Results of March and April 2008 TAKS Exit Level Retest
Online Comparability Study Analyses

07/16/08

Summary

The Texas Assessment of Knowledge and Skills (TAKS) has been offered online (in addition to paper) for certain grades and subject areas since 2005. The joint *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 1999, Standard 4.10) and the *Guidelines for Computer-Based Tests and Interpretations* (APA, 1986) both speak to the need for the evaluation of the score equivalence across computer-based and paper-based tests.

Traditionally, Texas uses a "matched samples comparability analyses", or MSCA (Way, Davis, & Fitzpatrick, 2006), design to study equivalence of the resulting raw score to scale score conversions between online and paper testing modes. In this design, a bootstrap sampling approach is used to select online and paper samples where each selected online student is matched to a paper student with the same gender, ethnicity, and level of proficiency on previous test scores. Once both the paper sample and online sample are selected, raw score to scale score conversions are calculated applying the Rasch measurement model. This sampling is repeated for a predetermined number of times. Conversion tables for each replication are retained and aggregated to obtain the final conversion and the standard error of linking. The equivalency between online and paper scale scores is then evaluated. A separate conversion table is recommended if the two sets of scores are considered not comparable. Between-mode item-level analyses and subgroup analyses for ethnic and gender subgroups are also performed.

As in the October 2007 exit level comparability studies, however, a modified strategy was implemented due to the large number of first-time testers participating in the March/April 2008 retest administrations. The first-time testers differed from the retesters in a few systematic ways. First of all, the first-time testers had higher mean raw scores than the retesters in the March/April administrations, and secondly there was a stronger association between the March/April 2008 raw scores and previous TAKS scores for the first-time testers.

To include the first-time testers in constructing the online score conversions, a modified MSCA approach was implemented by introducing student grade level in April 2007 as an additional matching variable. In other words, the online first-time testers were only matched to paper first-time testers, where the proficiency was defined by the grade 10 TAKS scores in April 2007. The online retesters were only matched to paper retesters, where the proficiency level was defined by the exit level TAKS scores in April 2007. Between-mode item-level analyses and subgroup analyses for ethnic and gender subgroups were performed in addition to the overall test-level analysis.

Using this revised methodology, the comparability analyses were carried out for the March 2008 TAKS exit level administrations in the following subject areas:
- English language arts (ELA);
- mathematics;
- science; and
- social studies.

With the criteria used to evaluate the equivalence (raw score differences that matter, scale score differences compared to two standard errors of linking, and raw score cuts at various performance levels), the comparability between online and paper administrations was assessed. For the March administrations, a mode effect between online and paper was found for mathematics, science and social studies. For mathematics and science, the online version was 1 point harder at both the 'Met the Standard' and the 'Commended' performance levels. For social studies, the online version was 1 point harder at the 'Met the Standard' level. For ELA, the mode of testing was found to have no impact – the cut scores were the same between online and paper at both the 'Met the Standard' and 'Commended' levels; no differences were found at any of the raw score points between online and paper.

The subgroup analysis results indicated that the online version appeared more difficult for all subgroups in the March administration. The only exception was ELA, where no mode effect was found between online and paper at the subgroup level. The item level analysis results indicated that across subject areas, most items exhibiting a significant mode effect were easier in the paper version.

For the April administrations, comparability analyses were carried out only for mathematics and science. For ELA and social studies, the sample sizes were too small for such analyses. For mathematics and science, the results indicated a mode effect

between online and paper. At the 'Met the Standard' level, the online version was one point harder for mathematics and was also one point harder for science. No differences were found at the 'Commended' level for either subject area.

The following table summarizes the comparability analysis results for each subject area tested.

Subject	Scale score conversions differ by more than two standard errors of linking throughout most of the raw score range		Raw scores differences greater than or equal to half of a raw score point throughout most of the raw score range		Cut Score Difference			
					Met the Standard		Commended	
	March	April	March	April	March	April	March	April
ELA		N/A		N/A		N/A		N/A
Math	X	X	X	X	X	X	X	
Science	X	X	X	X	X	X	X	
Social Studies	X	N/A	X	N/A	X	N/A		N/A

Based on the comparability analysis results, Pearson recommended using separate conversion tables for the March 2008 exit level online retest in mathematics, science and social studies and for the April 2008 exit level online retests in mathematics and science.

