	5
MA	GR10	T	22	241	5
MA	GR10	T	23	244	5
MA	GR10	T	24	247	6
MA	GR10	T	25	251	7
MA	GR10	T	26	256	8
MA	GR10	T	27	264	10
MA	GR10	T	28	272	14

Appendix F-2. Reading Raw to Scale Score Conversion

Subject	Grade	Strand	Raw Score	Scale Score	SEM
RL	GR3	T	0	155	14
RL	GR3	T	1	163	10
RL	GR3	T	2	171	8
RL	GR3	T	3	175	6
RL	GR3	T	4	179	6
RL	GR3	T	5	182	5
RL	GR3	T	6	185	5
RL	GR3	T	7	188	5
RL	GR3	T	8	190	5
RL	GR3	T	9	192	5
RL	GR3	T	10	194	5
RL	GR3	T	11	196	5
RL	GR3	T	12	198	5
RL	GR3	T	13	200	5
RL	GR3	T	14	202	5
RL	GR3	T	15	204	5
RL	GR3	T	16	207	5
RL	GR3	T	17	209	5
RL	GR3	T	18	211	5
RL	GR3	T	19	214	5
RL	GR3	T	20	216	5
RL	GR3	T	21	220	6
RL	GR3	T	22	224	7
RL	GR3	T	23	228	8
RL	GR3	T	24	236	10
RL	GR3	T	25	244	14

Appendix F-2. Reading Raw to Scale Score Conversion

Subject	Grade	Strand	Raw Score	Scale Score	SEM
RL	GR4	T	0	165	14
RL	GR4	T	1	173	10
RL	GR4	T	2	180	7
RL	GR4	T	3	185	6
RL	GR4	T	4	188	5
RL	GR4	T	5	191	5
RL	GR4	T	6	193	5
RL	GR4	T	7	195	4
RL	GR4	T	8	197	4
RL	GR4	T	9	199	4
RL	GR4	T	10	200	4
RL	GR4	T	11	202	4
RL	GR4	T	12	203	4
RL	GR4	T	13	205	4
RL	GR4	T	14	206	4
RL	GR4	T	15	208	4
RL	GR4	T	16	209	4
RL	GR4	T	17	211	4
RL	GR4	T	18	212	4
RL	GR4	T	19	214	4
RL	GR4	T	20	215	4
RL	GR4	T	21	217	4
RL	GR4	T	22	219	4
RL	GR4	T	23	221	4
RL	GR4	T	24	223	5
RL	GR4	T	25	225	5
RL	GR4	T	26	228	5
RL	GR4	T	27	231	6
RL	GR4	T	28	236	7
RL	GR4	T	29	243	10
RL	GR4	T	30	250	14

Appendix F-2. Reading Raw to Scale Score Conversion

Subject	Grade	Strand	Raw Score	Scale Score	SEM
RL	GR5	T	0	168	14
RL	GR5	T	1	175	10
RL	GR5	T	2	183	8
RL	GR5	T	3	187	6
RL	GR5	T	4	191	6
RL	GR5	T	5	194	5
RL	GR5	T	6	196	5
RL	GR5	T	7	199	5
RL	GR5	T	8	201	4
RL	GR5	T	9	203	4
RL	GR5	T	10	204	4
RL	GR5	T	11	206	4
RL	GR5	T	12	208	4
RL	GR5	T	13	209	4
RL	GR5	T	14	211	4
RL	GR5	T	15	212	4
RL	GR5	T	16	214	4
RL	GR5	T	17	215	4
RL	GR5	T	18	217	4
RL	GR5	T	19	219	4
RL	GR5	T	20	220	4
RL	GR5	T	21	222	4
RL	GR5	T	22	224	4
RL	GR5	T	23	226	5
RL	GR5	T	24	228	5
RL	GR5	T	25	230	5
RL	GR5	T	26	233	6
RL	GR5	T	27	237	6
RL	GR5	T	28	241	7
RL	GR5	T	29	249	10
RL	GR5	T	30	256	14

Appendix F-2. Reading Raw to Scale Score Conversion

Subject	Grade	Strand	Raw Score	Scale Score	SEM
RL	GR6	T	0	173	14
RL	GR6	T	1	180	10
RL	GR6	T	2	188	8
RL	GR6	T	3	193	6
RL	GR6	T	4	196	6
RL	GR6	T	5	199	5
RL	GR6	T	6	201	5
RL	GR6	T	7	204	5
RL	GR6	T	8	206	4
RL	GR6	T	9	207	4
RL	GR6	T	10	209	4
RL	GR6	T	11	211	4
RL	GR6	T	12	213	4
RL	GR6	T	13	214	4
RL	GR6	T	14	216	4
RL	GR6	T	15	217	4
RL	GR6	T	16	219	4
RL	GR6	T	17	220	4
RL	GR6	T	18	222	4
RL	GR6	T	19	224	4
RL	GR6	T	20	225	4
RL	GR6	T	21	227	4
RL	GR6	T	22	229	4
RL	GR6	T	23	231	5
RL	GR6	T	24	233	5
RL	GR6	T	25	235	5
RL	GR6	T	26	238	6
RL	GR6	T	27	242	6
RL	GR6	T	28	246	7
RL	GR6	T	29	254	10
RL	GR6	T	30	261	14

