

2008 TELPAS Reading Vertical Scaling Study Report

August 2008

*Prepared for the Texas Education Agency
by
Pearson
Psychometric Services*

Table of Contents

Background..... 3

Data Collection Design –TELPAS Reading Vertical Linking 3

TELPAS Reading Vertical Scaling Method 5

 Main Steps to Developing the TELPAS Reading Vertical Scales..... 9

 1. Calibration of Vertical Linking Items..... 9

 2. Diagnostic Information and Item Level Decisions..... 10

 3. Finding Vertical Linking Constants for Adjacent Grade Clusters. 11

 4. Defining the Base Grade Cluster..... 12

 5. Computation of the Final Vertical Linking Constant..... 12

 6. Selection of the Final Vertical Equating Constant 13

 7. Customization of Vertical Scale and Smoothing..... 15

Conclusions..... 17

 Appendix A 1. TELPAS Grades 2 and 3 Online and Paper Sample
 Percents..... 19

 Appendix A 2. TELPAS Grade Clusters 4–5 and 6–7 Online and Paper
 Sample Percents..... 20

 Appendix A 3. TELPAS Grade Clusters 8–9 and 10–12 Online and Paper
 Sample Percents..... 21

Background

In spring 2008 the Texas Education Agency (TEA) implemented the second version of the federally required Texas English Language Proficiency Assessment System (TELPAS) reading assessment. A vertical scale was developed for this assessment to facilitate the tracking of student growth in English reading acquisition from year to year. For this purpose, a vertical scaling study was conducted during the 2008 test administration. This report summarizes the study findings.

English language learners (ELLs) are assessed with TELPAS annually. In addition, unless ELLs are exempted for their first year in the United States, they take the TAKS tests in English or Spanish annually as well. ELLs are assessed with TELPAS until they are no longer considered to be limited English proficient based on meeting state-mandated exit criteria as determined by their language proficiency assessment committee (LPAC). The expectation for progress on TELPAS is that a student will show an increase of at least one proficiency level per year.

In spring 2008 the TELPAS reading test for grades 2–12 was administered online and on paper. Approximately two-thirds of the ELL population participated in the online administration and one-third took the test on paper. With few exceptions, future administrations of the TELPAS reading test will be online.

For study analyses, the online mode was considered the primary administration of the test because more than half of the population tested online during spring 2008 and because future administrations will be online. To conduct the study, vertical linking items were embedded into field-test positions for both the online and paper-pencil modes. Therefore the study resulted in a vertical scale for each testing mode: online and paper-pencil. The vertical scaling results were used during the last step of TELPAS reading proficiency level setting activities (in August 2008) to facilitate the evaluation of cut scores across adjacent grade clusters.

Data Collection Design –TELPAS Reading Vertical Linking

The data collection design for the TELPAS reading vertical scaling study used the embedded field-test positions in the operational 2008 TELPAS reading administrations. Using embedded field-test positions rather than collecting data through a separate vertical linking study was desirable because

students would have no knowledge of whether an item was an operational or vertical scaling anchor item. As a consequence, the students were assumed to have put the same effort into responding to the vertical linking items as the live items.

The TELPAS reading assessments for grades 2–12 consist of six grade-cluster tests. The data collection design was based on a common-item nonequivalent group design with dual grade-cluster common items. This design required designating either 4 or 8 of the regular field-test forms for the vertical scaling study. The data collection design is depicted in Figure 1. The design required using:

- (a) only previously field-tested items as anchor items
- (b) both lower grade-cluster items and upper grade-cluster items
- (c) embedded field-test positions

For example, some of the previously field-tested grade 3 items from the TELPAS reading item bank were used as anchor items and placed on both the grade 3 and 4–5 vertical linking forms. Likewise, some of the previously field-tested grade-cluster 4–5 items from the item bank were used as anchor items on both grade cluster 4–5 and grade 3 vertical linking forms. The number of vertical linking forms needed for this model is 8 for grades 3 through 8-9. Because the anchor items were only from adjacent grade clusters, grades 2 and 10–12 only had 4 forms each (see Figure 1).