Introduction

Online comparability studies were first implemented in Texas in 2005 covering the Texas Assessment of Knowledge and Skills (TAKS) in grade 8 reading, mathematics and social studies as well as exit level retests of TAKS in English language arts (ELA), mathematics, science and social studies. The 2006 comparability studies included TAKS grade 8 reading, mathematics, science, and social studies, grade 9 reading and mathematics, and all exit level July and October exit level retests. Further expansion for online testing in 2007 added TAKS grade 7 reading and mathematics and grade 10 ELA, mathematics, science, and social studies. In 2008 the online retest opportunities in March and April were offered the first time. This report summarizes the results from these two administrations.

Methodology

In 2005, TEA and Pearson devised a matched samples comparability analysis plan using a bootstrap approach in which students in the online group would be matched to students from the paper group on their previous TAKS test scores. Based on recommendations from the Texas Technical Advisory Committee (TTAC), additional demographic variables were considered as matching variables beginning in 2006. The decision was made to include previous TAKS scores, ethnicity, and gender as matching variables in the matched comparability analyses.

When the October 2007 exit level retest comparability studies were conducted, it was noted that a great number of students would have been excluded from the comparability analyses because of missing April 2007 TAKS exit level scores. Further investigation indicated that these students differed from the retesters in the following ways: their mean raw score was higher and their raw scores were more highly correlated to their previous TAKS scores (as indicated by higher r-square values) than those of the retesters in the October administration. Furthermore, these students were in grade 10 in April 2007. Excluding these students may have had an undesirable impact because they systematically differed from the retesters.

A similar situation occurred in both the March and April 2008 online retest administrations in that first-time testers as well as retesters participated in the administration (refer to Tables 1a and 1b). Although it is desirable to include the first-time testers in the study, when the first-time testing sample is too small, a problem could occur where not enough students would be assigned to certain matched groups. For this reason, prior to the test administration, Pearson developed a decision rule for whether or not to include first-time testers in the study. Specifically, first-time testers will be included if a), the online first-time testers have at least 500 students, or b), the online first-time testers account for at least 20% of all online testers. If the number of online first-time testers is less than 500, which accounts for less than 20% of all online testers, then these first-time testers will be excluded (from both modes) and only retesters will be included in the comparability analysis.

In order to include first-time testers in the comparability analyses if they met the sample size requirement, the same modifications were made to the comparability method as were made in October 2007. Because previous TAKS scores were based on grade 10 (and not exit level) performance for the first-time testers, a separate y-hat variable was created for those students. In addition, these first-time testers were matched with first-time testers only and not matched to the retesters in the bootstrapping process. Similarly, the retesters were matched with retesters only, not with the first-time testers.

Students must meet certain criteria to be eligible for inclusion in the comparability analysis, such as having valid test scores and score codes for the test, having valid previous primary TAKS scale scores, and having valid information in their gender and ethnicity variables. It should be mentioned that in April 2008 the total number of online ELA students was 477, where only 320 were eligible; and for social studies the total number tested was 317, with 201 eligible. These counts did not reach the level that would be sufficient for a comparability study, therefore for these two April tests the comparability analysis was not conducted.

The Stratified MSCA Method for March/April 2008 Exit Level Retests

The following is a list of steps in the 'modified' comparability analyses.

1) For the retest students who tested in the paper mode, their March (and separately for April) 2008 retest raw scores were regressed on their ELA, mathematics, science, and social studies scale scores from the April 2007 primary administration.

For students with TAKS exit level (XL) scores in April of 2007(retesters):

$$\hat{Y}_{predicted\_rawscore} = \beta_0 + \beta_1 X_{1(prev\_XL\_ELA)} + \beta_2 X_{2(prev\_XL\_Math)} + \beta_3 X_{3(prev\_XL\_Science)} + \beta_4 X_{4(prev\_XL\_SocialStudies)}$$

For students with TAKS grade 10 scores in April of 2007(first-time testers):

$$\hat{Y}_{predicted\_rawscore} = \beta_0 + \beta_1 X_{1(prev\_G10\_ELA)} + \beta_2 X_{2(prev\_G10\_Math)} + \beta_3 X_{3(prev\_G10\_Science)} + \beta_4 X_{4(prev\_G10\_SocialStudies)}$$

In other words, for the March (or April) 2008 exit level retest comparability analyses, the regression analysis was repeated separately for students with grade 10 TAKS scores in April 2007 and for students with exit level TAKS scores in April 2007.

2) The resulting regression weights were applied to all students (paper and online) to obtain an estimated raw score (y-hat) for each of the students. For the March (or April) 2008 exit level retest comparability analyses, online students' y-hat scores were generated using appropriate regression weights depending on whether they took grade 10 TAKS in April 2007 or they took exit level TAKS in April 2007. The regression weights for each group are shown in Table 1a for March 2008 and Table 1b for April 2008.

