Results of the 2008 TELPAS Reading Paper and Online Comparability Analyses

December 15, 2008

Table of Contents

LIST OF TABLES, FIGURES AND APPENDICES

SUMMARY

As part of the process of being transitioned to an exclusively-online assessment in 2009, the Texas English Language Proficiency Assessment System (TELPAS) reading tests for grades 2-12 were offered in both online and paper modes in 2008.

TELPAS comparability studies were conducted to evaluate the mode effect using the 2008 data. Results from this comparability study were used to adjust for mode differences, and any such adjustments were applied such that the performance standards remained consistent across modes of test administration.

A matched-samples design was used to study the between-mode comparability between online and paper administrations. Consistent with the recommendations from the Texas Technical Advisory Committee (TTAC), matching variables included estimated student ability (based on composite scores calculated by linearly combining the previous and current TELPAS scores) and demographic information. Details about the composite scores and matching procedures are provided in the methodology section of the report.

In the matched-samples analyses, a bootstrap sampling approach was conducted to select online and paper samples where each selected student who took the assessment on paper was matched to a student who took the assessment online with the same estimated ability level and demographic variables. Once both the paper and online samples were selected, raw-score-to-scale-score conversions were calculated for each sample using the Rasch measurement model. The raw score equivalents (and the raw cuts at various proficiency levels) between online and paper were evaluated to determine the comparability between modes.

Results from comparability studies conducted for the Texas Assessment of Knowledge and Skills (TAKS) indicated that students with previous test scores have some fundamental differences compared with those who do not have previous test scores, and the two groups of students (i.e., the group of students who have previous test scores and the group of students who do not have previous test scores) should not be matched across groups. Therefore, a stratified sampling approach was used in the current studies such that students testing on paper with 2007 scores were matched only with students testing online who also had 2007 scores. Students testing on paper who did not have 2007 scores were matched only with students testing online who similarly did not have 2007 scores.

Once the matching was completed, the equivalency between online and paper raw scores was then evaluated based on psychometric and practical criteria. A separate conversion table for the paper administration was recommended if the mode difference was great enough to meet the specified psychometric and practical criteria. A separate conversion table achieves the goal of adjusting scores for one mode while fixing the scores on the other mode, therefore adjusting for the difference between modes.

As indicated in the following table, the results show a mode effect in all grade clusters and at all cuts except at grade cluster 10–12 in favor of the paper version. In other words, the raw score cut tended to be higher on the paper version of the test indicating that the test was easier when administered in the paper form. However, this mode effect is largest at grade 2, decreases for students at the higher grade levels, and disappears at the high school level. The results of the study show a mode-by-grade-level interaction where TELPAS appears to be easier in the paper mode for students at lower grades, but of equal difficulty across modes at the high school level.

Cut Scores by Mode and Grade/Grade Cluster

Grade Cluster	Beginning/Intermediate		Intermediate/Advanced		Advanced/Advanced High	
	Online	Paper	Online	Paper	Online	Paper
2	551 (19)	555 (22)	615 (30)	618 (33)	670 (39)	673 (41)
3	597 (25)	599 (28)	648 (36)	651 (39)	699 (46)	702 (48)
4–5	610 (25)	611 (27)	668 (37)	668 (39)	718 (47)	722 (49)
6–7	613 (26)	616 (27)	674 (39)	676 (40)	731 (50)	734 (51)
8–9	633 (26)	636 (28)	681 (37)	684 (39)	738 (49)	738 (50)
10–12	644 (26)	644 (26)	704 (39)	704 (39)	757 (50)	757 (50)

(Note that in the above table, numbers preceding parentheses are *attainable* scale score cuts, whereas numbers inside parentheses are the corresponding raw score cuts. This table also appears as Table 6 in the report.)

Additional analyses assessing mode differences at the item level and the gender, migrant status, and economically-disadvantaged group levels were performed. These analyses revealed that, for most grade clusters, more items exhibited a mode effect favoring the paper mode, and fewer items exhibited a mode effect favoring the online mode. The only exception was for grade cluster 10–12, where there was the same number of items in each category. Results from the group analysis indicated that across all groups, there was a mode effect favoring paper administrations.

Based on the findings, Pearson recommended using separate conversion tables for students who tested on paper at grades 2 and 3; and grade clusters 4–5, 6–7, and 8–9. However, Pearson recommended using the online conversion table for students who tested on paper at grade cluster 10–12.

INTRODUCTION

A new version of the Texas English Language Proficiency Assessment System (TELPAS) reading tests for grades 2-12 was implemented in spring 2008. The former version, the TELPAS Reading Proficiency Tests in English (RPTE) was a paper-and-pencil test. The new assessment was developed to be administered as an online testing program. The new assessment was offered in both online and paper modes in 2008 so that a comparability study could be conducted and districts could have time to prepare fully for the transition to online testing. Any detected mode effect was to be adjusted for using the new scale, which was established after the August 2008 proficiency level setting.

METHODOLOGY

In 2005, TEA and Pearson devised a 'matched samples comparability analysis' (MSCA) plan for comparability analyses of Texas Assessment of Knowledge and Skills (TAKS) administered in both paper and online modes (Way, Davis, & Fitzpatrick, 2006). This plan used a bootstrap approach in which students in the online group (the non-primary mode for TAKS) would be matched to students from the paper group (the primary mode for TAKS) on their previous test scores. This matching was necessary because the study design did not include randomly equivalent groups, common students across modes, or common items across modes.

The 2008 TELPAS reading comparability study used a similar strategy. However, one important difference between TELPAS and TAKS is that the TELPAS primary administration mode was online in 2008, whereas for TAKS it has always been paper. In addition, due to the different student populations and nature of the tests, it was necessary to consider the possibility of using a different set of matching variables, both for the student ability level and their demographic composition.

A preliminary study was conducted using the 2007 data from the paper administration to evaluate which matching variables to include. The result was presented to the Texas Technical Advisory Committee (TTAC) in March 2008 for review and discussion. Following the recommendations from the TTAC, the decision was made to include the 2008 TELPAS listening, speaking, and writing ratings and (if available) the 2007 reading scores for ability matching.

The ability matching was based on a TELPAS composite score where the composite was the 2008 TELPAS reading raw scores (expressed below as $\hat{Y}_{\text{predicted\_rawscore}}$) predicted through a regression approach, where the variables used to predict the 2008 score were the TELPAS ratings from 2008 and either the 2007 TELPAS reading rating (for students who were grades 1 and 2 in 2007) or 2007 TELPAS reading scale scores (for students who were grades 3 through 11 in 2007). Once the regression models were established for the students testing on online (one for each grade/grade cluster), these same models were applied to students who took the 2008 TELPAS reading test on paper.

The models are specified below:

For current grades 2 and 3 students who had a reading rating in 2007:

$$\hat{Y}_{predicted\_rawscore} = \beta_0 + \beta_1 X_{1(2008\_ListeningR)} + \beta_2 X_{2(2008\_SpeakingR)} + \beta_3 X_{3(2008\_WritingR)} + \beta_4 X_{4(2007\_ReadingR)}$$

For current grades 2 and 3 student who did not have a reading rating in 2007:

$$\hat{Y}_{predicted\_rawscore} = \beta_0 + \beta_1 X_{1(2008\_ListeningR)} + \beta_2 X_{2(2008\_SpeakingR)} + \beta_3 X_{3(2008\_WritingR)}$$

For current grades 4 through 12 students who had a reading score in 2007:

$$\hat{Y}_{predicted\_rawscore} = \beta_0 + \beta_1 X_{1(2008\_ListeningR)} + \beta_2 X_{2(2008\_SpeakingR)} + \beta_3 X_{3(2008\_WritingR)} + \beta_4 X_{4(2007\_ReadingR)}$$

For current grades 4 through 12 students who did not have a reading score in 2007:

$$\hat{Y}_{predicted\_rawscore} = \beta_0 + \beta_1 X_{1(2008\_ListeningR)} + \beta_2 X_{2(2008\_SpeakingR)} + \beta_3 X_{3(2008\_WritingR)}$$

The composite score was then used to represent the student ability level and as a variable for matching. Additionally, some demographic information was considered as matching variables. These variables were: gender, migrant status, and economically disadvantaged status.

Furthermore, a modification was made to the traditional MSCA procedure. Specifically, students having 2007 reading scores were matched across mode, and students not having 2007 reading scores were matched across mode. In other words, matching was conducted **separately** for the two strata: students with 2007 RPTE reading scores and students without 2007 RPTE reading scores. Since the ability matching required full 2008 rating information, students not having scores on any of the 2008 ratings or not having demographic information were excluded from the analysis.

Table 1 depicts the matching variables used for each grade/grade cluster. Note that in the TELPAS reading assessment there are six grade clusters: grade 2, grade 3, grades 4–5, grades 6–7, grades 8–9, and grades 10–12. In addition, within each grade cluster there are two rows, each representing one student stratum: those with 2007 reading scores and those without 2007 reading scores. Teacher ratings of student performance in all four domains were available for grade 2 and grade 3 in 2007, so for these two grade levels the reading ratings (instead of raw scores on the RPTE) from 2007 were used as matching variables.

Phase I: Assessing Comparability through Raw Score Comparisons

In this phase, the goal was to address the question of whether the two test modes resulted in comparable raw scores (and raw score cuts based on the interim cuts, because the final standards had not been available by the time the comparability analyses were conducted). A bootstrap sampling approach was conducted to select online and paper samples where each selected paper student was matched to an online student with the same ability level (based on the composite score as explained previously) and demographic variables, specifically student gender, economic disadvantaged status and migrant status. The matching variables are listed in Table 1. Once both the paper and online matched samples were selected, raw score to scale score conversions were calculated using the Rasch measurement model. One hundred replications were used in the bootstrap sampling approach.

The bootstrap process can be summarized in the following steps:

1. Draw a sample of students testing on paper where the sample size is equal to the number of the paper starting sample (number of students taking the paper version of the test who were eligible for matching). Note that students must meet certain criteria to be eligible for inclusion in the comparability analysis, such as having complete information in their gender, migrant status, economically disadvantaged status, and scale score variables.
2. Draw a sample of students testing online of the same sample size and the same profile composition (the ability level, the gender, the economically disadvantaged status, and the migrant status) as the sample drawn in step 1.
3. Run calibration analyses on the paper sample from step 1. Retain the ability estimates.
4. Run calibration analyses on the online sample from step 2. Retain the item difficulty estimates.
5. Compute the "expected" paper raw scores based on the paper ability estimates from step 3 and online item difficulty estimates from step 4.
6. Using linear interpolation, obtain the paper (or online "equivalent") scale scores for the expected paper raw scores calculated from step 5. This would constitute an "alternate" conversion table for a given replication.
7. Repeat steps 1 through 6 - 100 times and average the values in the resulting conversion tables to get the final conversion table.

Note that a stratified sampling approach was used such that students testing on paper with 2007 scores were matched only with students testing online who also had 2007 scores. Students testing on paper who did not have 2007 scores were matched only with students testing online who similarly did not have 2007 scores. Based on the results from similar studies, these two groups of students (those with or without previous scores) may have some fundamental differences and should not be matched across groups.

It should be noted that the paper scale scores computed in step 6 were, like the raw score cuts, just interim and not the final scale scores. However, these interim scale scores were able to be used to revert to the ability estimates corresponding to each raw score point by the following equations:

$$(\text{Interim Scale Score}) = 48 \times (\text{Initial Ability}) + 575. \quad (1)$$

Here the scaling constants are 48 [the slope] and 575 [the intercept]. The scale scores can be linearly transformed back to the initial ability estimates:

$$(\text{Initial Ability}) = (1/48) \times (\text{Interim Scale Score}) + (-575/48). \quad (2)$$

The initial conversion tables, as well as the estimated standard errors of linking, were computed by averaging across replications. The equivalency between online and paper raw scores was then evaluated based on psychometric and practical criteria.

Based on previous research (Dorans & Feigenbaum, 1994; Dorans, Holland, Thayer, & Tateneni, 2003; Dorans & Lawrence, 1990) and feedback from the TTAC, three decision rules were used as criteria to evaluate comparability. Specifically, the criteria included the raw score cut differences, the magnitudes of scale score differences (statistically significant if greater than or equal to 2 standard errors of linking), and the magnitudes of raw score equivalents (practically significant if greater than or equal to 0.5). A separate conversion table for the paper administration would be recommended if the mode difference meets these specified psychometric and practical criteria. A separate conversion table, recommended when a mode effect is found, would achieve the goal of adjusting scores for one mode while holding the scores constant on the other mode.

Phase II: Finalizing Scale Score Conversions for the Paper Mode

The proficiency-level-setting process for TELPAS reading was completed in August of 2008. The proficiency level cut scores for the online version of the test were determined for each grade cluster. In addition, a vertical scale study was conducted earlier in 2008 and the vertical linking coefficients were calculated. To present the amount of growth with the scale scores, the vertical linking coefficients were applied to the initial ability estimates at each raw score point for each grade cluster, and the final paper scale scores were computed using the following equations:

First, apply the grade-appropriate vertical scale coefficient to the initial ability:

$$\text{(Scaled Ability)} = \text{(Initial Ability)} + \text{(Vertical Scale Coefficient)} \quad (3)$$

Then, apply the scaling constants to the scaled ability:

$$\text{(Paper Scale Score)} = 48 \times \text{(Scaled Ability)} + 575. \quad (4)$$

The scaling constants (48 [the slope] and 575 [the intercept]) linearly transform the scaled-ability estimates to scale scores that range between 0 and 1000.

The raw score cuts were then found by first locating the scale scores on the paper conversion tables that corresponded to the online scale score cuts. Procedurally, the first step was to find the scale scores that correspond to the online raw cut scores, which were determined during the proficiency-level-setting process. The second step was to find the attainable scale score from the paper conversion table that was equal to the online scale score. This attainable scale score became the equivalent cut score for the paper mode. The paper raw score associated with this scale score was then set as the cut score for the paper mode. However, if none of the attainable paper scale scores were the same as an online scale score, the lowest paper scale score greater than the online scale score would be used, and the raw score associated with this scale score would be set as the cut score for paper. As shown in Appendix 7, for example, the online scale score cut at the beginning/intermediate level was 551. However, this was not an attainable scale score for students taking the paper forms (the two attainable adjacent paper scale scores were 555 and 549). Since 555 was the lowest paper scale score greater than the online scale score cut, it was then set as the beginning/intermediate cut for the paper administration.

RESULTS

Table 2 presents stratification information of students by grade cluster. Excluding students due to their missing 2007 reading score would have resulted in the loss of a considerable number of students. By incorporating the stratified sampling approach, those students were retained.

Table 3 shows the demographic information for the initial samples in both the online and paper modes. Also included in the table is the average raw scores based on the unmatched samples. Results illustrate that for the initial samples (before any matching was performed) the paper sample scored higher in all grade clusters except for grades 6–7. However, since these two samples were not equivalent, no conclusions should be drawn about the mode differences based on these values at this point in the analyses.

Table 4 provides an overview of the final results of the comparability analysis for TELPAS reading in 2008. The final raw score cuts are in general higher for paper than for online, and the differences are greater in the lower grade clusters.

Recall that there were two phases to this comparability study. In phase I, the goal was to determine if a mode effect existed by comparing the raw scores both overall and at the various proficiency level cuts. In phase II, the results were combined from the TELPAS reading vertical study, the final scale scores (for the online mode) resulting from the TELPAS proficiency level setting, and the results from phase I to calculate the final scale score cuts for the paper mode. The findings are organized according to these phases in the following sections.

Phase I Results: Assessing Comparability Through Raw Score Comparisons

Figure 1 is the graphical presentation of the raw score cuts for the four proficiency levels across grade clusters from grade 2 through grade cluster 10–12. Generally the interim cuts at various proficiency levels are different between modes—except at the highest grade levels (10–12) and the intermediate/advanced cut for grade cluster 6–7, where there is no difference in cuts between online and paper. When differences are observed, the paper cut tends to be between 1 to 2 points higher than the online cut, although a 3-point difference (paper higher) was found at the beginning/intermediate cut for grade 3.

Appendix 1 presents information for the conversion table for grade 2 based on the interim cuts. Since the relationship between ability and raw score does not change, comparability can be evaluated even if a different set of raw score cuts are used. Based on the interim cuts, the online version was two raw score points harder at all three cuts. Appendices 2 to 6 present similar information for grade 3 through grade cluster 10–12. Note that the scale scores listed in these tables are not final as 1) the vertical scale was not applied, and 2) the online scale scores cuts were not finalized. Both of these issues are addressed in the phase II analyses.

Table 5 shows the trend in mode effect across grade cluster. The average mode effect, in raw score units, ranged from 0.2 to 1.8 with paper being easier for all grade clusters. It is interesting to note that, although the mode effect indicates that the paper mode appears easier, the average mode effect declines steadily from grades 2 to 6–7, and from grade clusters 8–9 to 10–12. However, there was an increase of mode effect at grade cluster 8–9. Figure 2 depicts this trend.

The magnitudes of the average mode effects were also evaluated by taking the total raw score into account. A one-point mode effect may be more practically significant when the total raw score is 50 than when the total raw score is 60. The last column in table 5 shows that the mode effect ranged from 0.31% to 3.67% of the total raw score.

Based on the above information, it appears that students taking the paper version needed a higher raw score to achieve the same proficiency level cuts. Since we match on student ability level in comparability analyses, this means that the paper version was easier. Therefore, given the same ability level, students testing on paper would need to score higher to achieve the same proficiency online. This holds true for all grade clusters except for grade cluster 10–12. The results were shared with TEA and the decision was made to use a different paper conversion table for each of the grade clusters except for grade cluster 10–12.

Phase II Results: Finalizing Scale Score Conversions for the Paper Mode

The TELPAS reading vertical scale study identified student growth from a given lower grade cluster to its adjacent upper grade cluster. The growth was applied to the initial ability estimates resulting from the phase I analyses to generate an ability estimate (the scaled ability) that reflects student growth when moving up a grade cluster. To convert this 'scaled' ability to a scale score, a set of scaling constants were applied such that the range of the scale scores will be between 0 and 1000. Once the scale scores were generated, it was straightforward to locate the cut scores for the paper mode by using the same sets of cuts for online, which were established in the standard-setting process.

This process of locating raw score cuts for the paper mode was repeated for grades 2 to 8–9. No paper conversion table was generated for grade cluster 10–12, since it was concluded that there was no mode effect for this grade cluster. As a result, the online conversion table and cuts were used to score students taking the paper assessment in this grade cluster.