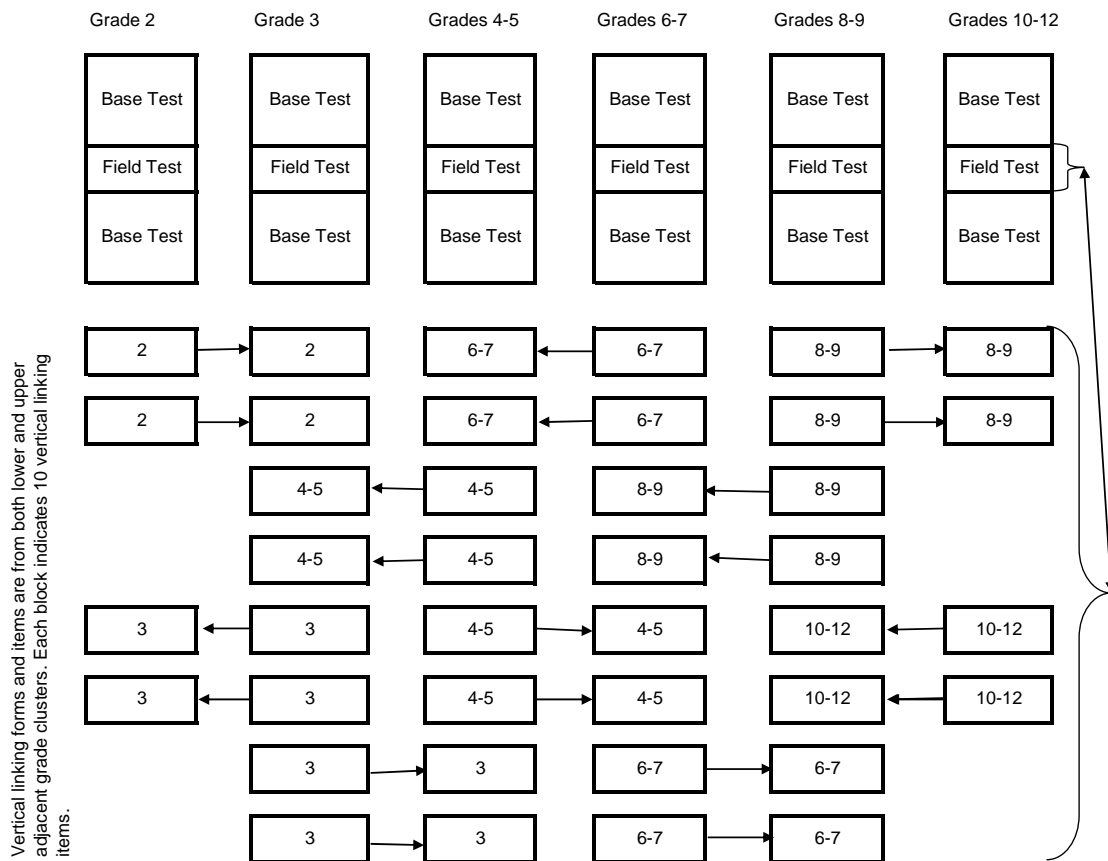


Figure 1. TELPAS Reading Vertical Scale Data Collection Design

TELPAS Reading Vertical Scaling Method

The TELPAS paper sample and online sample had similar demographic characteristics. No systematic differences between online and paper samples were observed (see Appendix A.1–A.3 for more information on the sample).

TELPAS Reading Vertical Scale Anchor Forms Descriptive Statistics

At least 10 items were allocated to vertical linking anchor items in a given form and other positions were allocated for field testing. The items used as anchor items were placed in the same or similar positions in adjacent grade clusters/levels. Both online and paper-pencil modes used the same data collection structure. Generally more than 2000 students were administered each form of the vertical scale linking forms. Tables 1–6 provide descriptive statistics for the vertical linking forms for both the online and paper modes. The statistics reflect raw scores from the operational assessment. The tables show that students at grades 2, 3, and grade cluster 4–5 had slightly higher mean raw scores on paper than online. In other grades, the online mode had slightly higher mean raw scores than the paper mode. Online comparability studies have been conducted for all grade levels.

GRADE	MODE	FORM	FREQS	MEANS	STD_DEV	MINIMUM	MAXIMUM
2	ONLINE	0	11893	31.607	9.932	1	49
2	ONLINE	51	2905	31.145	9.862	1	49
2	ONLINE	52	3010	31.707	10.058	1	49
2	ONLINE	53	3005	31.508	10.022	2	49
2	ONLINE	54	2973	32.059	9.762	1	49
2	PAPER	0	39251	32.912	10.074	2	49
2	PAPER	1	9732	32.868	10.044	2	49
2	PAPER	2	9815	33.002	10.165	2	49
2	PAPER	3	9880	32.822	10.017	2	49
2	PAPER	4	9824	32.955	10.068	3	49
FORM '0' refers to all vertical linking forms.							

2008 TELPAS Reading Vertical Scaling Study Report

Table 2: TELPAS READING VERTICAL SCALE FORMS DESCRIPTIVE STATISTICS FOR GRADE 3

GRADE	MODE	FORM	FREQS	MEANS	STD_DEV	MINIMUM	MAXIMUM
3	ONLINE	0	22126	41.13	11.318	1	58
3	ONLINE	51	2781	40.575	11.542	3	58
3	ONLINE	52	2714	41.295	11.435	8	58
3	ONLINE	53	2784	41.264	11.181	9	58
3	ONLINE	54	2745	41.11	11.232	4	58
3	ONLINE	55	2789	41.244	11.149	1	58
3	ONLINE	56	2699	40.942	11.486	1	58
3	ONLINE	57	2778	41.351	11.222	2	58
3	ONLINE	58	2836	41.257	11.291	7	58
3	PAPER	0	33480	42.546	11.225	1	58
3	PAPER	1	4147	42.328	11.405	5	58
3	PAPER	2	4238	42.352	11.434	7	58
3	PAPER	3	4264	42.618	11.079	9	58
3	PAPER	4	4176	42.887	11.026	1	58
3	PAPER	5	4086	42.953	11.094	4	58
3	PAPER	6	4198	42.61	11.136	6	58
3	PAPER	7	4155	42.283	11.307	4	58
3	PAPER	8	4216	42.349	11.3	5	58

FORM '0' refers to all vertical linking forms.

Table 3: TELPAS READING VERTICAL SCALE FORMS DESCRIPTIVE STATISTICS FOR GRADE CLUSTER 4-5

GRADE CLUSTER	MODE	FORM	FREQS	MEANS	STD_DEV	MINIMUM	MAXIMUM
4-5	ONLINE	0	31537	45.004	11.249	1	61
4-5	ONLINE	51	4561	44.992	11.377	3	61
4-5	ONLINE	52	4599	45.006	11.18	1	61
4-5	ONLINE	53	4479	44.927	11.195	8	61
4-5	ONLINE	54	4574	44.832	11.327	6	61
4-5	ONLINE	55	4453	45.064	11.13	7	61
4-5	ONLINE	56	4428	45.193	11.262	10	61

2008 TELPAS Reading Vertical Scaling Study Report

Table 3: TELPAS READING VERTICAL SCALE FORMS DESCRIPTIVE STATISTICS FOR GRADE CLUSTER 4-5

GRADE CLUSTER	MODE	FORM	FREQS	MEANS	STD_DEV	MINIMUM	MAXIMUM
4-5	ONLINE	57	4443	45.023	11.269	8	61
4-5	PAPER	0	35360	45.48	11.529	1	61
4-5	PAPER	1	4534	45.417	11.574	4	61
4-5	PAPER	2	4417	45.97	11.187	2	61
4-5	PAPER	3	4401	45.398	11.534	3	61
4-5	PAPER	4	4404	45.191	11.643	2	61
4-5	PAPER	5	4392	45.456	11.53	3	61
4-5	PAPER	6	4389	45.636	11.691	3	61
4-5	PAPER	7	4413	45.502	11.502	1	61
4-5	PAPER	8	4410	45.272	11.559	5	61

FORM '0' refers to all vertical linking forms.