3) Students in each group (first-time testers and retesters) were then broken into either 10 or 20 groups based on the estimated raw score. Due to the nature of the matched-sample bootstrap approach, we need to have enough students participating who are first-time testers to use the modified approach. The criteria for including first-time testers and breaking them into groups are shown in the following grid:

Proportion \ Sample Size	>=500	<500
>=20%	20 groups	10 groups
<20%	20 groups	exclude

Therefore, the first-time testers would be excluded from the analyses if their count was less than 500 **AND** was less than 20% of the total online students (first-time testers and retesters combined). In the March administration, both first-time testers and retesters were included in the analyses. For the April administration, however, the first-time testers were dropped from the comparability analyses as the result of applying the decision rules. Note that the same decision rules were applied to online retesters as well.

In order to modify the procedure such that retesters would only be matched to retesters and first-time testers would only be matched to first-time testers, the following steps were implemented:

I. Grouped the retest students (online plus paper) with previous exit level information into 10 (or 20) score groups based on their y-hat values. This produced a ranking variable ranging from 0 to 9 (or 19), for all retesters (students with exit level information).
II. Grouped the first-time students (online plus paper) with previous grade 10 information into 10 or 20 score groups based on their y-hat values. This also produced a ranking variable ranging from 0 to 9 (or 19), for all first-time testers (students with grade 10 information).
III. Added a constant to the exit level ranking from step I, such that the ranking wasn't duplicated between the two groups of students. For example, if 20 groups were used for both first-time testers and retesters, then the rank variable would range from 0 to 39, where 0-19 was for the first-time testers, and 20-39 was for the retesters.

IV. Separated the online group of students from the paper group of students to form the base of online and paper samples, and then proceeded with the bootstrap process.

Please note that if the students all had exit level TAKS score in April 2007, then there would have only had 20 groups for each subject area, which would have become the regular MSCA approach.

4) This grouping from step 3 could result in a 40 (previous score groups—20 for first-time testers plus 20 for retesters) by 4 (ethnicity groups) by 2 (gender groups) grid that was used in the matched sampling. Note that the total number of groups could be different from 40, depending on the number of groups for the first-time testers as well as the retesters.

5) To improve optimal matching, students with missing values on any of the matching variables were dropped from the study.

It should be noted that the March ELA (which took place in February) and April mathematics, science and social studies retests were post-eqauted tests because the primary tests were administered at the same time. The April ELA and March mathematics, science and social studies were pre-equated tests because no primary tests were administered.

Implementation of the Stratified MSCA Procedures

The modified matched samples comparability analysis method is described in the steps below. Please note that with the modified approach, the number of score groups could become 40 (instead of 20 as in the July 2007 comparability analyses).

1. All students eligible for matching were placed into score groups based on the regression of March (or April) 2008 retest raw score on the April 2007 scale scores as described earlier. Each student testing online with complete data was matched to a student from the available 2008 March (or April) paper TAKS data with an identical profile on the matching variables.
2. Online versus paper comparability analyses were performed using matched groups of students. The following steps were repeated 100 times (500 for ELA):
   a. A bootstrap sample of students was drawn from the online participants.
   b. A matched sample was drawn at random from the available March (or April) 2008 paper TAKS data.
   c. A raw score-to-raw score equating was carried out with the bootstrap samples.
   d. The raw score equivalents were transformed to scale scores using the appropriate paper pre- or post-equated score conversion tables and linear interpolation.
3. Online scale score conversions for each raw score were based on the average of the conversions calculated over each of the 100 replications (500 for ELA). These average scale score values comprised the alternate online conversion table.
4. The standard deviation of online scale score conversions at each raw score represented the conditional bootstrap standard errors of the linking.
5. To evaluate comparability, raw score points for which the difference between the online and paper scale score conversions exceeded two standard errors of the linking were noted. Also, the differences between the paper-raw-score equivalent and the corresponding paper raw score were calculated, and differences greater than 0.5 were noted.

To verify the comparability analysis results, two Pearson psychometricians programmed independent versions of the analysis. The online results based on these two analyses were compared to the paper results. The recommendations regarding the use of a separate online score conversion table were made based on a set of rules that TEA adopted in 2006. Pearson recommended the consideration of three pieces of information: the standard error of the linking, the magnitude of the raw score differences, and the rounding differences for cut scores.

- The standard error of the linking criterion was suggested by Dorans and Lawrence (1990): "To assess equivalence, it is convenient to compute the difference between the equating function and the identity transformation, and to divide this difference by the standard error of equating. If the resultant ratio falls within a bandwidth of plus or minus two,

then the equating function is deemed to be within sampling error of the identity function" (p. 247). In using this procedure, we paid special attention to differences in the range of scale scores around the "Met the Standard" and "Commended" score levels. Differences at the extremes of the scale are less important, given the purpose and primary uses of the TAKS tests. This standard error procedure is sensitive to sample size such that the standard errors will be greater when the sample sizes are smaller. Therefore, we also considered additional criteria.