The resulting final raw cut scores are listed in Table 6. Please note that this table has also been presented in the summary section of the report. For detailed information on the final paper conversion tables, refer to Appendices 7 to 12. The three proficiency level cuts have been color-coded: green for 'beginning/intermediate,' purple for 'intermediate/advanced,' and orange for 'advanced/advanced-high.'

Impact Data Analyses

The comparison of students reaching the Advanced High proficiency level (based on the final proficiency level cut points) shown in Table 4 indicates that applying the online conversion tables to the students testing on paper results in a higher percentage of students (who tested on paper) attaining the Advanced High proficiency level. This is true, with the exception of grade cluster 10–12, where there was no mode effect. The increased percentages of students (who tested on paper) attaining Advanced High also confirms that the paper version is easier; therefore, it is necessary to use a separate conversion table to adjust the paper raw score cuts so that they are higher than the online raw score cuts.

ADDITIONAL ANALYSES

To gather additional validity evidence supporting the decisions about modes of assessments and to be consistent with previous online comparability studies for other assessment programs, two sets of additional analyses were conducted—the group analysis and the item-level analysis. The group analysis compares mean differences of the total raw scores between the two testing modes across replications for male, female, migrant, non-migrant, economically disadvantaged, and non-economically disadvantaged student groups separately, whereas the item-level analysis compares mean differences of each item between the two testing modes across replications.

Group Analysis

The mean raw score differences (and the mean effect size, see Cohen [1992]) between the paper and online testing modes for each group are listed in Table 7. A significance test was performed for each matched group using the following equation:

$$Zdif = \frac{\overline{D}_{Diff}}{\sqrt{SE^2_{Diff}}}$$

where $\overline{D}_{Diff}$ is the grand mean of the differences between mean online and mean paper scores over the replications for each group; and SE_{diff} is the bootstrap standard error of the mean differences over the replications, also for each group.

The effect size between two group means at each replication was calculated using the following equation:

$$EffectSize = \frac{\overline{X}_{Group1} - \overline{X}_{Group2}}{\sqrt{\frac{(SD^2_{Group1} + SD^2_{Group2})}{2}}}$$

The effect sizes for the raw scores were based on the averages of the effect sizes over the replications.

As can be seen in Table 7, there was a consistent significant mode effect across gender and other groups for grade/grade clusters 2 to 8–9. With grade cluster 10–12, results indicated some significant mode effects; however, the magnitude of the mode effects was in general smaller than those in the other grade clusters.

Item-Level Analysis

The item-level analysis was performed in a similar manner to the group analysis. The mean item raw score differences across replications between the online and paper testing modes were computed, and the effect sizes were also calculated. Appendices 12 - 17 display the results of item-level comparison across replications for each grade cluster. The columns of the tables are as follows:

CBTPVAL: Mean item score for the students testing online across 100 replications.
PAPPVAL: Mean item score for the students testing on paper across 100 replications.
DIFPVAL: Mean item score differences between students testing online and students testing on paper across 100 replications
DIFSTD: Standard deviation of the mean differences across 100 replications
Z_DIF: Z statistic for the mean item score differences
ISIG: Items where the Z_DIF statistic was greater than 2 are noted by "*".
EFFECT_SIZE: Mean effect size over 100 replications.

Table 8 provides the frequency distribution of items that show no mode effect, items that show a mode effect favoring online, and items that show a mode effect favoring paper. For grade 2, of the 49 base test items, 37 (or 76%) items showed a paper-easier mode effect, 9 (or 18%) items showed an online-easier mode effect, and 3 (or 6%) items showed no mode effect. In general, it can be seen that across most grade levels, more items exhibited paper-easier mode effects. However, this trend is not evident at the high school level (grades 10–12), where half of the items showed no mode effect at all, and an equal number of items showed either paper-easier or online-easier mode effects. This is consistent with the conclusion that the paper version is easier since more items overall show a mode effect in favor of paper.

Figure 3 depicts the percentages for each mode-effect category by grade cluster. While the percentages of paper-easier items decrease by grade, the percentages of no-mode-effect items increase by grade. The percentages of online-easier items, however, remain relatively stable across grade levels. For most of the differences, the associated standard errors were relatively small—which might have contributed to the large number of items showing statistically significant differences between the testing modes. To help determine the practical significance of the differences, the average effect size for each difference was calculated and listed on the last column of the tables in appendices 12 to 17. It can be seen that the magnitude of the effect sizes ranged between 0 and 0.22.

Items with relatively large mode differences were further reviewed by the Pearson psychometric and content teams jointly. It was determined that many of the items exhibiting a mode effect in favor of paper were

- passage-based items;
- items with split-screens (needing scrolling); and
- items requiring students to reference back to a certain position in the reading passages.

No consistent pattern was found for items showing an online-easier mode effect.

CONCLUSION

The results of these matched-group comparability analyses suggest that students with the same estimated proficiency in TELPAS reading tend to score higher on the paper version of the assessment than the online version. It also shows that this mode effect seems to disappear at the highest grade cluster tested (grades 10–12). At the group level (i.e., gender, migrant status, and economically disadvantaged status), the paper versions consistently show higher average raw scores. At the item level, across most grade clusters, more items exhibit a paper-easier mode effect. The only exception was at grade cluster 10–12, where there are the same numbers of items in each category (paper-easier or online-easier).

In light of these findings, TEA determined that it would be appropriate to use separate conversion tables for students taking the TELPAS reading assessment on paper for grades 2 through 9. For students in grades 10 through 12 who took the assessment on paper, the online conversion table was used. When the proficiency level cuts were determined in summer 2008, the final cut scores for the paper mode were adjusted accordingly based on the final online cut scores.

DISCUSSION

Results indicated that the mode effect at the test level decreased as students' grade level increased. The only exception was between grades 6–7 and grades 8–9, where the mode effect increased slightly. The demographic information for these two grade clusters indicates that in grades 6–7, the majority of students have been in the U.S. for more than 4 years. In grades 8–9, however, there were relatively fewer students having been in U.S. schools for more than 4 years, and relatively more students having been in U.S. schools for 1 or 2 years. It is possible that the higher number of students with fewer years

in the U.S. might also have less experience with computers, which contributed to their lower scores online.

It was also observed that across most grade clusters, between 46% and 78% of items show a mode effect that favors the paper version. Except for grades 10–12, the number of items that do not show a mode effect accounts for less than 30%. This observation supports the use of separate paper conversion tables for most of the grade clusters. Since the online version of the test was found to be more challenging, a lower raw score was needed for students taking the test online to achieve the same proficiency level as students taking the test on paper.

REFERENCE

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1)*,* 155-159.

Dorans, N. J., & Feigenbaum, M. D. (1994). *Equating issues engendered by changes to the SAT and PSAT/NMSQT.* (ETS Research Memorandum 94-10). Princeton, NJ: Educational Testing Service.

Dorans, N. J., Holland, P. W., Thayer, D. T., & Tateneni, K. (2003). Invariance of score linking across gender groups for three Advanced Placement Program Examinations. In N. J. Dorans, (Ed.), *Population invariance of score linking: Theory and applications to Advanced Placement Program Examinations.* (ETS RR-03-27, pp. 79-118). Princeton, NJ: Educational Testing Service.

Dorans, N. J., & Lawrence, I. M. (1990). Checking the statistical equivalence of nearly identical test forms. *Applied Measurement in Education*, *3*, 245-254.

Way, W. D., Davis, L. L., & Fitzpatrick, S. (2006). *Score comparability of online and paper administrations of the Texas Assessment of Knowledge and Skills.* Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, CA. Available at http://www.pearsonedmeasurement.com/downloads/research/RR_06_01.pdf.

Table 1. Matching Variables Used for 2008 TELPAS Reading Comparability Studies

Grade Cluster	2007 Reading scores available?	2008 Listening Rating	2008 Speaking Rating	2008 Writing Rating	2007 RPTE Reading Scale Score	2007 RPTE Reading Rating*	Gender	Economically Disadvantaged Status	Migrant Status
2	No	√	√	√			√	√	√
	Yes	√	√	√		√	√	√	√
3	No	√	√	√			√	√	√
	Yes	√	√	√		√	√	√	√
4–5	No	√	√	√			√	√	√
	Yes	√	√	√	√		√	√	√
6–7	No	√	√	√			√	√	√
	Yes	√	√	√	√		√	√	√
8–9	No	√	√	√			√	√	√
	Yes	√	√	√	√		√	√	√
10–12	No	√	√	√			√	√	√
	Yes	√	√	√	√		√	√	√

Note. The grade 2 and grade 3 students included in the current analyses were in grade 1 and grade 2, respectively, in 2007. The 2007 RPTE reading ratings, therefore, were used for grades 2 and 3. The 2007 RPTE reading scale scores were used for the other grade clusters.

Table 2. Summary of Spring 2008 Sample Statistics and Regression Results, by Student Population and Grade Cluster

Grade Cluster	Student Population	Intercept	L_2008	S_2008	W_2008	R_2007	r-square	OnlineMean	OnlineN	PaperN	PaperMean
2	With 2007 Reading	13.28	1.17	0.46	4.68	1.60	0.50	31.89	52864	35331	33.30
	Without 2007 Reading	10.13	2.61	0.23	5.61	-	0.57	28.29	5049	3608	28.71
3	With 2007 Reading	16.97	1.65	0.58	4.04	3.18	0.52	41.56	49530	29846	43.16
	Without 2007 Reading	12.18	3.34	1.32	5.77	-	0.61	36.42	5346	3371	37.17
4–5	With 2007 Reading	-24.28	0.89	0.23	2.38	0.08	0.65	45.84	71676	30865	46.46
	Without 2007 Reading	12.34	3.61	0.91	6.57	-	0.66	38.96	8807	4478	38.77
6–7	With 2007 Reading	-26.26	1.04	-0.21	2.09	0.08	0.59	48.12	39941	17079	47.48
	Without 2007 Reading	15.82	3.02	1.25	6.39	-	0.63	41.63	7420	3863	41.21
8–9	With 2007 Reading	-35.37	1.05	-0.24	2.31	0.09	0.57	45.92	30890	12762	46.24
	Without 2007 Reading	14.55	3.08	1.04	6.76	-	0.59	40.22	9173	4397	40.24
10–12	With 2007 Reading	-38.89	1.07	-0.48	2.04	0.10	0.55	46.89	22943	12158	46.99
	Without 2007 Reading	19.74	2.39	0.56	6.32	-	0.44	46.76	4208	2052	46.22

Note: L_2008=Listening rating in 2008; S_2008=Speaking rating in 2008; W_2008=Writing rating in 2008; R_2007=Reading rating/scale scores in 2007

Table 3. Student Demographic and Descriptive Information for Spring 2008 TELPAS Reading Comparability Analyses

Grade Cluster	Number of Campuses		Number of Students		Mean Raw Score		Estimated Raw Score, or y-hat		Demographic Information (%)									
									Male		Hispanic		Migrant		Economically Disadvantaged		Two or Fewer Years in U.S. Schools	
	CBT	*PAP*	CBT	*PAP*	CBT	*PAP*	CBT	*PAP*	CBT	*PAP*	CBT	*PAP*	CBT	*PAP*	CBT	*PAP*	CBT	*PAP*
2	2226	1477	57913	38926	31.57	32.87	31.57	31.07	51	51	90	95	1	3	85	90	89	87
3	2313	1385	54876	33206	41.06	42.55	42.06	40.97	51	52	92	96	2	3	87	92	6	7
4–5	2593	1471	80483	35336	45.08	45.49	45.08	44.24	52	53	94	95	2	3	90	91	8	9
6–7	1378	961	47361	20937	47.10	46.33	47.10	45.84	54	55	94	95	2	4	89	91	13	14
8–9	1578	1044	40063	17151	44.61	44.80	44.61	43.95	56	58	93	93	3	4	84	86	21	22
10–12	793	566	27151	14198	46.87	46.88	46.87	46.73	54	55	90	91	3	4	79	81	14	12

Note: Information presented in Table 3 was based on the unmatched samples.

Table 4. Summary of the Spring 2008 TALPAS Reading Comparability Analyses

Grade Cluster	Sample Size		Raw Score Cuts		Pass Rates (%) Comparison**			Number of Raw Score Points with Meaningful Difference***	Decision (Conclusion)
	Online	Paper	Online	Paper	Online w/Online	Paper w/Paper	Paper w/Online		
2	57913	38926	19	22	30.2	27.3	26.2	34(49)	Alternate Conversion Table (Mode Effect)
			30	33	28.2	26.5	27.5		
			39	41	29.8	28.8	36.0		
3	54876	33206	25	28	20.1	18.0	15.6	51(58)	Alternate Conversion Table (Mode Effect)
			36	39	33.8	26.5	25.4		
			46	48	42.5	42.5	49.6		
4–5	80483	35336	25	27	15.7	15.7	14.1	47(61)	Alternate Conversion Table (Mode Effect)
			37	39	23.8	24.5	21.4		
			47	49	54.2	50.9	57.3		
6–7	47361	20937	26	27	14.4	17.6	16.6	22(63)	Alternate Conversion Table (Mode Effect)
			39	40	28.0	29.7	27.5		
			50	51	51.8	45.5	49.5		
8–9	40063	17151	26	28	17.9	17.4	15.8	45(63)	Alternate Conversion Table (Mode Effect)
			37	39	28.6	27.6	28.5		
			49	50	45.3	43.5	46.4		
10–12	27151	14198	26	26	16.1	16.2	16.2	0(64)	Online Conversion Table (No Mode Effect)
			39	39	30.9	29.6	29.6		
			50	50	48.3	48.9	48.9		

*: Raw score points corresponding to 'Beginning/Intermediate' (top), 'Intermediate/Advanced (middle), and 'Advanced/Advanced High' (Bottom) levels.

**: Pass rates based on different conversion tables.

***: Meaningful differences require both scale score statistical significance and raw score practical significance. Total maximum RS points shown in parentheses.

Table 5. Summary of Test-Level Mode Effect Across Grade Cluster

Grade Cluster	Total RS	Mean Mode Effect	Percent Total RS (%)
2	49	1.8	3.67
3	58	1.6	2.76
4–5	61	1.3	2.13
6–7	63	0.5	0.79
8–9	63	0.9	1.43
10–12	64	0.2	0.31

Note: The mean mode effect was computed in the raw score unit.

Table 6. Cut Scores by Mode and Grade/Grade Cluster

Grade Cluster	Beginning/Intermediate		Intermediate/Advanced		Advanced/Advanced High	
	Online	Paper	Online	Paper	Online	Paper
2	551 (19)	555 (22)	615 (30)	618 (33)	670 (39)	673 (41)
3	597 (25)	599 (28)	648 (36)	651 (39)	699 (46)	702 (48)
4–5	610 (25)	611 (27)	668 (37)	668 (39)	718 (47)	722 (49)
6–7	613 (26)	616 (27)	674 (39)	676 (40)	731 (50)	734 (51)
8–9	633 (26)	636 (28)	681 (37)	684 (39)	738 (49)	738 (50)
10–12	644 (26)	644 (26)	704 (39)	704 (39)	757 (50)	757 (50)

Note that in the above table, numbers preceding parentheses are *attainable* scale score cuts, whereas numbers inside parentheses are the corresponding raw score cuts.

Table 7. Summary of Group Analyses Across Grade Cluster

Grade Cluster	RS (Cbt-Pap)	Male	Female	Non-Migrant	Migrant	Non Economically Disadvantaged	Economically Disadvantaged
2	**-1.82**	**-1.79**	**-1.85**	**-1.83**	**-1.57**	**-1.43**	**-1.86**
3	**-1.64**	**-1.66**	**-1.61**	**-1.65**	**-1.32**	**-1.29**	**-1.67**
4–5	**-1.20**	**-1.14**	**-1.26**	**-1.2**	**-1.15**	**-1.06**	**-1.21**
6–7	**-0.47**	-0.18	**-0.81**	**-0.46**	**-0.74**	0.02	**-0.52**
8–9	**-0.83**	**-0.72**	**-0.98**	**-0.81**	**-1.3**	**-1.07**	**-0.79**
10–12	**-0.18**	-0.05	**-0.34**	-0.17	-0.44	-0.11	**-0.19**

Note that bold entries indicate statistically-significant mode differences at the 0.05 level.