Table 4: TELPAS READING VERTICAL SCALE FORMS DESCRIPTIVE STATISTICS FOR GRADE CLUSTER 6-7

GRADE CLUSTER	MODE	FORM	FREQS	MEANS	STD_DEV	MINIMUM	MAXIMUM
6-7	ONLINE	0	30048	46.985	11.034	1	63
6-7	ONLINE	52	3807	46.887	11.125	10	63
6-7	ONLINE	53	3840	46.647	11.035	6	63
6-7	ONLINE	54	3704	46.931	11.205	9	63
6-7	ONLINE	55	3651	47.184	10.909	2	63
6-7	ONLINE	56	3721	47.012	11.071	2	63
6-7	ONLINE	57	3747	47.153	11.05	6	63
6-7	ONLINE	58	3804	46.841	11.121	9	63
6-7	ONLINE	59	3774	47.24	10.742	1	63
6-7	PAPER	0	21227	46.256	11.258	2	63
6-7	PAPER	1	2697	46.063	11.329	2	63
6-7	PAPER	2	2673	46.151	11.284	8	63
6-7	PAPER	3	2643	46.158	11.223	4	63
6-7	PAPER	4	2648	46.463	11.212	7	63
6-7	PAPER	5	2624	46.208	11.309	8	63
6-7	PAPER	6	2657	46.554	11.304	5	63

2008 TELPAS Reading Vertical Scaling Study Report

Table 4: TELPAS READING VERTICAL SCALE FORMS DESCRIPTIVE STATISTICS FOR GRADE CLUSTER 6-7

GRADE CLUSTER	MODE	FORM	FREQS	MEANS	STD_DEV	MINIMUM	MAXIMUM
6-7	PAPER	7	2645	46.169	11.285	8	63
6-7	PAPER	8	2640	46.284	11.125	2	63
FORM '0' refers to all vertical linking forms.							

Table 5: TELPAS READING VERTICAL SCALE FORMS DESCRIPTIVE STATISTICS FOR GRADE CLUSTER 8-9

GRADE CLUSTER	MODE	FORM	FREQS	MEANS	STD_DEV	MINIMUM	MAXIMUM
8-9	ONLINE	0	16415	44.466	12.093	1	63
8-9	ONLINE	51	2035	44.032	12.218	1	63
8-9	ONLINE	52	2050	44.861	11.938	9	63
8-9	ONLINE	53	2080	44.276	12.175	10	63
8-9	ONLINE	54	2046	44.544	12.026	7	63
8-9	ONLINE	55	1981	44.504	12.097	7	63
8-9	ONLINE	56	2074	44.833	12.017	4	63
8-9	ONLINE	57	2099	44.222	12.217	2	63
8-9	ONLINE	58	2050	44.458	12.046	5	63
8-9	PAPER	0	17555	44.727	12.423	1	63
8-9	PAPER	1	2209	44.685	12.41	9	63
8-9	PAPER	2	2187	45.021	12.141	8	63
8-9	PAPER	3	2184	44.849	12.376	8	63
8-9	PAPER	4	2211	44.455	12.436	5	63
8-9	PAPER	5	2189	45.311	12.452	8	63
8-9	PAPER	6	2187	44.513	12.423	1	63
8-9	PAPER	7	2180	44.446	12.856	7	63
8-9	PAPER	8	2208	44.543	12.271	5	63
FORM '0' refers to all vertical linking forms.							

Table 6: TELPAS READING VERTICAL SCALE FORMS DESCRIPTIVE STATISTICS FOR GRADE CLUSTER 10–12

GRADE CLUSTER	MODE	FORM	FREQS	MEANS	STD_DEV	MINIMUM	MAXIMUM
10–12	ONLINE	0	9262	46.678	10.909	6	64
10–12	ONLINE	51	2295	46.493	10.862	10	64
10–12	ONLINE	52	2285	46.808	10.897	9	64
10–12	ONLINE	53	2298	46.668	10.917	9	64
10–12	ONLINE	54	2384	46.741	10.963	6	64
10–12	PAPER	0	14384	46.883	11.011	1	64
10–12	PAPER	1	3639	46.793	10.827	9	64
10–12	PAPER	2	3587	46.859	11.101	1	64
10–12	PAPER	3	3584	46.861	11.087	11	64
10–12	PAPER	4	3574	47.021	11.035	3	64

FORM '0' refers to all vertical linking forms.

Main Steps to Developing the TELPAS Reading Vertical Scales

1. Calibration of Vertical Linking Items

To develop a common scale across grade clusters, vertical linking items were calibrated on their grade level and off grade clusters with the live items. Based on feedback from the Texas Technical Advisory Committee (TTAC), within-grade cluster calibration of anchor items and live items was done with separate (rather than concurrent) calibrations.

A two-step calibration procedure was used similar to the typical embedded field-test calibrations in Texas. Using this method allowed us to define a scale with only on-grade-cluster items. The effect of off-grade-cluster items on the on-grade-cluster scale was controlled by conducting the equating with the two-step procedure. The first step was to calibrate on-grade-cluster items together using all available students. In the second step, the researcher calibrated both on-grade-level items and off-grade-level items together and computed an equating constant by computing the mean difference in on-grade-level item parameters between the first step calibration and the second step calibration. This mean difference, the mean/mean equating constant (Kolen & Brennan, 2004), was applied to all item difficulties. This calibration method was done form by form for all vertical linking forms.

2. Diagnostic Information and Item Level Decisions

Some of the item statistics were used as an indicator of model misfit. One indicator used was Rasch Mean Square Infits. If the Rasch Mean Square Infits were unexpectedly large or small, then the item was eliminated from the vertical linking set at grade level and both adjacent grade levels. There were only a few items that fell into these categories (21 items in both online and paper modes, see Table 7 for more information), which had either low (<.80) or high (>1.20) mean square infits.