- The magnitude of the raw score differences was evaluated using the criterion of differences that matter (DTM; Dorans & Feigenbaum, 1994). This was originally developed in the context of the SAT where scaled scores are reported in 10-point units. For a given raw score, if the resulting scales scores from the linking differed by fewer than 5 points, then the scale scores would ideally be rounded to the same value and would be considered equivalent. This process was adapted to other tests and the DTM was considered to be a half of a score unit for unrounded scores (Dorans, Holland, Thayer, & Tateneni, 2003). For the TAKS, the DTM was considered to be half of a raw score point. For a given proficiency level, if the corresponding raw scores from the linking differed by less than half of a raw score point, then the two could be considered equivalent.

- The third piece of information considered was the rounding differences for the cut scores. The raw score to scale score conversions for the paper and online tests were compared to see if they resulted in different raw score cut points across the two modes of test administration. Cuts were evaluated for both the "Met the Standard" and the "Commended Performance" levels. Pearson recommended that this information be used in conjunction with the magnitude of the raw score differences, and the statistical significance of the differences (based on the Dorans and Lawrence [1990] two standard errors of the linking).

In addition to those three pieces of information, subgroup and item-level analyses were conducted, and the impact of using the alternate score table was considered. Results from these analyses were used to inform decision-making in borderline cases. Using these pieces of information, overall psychometric judgment determined the recommendation for whether to use an alternate score table for the online TAKS administration.

Results

Tables 2a and 2b display the demographic information and descriptive statistics for the online and paper samples for each subject area tested. Within administration, the online and paper samples tended to be similar in terms of ethnicity proportions. Based on the sets of decision provided earlier, first-time testers were included in the March analyses, but in April they were excluded from both the math and science comparability analyses.

As can be seen, across all subject areas and administrations, the Hispanic students accounted for the majority of all the students tested. Higher proportion of Hispanic and African American students took the test in paper, whereas a higher proportion of white students took the tests online. It can also be seen from the tables that for mathematics and science, online students generally had lower raw scores than the paper students; whereas for ELA and social studies, online students generally had higher raw scores than the paper students---although for all subject areas the difference appeared to be very small.

Studies from previous comparability analyses on ELA retest administrations have consistently indicated an item type by mode interaction, namely students tested online tend to have higher essay scores, whereas students tested in paper tend to have higher scores on multiple-choice items. Table 3 provides the performance for March exit level ELA students by testing mode. Consistent with several previous ELA findings, students tested online appeared to perform better on the essay while students tested on paper performed better on the multiple choice items.

Table 4 provides an overview of the results of the March online and paper comparability studies. Test-level mode effects were detected for all subject areas tested except for ELA. Where a mode effect was detected, the 'Met the Standard' cut was lower (harder) online. For mathematics and science, the 'Commended' cut was lower for online as well.

Tables 5 to 8 detail the comparisons between online and paper for each March test. The columns of the tables are as follows:

RAW – Paper test raw score
CBT_RS – Equivalent raw scores on the online test based on the comparability linking. Note that a higher 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 100/500 replications.
PAP_SS – Paper test scale score conversions, based on the 2008 March TAKS mathematics, science, and social studies pre-equated scales, and ELA post-equated scales.
CBT_SS – Equivalent scale scores on the online test based on the comparability linking. Again, higher 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 100/500 replications.
RS_DIF – Difference between paper-raw-score equivalent and paper raw score.
SS_DIF – Difference between online scale score equivalent and paper 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 "*".
FINAL– Final recommended online scale score conversion.

* Note that the "*" in the SIG column indicates both statistical and practical significance, based on the recommendations from the TTAC. In addition, the scale scores that corresponded to the '1-sem' and '2-sem' performance cuts are underlined and italicized; whereas the scale scores that corresponded to the 'Met Standard' and 'Commended' performance cuts are highlighted and bolded.

For March ELA (see Table 5), the differences in raw scores were less than half of a raw score point and the scale score differences were not statistically significant throughout the entire scale. Student performance on the exit level ELA retest was shown to be not impacted by the test mode.

For both March mathematics (see Table 6) and March science (see Table 7), the online version of the test was more difficult. The online "Met the Standard" cut score was 1 point lower for both subjects, and the score differences were both statistically and practically significant across most of the scale range. The online "Commended" cut was also 1 point lower for both subjects.