Table 8. Summary of Item-Level Analyses Across Grade Cluster

Grade Cluster	Total Base Item Counts	Favor Paper	Favor Online	No Effect
2	49	37 (76%)	9 (18%)	3 (6%)
3	58	45 (78%)	5 (9%)	8 (14%)
4–5	61	39 (64%)	16 (26%)	6 (10%)
6–7	63	29 (46%)	16 (25%)	18 (29%)
8–9	63	33 (52%)	12 (19%)	18 (29%)
10–12	64	16 (25%)	16 (25%)	32 (50%)

Figure 1. Graphic Presentation of Interim Cut Scores by Mode at Each Grade Cluster

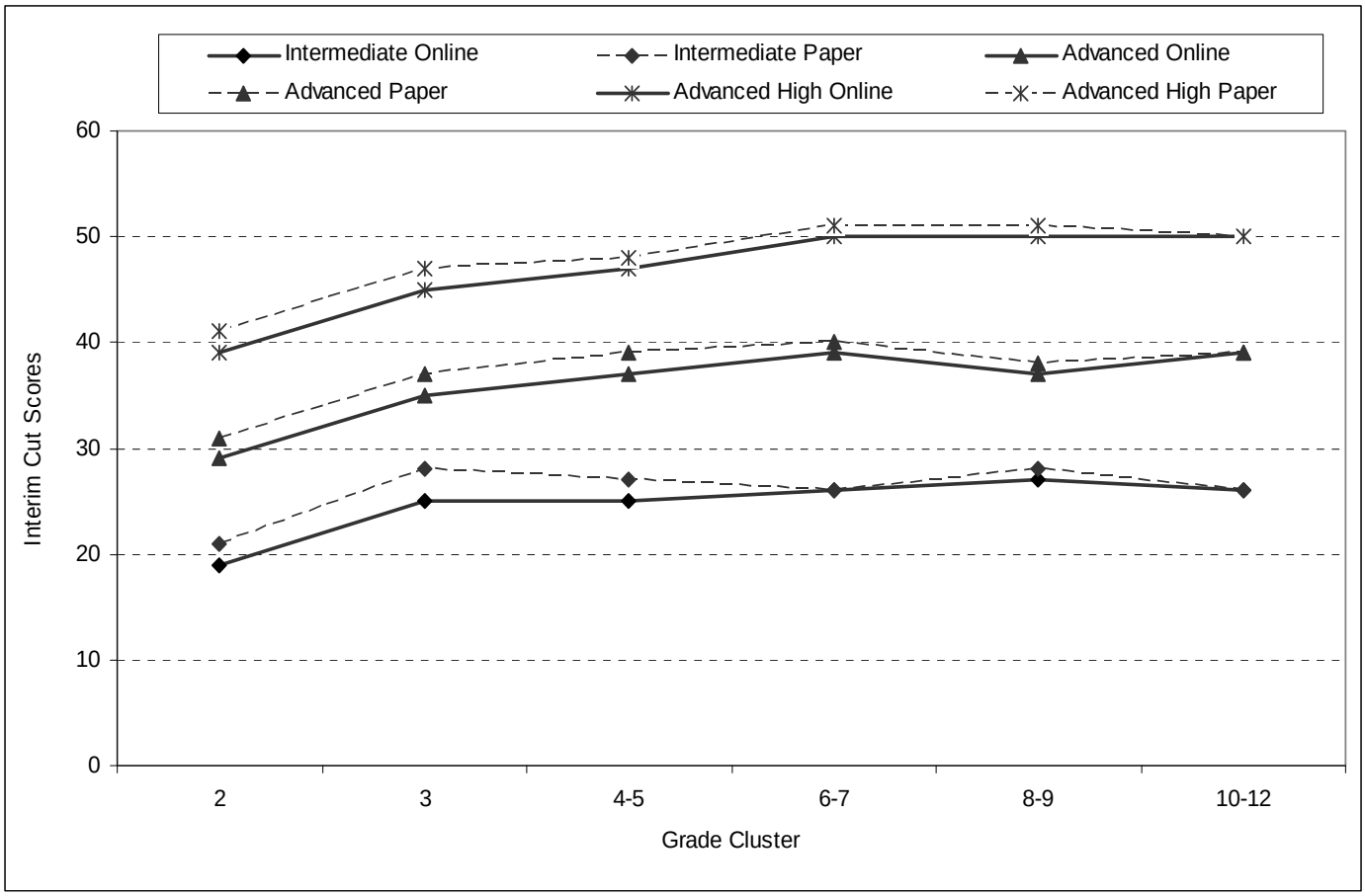


Figure 2. Graphic Presentation of the Mean Mode Effect by Grade Cluster

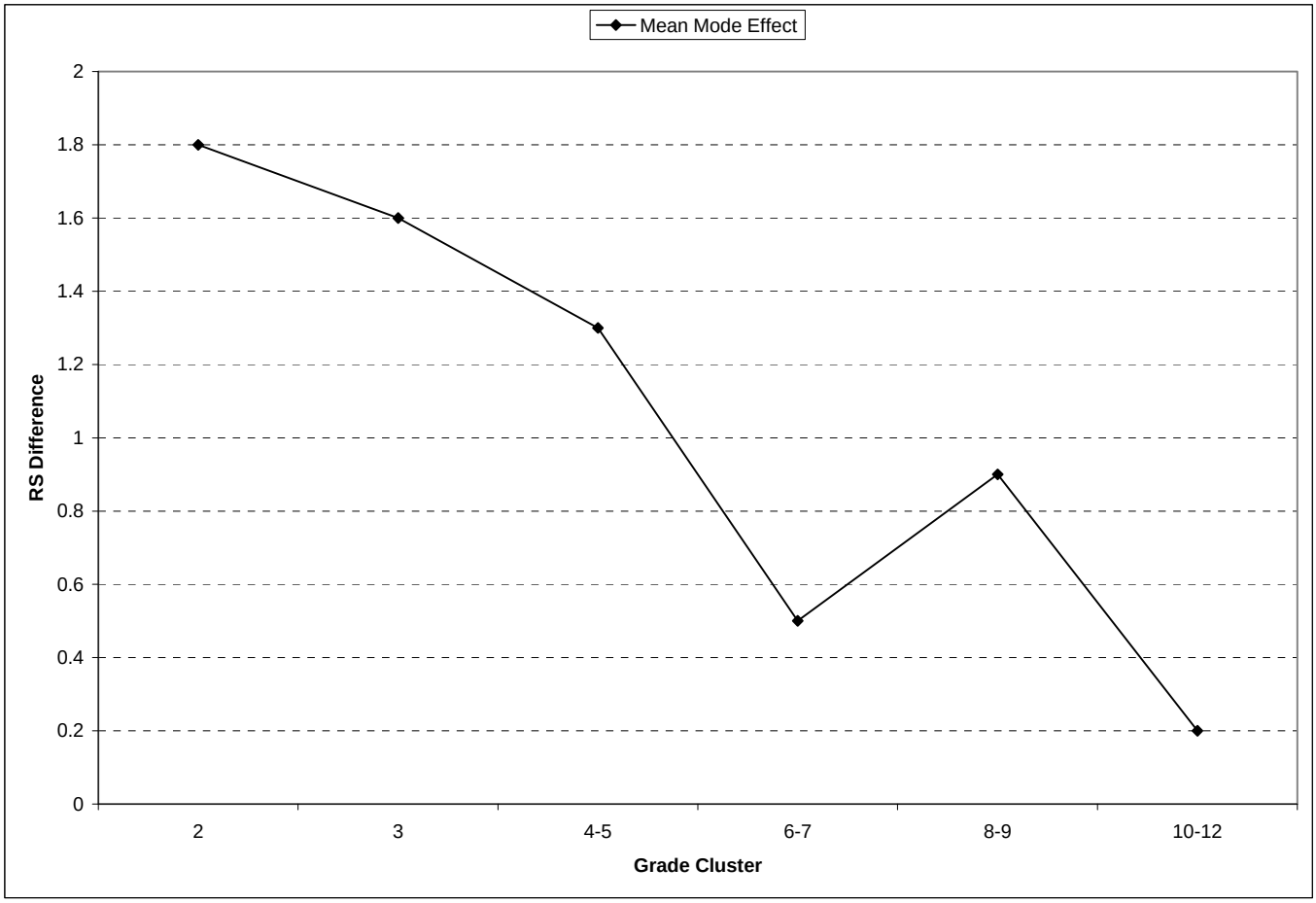
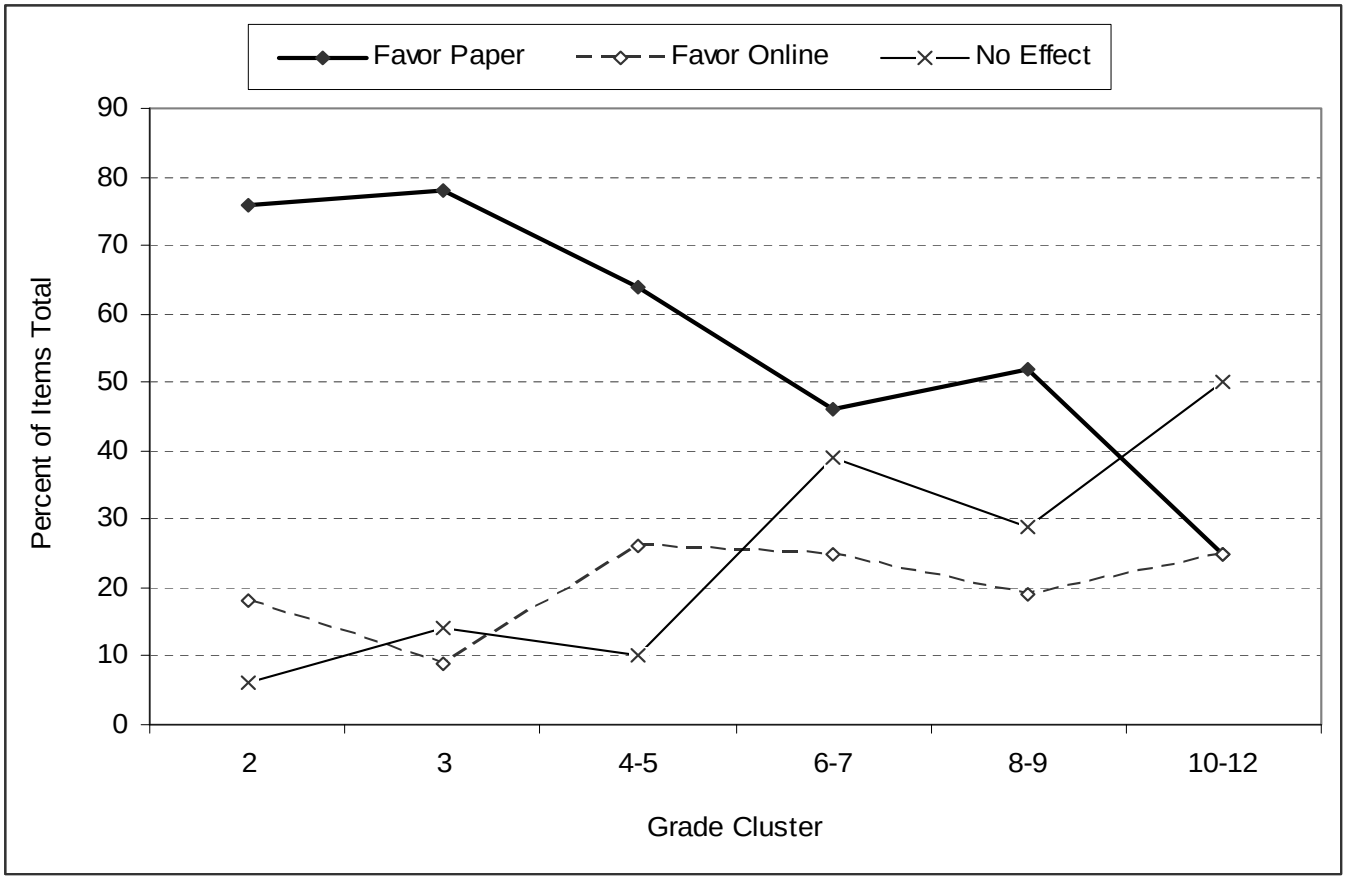


Figure 3. Percent of Items Favoring Each Mode, by Grade Cluster



Appendix 1. Summary of Phase One Comparability Analysis – Grade 2

RAW	PAP_RS	RS_SD	CBT_SS	PAP_SS	SS_SD	RS_DIF	SS_DIF	SIG
0	0.30	0.007	261.64	277.00	0.367	0.30	15.36	
1	0.99	0.020	312.84	312.14	0.973	-0.01	-0.70	
2	1.94	0.031	352.84	350.46	1.244	-0.06	-2.38	
3	2.87	0.037	379.11	375.63	0.979	-0.13	-3.48	
4	3.77	0.040	399.73	395.07	0.827	-0.23	-4.67	
5	4.66	0.041	417.16	411.30	0.718	-0.34	-5.86	
6	5.54	0.042	432.45	425.45	0.635	-0.46	-7.01	
7	6.41	0.042	446.16	438.09	0.569	-0.59	-8.07	*
8	7.27	0.042	458.61	449.58	0.519	-0.73	-9.03	*
9	8.13	0.042	470.01	460.14	0.479	-0.87	-9.87	*
10	8.99	0.043	480.53	469.90	0.471	-1.01	-10.63	*
11	9.85	0.044	490.30	478.95	0.460	-1.15	-11.35	*
12	10.71	0.045	499.41	487.46	0.439	-1.29	-11.95	*
13	11.57	0.046	507.97	495.53	0.423	-1.43	-12.44	*
14	12.44	0.048	516.04	503.20	0.411	-1.56	-12.85	*
15	13.32	0.050	523.69	510.53	0.402	-1.68	-13.17	*
16	14.20	0.052	530.98	517.56	0.394	-1.80	-13.42	*
17	15.09	0.053	537.95	524.34	0.389	-1.91	-13.62	*
18	15.99	0.055	544.65	530.88	0.395	-2.01	-13.77	*
19	16.89	0.057	551.12	537.22	0.395	-2.11	-13.90	*
20	17.81	0.058	557.39	543.40	0.391	-2.19	-14.00	*
21	**18.74**	0.060	563.49	549.44	0.387	-2.26	-14.06	*
22	19.68	0.061	569.45	555.37	0.384	-2.32	-14.08	*
23	20.62	0.062	575.29	561.21	0.381	-2.38	-14.08	*
24	21.58	0.063	581.04	566.97	0.378	-2.42	-14.07	*
25	22.55	0.064	586.71	572.67	0.375	-2.45	-14.04	*
26	23.53	0.065	592.33	578.34	0.373	-2.47	-14.00	*
27	24.52	0.065	597.92	583.98	0.370	-2.48	-13.93	*
28	25.52	0.065	603.49	589.62	0.368	-2.48	-13.87	*
29	26.53	0.065	609.06	595.27	0.365	-2.47	-13.79	*
30	27.54	0.065	614.66	600.95	0.363	-2.46	-13.71	*
31	**28.57**	0.065	620.31	606.68	0.361	-2.43	-13.63	*
32	29.61	0.064	626.01	612.48	0.358	-2.39	-13.54	*
33	30.66	0.063	631.81	618.37	0.356	-2.34	-13.44	*
34	31.71	0.062	637.72	624.38	0.353	-2.29	-13.33	*
35	32.78	0.061	643.78	630.54	0.351	-2.22	-13.24	*
36	33.86	0.059	650.01	636.87	0.347	-2.14	-13.14	*
37	34.94	0.057	656.48	643.43	0.346	-2.06	-13.05	*
38	36.04	0.055	663.22	650.27	0.350	-1.96	-12.95	*
39	37.15	0.052	670.30	657.47	0.351	-1.85	-12.83	*
40	38.27	0.049	677.82	665.10	0.348	-1.73	-12.71	*
41	**39.40**	0.046	685.88	673.28	0.346	-1.60	-12.59	*
42	40.54	0.042	694.64	682.18	0.341	-1.46	-12.46	*
43	41.70	0.038	704.36	692.00	0.335	-1.30	-12.36	*
44	42.87	0.034	715.37	703.10	0.328	-1.13	-12.27	*
45	44.06	0.029	728.30	716.14	0.369	-0.94	-12.16	*
46	45.27	0.023	744.29	732.55	0.364	-0.73	-11.74	*
47	46.49	0.016	765.90	754.88	0.349	-0.51	-11.02	*
48	47.73	0.009	801.26	791.87	0.305	-0.27	-9.39	
49	48.62	0.003	849.64	831.16	0.131	-0.38	-18.48	

RAW - Online test raw score
PAP_RS - Equivalent raw scores on the online test based on the comparability linking. Note that a lower equivalent raw score indicates that the online version of the test was more difficult.
RS_SD - Standard deviation of the equivalent raw scores over the replications.
CBT_SS - Online test scale score conversions, based on the 2008 TELPAS post-equated scales.
PAP_SS - Equivalent scale scores on the online test based on the comparability linking. Again, lower equivalent scale scores indicate that the online version of the test was more difficult.

SS_SD - Standard deviation of the equivalent scale scores over the replications.
RS_DIF - Difference between paper raw score equivalent and online raw score.
SS_DIF - Difference between paper scale score equivalent and online scale score.
SIG^- Raw score points where scale score differences exceed two standard errors of the linking and where the difference in raw scores is greater than half a point are noted by "*".

^ Note that the "*" in the SIG column now indicates both statistical and practical significance, based on recommendations from the Texas Technical Advisory Committee.

Appendix 2. Summary of Phase One Comparability Analysis – Grade 3

RAW	PAP_RS	RS_SD	CBT_SS	PAP_SS	SS_SD	RS_DIF	SS_DIF	SIG
0	0.26	0.006	274.15	287.14	0.302	0.26	12.99	
1	0.87	0.018	323.98	317.61	0.894	-0.13	-6.37	
2	1.75	0.031	361.64	352.24	1.149	-0.25	-9.40	
3	2.63	0.039	385.46	376.69	0.937	-0.37	-8.77	
4	3.51	0.046	403.55	394.76	0.825	-0.49	-8.79	
5	4.40	0.050	418.46	409.47	0.747	-0.60	-8.99	*
6	5.28	0.053	431.32	422.05	0.687	-0.72	-9.26	*
7	6.16	0.056	442.71	433.16	0.638	-0.84	-9.55	*
8	7.04	0.058	453.01	443.16	0.613	-0.96	-9.85	*
9	7.93	0.060	462.44	452.26	0.614	-1.07	-10.17	*
10	8.81	0.062	471.15	460.67	0.584	-1.19	-10.49	*
11	9.70	0.064	479.28	468.53	0.556	-1.30	-10.75	*
12	10.59	0.066	486.91	475.93	0.533	-1.41	-10.98	*
13	11.48	0.068	494.11	482.93	0.515	-1.52	-11.18	*
14	12.37	0.069	500.94	489.60	0.500	-1.63	-11.34	*
15	13.27	0.071	507.45	495.98	0.488	-1.73	-11.47	*
16	14.18	0.073	513.68	502.09	0.478	-1.82	-11.59	*
17	15.09	0.075	519.67	507.98	0.472	-1.91	-11.69	*
18	16.00	0.077	525.44	513.67	0.471	-2.00	-11.77	*
19	16.92	0.079	531.02	519.17	0.469	-2.08	-11.85	*
20	17.84	0.080	536.43	524.52	0.464	-2.16	-11.90	*
21	18.77	0.082	541.69	529.74	0.457	-2.23	-11.95	*
22	19.71	0.083	546.82	534.84	0.452	-2.29	-11.98	*
23	20.65	0.085	551.84	539.84	0.446	-2.35	-12.00	*
24	21.60	0.086	556.76	544.76	0.441	-2.40	-12.01	*
25	22.55	0.087	561.60	549.59	0.437	-2.45	-12.00	*
26	23.51	0.088	566.36	554.37	0.432	-2.49	-11.99	*
27	24.48	0.089	571.05	559.09	0.429	-2.52	-11.96	*
28	**25.46**	0.089	575.70	563.77	0.424	-2.54	-11.93	*
29	26.44	0.090	580.31	568.42	0.422	-2.56	-11.89	*
30	27.43	0.090	584.88	573.04	0.418	-2.57	-11.84	*
31	28.42	0.090	589.44	577.66	0.415	-2.58	-11.78	*
32	29.43	0.090	593.99	582.27	0.412	-2.57	-11.72	*
33	30.44	0.090	598.53	586.88	0.409	-2.56	-11.65	*
34	31.46	0.089	603.09	591.52	0.406	-2.54	-11.57	*
35	32.48	0.089	607.66	596.18	0.403	-2.52	-11.48	*
36	33.52	0.088	612.27	600.89	0.401	-2.48	-11.39	*
37	**34.56**	0.087	616.93	605.64	0.397	-2.44	-11.29	*
38	35.61	0.086	621.64	610.46	0.395	-2.39	-11.18	*
39	36.66	0.084	626.42	615.35	0.392	-2.34	-11.07	*
40	37.72	0.083	631.29	620.34	0.389	-2.28	-10.95	*
41	38.80	0.081	636.27	625.44	0.386	-2.20	-10.83	*
42	39.87	0.078	641.38	630.67	0.382	-2.13	-10.71	*
43	40.96	0.076	646.64	636.06	0.381	-2.04	-10.57	*
44	42.05	0.073	652.07	641.64	0.383	-1.95	-10.44	*
45	43.15	0.070	657.72	647.44	0.383	-1.85	-10.29	*
46	44.25	0.067	663.63	653.50	0.380	-1.75	-10.13	*
47	**45.36**	0.064	669.84	659.87	0.377	-1.64	-9.97	*
48	46.48	0.060	676.43	666.62	0.373	-1.52	-9.81	*
49	47.61	0.056	683.48	673.84	0.368	-1.39	-9.64	*
50	48.74	0.051	691.11	681.63	0.362	-1.26	-9.48	*
51	49.87	0.047	699.47	690.14	0.356	-1.13	-9.33	*
52	51.02	0.041	708.81	699.64	0.374	-0.98	-9.18	*
53	52.17	0.036	719.49	710.58	0.384	-0.83	-8.91	*
54	53.32	0.030	732.12	723.54	0.378	-0.68	-8.58	*
55	54.48	0.023	747.84	739.69	0.368	-0.52	-8.15	*
56	55.65	0.016	769.21	761.70	0.350	-0.35	-7.51	
57	56.82	0.009	804.39	798.11	0.302	-0.18	-6.27	
58	57.65	0.003	852.66	835.57	0.129	-0.35	-17.09	

RAW - Online test raw score
PAP_RS - Equivalent raw scores on the online test based on the comparability linking.
  Note that a lower equivalent raw score indicates that the online version of the test
  was more difficult.
RS_SD - Standard deviation of the equivalent raw scores over the replications.
CBT_SS - Online test scale score conversions, based on the 2008 TELPAS post-equated scales.
PAP_SS - Equivalent scale scores on the online test based on the comparability linking.
  Again, lower equivalent scale scores indicate that the online version of the test was
  more difficult.
SS_SD - Standard deviation of the equivalent scale scores over the replications.
RS_DIF - Difference between paper raw score equivalent and online raw score.
SS_DIF - Difference between paper scale score equivalent and online scale score.
SIG^- Raw score points where scale score differences exceed two standard errors
  of the linking and where the difference in raw scores is greater than half a point
  are noted by "*".