Table 7. Eliminated items due to being outliers and misfit for each pair of adjacent grade/grade cluster

Grades	Online Mode		Paper Mode	
	# Outliers	# Misfit	# Outliers	# Misfit
02-03	1	3	1	4
03-04	0	6	0	7
04-06	0	1	0	1
06-08	0	6	1	3
08-10	0	5	0	6
Total	1	21	2	21

TELPAS Online Vertical Linking Items

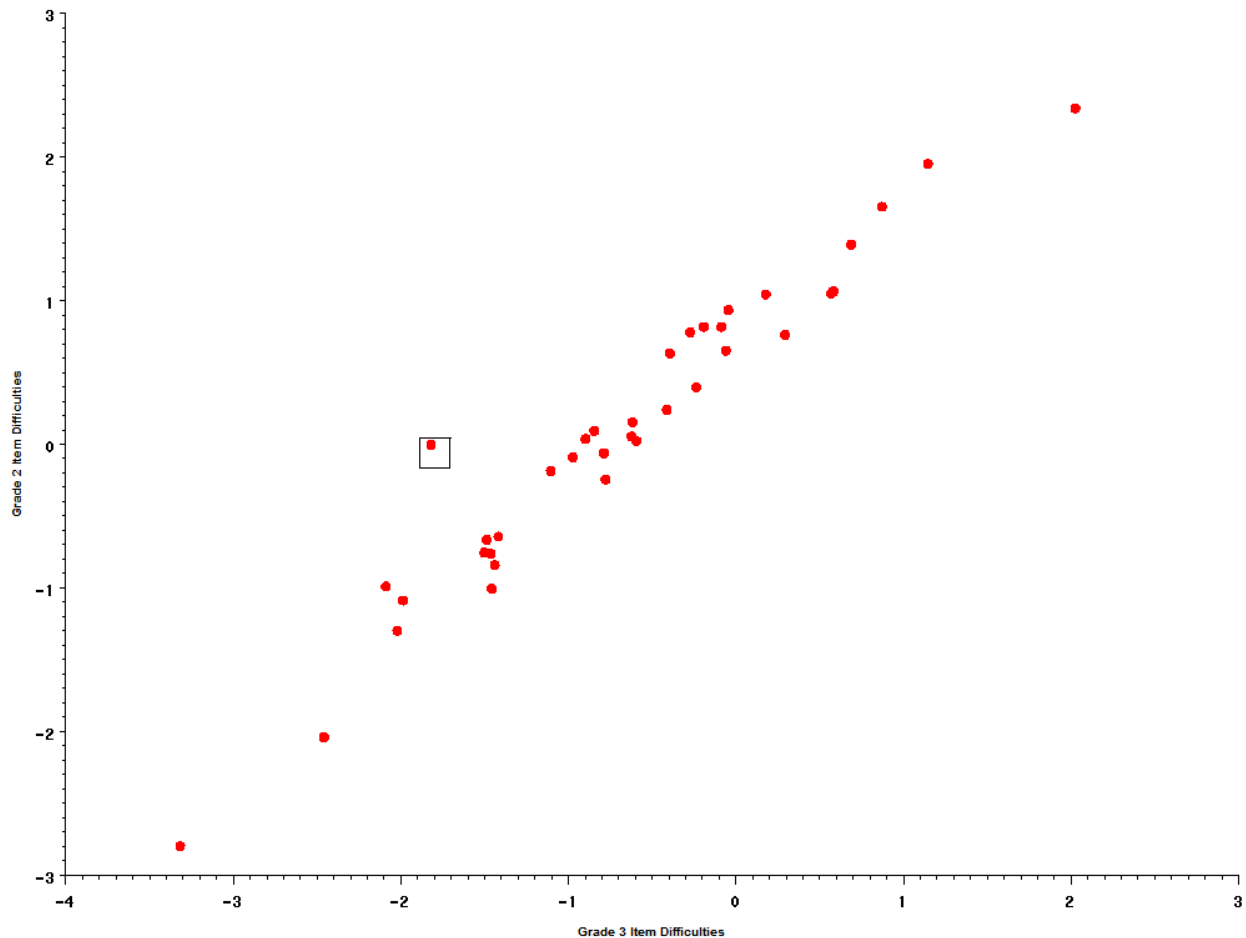


Figure 2. TELPAS reading online item difficulties for grade 2 and grade 3

Another indicator was outliers. The Rasch item difficulties for the vertical linking items from on-grade and off-grade clusters were plotted against each other, and some of the outliers were eliminated (1 item total for the online mode and 2 items for paper mode) from the computation of the vertical equating constant. The online outlier is indicated in Figure 2 with a square around the red dot.

3. Finding Vertical Linking Constants for Adjacent Grade Clusters

The Mean/Mean method (Kolen & Brennan, 2004) was used to compute vertical linking constants between the adjacent grade clusters. Vertical linking constants were computed in 3 ways in each of the 2 modes:

- 1) Online (see Table 8 for the online constants)
 - a. using only the lower grade cluster anchor items,
 - b. using only the upper grade cluster anchor items, and
 - c. using the combined anchors set (both lower and upper grade cluster anchor items).

Table 8. TELPAS reading online vertical linking constants by anchor item types

Adjacent Grade Clusters	Equating Constant Using All Items	Equating Constant Using Only Lower Grade items	Equating Constant Using Only Upper Grade items
0203	0.73978	0.74218	0.73711
0304	0.42906	0.49047	0.39836
0406	0.07908	0.05459	0.10485
0608	0.36819	0.38257	0.34999
0810	0.28554	0.27087	0.30106

- 2) Paper (see Table 9 for the online constants)
 - a. using only the lower grade cluster anchor items,
 - b. using only the upper grade cluster anchor items, and
 - c. using the combined anchors set (both lower and upper grade cluster anchor items).

Table 9. TELPAS reading paper vertical linking constants by anchor item types

Adjacent Grade Clusters	Equating Constant Using All Items	Equating Constant Using Only Lower Grade items	Equating Constant Using Only Upper Grade items
0203	0.83388	0.85742	0.80895
0304	0.35077	0.40246	0.30213
0406	0.07893	0.02444	0.13070

2008 TELPAS Reading Vertical Scaling Study Report

Adjacent Grade Clusters	Equating Constant Using All Items	Equating Constant Using Only Lower Grade items	Equating Constant Using Only Upper Grade items
0608	0.30774	0.36140	0.24775
0810	0.35871	0.31904	0.40582

Estimating vertical linking constants in three different ways and in two modes provided options to create the vertical scale in different ways; however, there did not appear to be meaningful differences across the modes and sources of the linking items (either from the lower grades or higher grades). Both testing modes produced comparable adjacent grade level vertical equating constants. Also, the direction of the vertical linking items had no major differences among adjacent vertical linking constants (see Table 8 and 9). Since the final linking constants (section 5 below) would be calculated from these interim values, it would be reasonable to believe that the final constants would be similar between the testing modes and therefore one vertical scale would suffice.

4. Defining the Base Grade Cluster

Pearson decided to use the lowest grade cluster as the base grade for the vertical scale. The base grade refers to the anchoring point of the vertical scale. Therefore, the base grades for the TELPAS reading vertical scale was grade 2.