Appendix F-2. Reading Raw to Scale Score Conversion

Subject	Grade	Strand	Raw Score	Scale Score	SEM
RL	GR7	T	0	180	14
RL	GR7	T	1	188	10
RL	GR7	T	2	195	8
RL	GR7	T	3	200	6
RL	GR7	T	4	204	6
RL	GR7	T	5	207	5
RL	GR7	T	6	209	5
RL	GR7	T	7	212	5
RL	GR7	T	8	214	5
RL	GR7	T	9	216	4
RL	GR7	T	10	218	4
RL	GR7	T	11	220	4
RL	GR7	T	12	221	4
RL	GR7	T	13	223	4
RL	GR7	T	14	225	4
RL	GR7	T	15	227	4
RL	GR7	T	16	228	4
RL	GR7	T	17	230	4
RL	GR7	T	18	232	4
RL	GR7	T	19	234	4
RL	GR7	T	20	235	4
RL	GR7	T	21	237	5
RL	GR7	T	22	240	5
RL	GR7	T	23	242	5
RL	GR7	T	24	244	5
RL	GR7	T	25	247	5
RL	GR7	T	26	250	6
RL	GR7	T	27	254	7
RL	GR7	T	28	259	8
RL	GR7	T	29	267	11
RL	GR7	T	30	274	15



Appendix F-2. Reading Raw to Scale Score Conversion

Subject	Grade	Strand	Raw Score	Scale Score	SEM
RL	GR8	T	0	176	14
RL	GR8	T	1	183	10
RL	GR8	T	2	191	8
RL	GR8	T	3	196	6
RL	GR8	T	4	199	6
RL	GR8	T	5	203	5
RL	GR8	T	6	205	5
RL	GR8	T	7	208	5
RL	GR8	T	8	210	5
RL	GR8	T	9	212	4
RL	GR8	T	10	214	4
RL	GR8	T	11	216	4
RL	GR8	T	12	218	4
RL	GR8	T	13	219	4
RL	GR8	T	14	221	4
RL	GR8	T	15	223	4
RL	GR8	T	16	224	4
RL	GR8	T	17	226	4
RL	GR8	T	18	228	4
RL	GR8	T	19	230	4
RL	GR8	T	20	232	4
RL	GR8	T	21	233	4
RL	GR8	T	22	236	5
RL	GR8	T	23	238	5
RL	GR8	T	24	240	5
RL	GR8	T	25	243	5
RL	GR8	T	26	246	6
RL	GR8	T	27	250	6
RL	GR8	T	28	254	8
RL	GR8	T	29	262	10
RL	GR8	T	30	270	14

Appendix F-2. Reading Raw to Scale Score Conversion

Subject	Grade	Strand	Raw Score	Scale Score	SEM
RL	GR10	T	0	189	14
RL	GR10	T	1	197	10
RL	GR10	T	2	204	7
RL	GR10	T	3	209	6
RL	GR10	T	4	212	6
RL	GR10	T	5	215	5
RL	GR10	T	6	217	5
RL	GR10	T	7	220	5
RL	GR10	T	8	222	4
RL	GR10	T	9	223	4
RL	GR10	T	10	225	4
RL	GR10	T	11	227	4
RL	GR10	T	12	228	4
RL	GR10	T	13	230	4
RL	GR10	T	14	232	4
RL	GR10	T	15	233	4
RL	GR10	T	16	235	4
RL	GR10	T	17	236	4
RL	GR10	T	18	238	4
RL	GR10	T	19	239	4
RL	GR10	T	20	241	4
RL	GR10	T	21	243	4
RL	GR10	T	22	245	4
RL	GR10	T	23	247	5
RL	GR10	T	24	249	5
RL	GR10	T	25	251	5
RL	GR10	T	26	254	6
RL	GR10	T	27	257	6
RL	GR10	T	28	262	7
RL	GR10	T	29	270	10
RL	GR10	T	30	277	14