^ Note that the "*" in the SIG column now indicates both statistical and practical
  significance, based on recommendations from the Texas Technical Advisory Committee.

Appendix 3. Summary of Phase One Comparability Analysis – Grade Cluster 4–5

RAW	PAP_RS	RS_SD	CBT_SS	PAP_SS	SS_SD	RS_DIF	SS_DIF	SIG
0	0.31	0.007	265.36	280.50	0.332	0.31	15.14	
1	1.02	0.020	314.65	315.20	0.803	0.02	0.55	
2	2.01	0.036	351.48	351.60	1.015	0.01	0.12	
3	2.99	0.047	374.56	374.21	0.982	-0.01	-0.35	
4	3.95	0.055	392.02	391.14	0.942	-0.05	-0.88	
5	4.90	0.062	406.42	404.98	0.889	-0.10	-1.44	
6	5.84	0.066	418.89	416.88	0.829	-0.16	-2.01	
7	6.77	0.070	430.03	427.45	0.777	-0.23	-2.58	
8	7.69	0.072	440.18	437.04	0.731	-0.31	-3.14	
9	8.61	0.073	449.58	445.89	0.691	-0.39	-3.69	
10	9.52	0.074	458.38	454.15	0.654	-0.48	-4.23	
11	10.43	0.075	466.68	461.94	0.621	-0.57	-4.74	*
12	11.34	0.075	474.56	469.33	0.590	-0.66	-5.22	*
13	12.24	0.075	482.07	476.38	0.563	-0.76	-5.69	*
14	13.15	0.075	489.25	483.14	0.537	-0.85	-6.11	*
15	14.06	0.074	496.15	489.65	0.518	-0.94	-6.50	*
16	14.97	0.074	502.79	495.92	0.506	-1.03	-6.87	*
17	15.88	0.074	509.19	501.99	0.491	-1.12	-7.20	*
18	16.79	0.074	515.38	507.88	0.473	-1.21	-7.51	*
19	17.71	0.074	521.38	513.61	0.456	-1.29	-7.77	*
20	18.64	0.074	527.20	519.20	0.442	-1.36	-8.00	*
21	19.57	0.074	532.86	524.67	0.429	-1.43	-8.19	*
22	20.50	0.074	538.38	530.02	0.417	-1.50	-8.36	*
23	21.44	0.074	543.76	535.26	0.409	-1.56	-8.50	*
24	22.38	0.074	549.02	540.41	0.399	-1.62	-8.60	*
25	23.33	0.074	554.17	545.48	0.391	-1.67	-8.70	*
26	24.28	0.075	559.23	550.47	0.384	-1.72	-8.77	*
27	**25.24**	0.075	564.21	555.39	0.378	-1.76	-8.82	*
28	26.21	0.075	569.12	560.25	0.373	-1.79	-8.86	*
29	27.18	0.075	573.96	565.07	0.367	-1.82	-8.89	*
30	28.15	0.075	578.75	569.85	0.363	-1.85	-8.90	*
31	29.13	0.075	583.50	574.60	0.358	-1.87	-8.90	*
32	30.12	0.075	588.20	579.32	0.355	-1.88	-8.89	*
33	31.11	0.075	592.89	584.02	0.351	-1.89	-8.87	*
34	32.11	0.074	597.56	588.72	0.348	-1.89	-8.85	*
35	33.11	0.074	602.23	593.42	0.345	-1.89	-8.81	*
36	34.12	0.073	606.90	598.12	0.342	-1.88	-8.78	*
37	35.13	0.073	611.58	602.85	0.339	-1.87	-8.73	*
38	36.15	0.072	616.29	607.60	0.337	-1.85	-8.69	*
39	**37.17**	0.071	621.03	612.39	0.335	-1.83	-8.64	*
40	38.20	0.070	625.81	617.23	0.332	-1.80	-8.58	*
41	39.23	0.069	630.66	622.13	0.330	-1.77	-8.53	*
42	40.27	0.068	635.58	627.10	0.328	-1.73	-8.47	*
43	41.31	0.066	640.58	632.17	0.326	-1.69	-8.41	*
44	42.35	0.065	645.69	637.34	0.323	-1.65	-8.35	*
45	43.40	0.063	650.93	642.63	0.322	-1.60	-8.30	*
46	44.45	0.061	656.31	648.07	0.320	-1.55	-8.24	*
47	45.51	0.059	661.87	653.69	0.318	-1.49	-8.18	*
48	**46.58**	0.057	667.64	659.51	0.316	-1.42	-8.13	*
49	47.64	0.054	673.66	665.58	0.314	-1.36	-8.08	*
50	48.72	0.052	679.99	671.95	0.312	-1.28	-8.04	*
51	49.79	0.049	686.68	678.69	0.310	-1.21	-8.00	*
52	50.88	0.046	693.82	685.86	0.307	-1.12	-7.96	*
53	51.97	0.042	701.54	693.59	0.307	-1.03	-7.95	*
54	53.06	0.039	709.98	702.07	0.326	-0.94	-7.90	*
55	54.17	0.035	719.38	711.55	0.326	-0.83	-7.83	*
56	55.28	0.030	730.11	722.37	0.325	-0.72	-7.74	*
57	56.40	0.025	742.77	735.17	0.322	-0.60	-7.60	*
58	57.53	0.020	758.51	751.13	0.316	-0.47	-7.38	
59	58.67	0.014	779.87	772.90	0.302	-0.33	-6.97	

RAW	PAP_RS	RS_SD	CBT_SS	PAP_SS	SS_SD	RS_DIF	SS_DIF	SIG
60	59.83	0.007	815.02	809.03	0.262	-0.17	-6.00	
61	60.65	0.002	863.26	846.24	0.113	-0.35	-17.02	

RAW - Online test raw score
PAP_RS - Equivalent raw scores on the online test based on the comparability linking.
  Note that a lower equivalent raw score indicates that the online version of the test
  was more difficult.
RS_SD - Standard deviation of the equivalent raw scores over the replications.
CBT_SS - Online test scale score conversions, based on the 2008 TELPAS post-equated scales.
PAP_SS - Equivalent scale scores on the online test based on the comparability linking.
  Again, lower equivalent scale scores indicate that the online version of the test was
  more difficult.
SS_SD - Standard deviation of the equivalent scale scores over the replications.
RS_DIF - Difference between paper raw score equivalent and online raw score.
SS_DIF - Difference between paper scale score equivalent and online scale score.
SIG^- Raw score points where scale score differences exceed two standard errors
  of the linking and where the difference in raw scores is greater than half a point
  are noted by "*".

^ Note that the "*" in the SIG column now indicates both statistical and practical
  significance, based on recommendations from the Texas Technical Advisory Committee.

Appendix 4. Summary of Phase One Comparability Analysis – Grade Cluster 6–7

RAW	PAP_RS	RS_SD	CBT_SS	PAP_SS	SS_SD	RS_DIF	SS_DIF	SIG
0	0.34	0.010	263.94	280.65	0.513	0.34	16.71	
1	1.10	0.029	313.57	317.16	1.071	0.10	3.59	
2	2.14	0.047	350.79	354.03	1.093	0.14	3.25	
3	3.15	0.060	374.07	376.72	1.046	0.15	2.65	
4	4.14	0.069	391.59	393.66	0.993	0.14	2.07	
5	5.13	0.076	405.93	407.48	0.948	0.13	1.55	
6	6.10	0.082	418.25	419.34	0.908	0.10	1.09	
7	7.07	0.086	429.15	429.84	0.871	0.07	0.69	
8	8.04	0.089	439.02	439.36	0.835	0.04	0.34	
9	9.01	0.091	448.10	448.13	0.800	0.01	0.04	
10	9.97	0.093	456.54	456.31	0.766	-0.03	-0.23	
11	10.94	0.093	464.47	463.99	0.732	-0.06	-0.47	
12	11.91	0.094	471.96	471.28	0.701	-0.09	-0.68	
13	12.88	0.094	479.11	478.23	0.671	-0.12	-0.88	
14	13.85	0.094	485.93	484.89	0.642	-0.15	-1.05	
15	14.82	0.094	492.48	491.29	0.616	-0.18	-1.20	
16	15.79	0.094	498.80	497.47	0.594	-0.21	-1.34	
17	16.76	0.094	504.91	503.44	0.573	-0.24	-1.47	
18	17.73	0.094	510.83	509.25	0.555	-0.27	-1.58	
19	18.71	0.094	516.58	514.89	0.539	-0.29	-1.69	
20	19.68	0.094	522.19	520.40	0.524	-0.32	-1.79	
21	20.66	0.094	527.65	525.77	0.512	-0.34	-1.88	
22	21.63	0.094	533.00	531.03	0.502	-0.37	-1.97	
23	22.61	0.094	538.23	536.19	0.493	-0.39	-2.05	
24	23.59	0.094	543.36	541.25	0.484	-0.41	-2.12	
25	24.57	0.095	548.41	546.22	0.478	-0.43	-2.19	
26	**25.55**	0.095	553.37	551.12	0.471	-0.45	-2.25	
27	26.53	0.095	558.27	555.96	0.467	-0.47	-2.31	
28	27.51	0.096	563.10	560.73	0.462	-0.49	-2.36	
29	28.49	0.096	567.87	565.46	0.459	-0.51	-2.41	*
30	29.48	0.096	572.60	570.14	0.455	-0.52	-2.46	*
31	30.47	0.096	577.29	574.79	0.452	-0.53	-2.50	*
32	31.45	0.096	581.96	579.41	0.450	-0.55	-2.55	*
33	32.44	0.096	586.59	584.01	0.447	-0.56	-2.59	*
34	33.43	0.096	591.21	588.59	0.446	-0.57	-2.62	*
35	34.42	0.096	595.83	593.17	0.444	-0.58	-2.66	*
36	35.42	0.096	600.44	597.76	0.442	-0.58	-2.69	*
37	36.41	0.095	605.07	602.35	0.441	-0.59	-2.72	*
38	37.41	0.095	609.71	606.96	0.439	-0.59	-2.75	*
39	38.40	0.094	614.37	611.60	0.438	-0.60	-2.77	*
40	**39.40**	0.093	619.07	616.27	0.436	-0.60	-2.80	*
41	40.40	0.092	623.81	620.98	0.434	-0.60	-2.82	*
42	41.41	0.090	628.60	625.76	0.434	-0.59	-2.85	*
43	42.41	0.089	633.46	630.60	0.432	-0.59	-2.87	*
44	43.42	0.087	638.40	635.52	0.431	-0.58	-2.89	*
45	44.42	0.085	643.44	640.54	0.430	-0.58	-2.91	*
46	45.43	0.083	648.59	645.67	0.428	-0.57	-2.92	*
47	46.44	0.081	653.86	650.93	0.427	-0.56	-2.94	*
48	47.46	0.079	659.30	656.34	0.427	-0.54	-2.95	*
49	48.47	0.076	664.90	661.94	0.424	-0.53	-2.96	*
50	49.49	0.073	670.73	667.75	0.424	-0.51	-2.98	*
51	**50.51**	0.070	676.81	673.83	0.423	-0.49	-2.99	
52	51.53	0.066	683.19	680.20	0.422	-0.47	-2.99	
53	52.56	0.062	689.95	686.95	0.420	-0.44	-3.00	
54	53.58	0.058	697.15	694.15	0.419	-0.42	-2.99	
55	54.62	0.054	704.92	701.93	0.418	-0.38	-2.99	
56	55.65	0.049	713.42	710.44	0.416	-0.35	-2.98	
57	56.69	0.044	722.88	719.92	0.413	-0.31	-2.96	
58	57.73	0.038	733.66	730.73	0.409	-0.27	-2.93	
59	58.77	0.032	746.38	743.50	0.403	-0.23	-2.88	

RAW	PAP_RS	RS_SD	CBT_SS	PAP_SS	SS_SD	RS_DIF	SS_DIF	SIG
60	59.82	0.025	762.18	759.38	0.393	-0.18	-2.80	
61	60.88	0.017	783.59	780.96	0.371	-0.12	-2.64	
62	61.94	0.009	818.79	816.53	0.320	-0.06	-2.26	
63	62.68	0.003	867.06	851.62	0.136	-0.32	-15.44	

RAW - Online test raw score
PAP_RS - Equivalent raw scores on the online test based on the comparability linking.
  Note that a lower equivalent raw score indicates that the online version of the test
  was more difficult.
RS_SD - Standard deviation of the equivalent raw scores over the replications.
CBT_SS - Online test scale score conversions, based on the 2008 TELPAS post-equated scales.
PAP_SS - Equivalent scale scores on the online test based on the comparability linking.
  Again, lower equivalent scale scores indicate that the online version of the test was
  more difficult.
SS_SD - Standard deviation of the equivalent scale scores over the replications.
RS_DIF - Difference between paper raw score equivalent and online raw score.
SS_DIF - Difference between paper scale score equivalent and online scale score.
SIG^- Raw score points where scale score differences exceed two standard errors
  of the linking and where the difference in raw scores is greater than half a point
  are noted by "*".

^ Note that the "*" in the SIG column now indicates both statistical and practical
  significance, based on recommendations from the Texas Technical Advisory Committee.

Appendix 5. Summary of Phase One Comparability Analysis – Grade Cluster 8–9

RAW	PAP_RS	RS_SD	CBT_SS	PAP_SS	SS_SD	RS_DIF	SS_DIF	SIG
0	0.30	0.007	273.34	287.91	0.336	0.30	14.57	
1	0.99	0.021	322.32	321.66	0.940	-0.01	-0.65	
2	1.97	0.038	358.66	357.35	1.252	-0.03	-1.32	
3	2.94	0.052	381.25	379.76	1.123	-0.06	-1.49	
4	3.90	0.063	398.22	396.46	1.048	-0.10	-1.75	
5	4.85	0.071	412.11	410.06	0.987	-0.15	-2.05	
6	5.80	0.078	424.07	421.71	0.932	-0.20	-2.36	
7	6.75	0.083	434.69	432.02	0.884	-0.25	-2.67	
8	7.69	0.087	444.33	441.34	0.841	-0.31	-2.99	
9	8.63	0.090	453.21	449.92	0.802	-0.37	-3.29	
10	9.57	0.093	461.48	457.90	0.768	-0.43	-3.59	
11	10.50	0.095	469.26	465.39	0.737	-0.50	-3.86	
12	11.44	0.096	476.60	472.48	0.709	-0.56	-4.13	*
13	12.37	0.098	483.60	479.23	0.684	-0.63	-4.37	*
14	13.31	0.099	490.27	485.68	0.661	-0.69	-4.59	*
15	14.25	0.100	496.67	491.87	0.642	-0.75	-4.80	*
16	15.19	0.101	502.83	497.84	0.624	-0.81	-4.99	*
17	16.13	0.102	508.76	503.61	0.608	-0.87	-5.15	*
18	17.08	0.104	514.49	509.19	0.597	-0.92	-5.29	*
19	18.02	0.105	520.04	514.62	0.588	-0.98	-5.42	*
20	18.98	0.106	525.43	519.90	0.580	-1.02	-5.53	*
21	19.93	0.107	530.68	525.05	0.572	-1.07	-5.62	*
22	20.89	0.108	535.79	530.09	0.564	-1.11	-5.70	*
23	21.85	0.109	540.78	535.02	0.557	-1.15	-5.76	*
24	22.82	0.110	545.67	539.86	0.550	-1.18	-5.81	*
25	23.79	0.111	550.45	544.62	0.544	-1.21	-5.84	*
26	24.76	0.113	555.16	549.30	0.538	-1.24	-5.86	*
27	25.74	0.113	559.78	553.91	0.534	-1.26	-5.87	*
28	**26.72**	0.114	564.34	558.46	0.528	-1.28	-5.87	*
29	27.70	0.115	568.84	562.97	0.525	-1.30	-5.87	*
30	28.69	0.116	573.29	567.43	0.521	-1.31	-5.86	*
31	29.68	0.116	577.70	571.86	0.518	-1.32	-5.84	*
32	30.67	0.117	582.07	576.26	0.515	-1.33	-5.81	*
33	31.67	0.117	586.42	580.64	0.512	-1.33	-5.78	*
34	32.67	0.117	590.75	585.00	0.510	-1.33	-5.75	*
35	33.68	0.117	595.07	589.36	0.508	-1.32	-5.72	*
36	34.69	0.117	599.39	593.72	0.507	-1.31	-5.67	*
37	35.70	0.117	603.71	598.08	0.505	-1.30	-5.63	*
38	**36.71**	0.117	608.05	602.46	0.504	-1.29	-5.59	*
39	37.73	0.116	612.42	606.87	0.502	-1.27	-5.55	*
40	38.75	0.115	616.81	611.31	0.502	-1.25	-5.50	*
41	39.77	0.114	621.25	615.79	0.500	-1.23	-5.46	*
42	40.79	0.113	625.74	620.33	0.500	-1.21	-5.41	*
43	41.82	0.111	630.30	624.93	0.499	-1.18	-5.37	*
44	42.85	0.109	634.94	629.62	0.498	-1.15	-5.32	*
45	43.88	0.107	639.68	634.40	0.498	-1.12	-5.28	*
46	44.92	0.104	644.53	639.30	0.497	-1.08	-5.23	*
47	45.96	0.102	649.52	644.33	0.498	-1.04	-5.18	*
48	47.00	0.099	654.66	649.52	0.500	-1.00	-5.14	*
49	48.05	0.095	659.99	654.90	0.502	-0.95	-5.08	*
50	49.09	0.091	665.54	660.50	0.504	-0.91	-5.04	*
51	**50.14**	0.087	671.36	666.38	0.507	-0.86	-4.98	*
52	51.20	0.083	677.49	672.57	0.506	-0.80	-4.92	*
53	52.25	0.078	684.01	679.15	0.505	-0.75	-4.86	*
54	53.31	0.072	690.99	686.21	0.504	-0.69	-4.78	*
55	54.38	0.066	698.57	693.86	0.502	-0.62	-4.71	*
56	55.45	0.060	706.90	702.28	0.499	-0.55	-4.62	*
57	56.52	0.053	716.23	711.71	0.495	-0.48	-4.51	
58	57.59	0.046	726.92	722.53	0.489	-0.41	-4.39	
59	58.67	0.038	739.58	735.36	0.480	-0.33	-4.22	