5. Computation of the Final Vertical Linking Constant

As indicated earlier, the Mean/Mean equating procedure was used to find the linking constants between adjacent grade clusters. After finding the vertical linking constant between adjacent grade clusters, a cumulative linking constant was defined from the base grade cluster to the upper grade cluster for any grade cluster that was not adjacent to the base grade. For example, at grade 3, the *vertical linking constant* (between grades 2 and 3) was the difference between the mean vertical linking item difficulties for grade 2 and grade 3. On the other hand, the *final vertical linking constant* (between grade cluster 4-5 and the base grade cluster, grade 2) was the vertical linking constant between grades 2 and 3 *plus* the vertical linking constant between grade clusters 3 and 4-5 (see Table 10 below for the formulas).

Table 10. Final Vertical Linking Constants for TELPAS Reading

Grade/ Grade Cluster	Vertical Linking Constant	Final Vertical Linking Constant
2	0	0
3	$VC_{23} = VL_3 - VL_2$	$0 + VC_{23}$
4-5	$VC_{34} = VL_4 - VL_3$	$VC_{23} + VC_{34}$

2008 TELPAS Reading Vertical Scaling Study Report

6-7	$VC_{46} = VL_6 - VL_4$	$VC_{23} + VC_{34} + VC_{46}$
8-9	$VC_{68} = VL_8 - VL_6$	$VC_{23} + VC_{34} + VC_{46} + VC_{68}$
10-12	$VC_{810} = VL_{10} - VL_8$	$VC_{23} + VC_{34} + VC_{46} + VC_{68} + VC_{810}$

The final vertical linking constant for TELPAS reading was computed as described above. Recall that grade clusters were used in place of grades in the upper grade levels. The final vertical equating constants are shown in Table 11 for the online mode and in Table 12 for the paper mode. Final paper and online vertical equating constants were very similar. One to one comparison of the vertical scale is difficult but helpful here because those vertical scaling constants did not consider actual testing mode effects.

Table 11. Final Vertical Linking Constants for TELPAS Online Reading

Adjacent Grade Clusters	Final Equating Constant Using All Items	Final Equating Constant Using Only Lower Grade items	Final Equating Constant Using Only Upper Grade items
0203	0.73978	0.74218	0.73711
0304	1.16885	1.23265	1.13547
0406	1.24793	1.28724	1.24032
0608	1.61612	1.66981	1.59031
0810	1.90165	1.94068	1.89137

Table 12. Final Vertical Linking Constants for TELPAS Paper Reading

Adjacent Grade Clusters	Final Equating Constant Using All Items	Final Equating Constant Using Only Lower Grade items	Final Equating Constant Using Only Upper Grade items
0203	0.83388	0.85742	0.80895
0304	1.18466	1.25989	1.11108
0406	1.26359	1.28433	1.24178
0608	1.57132	1.64573	1.48953
0810	1.93003	1.96477	1.89536

6. Selection of the Final Vertical Equating Constant

The final vertical equating constants were plotted in Figure 3 (Online) and Figure 4 (Paper). The online final vertical scaling constants were less variable than their paper counterparts. Although all three methods produce similar results, the final vertical equating constant computed from all items is the best practice because the developed vertical scale will have a more stable vertical linking constant.

Final equating constants for the online mode were chosen to serve as the vertical scale for both online and paper modes since there appeared to be no major differences between the two sets of constants, and also that in the future the TELPAS reading assessment will be administered exclusively online. The application of the final vertical linking constant would take place after horizontal equating is done for the operational items. The same final vertical linking constants will be used for future TELPAS reading administrations.