Appendix F-3. Science Raw to Scale Score Conversion

Subject	Grade	Strand	Raw Score	Scale Score	SEM
SC	GR5	T	0	173	14
SC	GR5	T	1	180	10
SC	GR5	T	2	188	7
SC	GR5	T	3	193	6
SC	GR5	T	4	196	6
SC	GR5	T	5	199	5
SC	GR5	T	6	201	5
SC	GR5	T	7	204	5
SC	GR5	T	8	206	4
SC	GR5	T	9	208	4
SC	GR5	T	10	209	4
SC	GR5	T	11	211	4
SC	GR5	T	12	213	4
SC	GR5	T	13	215	4
SC	GR5	T	14	216	4
SC	GR5	T	15	218	4
SC	GR5	T	16	219	4
SC	GR5	T	17	221	4
SC	GR5	T	18	223	4
SC	GR5	T	19	224	4
SC	GR5	T	20	226	4
SC	GR5	T	21	228	4
SC	GR5	T	22	230	4
SC	GR5	T	23	232	5
SC	GR5	T	24	234	5
SC	GR5	T	25	237	5
SC	GR5	T	26	240	6
SC	GR5	T	27	243	6
SC	GR5	T	28	248	8
SC	GR5	T	29	255	10
SC	GR5	T	30	263	14

Appendix F-3. Science Raw to Scale Score Conversion

Subject	Grade	Strand	Raw Score	Scale Score	SEM
SC	GR8	T	0	184	14
SC	GR8	T	1	191	10
SC	GR8	T	2	199	8
SC	GR8	T	3	204	6
SC	GR8	T	4	207	6
SC	GR8	T	5	210	5
SC	GR8	T	6	213	5
SC	GR8	T	7	215	5
SC	GR8	T	8	217	4
SC	GR8	T	9	219	4
SC	GR8	T	10	221	4
SC	GR8	T	11	222	4
SC	GR8	T	12	224	4
SC	GR8	T	13	226	4
SC	GR8	T	14	227	4
SC	GR8	T	15	229	4
SC	GR8	T	16	230	4
SC	GR8	T	17	232	4
SC	GR8	T	18	234	4
SC	GR8	T	19	235	4
SC	GR8	T	20	237	4
SC	GR8	T	21	239	4
SC	GR8	T	22	240	4
SC	GR8	T	23	242	5
SC	GR8	T	24	245	5
SC	GR8	T	25	247	5
SC	GR8	T	26	250	6
SC	GR8	T	27	253	6
SC	GR8	T	28	258	7
SC	GR8	T	29	265	10
SC	GR8	T	30	273	14

Appendix F-3. Science Raw to Scale Score Conversion

Subject	Grade	Strand	Raw Score	Scale Score	SEM
SC	GR10	T	0	191	14
SC	GR10	T	1	198	10
SC	GR10	T	2	206	7
SC	GR10	T	3	210	6
SC	GR10	T	4	214	6
SC	GR10	T	5	217	5
SC	GR10	T	6	219	5
SC	GR10	T	7	221	5
SC	GR10	T	8	223	4
SC	GR10	T	9	225	4
SC	GR10	T	10	227	4
SC	GR10	T	11	228	4
SC	GR10	T	12	230	4
SC	GR10	T	13	231	4
SC	GR10	T	14	233	4
SC	GR10	T	15	234	4
SC	GR10	T	16	236	4
SC	GR10	T	17	238	4
SC	GR10	T	18	239	4
SC	GR10	T	19	241	4
SC	GR10	T	20	243	4
SC	GR10	T	21	244	4
SC	GR10	T	22	246	4
SC	GR10	T	23	248	5
SC	GR10	T	24	251	5
SC	GR10	T	25	253	5
SC	GR10	T	26	256	6
SC	GR10	T	27	260	6
SC	GR10	T	28	265	8
SC	GR10	T	29	272	10
SC	GR10	T	30	280	14

## References

- Choi, S. W. and Tinkler, T. V. (2002). Evaluating comparability of paper and pencil and computer based assessment in a K-12 setting. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA.
- Doran, H., and Cohen, J. (2006) Oregon Technical Report I. Technical Report for the Oregon Department of Education. Washington DC: American Institutes for Research.
- Kingsbury, G. G. and Hauser, C. (2007). Individual validity in the context of an adaptive test. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- Kolen, M. J. and Brennan, R. L. (2004) *Test equating, scaling, and linking: Methods and practices*. (2<sup>nd</sup> Ed.). New York: Springer-Verlag.
- McCall, M.S. (2005). Technical review of the Oregon Department of Education assessment system. Northwest Evaluation Association.
- Nunnally, J. C. and Bernstein, I. (1994). *Psychometric Theory*, 3<sup>rd</sup> Ed. McGraw-Hill.
- Oregon Department of Education (2007). Oregon's Statewide Assessment Technical Report, Volume 4: Reliability and Validity. Salem, OR: Author.
- Wise, S. L., Bhola, D. S., and Yang, S. (2006). Taking the time to improve the validity of low-stakes tests: The effort-moderating CBT. *Educational measurement: Issues and practice*, 25, 21-30.

## **Appendix A**

### **Letter of Invitation to Participate in Study**

## **Appendix B**

### **Instructions to Participating Schools**



## **Appendix C**

### **Number of Items by Strand**

## **Appendix D**

### **Winsteps Control File (Sample)**

## Appendix E

### Winsteps Output (Sample)

## **Appendix F**

### **Raw to Scale Score Tables for Paper Tests**