RAW	PAP_RS	RS_SD	CBT_SS	PAP_SS	SS_SD	RS_DIF	SS_DIF	SIG
60	59.75	0.030	755.37	751.37	0.466	-0.25	-4.01	
61	60.83	0.021	776.86	773.18	0.441	-0.17	-3.68	
62	61.91	0.011	812.21	809.16	0.382	-0.09	-3.06	
63	62.67	0.003	860.63	844.84	0.164	-0.33	-15.79	

RAW - Online test raw score
PAP_RS - Equivalent raw scores on the online test based on the comparability linking.
  Note that a lower equivalent raw score indicates that the online version of the test
  was more difficult.
RS_SD - Standard deviation of the equivalent raw scores over the replications.
CBT_SS - Online test scale score conversions, based on the 2008 TELPAS post-equated scales.
PAP_SS - Equivalent scale scores on the online test based on the comparability linking.
  Again, lower equivalent scale scores indicate that the online version of the test was
  more difficult.
SS_SD - Standard deviation of the equivalent scale scores over the replications.
RS_DIF - Difference between paper raw score equivalent and online raw score.
SS_DIF - Difference between paper scale score equivalent and online scale score.
SIG^- Raw score points where scale score differences exceed two standard errors
  of the linking and where the difference in raw scores is greater than half a point
  are noted by "*".

^ Note that the "*" in the SIG column now indicates both statistical and practical
  significance, based on recommendations from the Texas Technical Advisory Committee.

Appendix 6. Summary of Phase One Comparability Analysis – Grade Cluster 10–12

RAW	PAP_RS	RS_SD	CBT_SS	PAP_SS	SS_SD	RS_DIF	SS_DIF	SIG
0	0.35	0.013	263.09	280.12	0.638	0.35	17.03	
1	1.13	0.038	312.40	317.29	1.383	0.13	4.90	
2	2.22	0.063	349.25	354.37	1.449	0.22	5.12	
3	3.28	0.081	372.29	377.16	1.401	0.28	4.87	
4	4.31	0.093	389.68	394.18	1.331	0.31	4.50	
5	5.33	0.102	403.98	408.07	1.259	0.33	4.10	
6	6.34	0.108	416.31	420.00	1.190	0.34	3.69	
7	7.33	0.113	427.28	430.57	1.128	0.33	3.30	
8	8.32	0.116	437.24	440.17	1.071	0.32	2.92	
9	9.30	0.119	446.45	449.00	1.019	0.30	2.56	
10	10.27	0.120	455.03	457.24	0.972	0.27	2.21	
11	11.25	0.121	463.10	464.99	0.930	0.25	1.89	
12	12.22	0.122	470.75	472.34	0.892	0.22	1.59	
13	13.19	0.123	478.03	479.33	0.858	0.19	1.30	
14	14.16	0.124	484.98	486.02	0.828	0.16	1.04	
15	15.12	0.124	491.66	492.45	0.802	0.12	0.79	
16	16.09	0.125	498.08	498.64	0.779	0.09	0.57	
17	17.06	0.126	504.27	504.63	0.760	0.06	0.36	
18	18.03	0.126	510.27	510.43	0.742	0.03	0.16	
19	19.00	0.127	516.08	516.06	0.725	-0.00	-0.01	
20	19.97	0.127	521.72	521.55	0.711	-0.03	-0.17	
21	20.94	0.128	527.21	526.89	0.697	-0.06	-0.32	
22	21.92	0.129	532.56	532.11	0.684	-0.08	-0.45	
23	22.89	0.129	537.80	537.23	0.673	-0.11	-0.57	
24	23.87	0.130	542.92	542.24	0.661	-0.13	-0.68	
25 →	**24.85**	0.130	547.94	547.16	0.649	-0.15	-0.77	
26	25.82	0.130	552.87	552.01	0.641	-0.18	-0.86	
27	26.81	0.130	557.72	556.78	0.631	-0.19	-0.94	
28	27.79	0.130	562.50	561.48	0.621	-0.21	-1.01	
29	28.77	0.130	567.20	566.13	0.612	-0.23	-1.08	
30	29.76	0.130	571.86	570.73	0.605	-0.24	-1.14	
31	30.74	0.130	576.47	575.28	0.597	-0.26	-1.19	
32	31.73	0.129	581.03	579.80	0.589	-0.27	-1.23	
33	32.72	0.129	585.56	584.29	0.583	-0.28	-1.27	
34	33.71	0.128	590.07	588.76	0.576	-0.29	-1.31	
35	34.70	0.127	594.56	593.21	0.571	-0.30	-1.35	
36	35.69	0.126	599.03	597.65	0.565	-0.31	-1.38	
37 →	36.69	0.125	603.49	602.09	0.558	-0.31	-1.40	
38 →	**37.68**	0.124	607.96	606.54	0.554	-0.32	-1.43	
39	38.68	0.122	612.44	610.99	0.548	-0.32	-1.45	
40	39.67	0.121	616.94	615.48	0.544	-0.33	-1.47	
41	40.67	0.119	621.47	619.99	0.539	-0.33	-1.48	
42	41.67	0.117	626.03	624.53	0.534	-0.33	-1.50	
43	42.67	0.115	630.65	629.13	0.531	-0.33	-1.51	
44	43.67	0.113	635.32	633.80	0.526	-0.33	-1.52	
45	44.68	0.110	640.06	638.53	0.523	-0.32	-1.53	
46	45.68	0.107	644.90	643.36	0.519	-0.32	-1.54	
47	46.69	0.104	649.84	648.29	0.515	-0.31	-1.54	
48 →	**47.69**	0.101	654.90	653.35	0.512	-0.31	-1.55	
49	48.70	0.098	660.11	658.56	0.508	-0.30	-1.55	
50	49.71	0.094	665.50	663.95	0.505	-0.29	-1.55	
51	50.72	0.090	671.10	669.55	0.501	-0.28	-1.55	
52	51.74	0.085	676.95	675.40	0.498	-0.26	-1.55	
53	52.75	0.080	683.10	681.56	0.494	-0.25	-1.55	
54	53.76	0.075	689.62	688.08	0.490	-0.24	-1.54	
55	54.78	0.070	696.59	695.06	0.486	-0.22	-1.53	
56	55.80	0.064	704.12	702.60	0.482	-0.20	-1.52	
57	56.82	0.058	712.39	710.89	0.476	-0.18	-1.50	
58	57.84	0.051	721.63	720.14	0.470	-0.16	-1.49	
59	58.86	0.044	732.19	730.73	0.463	-0.14	-1.46	

RAW	PAP_RS	RS_SD	CBT_SS	PAP_SS	SS_SD	RS_DIF	SS_DIF	SIG
60	59.89	0.036	744.68	743.26	0.452	-0.11	-1.42	
61	60.91	0.028	760.27	758.90	0.437	-0.09	-1.37	
62	61.94	0.019	781.48	780.20	0.410	-0.06	-1.28	
63	62.97	0.010	816.48	815.40	0.350	-0.03	-1.08	
64	63.69	0.003	864.64	849.74	0.148	-0.31	-14.90	

RAW - Online test raw score
PAP_RS - Equivalent raw scores on the online test based on the comparability linking. Note that a lower equivalent raw score indicates that the online version of the test was more difficult.
RS_SD - Standard deviation of the equivalent raw scores over the replications.
CBT_SS - Online test scale score conversions, based on the 2008 TELPAS post-equated scales.
PAP_SS - Equivalent scale scores on the online test based on the comparability linking. Again, lower equivalent scale scores indicate that the online version of the test was more difficult.
SS_SD - Standard deviation of the equivalent scale scores over the replications.
RS_DIF - Difference between paper raw score equivalent and online raw score.
SS_DIF - Difference between paper scale score equivalent and online scale score.
SIG^- Raw score points where scale score differences exceed two standard errors of the linking and where the difference in raw scores is greater than half a point are noted by "*".

^ Note that the "*" in the SIG column now indicates both statistical and practical significance, based on recommendations from the Texas Technical Advisory Committee.

Appendix 7. Final Results of Comparability Analysis – Grade 2

RAW	PAP_RS	RS_SD	CBT_SS	PAP_SS	SS_SD	RS_DIF	SS_DIF	SIG
0	0.30	0.007	261.64	277.00	0.367	0.30	15.36	
1	0.99	0.020	312.84	312.14	0.973	-0.01	-0.70	
2	1.94	0.031	352.84	350.46	1.244	-0.06	-2.38	
3	2.87	0.037	379.11	375.63	0.979	-0.13	-3.48	
4	3.77	0.040	399.73	395.07	0.827	-0.23	-4.67	
5	4.66	0.041	417.16	411.30	0.718	-0.34	-5.86	
6	5.54	0.042	432.45	425.45	0.635	-0.46	-7.01	
7	6.41	0.042	446.16	438.09	0.569	-0.59	-8.07	*
8	7.27	0.042	458.61	449.58	0.519	-0.73	-9.03	*
9	8.13	0.042	470.01	460.14	0.479	-0.87	-9.87	*
10	8.99	0.043	480.53	469.90	0.471	-1.01	-10.63	*
11	9.85	0.044	490.30	478.95	0.460	-1.15	-11.35	*
12	10.71	0.045	499.41	487.46	0.439	-1.29	-11.95	*
13	11.57	0.046	507.97	495.53	0.423	-1.43	-12.44	*
14	12.44	0.048	516.04	503.20	0.411	-1.56	-12.85	*
15	13.32	0.050	523.69	510.53	0.402	-1.68	-13.17	*
16	14.20	0.052	530.98	517.56	0.394	-1.80	-13.42	*
17	15.09	0.053	537.95	524.34	0.389	-1.91	-13.62	*
18	15.99	0.055	544.65	530.88	0.395	-2.01	-13.77	*
19	16.89	0.057	551.12	537.22	0.395	-2.11	-13.90	*
20	17.81	0.058	557.39	543.40	0.391	-2.19	-14.00	*
21	18.74	0.060	563.49	549.44	0.387	-2.26	-14.06	*
22	19.68	0.061	569.45	**555.37**	0.384	-2.32	-14.08	*
23	20.62	0.062	575.29	561.21	0.381	-2.38	-14.08	*
24	21.58	0.063	581.04	566.97	0.378	-2.42	-14.07	*
25	22.55	0.064	586.71	572.67	0.375	-2.45	-14.04	*
26	23.53	0.065	592.33	578.34	0.373	-2.47	-14.00	*
27	24.52	0.065	597.92	583.98	0.370	-2.48	-13.93	*
28	25.52	0.065	603.49	589.62	0.368	-2.48	-13.87	*
29	26.53	0.065	609.06	595.27	0.365	-2.47	-13.79	*
30	27.54	0.065	614.66	600.95	0.363	-2.46	-13.71	*
31	28.57	0.065	620.31	606.68	0.361	-2.43	-13.63	*
32	29.61	0.064	626.01	612.48	0.358	-2.39	-13.54	*
33	30.66	0.063	631.81	**618.37**	0.356	-2.34	-13.44	*
34	31.71	0.062	637.72	624.38	0.353	-2.29	-13.33	*
35	32.78	0.061	643.78	630.54	0.351	-2.22	-13.24	*
36	33.86	0.059	650.01	636.87	0.347	-2.14	-13.14	*
37	34.94	0.057	656.48	643.43	0.346	-2.06	-13.05	*
38	36.04	0.055	663.22	650.27	0.350	-1.96	-12.95	*
39	37.15	0.052	670.30	657.47	0.351	-1.85	-12.83	*
40	38.27	0.049	677.82	665.10	0.348	-1.73	-12.71	*
41	39.40	0.046	685.88	**673.28**	0.346	-1.60	-12.59	*
42	40.54	0.042	694.64	682.18	0.341	-1.46	-12.46	*
43	41.70	0.038	704.36	692.00	0.335	-1.30	-12.36	*
44	42.87	0.034	715.37	703.10	0.328	-1.13	-12.27	*
45	44.06	0.029	728.30	716.14	0.369	-0.94	-12.16	*
46	45.27	0.023	744.29	732.55	0.364	-0.73	-11.74	*
47	46.49	0.016	765.90	754.88	0.349	-0.51	-11.02	*
48	47.73	0.009	801.26	791.87	0.305	-0.27	-9.39	
49	48.62	0.003	849.64	831.16	0.131	-0.38	-18.48	

RAW - Online test raw score
PAP_RS - Equivalent raw scores on the online test based on the comparability linking.
  Note that a lower equivalent raw score indicates that the online version of the test
  was more difficult.
RS_SD - Standard deviation of the equivalent raw scores over the replications.
CBT_SS - Online test scale score conversions, based on the 2008 TELPAS post-equated scales.
PAP_SS - Equivalent scale scores on the online test based on the comparability linking.
  Again, lower equivalent scale scores indicate that the online version of the test was
  more difficult.

SS_SD - Standard deviation of the equivalent scale scores over the replications.
RS_DIF - Difference between paper raw score equivalent and online raw score.
SS_DIF - Difference between paper scale score equivalent and online scale score.
SIG^- Raw score points where scale score differences exceed two standard errors of the linking and where the difference in raw scores is greater than half a point are noted by "*".

^ Note that the "*" in the SIG column now indicates both statistical and practical significance, based on recommendations from the Texas Technical Advisory Committee.

Appendix 8. Final Results of Comparability Analysis – Grade 3

RAW	PAP_RS	RS_SD	CBT_SS	PAP_SS	SS_SD	RS_DIF	SS_DIF	SIG
0	0.26	0.006	309.66	322.61	0.302	0.26	12.99	
1	0.87	0.018	359.49	353.01	0.894	-0.13	-6.37	
2	1.75	0.031	397.15	387.74	1.149	-0.25	-9.40	
3	2.63	0.039	420.97	412.16	0.937	-0.37	-8.77	
4	3.51	0.046	439.06	430.19	0.825	-0.49	-8.79	
5	4.40	0.050	453.97	445.02	0.747	-0.60	-8.99	*
6	5.28	0.053	466.83	457.57	0.687	-0.72	-9.26	*
7	6.16	0.056	478.22	468.65	0.638	-0.84	-9.55	*
8	7.04	0.058	488.52	478.63	0.613	-0.96	-9.85	*
9	7.93	0.060	497.94	487.80	0.614	-1.07	-10.17	*
10	8.81	0.062	506.66	496.15	0.584	-1.19	-10.49	*
11	9.70	0.064	514.79	504.05	0.556	-1.30	-10.75	*
12	10.59	0.066	522.42	511.46	0.533	-1.41	-10.98	*
13	11.48	0.068	529.62	518.45	0.515	-1.52	-11.18	*
14	12.37	0.069	536.45	525.08	0.500	-1.63	-11.34	*
15	13.27	0.071	542.96	531.46	0.488	-1.73	-11.47	*
16	14.18	0.073	549.19	537.62	0.478	-1.82	-11.59	*
17	15.09	0.075	555.18	543.52	0.472	-1.91	-11.69	*
18	16.00	0.077	560.95	549.19	0.471	-2.00	-11.77	*
19	16.92	0.079	566.53	554.70	0.469	-2.08	-11.85	*
20	17.84	0.080	571.94	560.03	0.464	-2.16	-11.90	*
21	18.77	0.082	577.20	565.24	0.457	-2.23	-11.95	*
22	19.71	0.083	582.33	570.37	0.452	-2.29	-11.98	*
23	20.65	0.085	587.35	575.36	0.446	-2.35	-12.00	*
24	21.60	0.086	592.27	580.28	0.441	-2.40	-12.01	*
25	22.55	0.087	597.11	585.09	0.437	-2.45	-12.00	*
26	23.51	0.088	601.86	589.86	0.432	-2.49	-11.99	*
27	24.48	0.089	606.56	594.59	0.429	-2.52	-11.96	*
28	25.46	0.089	611.21	**599.30**	0.424	-2.54	-11.93	*
29	26.44	0.090	615.82	603.93	0.422	-2.56	-11.89	*
30	27.43	0.090	620.39	608.56	0.418	-2.57	-11.84	*
31	28.42	0.090	624.95	613.15	0.415	-2.58	-11.78	*
32	29.43	0.090	629.50	617.79	0.412	-2.57	-11.72	*
33	30.44	0.090	634.04	622.40	0.409	-2.56	-11.65	*
34	31.46	0.089	638.60	627.04	0.406	-2.54	-11.57	*
35	32.48	0.089	643.17	631.68	0.403	-2.52	-11.48	*
36	33.52	0.088	647.78	636.41	0.401	-2.48	-11.39	*
37	34.56	0.087	652.44	641.16	0.397	-2.44	-11.29	*
38	35.61	0.086	657.15	645.98	0.395	-2.39	-11.18	*
39	36.66	0.084	661.93	**650.85**	0.392	-2.34	-11.07	*
40	37.72	0.083	666.80	655.83	0.389	-2.28	-10.95	*
41	38.80	0.081	671.78	660.97	0.386	-2.20	-10.83	*
42	39.87	0.078	676.89	666.17	0.382	-2.13	-10.71	*
43	40.96	0.076	682.14	671.58	0.381	-2.04	-10.57	*
44	42.05	0.073	687.58	677.15	0.383	-1.95	-10.44	*
45	43.15	0.070	693.23	682.96	0.383	-1.85	-10.29	*
46	44.25	0.067	699.14	689.00	0.380	-1.75	-10.13	*
47	45.36	0.064	705.35	695.36	0.377	-1.64	-9.97	*
48	46.48	0.060	711.94	**702.12**	0.373	-1.52	-9.81	*
49	47.61	0.056	718.99	709.37	0.368	-1.39	-9.64	*
50	48.74	0.051	726.62	717.16	0.362	-1.26	-9.48	*
51	49.87	0.047	734.98	725.63	0.356	-1.13	-9.33	*
52	51.02	0.041	744.32	735.17	0.374	-0.98	-9.18	*
53	52.17	0.036	755.00	746.14	0.384	-0.83	-8.91	*
54	53.32	0.030	767.63	759.04	0.378	-0.68	-8.58	*
55	54.48	0.023	783.35	775.18	0.368	-0.52	-8.15	*
56	55.65	0.016	804.72	797.24	0.350	-0.35	-7.51	
57	56.82	0.009	839.90	833.56	0.302	-0.18	-6.27	
58	57.65	0.003	888.17	871.27	0.129	-0.35	-17.09	