TELPAS Online Final Vertical Equating Constant

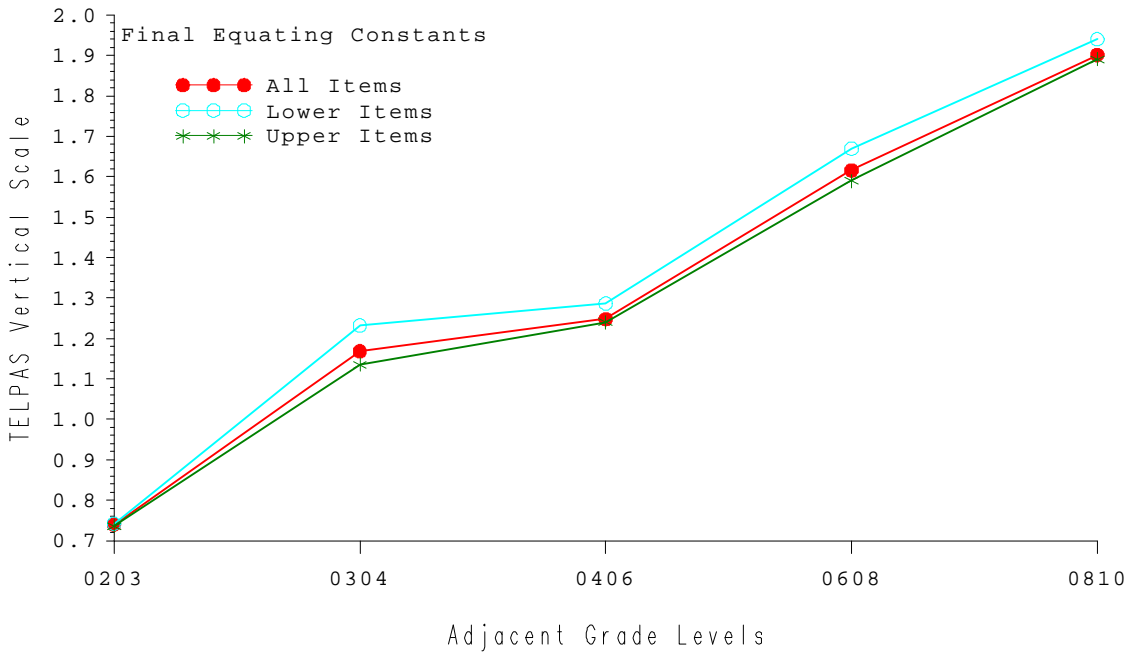


Figure 3. TELPAS reading online final vertical equating constants

TELPAS Paper Final Vertical Equating Constant

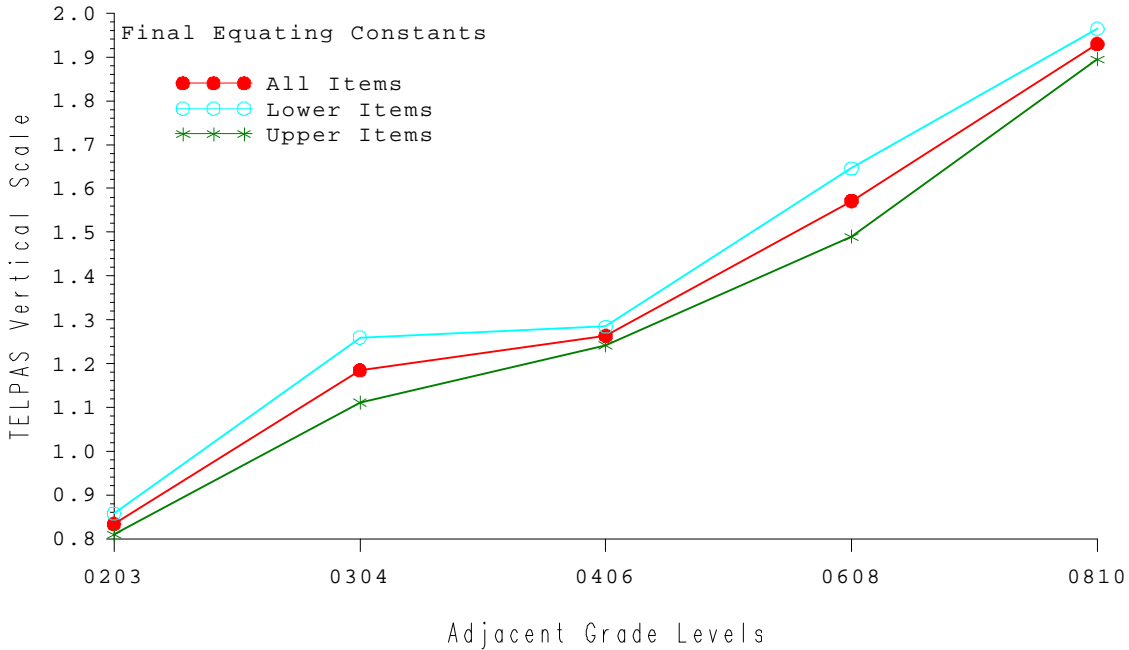


Figure 4. TELPAS reading paper final vertical equating constants

7. Customization of Vertical Scale and Smoothing

It might be necessary to place the newly developed vertical scale on another scale with properties that are easier to communicate to stakeholders, such as parents, teachers and school administrators.

The trends of TELPAS reading final cut scores were mapped for all grade levels and shown in Figure 5. Overall, the ELL English learning growth patterns were acceptable. The final cut scores demonstrate a continuing growth pattern from one grade level to another grade level. The slope of the growth pattern was less steep between the 4-5 and 6-7 grade clusters. A similar growth pattern was observed in the development of a vertical scale for the previous assessment, the Reading Proficiency Tests in English (RPTE) in 2000.