RAW - Online test raw score
PAP_RS - Equivalent raw scores on the online test based on the comparability linking.
  Note that a lower equivalent raw score indicates that the online version of the test
  was more difficult.
RS_SD - Standard deviation of the equivalent raw scores over the replications.
CBT_SS - Online test scale score conversions, based on the 2008 TELPAS post-equated scales.
PAP_SS - Equivalent scale scores on the online test based on the comparability linking.
  Again, lower equivalent scale scores indicate that the online version of the test was
  more difficult.
SS_SD - Standard deviation of the equivalent scale scores over the replications.
RS_DIF - Difference between paper raw score equivalent and online raw score.
SS_DIF - Difference between paper scale score equivalent and online scale score.
SIG^- Raw score points where scale score differences exceed two standard errors
  of the linking and where the difference in raw scores is greater than half a point
  are noted by "*".

^ Note that the "*" in the SIG column now indicates both statistical and practical
  significance, based on recommendations from the Texas Technical Advisory Committee.

Appendix 9. Final Results of Comparability Analysis – Grade Cluster 4–5

RAW	PAP_RS	RS_SD	CBT_SS	PAP_SS	SS_SD	RS_DIF	SS_DIF	SIG
0	0.31	0.007	321.47	336.75	0.332	0.31	15.14	
1	1.02	0.020	370.75	371.49	0.803	0.02	0.55	
2	2.01	0.036	407.59	407.82	1.015	0.01	0.12	
3	2.99	0.047	430.67	430.44	0.982	-0.01	-0.35	
4	3.95	0.055	448.12	447.25	0.942	-0.05	-0.88	
5	4.90	0.062	462.52	461.08	0.889	-0.10	-1.44	
6	5.84	0.066	474.99	473.00	0.829	-0.16	-2.01	
7	6.77	0.070	486.13	483.57	0.777	-0.23	-2.58	
8	7.69	0.072	496.29	493.14	0.731	-0.31	-3.14	
9	8.61	0.073	505.69	502.02	0.691	-0.39	-3.69	
10	9.52	0.074	514.48	510.26	0.654	-0.48	-4.23	
11	10.43	0.075	522.79	518.05	0.621	-0.57	-4.74	*
12	11.34	0.075	530.66	525.46	0.590	-0.66	-5.22	*
13	12.24	0.075	538.17	532.46	0.563	-0.76	-5.69	*
14	13.15	0.075	545.36	539.25	0.537	-0.85	-6.11	*
15	14.06	0.074	552.26	545.77	0.518	-0.94	-6.50	*
16	14.97	0.074	558.89	552.05	0.506	-1.03	-6.87	*
17	15.88	0.074	565.30	558.10	0.491	-1.12	-7.20	*
18	16.79	0.074	571.49	563.95	0.473	-1.21	-7.51	*
19	17.71	0.074	577.48	569.69	0.456	-1.29	-7.77	*
20	18.64	0.074	583.31	575.33	0.442	-1.36	-8.00	*
21	19.57	0.074	588.96	580.80	0.429	-1.43	-8.19	*
22	20.50	0.074	594.48	586.13	0.417	-1.50	-8.36	*
23	21.44	0.074	599.86	591.39	0.409	-1.56	-8.50	*
24	22.38	0.074	605.12	596.53	0.399	-1.62	-8.60	*
25	23.33	0.074	610.28	601.60	0.391	-1.67	-8.70	*
26	24.28	0.075	615.34	606.57	0.384	-1.72	-8.77	*
27	25.24	0.075	620.31	**611.49**	0.378	-1.76	-8.82	*
28	26.21	0.075	625.22	616.38	0.373	-1.79	-8.86	*
29	27.18	0.075	630.07	621.20	0.367	-1.82	-8.89	*
30	28.15	0.075	634.85	625.95	0.363	-1.85	-8.90	*
31	29.13	0.075	639.60	630.69	0.358	-1.87	-8.90	*
32	30.12	0.075	644.31	635.42	0.355	-1.88	-8.89	*
33	31.11	0.075	649.00	640.12	0.351	-1.89	-8.87	*
34	32.11	0.074	653.67	644.83	0.348	-1.89	-8.85	*
35	33.11	0.074	658.34	649.51	0.345	-1.89	-8.81	*
36	34.12	0.073	663.01	654.23	0.342	-1.88	-8.78	*
37	35.13	0.073	667.69	658.94	0.339	-1.87	-8.73	*
38	36.15	0.072	672.39	663.71	0.337	-1.85	-8.69	*
39	37.17	0.071	677.13	**668.49**	0.335	-1.83	-8.64	*
40	38.20	0.070	681.92	673.34	0.332	-1.80	-8.58	*
41	39.23	0.069	686.76	678.23	0.330	-1.77	-8.53	*
42	40.27	0.068	691.68	683.23	0.328	-1.73	-8.47	*
43	41.31	0.066	696.68	688.29	0.326	-1.69	-8.41	*
44	42.35	0.065	701.79	693.43	0.323	-1.65	-8.35	*
45	43.40	0.063	707.03	698.73	0.322	-1.60	-8.30	*
46	44.45	0.061	712.41	704.15	0.320	-1.55	-8.24	*
47	45.51	0.059	717.98	709.78	0.318	-1.49	-8.18	*
48	46.58	0.057	723.74	715.64	0.316	-1.42	-8.13	*
49	47.64	0.054	729.77	**721.67**	0.314	-1.36	-8.08	*
50	48.72	0.052	736.10	728.08	0.312	-1.28	-8.04	*
51	49.79	0.049	742.79	734.77	0.310	-1.21	-8.00	*
52	50.88	0.046	749.93	741.98	0.307	-1.12	-7.96	*
53	51.97	0.042	757.64	749.71	0.307	-1.03	-7.95	*
54	53.06	0.039	766.08	758.15	0.326	-0.94	-7.90	*
55	54.17	0.035	775.48	767.68	0.326	-0.83	-7.83	*
56	55.28	0.030	786.21	778.49	0.325	-0.72	-7.74	*
57	56.40	0.025	798.87	791.28	0.322	-0.60	-7.60	*
58	57.53	0.020	814.61	807.22	0.316	-0.47	-7.38	
59	58.67	0.014	835.97	828.92	0.302	-0.33	-6.97	

RAW	PAP_RS	RS_SD	CBT_SS	PAP_SS	SS_SD	RS_DIF	SS_DIF	SIG
60	59.83	0.007	871.13	865.15	0.262	-0.17	-6.00	
61	60.65	0.002	919.37	902.48	0.113	-0.35	-17.02	

RAW - Online test raw score
PAP_RS - Equivalent raw scores on the online test based on the comparability linking.
  Note that a lower equivalent raw score indicates that the online version of the test
  was more difficult.
RS_SD - Standard deviation of the equivalent raw scores over the replications.
CBT_SS - Online test scale score conversions, based on the 2008 TELPAS post-equated scales.
PAP_SS - Equivalent scale scores on the online test based on the comparability linking.
  Again, lower equivalent scale scores indicate that the online version of the test was
  more difficult.
SS_SD - Standard deviation of the equivalent scale scores over the replications.
RS_DIF - Difference between paper raw score equivalent and online raw score.
SS_DIF - Difference between paper scale score equivalent and online scale score.
SIG^- Raw score points where scale score differences exceed two standard errors
  of the linking and where the difference in raw scores is greater than half a point
  are noted by "*".

^ Note that the "*" in the SIG column now indicates both statistical and practical
  significance, based on recommendations from the Texas Technical Advisory Committee.

Appendix 10. Final Results of Comparability Analysis – Grade Cluster 6–7

RAW	PAP_RS	RS_SD	CBT_SS	PAP_SS	SS_SD	RS_DIF	SS_DIF	SIG
0	0.34	0.010	323.84	340.71	0.513	0.34	16.71	
1	1.10	0.029	373.47	377.19	1.071	0.10	3.59	
2	2.14	0.047	410.69	413.95	1.093	0.14	3.25	
3	3.15	0.060	433.97	436.60	1.046	0.15	2.65	
4	4.14	0.069	451.49	453.50	0.993	0.14	2.07	
5	5.13	0.076	465.84	467.44	0.948	0.13	1.55	
6	6.10	0.082	478.15	479.24	0.908	0.10	1.09	
7	7.07	0.086	489.05	489.74	0.871	0.07	0.69	
8	8.04	0.089	498.92	499.28	0.835	0.04	0.34	
9	9.01	0.091	508.00	508.08	0.800	0.01	0.04	
10	9.97	0.093	516.44	516.19	0.766	-0.03	-0.23	
11	10.94	0.093	524.37	523.89	0.732	-0.06	-0.47	
12	11.91	0.094	531.86	531.19	0.701	-0.09	-0.68	
13	12.88	0.094	539.01	538.15	0.671	-0.12	-0.88	
14	13.85	0.094	545.83	544.81	0.642	-0.15	-1.05	
15	14.82	0.094	552.38	551.20	0.616	-0.18	-1.20	
16	15.79	0.094	558.71	557.38	0.594	-0.21	-1.34	
17	16.76	0.094	564.81	563.35	0.573	-0.24	-1.47	
18	17.73	0.094	570.73	569.13	0.555	-0.27	-1.58	
19	18.71	0.094	576.48	574.82	0.539	-0.29	-1.69	
20	19.68	0.094	582.09	580.29	0.524	-0.32	-1.79	
21	20.66	0.094	587.55	585.69	0.512	-0.34	-1.88	
22	21.63	0.094	592.90	590.92	0.502	-0.37	-1.97	
23	22.61	0.094	598.13	596.09	0.493	-0.39	-2.05	
24	23.59	0.094	603.26	601.16	0.484	-0.41	-2.12	
25	24.57	0.095	608.31	606.14	0.478	-0.43	-2.19	
26	25.55	0.095	613.27	611.04	0.471	-0.45	-2.25	
27	26.53	0.095	618.17	**615.87**	0.467	-0.47	-2.31	
28	27.51	0.096	623.00	620.63	0.462	-0.49	-2.36	
29	28.49	0.096	627.77	625.34	0.459	-0.51	-2.41	*
30	29.48	0.096	632.51	630.04	0.455	-0.52	-2.46	*
31	30.47	0.096	637.20	634.71	0.452	-0.53	-2.50	*
32	31.45	0.096	641.86	639.29	0.450	-0.55	-2.55	*
33	32.44	0.096	646.49	643.90	0.447	-0.56	-2.59	*
34	33.43	0.096	651.12	648.48	0.446	-0.57	-2.62	*
35	34.42	0.096	655.73	653.05	0.444	-0.58	-2.66	*
36	35.42	0.096	660.35	657.67	0.442	-0.58	-2.69	*
37	36.41	0.095	664.97	662.24	0.441	-0.59	-2.72	*
38	37.41	0.095	669.61	666.87	0.439	-0.59	-2.75	*
39	38.40	0.094	674.27	671.47	0.438	-0.60	-2.77	*
40	39.40	0.093	678.97	**676.15**	0.436	-0.60	-2.80	*
41	40.40	0.092	683.71	680.86	0.434	-0.60	-2.82	*
42	41.41	0.090	688.50	685.67	0.434	-0.59	-2.85	*
43	42.41	0.089	693.36	690.50	0.432	-0.59	-2.87	*
44	43.42	0.087	698.30	695.44	0.431	-0.58	-2.89	*
45	44.42	0.085	703.34	700.42	0.430	-0.58	-2.91	*
46	45.43	0.083	708.49	705.56	0.428	-0.57	-2.92	*
47	46.44	0.081	713.76	710.81	0.427	-0.56	-2.94	*
48	47.46	0.079	719.20	716.26	0.427	-0.54	-2.95	*
49	48.47	0.076	724.80	721.83	0.424	-0.53	-2.96	*
50	49.49	0.073	730.63	727.66	0.424	-0.51	-2.98	*
51	50.51	0.070	736.71	**733.73**	0.423	-0.49	-2.99	
52	51.53	0.066	743.09	740.09	0.422	-0.47	-2.99	
53	52.56	0.062	749.85	746.87	0.420	-0.44	-3.00	
54	53.58	0.058	757.05	754.02	0.419	-0.42	-2.99	
55	54.62	0.054	764.82	761.87	0.418	-0.38	-2.99	
56	55.65	0.049	773.32	770.35	0.416	-0.35	-2.98	
57	56.69	0.044	782.78	779.85	0.413	-0.31	-2.96	
58	57.73	0.038	793.56	790.65	0.409	-0.27	-2.93	
59	58.77	0.032	806.28	803.36	0.403	-0.23	-2.88	

RAW	PAP_RS	RS_SD	CBT_SS	PAP_SS	SS_SD	RS_DIF	SS_DIF	SIG
60	59.82	0.025	822.08	819.23	0.393	-0.18	-2.80	
61	60.88	0.017	843.49	840.92	0.371	-0.12	-2.64	
62	61.94	0.009	878.69	876.58	0.320	-0.06	-2.26	
63	62.68	0.003	926.96	911.51	0.136	-0.32	-15.44	

RAW - Online test raw score
PAP_RS - Equivalent raw scores on the online test based on the comparability linking.
  Note that a lower equivalent raw score indicates that the online version of the test
  was more difficult.
RS_SD - Standard deviation of the equivalent raw scores over the replications.
CBT_SS - Online test scale score conversions, based on the 2008 TELPAS post-equated scales.
PAP_SS - Equivalent scale scores on the online test based on the comparability linking.
  Again, lower equivalent scale scores indicate that the online version of the test was
  more difficult.
SS_SD - Standard deviation of the equivalent scale scores over the replications.
RS_DIF - Difference between paper raw score equivalent and online raw score.
SS_DIF - Difference between paper scale score equivalent and online scale score.
SIG^- Raw score points where scale score differences exceed two standard errors
  of the linking and where the difference in raw scores is greater than half a point
  are noted by "*".

^ Note that the "*" in the SIG column now indicates both statistical and practical
  significance, based on recommendations from the Texas Technical Advisory Committee.

Appendix 11. Final Results of Comparability Analysis – Grade Cluster 8–9

RAW	PAP_RS	RS_SD	CBT_SS	PAP_SS	SS_SD	RS_DIF	SS_DIF	SIG
0	0.30	0.007	350.91	365.61	0.336	0.30	14.57	
1	0.99	0.021	399.89	399.40	0.940	-0.01	-0.65	
2	1.97	0.038	436.24	435.15	1.252	-0.03	-1.32	
3	2.94	0.052	458.82	457.47	1.123	-0.06	-1.49	
4	3.90	0.063	475.79	474.09	1.048	-0.10	-1.75	
5	4.85	0.071	489.69	487.60	0.987	-0.15	-2.05	
6	5.80	0.078	501.65	499.26	0.932	-0.20	-2.36	
7	6.75	0.083	512.26	509.61	0.884	-0.25	-2.67	
8	7.69	0.087	521.90	518.92	0.841	-0.31	-2.99	
9	8.63	0.090	530.78	527.50	0.802	-0.37	-3.29	
10	9.57	0.093	539.06	535.50	0.768	-0.43	-3.59	
11	10.50	0.095	546.83	542.94	0.737	-0.50	-3.86	
12	11.44	0.096	554.18	550.06	0.709	-0.56	-4.13	*
13	12.37	0.098	561.17	556.77	0.684	-0.63	-4.37	*
14	13.31	0.099	567.84	563.24	0.661	-0.69	-4.59	*
15	14.25	0.100	574.25	569.44	0.642	-0.75	-4.80	*
16	15.19	0.101	580.40	575.42	0.624	-0.81	-4.99	*
17	16.13	0.102	586.33	581.17	0.608	-0.87	-5.15	*
18	17.08	0.104	592.06	586.79	0.597	-0.92	-5.29	*
19	18.02	0.105	597.61	592.17	0.588	-0.98	-5.42	*
20	18.98	0.106	603.00	597.50	0.580	-1.02	-5.53	*
21	19.93	0.107	608.25	602.63	0.572	-1.07	-5.62	*
22	20.89	0.108	613.36	607.67	0.564	-1.11	-5.70	*
23	21.85	0.109	618.35	612.60	0.557	-1.15	-5.76	*
24	22.82	0.110	623.24	617.46	0.550	-1.18	-5.81	*
25	23.79	0.111	628.03	622.21	0.544	-1.21	-5.84	*
26	24.76	0.113	632.73	626.88	0.538	-1.24	-5.86	*
27	25.74	0.113	637.35	631.51	0.534	-1.26	-5.87	*
28	26.72	0.114	641.91	636.06	0.528	-1.28	-5.87	*
29	27.70	0.115	646.41	640.54	0.525	-1.30	-5.87	*
30	28.69	0.116	650.86	645.02	0.521	-1.31	-5.86	*
31	29.68	0.116	655.27	649.44	0.518	-1.32	-5.84	*
32	30.67	0.117	659.64	653.82	0.515	-1.33	-5.81	*
33	31.67	0.117	663.99	658.20	0.512	-1.33	-5.78	*
34	32.67	0.117	668.32	662.56	0.510	-1.33	-5.75	*
35	33.68	0.117	672.65	666.94	0.508	-1.32	-5.72	*
36	34.69	0.117	676.96	671.31	0.507	-1.31	-5.67	*
37	35.70	0.117	681.29	675.67	0.505	-1.30	-5.63	*
38	36.71	0.117	685.63	680.03	0.504	-1.29	-5.59	*
39	37.73	0.116	689.99	684.45	0.502	-1.27	-5.55	*
40	38.75	0.115	694.38	688.90	0.502	-1.25	-5.50	*
41	39.77	0.114	698.82	693.37	0.500	-1.23	-5.46	*
42	40.79	0.113	703.31	697.89	0.500	-1.21	-5.41	*
43	41.82	0.111	707.87	702.51	0.499	-1.18	-5.37	*
44	42.85	0.109	712.52	707.19	0.498	-1.15	-5.32	*
45	43.88	0.107	717.25	711.96	0.498	-1.12	-5.28	*
46	44.92	0.104	722.10	716.87	0.497	-1.08	-5.23	*
47	45.96	0.102	727.09	721.91	0.498	-1.04	-5.18	*
48	47.00	0.099	732.23	727.09	0.500	-1.00	-5.14	*
49	48.05	0.095	737.56	732.50	0.502	-0.95	-5.08	*
50	49.09	0.091	743.12	738.06	0.504	-0.91	-5.04	*
51	50.14	0.087	748.93	743.93	0.507	-0.86	-4.98	*
52	51.20	0.083	755.07	750.16	0.506	-0.80	-4.92	*
53	52.25	0.078	761.58	756.70	0.505	-0.75	-4.86	*
54	53.31	0.072	768.57	763.75	0.504	-0.69	-4.78	*
55	54.38	0.066	776.14	771.45	0.502	-0.62	-4.71	*
56	55.45	0.060	784.48	779.89	0.499	-0.55	-4.62	*
57	56.52	0.053	793.80	789.32	0.495	-0.48	-4.51	
58	57.59	0.046	804.49	800.11	0.489	-0.41	-4.39	
59	58.67	0.038	817.16	812.98	0.480	-0.33	-4.22	

RAW	PAP_RS	RS_SD	CBT_SS	PAP_SS	SS_SD	RS_DIF	SS_DIF	SIG
60	59.75	0.030	832.95	829.00	0.466	-0.25	-4.01	
61	60.83	0.021	854.43	850.78	0.441	-0.17	-3.68	
62	61.91	0.011	889.78	886.60	0.382	-0.09	-3.06	
63	62.67	0.003	938.21	922.23	0.164	-0.33	-15.79	

RAW - Online test raw score
PAP_RS - Equivalent raw scores on the online test based on the comparability linking.
  Note that a lower equivalent raw score indicates that the online version of the test
  was more difficult.
RS_SD - Standard deviation of the equivalent raw scores over the replications.
CBT_SS - Online test scale score conversions, based on the 2008 TELPAS post-equated scales.
PAP_SS - Equivalent scale scores on the online test based on the comparability linking.
  Again, lower equivalent scale scores indicate that the online version of the test was
  more difficult.
SS_SD - Standard deviation of the equivalent scale scores over the replications.
RS_DIF - Difference between paper raw score equivalent and online raw score.
SS_DIF - Difference between paper scale score equivalent and online scale score.
SIG^- Raw score points where scale score differences exceed two standard errors
  of the linking and where the difference in raw scores is greater than half a point
  are noted by "*".

^ Note that the "*" in the SIG column now indicates both statistical and practical
  significance, based on recommendations from the Texas Technical Advisory Committee.

Appendix 12. Summary of Item-Level Analysis – Grade 2

inum	cbtpval	pappval	difpval	difstd	z_dif	isig	effect_size
1	0.98636	0.98448	0.00188	.000788564	2.3837	*	0.01568
2	0.98836	0.97507	0.01329	.000981242	13.5452	*	0.09932
3	0.96519	0.95992	0.00527	.001044975	5.0459	*	0.02777
4	0.81708	0.85180	-0.03471	.002499755	-13.8860	*	-0.09350
5	0.90428	0.91100	-0.00672	.001770270	-3.7982	*	-0.02322
6	0.97920	0.97305	0.00614	.001079391	5.6916	*	0.04025
7	0.89978	0.89785	0.00193	.002024838	0.9512		0.00639
8	0.92248	0.91549	0.00699	.001726267	4.0478	*	0.02562
9	0.97837	0.97627	0.00210	.000909014	2.3075	*	0.01408
10	0.78309	0.79501	-0.01192	.002381339	-5.0046	*	-0.02921
11	0.95497	0.95653	-0.00156	.001360327	-1.1441		-0.00757
12	0.90058	0.89105	0.00953	.002139625	4.4520	*	0.03119
13	0.60923	0.62747	-0.01825	.003054431	-5.9743	*	-0.03757
14	0.40988	0.46082	-0.05093	.003688496	-13.8088	*	-0.10287
15	0.47695	0.49584	-0.01889	.003221313	-5.8652	*	-0.03781
16	0.59189	0.62812	-0.03623	.003012232	-12.0281	*	-0.07433
17	0.42271	0.43157	-0.00886	.003028198	-2.9255	*	-0.01791
18	0.69595	0.72099	-0.02504	.002735060	-9.1558	*	-0.05512
19	0.60303	0.61168	-0.00865	.003429303	-2.5223	*	-0.01771
20	0.60636	0.67487	-0.06851	.003402204	-20.1369	*	-0.14315
21	0.45412	0.43522	0.01890	.003265113	5.7879	*	0.03804
22	0.37694	0.43047	-0.05352	.003402145	-15.7325	*	-0.10926
23	0.75055	0.83941	-0.08886	.002477131	-35.8720	*	-0.22145
24	0.65800	0.72329	-0.06529	.002688514	-24.2851	*	-0.14161
25	0.72552	0.81524	-0.08972	.002705401	-33.1624	*	-0.21454
26	0.79558	0.81968	-0.02410	.002885895	-8.3517	*	-0.06117
27	0.66985	0.71993	-0.05008	.002923590	-17.1294	*	-0.10892
28	0.33587	0.40324	-0.06736	.003603168	-18.6953	*	-0.13990

inum	cbtpval	pappval	difpval	difstd	z_dif	isig	effect_size
29	0.37960	0.43617	-0.05657	.003061751	-18.4777	*	-0.11531
30	0.52021	0.56633	-0.04612	.003415526	-13.5039	*	-0.09269
31	0.43024	0.44204	-0.01180	.002748679	-4.2931	*	-0.02380
32	0.59996	0.60040	-0.00044	.002895222	-0.1534		-0.00091
33	0.30705	0.30064	0.00641	.003052548	2.0987	*	0.01393
34	0.32947	0.35941	-0.02995	.003586533	-8.3493	*	-0.06305
35	0.60271	0.68860	-0.08589	.003124637	-27.4889	*	-0.18030
36	0.65605	0.73468	-0.07863	.003145371	-24.9974	*	-0.17146
37	0.73870	0.77434	-0.03563	.003072638	-11.5975	*	-0.08310
38	0.71771	0.76397	-0.04625	.002543151	-18.1875	*	-0.10571
39	0.67771	0.73203	-0.05432	.003106667	-17.4843	*	-0.11930
40	0.32022	0.41541	-0.09519	.003416508	-27.8620	*	-0.19838
41	0.32350	0.40326	-0.07976	.003167170	-25.1848	*	-0.16642
42	0.39040	0.49516	-0.10476	.003299690	-31.7478	*	-0.21209
43	0.45977	0.53144	-0.07167	.003332628	-21.5044	*	-0.14371
44	0.55090	0.61024	-0.05934	.003581167	-16.5712	*	-0.12048
45	0.40139	0.44199	-0.04060	.003426834	-11.8464	*	-0.08228
46	0.56368	0.64407	-0.08039	.003059775	-26.2741	*	-0.16493
47	0.70153	0.73456	-0.03303	.002987132	-11.0562	*	-0.07345
48	0.55484	0.62596	-0.07112	.003123003	-22.7737	*	-0.14501
49	0.55893	0.64109	-0.08215	.003045790	-26.9732	*	-0.16829

Appendix 13. Summary of Item-Level Analysis – Grade 3

inum	cbtpval	pappval	difpval	difstd	z_dif	isig	effect_size
1	0.95688	0.96300	-0.006126	.001290643	-4.7462	*	-0.03125
2	0.96402	0.95748	0.006547	.001269676	5.1567	*	0.03373
3	0.95282	0.95687	-0.004051	.001328566	-3.0492	*	-0.01950
4	0.98701	0.98669	0.000320	.000851735	0.3758		0.00281
5	0.98352	0.97914	0.004388	.000999449	4.3902	*	0.03242
6	0.84326	0.84565	-0.002393	.002180343	-1.0976		-0.00660
7	0.83035	0.85871	-0.028364	.002780325	-10.2016	*	-0.07833
8	0.97650	0.97455	0.001944	.001101089	1.7657		0.01257
9	0.55895	0.53666	0.022290	.003108873	7.1697	*	0.04480
10	0.97866	0.97793	0.000721	.000919221	0.7843		0.00496
11	0.72059	0.73153	-0.010945	.003133775	-3.4925	*	-0.02454
12	0.86237	0.86649	-0.004122	.002217467	-1.8588		-0.01204
13	0.57330	0.59732	-0.024022	.003677838	-6.5315	*	-0.04877
14	0.95642	0.95388	0.002539	.001653201	1.5360		0.01226
15	0.73635	0.74847	-0.012117	.003048526	-3.9746	*	-0.02771
16	0.65254	0.69614	-0.043601	.003462431	-12.5926	*	-0.09314
17	0.55870	0.57127	-0.012575	.003701421	-3.3973	*	-0.02537
18	0.73354	0.80765	-0.074114	.003119272	-23.7600	*	-0.17696
19	0.89162	0.91121	-0.019584	.002131202	-9.1890	*	-0.06572
20	0.88485	0.91428	-0.029431	.002139000	-13.7590	*	-0.09803
21	0.85583	0.88498	-0.029152	.002238285	-13.0244	*	-0.08688
22	0.86614	0.87273	-0.006592	.002363130	-2.7895	*	-0.01956
23	0.63921	0.69919	-0.059982	.003366210	-17.8189	*	-0.12775
24	0.60177	0.61871	-0.016945	.003861978	-4.3877	*	-0.03475
25	0.50628	0.49624	0.010046	.003072441	3.2697	*	0.02009
26	0.62148	0.71366	-0.092186	.003257414	-28.3005	*	-0.19663
27	0.45963	0.49843	-0.038798	.003191746	-12.1557	*	-0.07772
28	0.78866	0.82647	-0.037813	.002679255	-14.1133	*	-0.09603

inum	cbtpval	pappval	difpval	difstd	z_dif	isig	effect_size
29	0.72722	0.78075	-0.053529	.002836179	-18.8738	*	-0.12453
30	0.67049	0.71495	-0.044457	.003220101	-13.8060	*	-0.09647
31	0.82346	0.84435	-0.020888	.002384773	-8.7591	*	-0.05614
32	0.68677	0.70877	-0.022000	.002868265	-7.6700	*	-0.04792
33	0.53612	0.59238	-0.056259	.003692833	-15.2348	*	-0.11364
34	0.47937	0.50330	-0.023929	.003622287	-6.6062	*	-0.04788
35	0.36987	0.43626	-0.066390	.003480465	-19.0749	*	-0.13566
36	0.48208	0.43380	0.048274	.003573583	13.5086	*	0.09701
37	0.67043	0.71351	-0.043086	.003265414	-13.1947	*	-0.09343
38	0.31529	0.31394	0.001353	.003384504	0.3997		0.00291
39	0.57646	0.61912	-0.042659	.003530877	-12.0818	*	-0.08708
40	0.71589	0.75651	-0.040614	.003512109	-11.5639	*	-0.09225
41	0.67935	0.73368	-0.054327	.002878890	-18.8709	*	-0.11952
42	0.73351	0.77627	-0.042755	.003179572	-13.4467	*	-0.09952
43	0.61864	0.65175	-0.033110	.003317776	-9.9795	*	-0.06882
44	0.66378	0.72253	-0.058747	.003085776	-19.0380	*	-0.12764
45	0.58564	0.67698	-0.091347	.003425682	-26.6653	*	-0.19020
46	0.46188	0.47686	-0.014982	.003734857	-4.0113	*	-0.03002
47	0.38148	0.44244	-0.060956	.003368232	-18.0972	*	-0.12409
48	0.53516	0.57323	-0.038073	.003860637	-9.8618	*	-0.07665
49	0.47067	0.56775	-0.097082	.003095031	-31.3671	*	-0.19523
50	0.65229	0.69141	-0.039124	.003397224	-11.5164	*	-0.08340
51	0.62775	0.71471	-0.086964	.003531559	-24.6247	*	-0.18592
52	0.53163	0.54396	-0.012324	.003437699	-3.5850	*	-0.02472
53	0.75992	0.80087	-0.040949	.002824742	-14.4964	*	-0.09903
54	0.86254	0.89519	-0.032654	.002674355	-12.2102	*	-0.10020
55	0.80232	0.83231	-0.029988	.002813694	-10.6578	*	-0.07766
56	0.88644	0.89246	-0.006019	.002077102	-2.8977	*	-0.01919

inum	cbtpval	pappval	difpval	difstd	z_dif	isig	effect_size
57	0.77587	0.80495	-0.029085	.002705218	-10.7515	*	-0.07150
58	0.89260	0.89267	-0.000076	.002107464	-0.0360		-0.00024

Appendix 14. Summary of Item-Level Analysis – Grade Cluster 4–5

inum	cbtpval	pappval	difpval	difstd	z_dif	isig	effect_size
1	0.98727	0.98418	0.003093	.000913222	3.3874	*	0.02609
2	0.97556	0.97023	0.005331	.001196003	4.4572	*	0.03283
3	0.87440	0.88503	-0.010631	.002074705	-5.1239	*	-0.03268
4	0.97620	0.97146	0.004740	.000988726	4.7945	*	0.02971
5	0.98514	0.97736	0.007788	.000856626	9.0913	*	0.05744
6	0.87636	0.88650	-0.010137	.002121946	-4.7772	*	-0.03136
7	0.98131	0.97444	0.006869	.001089104	6.3075	*	0.04670
8	0.49247	0.51497	-0.022505	.003597306	-6.2559	*	-0.04502
9	0.92156	0.91643	0.005129	.001636696	3.1336	*	0.01880
10	0.59530	0.61678	-0.021474	.003498176	-6.1387	*	-0.04396
11	0.97050	0.95900	0.011495	.001167640	9.8442	*	0.06236
12	0.43769	0.43981	-0.002128	.003060676	-0.6953		-0.00429
13	0.96783	0.96413	0.003695	.001357496	2.7218	*	0.02038
14	0.67359	0.70680	-0.033211	.003058136	-10.8600	*	-0.07187
15	0.80052	0.79944	0.001088	.002506849	0.4341		0.00272
16	0.71103	0.73381	-0.022781	.002569932	-8.8644	*	-0.05089
17	0.73986	0.74950	-0.009642	.002741734	-3.5169	*	-0.02211
18	0.58710	0.58210	0.004997	.003405334	1.4673		0.01014
19	0.96165	0.95885	0.002802	.001472834	1.9026		0.01434
20	0.95979	0.95266	0.007129	.001402524	5.0828	*	0.03484
21	0.95225	0.94588	0.006376	.001545382	4.1260	*	0.02901
22	0.96613	0.96347	0.002657	.001166463	2.2776	*	0.01442
23	0.91037	0.91854	-0.008166	.001971089	-4.1428	*	-0.02920
24	0.90373	0.91454	-0.010804	.001796346	-6.0146	*	-0.03760
25	0.92893	0.93488	-0.005949	.001754943	-3.3901	*	-0.02362
26	0.79842	0.81419	-0.015771	.002744338	-5.7467	*	-0.03992
27	0.53608	0.54470	-0.008614	.003404407	-2.5304	*	-0.01729
28	0.77033	0.76172	0.008603	.002880181	2.9868	*	0.02032

inum	cbtpval	pappval	difpval	difstd	z_dif	isig	effect_size
29	0.45207	0.51579	-0.063728	.003474216	-18.3432	*	-0.12778
30	0.70900	0.77241	-0.063412	.003357944	-18.8840	*	-0.14507
31	0.59439	0.62672	-0.032328	.003101046	-10.4249	*	-0.06633
32	0.62083	0.64833	-0.027499	.003537820	-7.7730	*	-0.05713
33	0.57086	0.54922	0.021643	.003026730	7.1506	*	0.04361
34	0.74637	0.73534	0.011026	.002851572	3.8666	*	0.02517
35	0.62867	0.70543	-0.076753	.002991981	-25.6529	*	-0.16341
36	0.45813	0.46382	-0.005695	.003391253	-1.6793		-0.01143
37	0.66625	0.72161	-0.055357	.002901704	-19.0775	*	-0.12034
38	0.68124	0.73232	-0.051078	.002885009	-17.7047	*	-0.11238
39	0.79579	0.80973	-0.013938	.002607546	-5.3454	*	-0.03504
40	0.57454	0.60628	-0.031737	.003246744	-9.7749	*	-0.06457
41	0.37443	0.40393	-0.029501	.003410025	-8.6514	*	-0.06054
42	0.45612	0.48469	-0.028565	.003772352	-7.5723	*	-0.05726
43	0.72770	0.74500	-0.017297	.003055012	-5.6618	*	-0.03926
44	0.37216	0.46118	-0.089028	.003235092	-27.5195	*	-0.18132
45	0.56596	0.56674	-0.000781	.003315938	-0.2356		-0.00158
46	0.83970	0.86178	-0.022082	.002392855	-9.2284	*	-0.06200
47	0.76067	0.80550	-0.044834	.003021527	-14.8382	*	-0.10895
48	0.86646	0.90372	-0.037259	.002121317	-17.5640	*	-0.11703
49	0.79382	0.84368	-0.049868	.002284711	-21.8266	*	-0.12972
50	0.75646	0.74912	0.007337	.002696409	2.7211	*	0.01701
51	0.46933	0.54061	-0.071278	.003508472	-20.3161	*	-0.14293
52	0.55309	0.59104	-0.037958	.003284656	-11.5561	*	-0.07677
53	0.44653	0.45806	-0.011530	.003687997	-3.1264	*	-0.02317
54	0.40496	0.46093	-0.055968	.003426108	-16.3358	*	-0.11314
55	0.60991	0.59774	0.012168	.003193357	3.8103	*	0.02488
56	0.66757	0.71039	-0.042822	.002862041	-14.9622	*	-0.09261
57	0.75584	0.81616	-0.060318	.002816120	-21.4190	*	-0.14748

inum	cbtpval	pappval	difpval	difstd	z_dif	isig	effect_size
58	0.85402	0.86493	-0.010903	.002285452	-4.7705	*	-0.03137
59	0.82729	0.84896	-0.021674	.002580928	-8.3978	*	-0.05887
60	0.66255	0.71206	-0.049505	.002975870	-16.6354	*	-0.10694
61	0.78574	0.83402	-0.048281	.002633663	-18.3321	*	-0.12327