TELPAS Reading Vertical Scale Trends with the Final Cuts

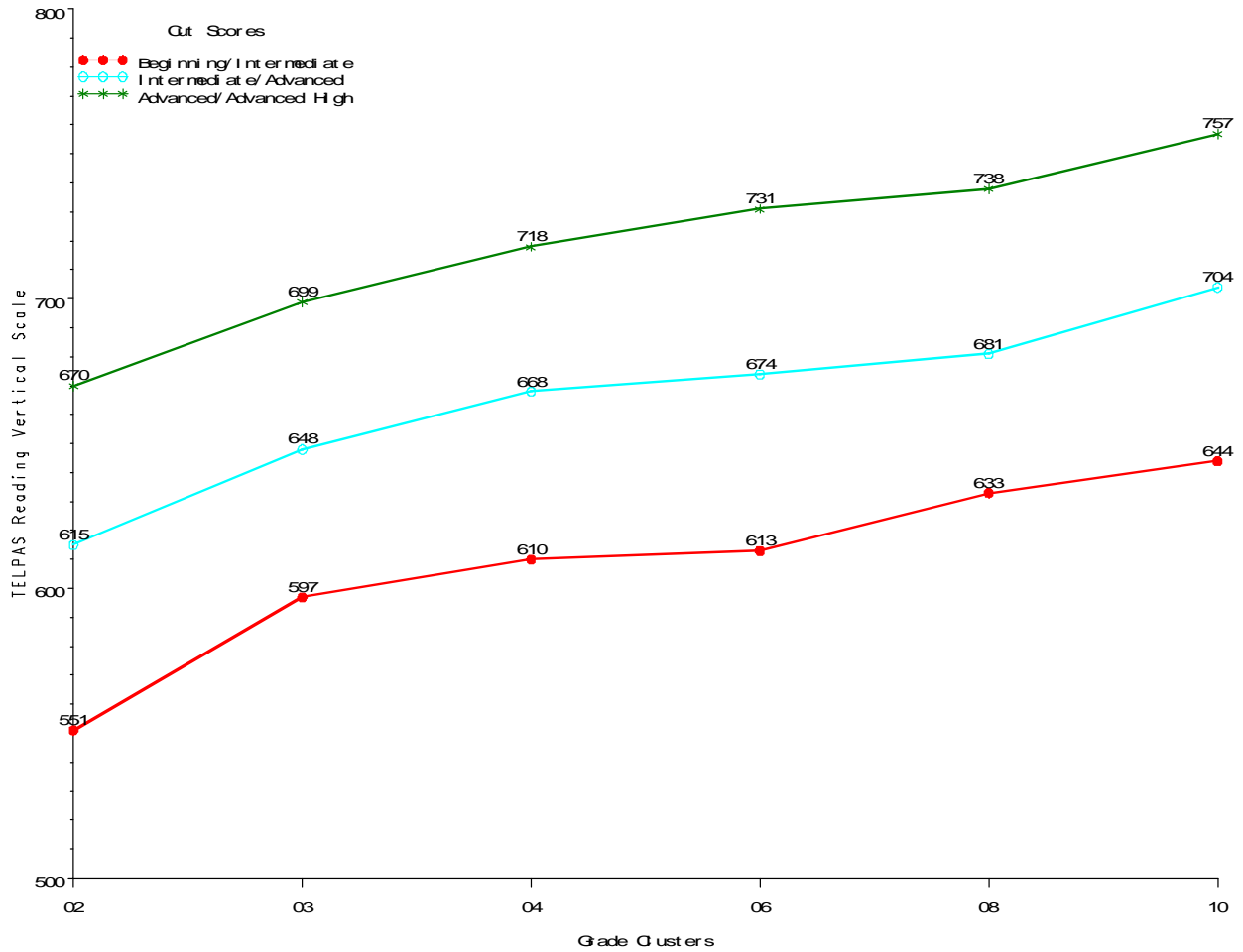


Figure 5. TELPAS online final cut score trends across grade clusters

The final TELPAS reading vertical equating constants were applied to the 2008 TELPAS online data. The change in the scale score to percentile score relationship was analyzed across grades (see Figure 6). The current vertical scale clearly separates grade 2 students from grade 3, grade 3 students from grade 4-5, and grade 8-9 students from grade 10-12. The differences between the 4-5 and 6-7 grade clusters are less clear. In other words, the 50th percentile rank corresponds different ability scores for grades 2, 3, and 10-12 (see figure 6). However grades 4-5, 6-7, and 8-9 the 50th percentile rank corresponds to similar ability scores.

TELPAS Online Vertical Scale Cumulative Ability Distribution

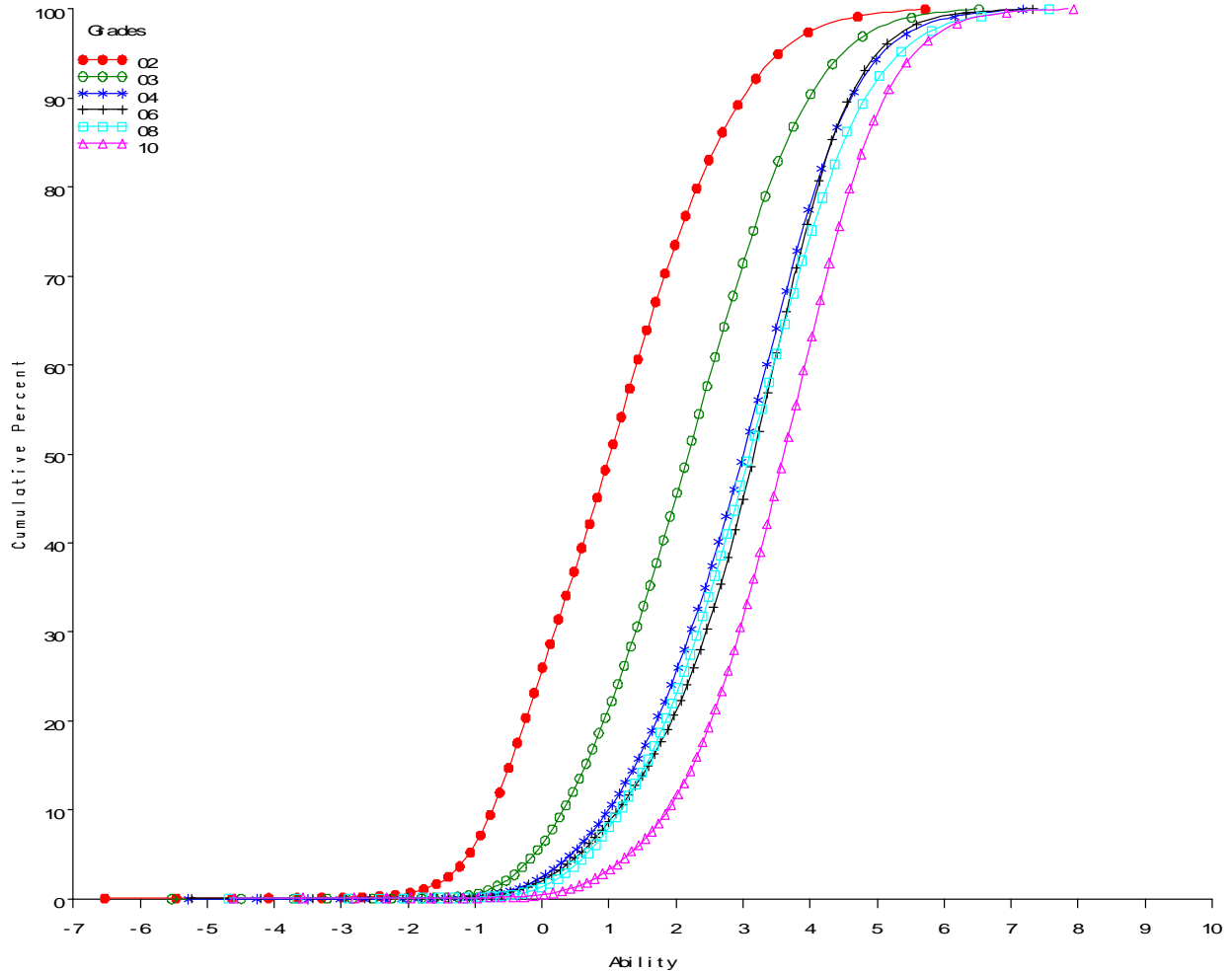


Figure 6. TELPAS reading online cumulative percent score distributions (note that the 'Grades' in the legend represents the grade level for grades 2 and 3, and represents the lowest grade level for the other grade clusters.)

Conclusions

The 2008 TELPAS reading vertical study resulted in a reasonable scale that can be used to monitor students' English language proficiency development as they receive increased amounts of English language instruction in Texas. The increase in difficulty of the cut points distinguishing proficiency levels across grade clusters was greater in the lower grades and grade clusters than in the upper grade clusters. The increase from grade 2 to grade 3 or from grade 2 to grade cluster 4-5 is generally greater than the increase from grade cluster 4-5 to the upper grade clusters. The steeper increase in the lower grades and grade clusters makes sense for several reasons. First, younger students tend to learn a non-native language faster than older students, so the learning curve tends to be steeper. Second, the student

population taking TELPAS reading changes across the grade clusters. Specifically, the proportion of students who have five or more years in special language programs and therefore must take TELPAS reading is greater in the upper grade clusters than in the lower grade clusters. Students with more than five years in the programs tend to be students who have more difficulty learning a new language and succeeding in school. The performance of these lower-performing students may have contributed to the less steep pattern in the vertical scale in the upper grade clusters.