Appendix 15. Summary of Item-Level Analysis – Grade Cluster 6–7

inum	cbtpval	pappval	difpval	difstd	z_dif	isig	effect_size
1	0.96278	0.95941	0.003366	.001639264	2.0535	*	0.01740
2	0.95422	0.96277	-0.008551	.001740985	-4.9115	*	-0.04288
3	0.95186	0.94945	0.002408	.002011262	1.1971		0.01112
4	0.96763	0.96849	-0.000857	.001717599	-0.4991		-0.00486
5	0.97763	0.97565	0.001983	.001358534	1.4597		0.01313
6	0.99217	0.98714	0.005030	.000959864	5.2407	*	0.04973
7	0.47532	0.44433	0.030993	.004160751	7.4488	*	0.06222
8	0.98414	0.97938	0.004759	.001169434	4.0691	*	0.03555
9	0.98061	0.97745	0.003161	.001377231	2.2951	*	0.02206
10	0.85879	0.86749	-0.008701	.002987954	-2.9121	*	-0.02531
11	0.92628	0.91830	0.007977	.001897137	4.2046	*	0.02980
12	0.97003	0.96475	0.005283	.001605998	3.2892	*	0.02974
13	0.91165	0.91258	-0.000924	.002124166	-0.4351		-0.00326
14	0.75344	0.77897	-0.025533	.003960771	-6.4465	*	-0.06036
15	0.52106	0.52897	-0.007917	.005317897	-1.4887		-0.01585
16	0.64913	0.63468	0.014450	.004427321	3.2639	*	0.03015
17	0.80868	0.81100	-0.002321	.003832699	-0.6055		-0.00591
18	0.64182	0.65445	-0.012629	.004427976	-2.8522	*	-0.02645
19	0.81223	0.85843	-0.046200	.003518386	-13.1310	*	-0.12481
20	0.76776	0.77535	-0.007591	.003324510	-2.2833	*	-0.01807
21	0.90795	0.90741	0.000543	.002354844	0.2306		0.00188
22	0.81752	0.83159	-0.014074	.003143832	-4.4766	*	-0.03701
23	0.86473	0.86805	-0.003315	.002686974	-1.2338		-0.00974
24	0.90965	0.91115	-0.001499	.002626764	-0.5706		-0.00524
25	0.71144	0.70380	0.007644	.004207326	1.8168		0.01681
26	0.71061	0.75035	-0.039748	.003831808	-10.3731	*	-0.08967
27	0.70948	0.72644	-0.016966	.003673297	-4.6186	*	-0.03770
28	0.70571	0.70489	0.000827	.003876974	0.2134		0.00182

inum	cbtpval	pappval	difpval	difstd	z_dif	isig	effect_size
29	0.63084	0.65344	-0.022602	.004542116	-4.9760	*	-0.04716
30	0.56298	0.55814	0.004842	.004352122	1.1125		0.00976
31	0.37068	0.39498	-0.024303	.004680798	-5.1920	*	-0.05002
32	0.38584	0.40558	-0.019734	.004817318	-4.0966	*	-0.04037
33	0.50632	0.51769	-0.011372	.005418224	-2.0989	*	-0.02275
34	0.39739	0.42515	-0.027763	.004159377	-6.6748	*	-0.05645
35	0.59531	0.61375	-0.018438	.003991560	-4.6193	*	-0.03771
36	0.96376	0.96582	-0.002056	.001680602	-1.2235		-0.01115
37	0.95358	0.96462	-0.011041	.001974850	-5.5907	*	-0.05574
38	0.92440	0.92906	-0.004661	.002318897	-2.0099	*	-0.01789
39	0.86850	0.87329	-0.004784	.002814325	-1.7000		-0.01426
40	0.90792	0.90993	-0.002010	.002485358	-0.8089		-0.00699
41	0.78881	0.79803	-0.009220	.003673170	-2.5101	*	-0.02278
42	0.71586	0.73679	-0.020926	.004266502	-4.9047	*	-0.04694
43	0.61516	0.65291	-0.037756	.004223706	-8.9390	*	-0.07844
44	0.66656	0.67762	-0.011052	.004491527	-2.4607	*	-0.02354
45	0.68400	0.65614	0.027867	.004522303	6.1621	*	0.05930
46	0.65539	0.68404	-0.028651	.004144608	-6.9129	*	-0.06094
47	0.90812	0.89783	0.010296	.002704840	3.8064	*	0.03480
48	0.80899	0.79623	0.012760	.003014025	4.2336	*	0.03207
49	0.77205	0.77027	0.001786	.003385893	0.5276		0.00425
50	0.48816	0.49975	-0.011599	.004156415	-2.7906	*	-0.02320
51	0.62187	0.63396	-0.012097	.004264780	-2.8364	*	-0.02503
52	0.48888	0.46995	0.018926	.004895253	3.8662	*	0.03789
53	0.41049	0.45752	-0.047029	.004759688	-9.8807	*	-0.09500
54	0.38649	0.38489	0.001603	.004505451	0.3558		0.00329
55	0.58199	0.55863	0.023357	.004749520	4.9178	*	0.04720
56	0.47591	0.49491	-0.018993	.005136945	-3.6974	*	-0.03801
57	0.38225	0.37526	0.006984	.004955248	1.4095		0.01440

inum	cbtpval	pappval	difpval	difstd	z_dif	isig	effect_size
58	0.75142	0.74144	0.009984	.003612444	2.7638	*	0.02295
59	0.82412	0.81206	0.012057	.003355323	3.5934	*	0.03126
60	0.46136	0.53400	-0.072632	.003547516	-20.4741	*	-0.14565
61	0.76641	0.75923	0.007186	.003451872	2.0819	*	0.01690
62	0.55475	0.61308	-0.058325	.004617616	-12.6310	*	-0.11854
63	0.78381	0.80325	-0.019442	.003671993	-5.2946	*	-0.04804

Appendix 16. Summary of Item-Level Analysis – Grade Cluster 8–9

inum	cbtpval	pappval	difpval	difstd	z_dif	isig	effect_size
1	0.95069	0.95349	-0.002803	.001999513	-1.4020		-0.01315
2	0.97629	0.97215	0.004144	.001673229	2.4769	*	0.02614
3	0.96644	0.96605	0.000395	.001651121	0.2394		0.00219
4	0.97685	0.97260	0.004249	.001484642	2.8622	*	0.02705
5	0.90751	0.90737	0.000141	.002481323	0.0566		0.00049
6	0.77194	0.77562	-0.003680	.003553615	-1.0355		-0.00879
7	0.95639	0.94947	0.006928	.001970758	3.5153	*	0.03271
8	0.96818	0.96352	0.004661	.002143917	2.1740	*	0.02566
9	0.64253	0.68156	-0.039024	.003801166	-10.2663	*	-0.08258
10	0.95507	0.94778	0.007291	.001977239	3.6872	*	0.03391
11	0.71270	0.72135	-0.008644	.004320235	-2.0008	*	-0.01919
12	0.61823	0.62534	-0.007111	.004652878	-1.5283		-0.01466
13	0.69793	0.74408	-0.046147	.004101474	-11.2512	*	-0.10303
14	0.78588	0.80553	-0.019643	.004065258	-4.8318	*	-0.04873
15	0.57561	0.60970	-0.034089	.005056920	-6.7411	*	-0.06942
16	0.78257	0.83967	-0.057100	.004030332	-14.1675	*	-0.14628
17	0.96656	0.96166	0.004901	.001896618	2.5839	*	0.02633
18	0.96959	0.96120	0.008396	.001857018	4.5212	*	0.04593
19	0.89078	0.91068	-0.019896	.002482942	-8.0132	*	-0.06657
20	0.91304	0.91796	-0.004914	.002458355	-1.9989	*	-0.01767
21	0.92084	0.91708	0.003762	.002451044	1.5350		0.01381
22	0.75950	0.78047	-0.020970	.004154792	-5.0472	*	-0.04984
23	0.63034	0.65768	-0.027347	.003927784	-6.9623	*	-0.05714
24	0.77493	0.77316	0.001772	.004565308	0.3881		0.00424
25	0.50749	0.54854	-0.041045	.005044670	-8.1364	*	-0.08229
26	0.55740	0.55665	0.000747	.005565964	0.1342		0.00151
27	0.59680	0.61082	-0.014021	.005040605	-2.7817	*	-0.02867
28	0.40759	0.43181	-0.024215	.004746120	-5.1020	*	-0.04908

inum	cbtpval	pappval	difpval	difstd	z_dif	isig	effect_size
29	0.67182	0.67782	-0.005990	.004519304	-1.3255		-0.01279
30	0.53699	0.54843	-0.011442	.004754777	-2.4065	*	-0.02297
31	0.42534	0.49265	-0.067310	.005581370	-12.0598	*	-0.13539
32	0.49106	0.52530	-0.034242	.005403766	-6.3368	*	-0.06854
33	0.40105	0.41634	-0.015296	.005298687	-2.8867	*	-0.03112
34	0.52657	0.56975	-0.043179	.004622315	-9.3415	*	-0.08685
35	0.53294	0.54498	-0.012035	.006119869	-1.9666	*	-0.02415
36	0.49930	0.53225	-0.032950	.005401493	-6.1002	*	-0.06597
37	0.86655	0.86781	-0.001262	.003055387	-0.4131		-0.00371
38	0.81825	0.82503	-0.006772	.003744911	-1.8084		-0.01768
39	0.80164	0.80570	-0.004059	.003645152	-1.1136		-0.01022
40	0.79148	0.80617	-0.014690	.004268664	-3.4412	*	-0.03665
41	0.67923	0.70470	-0.025467	.004389795	-5.8015	*	-0.05518
42	0.63430	0.68009	-0.045792	.003926786	-11.6615	*	-0.09659
43	0.38303	0.37493	0.008099	.004561803	1.7754		0.01669
44	0.23257	0.24094	-0.008370	.004233904	-1.9768	*	-0.01970
45	0.46858	0.49701	-0.028434	.005455812	-5.2118	*	-0.05693
46	0.55110	0.54684	0.004263	.004397298	0.9695		0.00857
47	0.47474	0.47420	0.000544	.004996290	0.1089		0.00109
48	0.88503	0.89723	-0.012205	.003330817	-3.6643	*	-0.03919
49	0.81129	0.82269	-0.011400	.003268642	-3.4877	*	-0.02949
50	0.92409	0.91483	0.009255	.002595305	3.5662	*	0.03401
51	0.79846	0.82190	-0.023441	.003891752	-6.0232	*	-0.05979
52	0.51557	0.55425	-0.038680	.005449206	-7.0983	*	-0.07761
53	0.64067	0.60912	0.031552	.004680260	6.7414	*	0.06521
54	0.59442	0.56802	0.026394	.004851665	5.4403	*	0.05352
55	0.60949	0.59440	0.015092	.004197551	3.5955	*	0.03084
56	0.72233	0.69252	0.029804	.004110695	7.2504	*	0.06555
57	0.54037	0.54239	-0.002021	.005533203	-0.3652		-0.00406

inum	cbtpval	pappval	difpval	difstd	z_dif	isig	effect_size
58	0.60479	0.61186	-0.007068	.005562114	-1.2708		-0.01448
59	0.46535	0.46859	-0.003239	.005213147	-0.6213		-0.00649
60	0.78345	0.79580	-0.012353	.003600463	-3.4308	*	-0.03031
61	0.61315	0.67864	-0.065492	.005011123	-13.0693	*	-0.13727
62	0.77202	0.82004	-0.048024	.004195529	-11.4465	*	-0.11939
63	0.77322	0.82278	-0.049551	.003793234	-13.0630	*	-0.12365

Appendix 17. Summary of Item-Level Analysis – Grade Cluster 10–12

inum	cbtpval	pappval	difpval	difstd	z_dif	isig	effect_size
1	0.96885	0.96606	0.002783	.001953331	1.4250		0.01570
2	0.96264	0.96916	-0.006524	.002064069	-3.1608	*	-0.03597
3	0.98071	0.98039	0.000324	.001662974	0.1948		0.00231
4	0.98730	0.98229	0.005011	.001462953	3.4250	*	0.04092
5	0.96469	0.95496	0.009736	.002128870	4.5733	*	0.04958
6	0.99032	0.98460	0.005721	.001435494	3.9855	*	0.05147
7	0.74553	0.76355	-0.018022	.004632060	-3.8906	*	-0.04189
8	0.98271	0.97501	0.007700	.001866876	4.1244	*	0.05353
9	0.89635	0.90808	-0.011740	.003643309	-3.2223	*	-0.03954
10	0.64715	0.66415	-0.017005	.005827555	-2.9181	*	-0.03579
11	0.90080	0.89692	0.003886	.002919476	1.3310		0.01289
12	0.74481	0.75284	-0.008029	.004718916	-1.7015		-0.01852
13	0.81107	0.86794	-0.056875	.003856062	-14.7495	*	-0.15541
14	0.63296	0.67753	-0.044569	.005167911	-8.6242	*	-0.09388
15	0.54554	0.53735	0.008189	.005871530	1.3947		0.01644
16	0.62215	0.62978	-0.007629	.005177350	-1.4734		-0.01577
17	0.57053	0.60182	-0.031284	.006168412	-5.0716	*	-0.06355
18	0.96451	0.95942	0.005098	.001988500	2.5637	*	0.02666
19	0.97037	0.97217	-0.001800	.001859709	-0.9677		-0.01079
20	0.97166	0.96615	0.005516	.001794088	3.0747	*	0.03179
21	0.98086	0.97473	0.006130	.001428937	4.2897	*	0.04160
22	0.92472	0.91855	0.006170	.002582271	2.3893	*	0.02295
23	0.90372	0.90516	-0.001433	.003138252	-0.4565		-0.00486
24	0.95295	0.94208	0.010869	.002000135	5.4342	*	0.04875
25	0.85101	0.84911	0.001897	.003767829	0.5034		0.00532
26	0.87113	0.87823	-0.007106	.003576025	-1.9871	*	-0.02146
27	0.89251	0.88888	0.003633	.003354050	1.0831		0.01166
28	0.71939	0.72865	-0.009257	.004766274	-1.9422		-0.02071

inum	cbtpval	pappval	difpval	difstd	z_dif	isig	effect_size
29	0.47996	0.47770	0.002261	.005711665	0.3958		0.00453
30	0.46699	0.49456	-0.027567	.005314442	-5.1872	*	-0.05520
31	0.64798	0.65672	-0.008739	.004576235	-1.9096		-0.01835
32	0.53230	0.52770	0.004608	.005728281	0.8044		0.00923
33	0.60077	0.61874	-0.017963	.005625687	-3.1930	*	-0.03683
34	0.49602	0.51781	-0.021790	.005704197	-3.8201	*	-0.04360
35	0.86694	0.87002	-0.003077	.003304192	-0.9313		-0.00910
36	0.85948	0.86430	-0.004820	.004179463	-1.1532		-0.01397
37	0.89908	0.89225	0.006828	.003364174	2.0295	*	0.02234
38	0.84285	0.84187	0.000987	.003830710	0.2578		0.00272
39	0.81042	0.81719	-0.006777	.004731245	-1.4324		-0.01740
40	0.46964	0.46495	0.004681	.006141974	0.7621		0.00938
41	0.69158	0.69822	-0.006639	.004721789	-1.4060		-0.01441
42	0.54156	0.55519	-0.013632	.005650292	-2.4127	*	-0.02740
43	0.43821	0.44985	-0.011637	.005211930	-2.2327	*	-0.02342
44	0.45650	0.48404	-0.027543	.004769491	-5.7747	*	-0.05521
45	0.52032	0.51137	0.008947	.004959228	1.8041		0.01790
46	0.41328	0.41500	-0.001726	.005762246	-0.2996		-0.00351
47	0.42989	0.43645	-0.006569	.005362994	-1.2249		-0.01326
48	0.32017	0.32218	-0.002017	.005250942	-0.3842		-0.00432
49	0.52350	0.54578	-0.022271	.005348643	-4.1638	*	-0.04466
50	0.68972	0.68730	0.002414	.004487410	0.5379		0.00522
51	0.71099	0.70919	0.001808	.004977912	0.3632		0.00399
52	0.61769	0.63106	-0.013368	.006043344	-2.2120	*	-0.02760
53	0.65808	0.62056	0.037517	.005475185	6.8522	*	0.07820
54	0.60140	0.60538	-0.003975	.005422548	-0.7330		-0.00812
55	0.47680	0.46734	0.009452	.005023893	1.8814		0.01893
56	0.74646	0.72643	0.020025	.004849067	4.1297	*	0.04548
57	0.65454	0.64530	0.009244	.004890129	1.8904		0.01938

inum	cbtpval	pappval	difpval	difstd	z_dif	isig	effect_size
58	0.52739	0.51443	0.012965	.005946757	2.1801	*	0.02596
59	0.74832	0.73841	0.009906	.004219912	2.3475	*	0.02269
60	0.68198	0.68041	0.001571	.005118001	0.3070		0.00337
61	0.86221	0.84423	0.017982	.003371507	5.3336	*	0.05085
62	0.86034	0.84816	0.012176	.003878488	3.1395	*	0.03453
63	0.79283	0.79515	-0.002324	.004527565	-0.5134		-0.00575
64	0.81672	0.81701	-0.000294	.004112459	-0.0716		-0.00075