The increase in vertical scaling constants between grade clusters was smallest between the scales at grade clusters 4-5 and 6-7. This result was also found in the vertical scale study for the RPTE and likely reflects the population change over these grade clusters.

Evidence supporting findings from the TELPAS reading vertical scaling study came from the consistency in results across the three sets of analyses in the current study and from consistency in findings from the current study and the study completed on the RPTE assessment. Results were similar across analyses of data from the different study designs (i.e., items taken by students higher and lower grade cluster versus items taken by students only in higher grade cluster) and from the different modes of testing (i.e., online and paper). As indicated in the methods section, two different sets of anchor items were used to link all grade clusters such that items were used from the higher adjacent and lower adjacent grade clusters. Both upper and lower grade items produced very consistent growth patterns as grade cluster increased. The second consistent finding was that the mode of testing (online versus paper-pencil) also produced comparable growth patterns. The last consistency came from the comparison of the previous vertical scale developed in 2000 with the current one. In both cases there was large growth at the elementary grades and less growth between middle and high school grades. This is indicated by the grade-to-grade equating constants shown on tables 8 and 9. The magnitudes for growth (represented by the values of the equating constants) are greater at the lower grade/grade cluster (from grade 2 to grade 3 and from grade 3 to grade cluster 4-5) but lower at the other upper grade clusters.

References;

Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices* (2nd ed.). New York: Springer-Verlag.

Appendix A 1. TELPAS Grades 2 and 3 Online and Paper Sample Percents.

Grade 2

Variable	Value	Online (%)	Paper (%)
Gender	Female	48.52	48.64
	Male	51.42	51.21
Migrant		1.4	2.62
Bilingual		66.11	73.67
Years in US Schools	0	0.45	0.6
	1	2.53	2.96
	2	85.28	83.2
	3	9.74	11.41
	4	0.59	0.96
	5	0.05	0.13
Listening Rating	1	9.72	10.94
	2	28.11	29.43
	3	34.09	34.33
	4	27.04	24.88
Speaking Rating	1	15.23	16.72
	2	30.65	31.73
	3	30.87	30.9
	4	22.2	20.22
Writing Rating	1	24.59	26.56
	2	35.55	36.61
	3	24.33	23.42
	4	14.31	12.71

Grade 3

Variable	Value	Online (%)	Paper (%)
Gender	Female	48.54	48.3
	Male	51.34	51.58
Migrant		1.67	2.72
Bilingual		64.85	71.95
Years in US Schools	0	0.52	0.47
	1	2.34	2.32
	2	3.59	4.08
	3	79.42	77.21
	4	12.16	13.7
	5	0.97	1.59
Listening Rating	1	6.21	6.82
	2	21.52	22.27
	3	35.35	35.24
	4	36.21	35.41
Speaking Rating	1	9.78	10.68
	2	25.83	26.3
	3	34.3	33.68
	4	29.37	29.07
Writing Rating	1	15.94	16.87
	2	34.92	34.84
	3	29.36	29.18
	4	18.88	18.59

Appendix A 2. TELPAS Grade Clusters 4-5 and 6-7 Online and Paper Sample Percents.

Grades 4-5

Variable	Value	Online (%)	Paper (%)
Gender	Female	47.67	47.14
	Male	52.26	52.8
Migrant		2.09	2.68
Bilingual		66.72	62.13
Years in US Schools	0	0.62	0.72
	1	3.09	3.43
	2	4.34	4.66
	3	5.32	5.56
	4	43.03	43.16
	5	42.35	41.66
Listening Rating	1	5.17	6.47
	2	16.35	18.08
	3	34.2	35.61
	4	43.32	39.54
Speaking Rating	1	7.84	9.11
	2	20.85	22.31
	3	34.86	35.6
	4	35.49	32.68
Writing Rating	1	10.98	12.64
	2	28.96	30.18
	3	33.65	34.29
	4	25.27	22.34

Grades 6-7

Variable	Value	Online (%)	Paper (%)
Gender	Female	45.61	44.98
	Male	54.33	54.97
Migrant		2.4	3.5
Bilingual		9.37	6.32
Years in US Schools	0	1.01	1.13
	1	5.31	6.06
	2	6.47	7.27
	3	7.06	7.6
	4	6.25	6.61
	5	71.96	70.56
Listening Rating	1	5.89	7.22
	2	18.77	21.08
	3	34.21	35.63
	4	39.39	35.27
Speaking Rating	1	7.99	9.86
	2	21.33	24.62
	3	34.68	34.88
	4	34.26	29.83
Writing Rating	1	9.46	11.92
	2	30.46	33.06
	3	34.75	34.32
	4	23.41	19.33

Appendix A 3. TELPAS Grade Clusters 8-9 and 10-12 Online and Paper Sample Percents.

Grades 8-9

Variable	Value	Online (%)	Paper (%)
Gender	Female	43.71	42.09
	Male	56.23	57.76
Migrant		2.59	3.53
Bilingual		0.28	0.14
Years in US Schools	0	1.88	2.62
	1	9.35	9.4
	2	9.79	9.91
	3	9.23	8.91
	4	7.1	7.3
	5	60.81	60.86
Listening Rating	1	8.17	8.71
	2	20.96	23.21
	3	33.87	34.18
	4	35.08	32.32
Speaking Rating	1	11.87	12.78
	2	23.31	26.31
	3	32.42	31.96
	4	30.46	27.35
Writing Rating	1	11.18	12.28
	2	31.5	33.34
	3	33.42	32.4
	4	21.5	19.73

Grades 10-12

Variable	Value	Online (%)	Paper (%)
Gender	Female	46.37	44.48
	Male	53.61	55.43
Migrant		2.51	4.08
Bilingual		0.04	0.01
Years in US Schools	0	0.88	0.98
	1	3.96	2.92
	2	9.26	8.29
	3	11.86	11.88
	4	11.57	10.68
	5	60.16	64.51
Listening Rating	1	3.34	3.3
	2	17.25	18.67
	3	36.83	37.3
	4	39.77	39.71
Speaking Rating	1	5.56	6.08
	2	22.22	24.44
	3	35.57	36.02
	4	33.82	32.39
Writing Rating	1	4.43	5.3
	2	27.39	29.57
	3	38.05	38.83
	4	26.8	24.92