For March social studies (see Table 8), the 'Met the Standard' cut was one point lower online. Although the score differences were marked as both practically and statistically significant at only a few raw score levels (refer to Table 8, SIG column), the differences in raw scores and scale scores at about half of raw score levels were practically significant. If the online sample size had been larger, the differences might have also been statistically significant.

Table 9 provides an overview of the results of the April online and paper comparability studies. Test-level mode effects were detected for the two subject areas analyzed: mathematics and science. Specifically, the 'Met the Standard' cut was lower (harder) online.

Tables 10 and 11 detail the comparisons between online and paper versions of the April mathematics and science, respectively. The columns of the tables are as defined above, with the exception that the paper test scale score conversions (PAP_SS) were post-equated. For both April mathematics (see Table 10) and April science (see Table 11), the online versions of the test were more difficult. The online "Met the Standard" cut score was 1 point lower for both subjects, and the score differences were both statistically and practically significant across most of the scale range. The online "Commended" cut was the same as the paper cut for both subjects.

For all subject areas in both March and April, large differences occur at the lowest and highest scaled score points because WINSTEPS (the IRT calibration software used in the study) does not estimate abilities for zero and perfect scores. These differences are not meaningful. Therefore, for each test the online conversion table has the scale score at the extreme ends set to the paper, as has been consistently done in the past.

Impact Data Analyses

The pass-rate comparison shown in Table 4 indicates that applying the paper conversion tables to the online students resulted in a lower percentage of students achieving the 'Met the Standard' level for math, science and social studies in the March administration. Using the alternate (online) conversion tables resulted in passing rates for the online students that were more similar to passing rates for paper students. Similar information for the April administration is presented in Table 9. Applying the paper conversion tables to the online students resulted in a lower percentage of students achieving the 'Met the Standard' level for mathematics and science. Using the alternate (online) conversion tables resulted in more similar pass-rates across modes.

Additional Analysis

As consistent with previous online comparability studies, two sets of additional analyses were conducted: the subgroup analysis and the item-level analysis. The subgroup analysis compares mean differences of the total raw scores between the two testing modes across replications for male, female, White, Hispanic and African American student groups separately, whereas the item-level analysis compares mean differences of each item between the two testing modes across replications.

Subgroup Analysis

The mean raw score differences (and the mean effect size, see Cohen [1992]) between the online and paper testing modes for each subgroup for each subject area tested in March are listed in Table 12. A significance test was performed for each 'matched' subgroup 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 subgroup; and SE_{diff} is the bootstrap standard error of the mean differences over the replications, also for each subgroup.

The effect size between two group means <u>at each replication</u> was calculated by 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 12, there was a consistent significant mode effect across gender and ethnicity groups for March mathematics and science (online harder). For ELA and social studies, no significant mode effects were present for any group.

Table 13 lists the subgroup-level analysis results for April mathematics and science. Again, the online mode is consistently harder for both subjects across all groups.

Item-Level Analysis

Item-level analysis was performed in a similar way to the subgroup analysis. The mean item raw score differences across replications between the online and paper testing modes for each subject area were computed, and the effect sizes were also calculated. Tables 14 to 17 display the results of item-level comparison across replications for each subject area tested in March. The columns of the tables are as follows:

CBT_PVAL: Mean item score for the online students across 100/500 replications.
PAP_PVAL: Mean item score for the paper students across 100/500 replications.
DIF_PVAL: Mean item score differences between online and paper students across 100/500 replications
DIF_STD: Standard deviation of the mean differences across 100/500 replications
Z_DIF: Z statistic for the mean item score differences
SIG: Items where the Z_DIF statistic was greater than 2 are noted by "*".
EFFECT_SIZE: Mean effect size over 100/500 replications.

As can be seen from the tables, when there was a significant item-level mode difference, the paper group tended to score higher for all subject areas. Specifically for math, the paper group scored higher on 17 items of the 21 items exhibiting a mode effect. For science, the paper group scored higher on 14 of the 15 items exhibiting a mode effect. For both ELA and social studies, few items exhibit a significant mode effect. In ELA, only 3 items exhibit a mode effect, with two of them in favor of paper. Finally, for social studies, the paper group scored higher on 4 of the 5 items exhibiting a mode effect.

It should also be noted that 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 difference 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 these tables. It can be seen that the magnitude of the effect sizes ranged between 0 and 0.24.

For the April administration, the item-level analysis results are shown for math in Table 18 and for science in Table 19. For math, all 16 items showing a mode effect were in favor of paper. Similarly, although only 5 items exhibited a mode effect for science, the mode effect for all was in favor of paper. The effect sizes were between 0 and 0.26.

Summary and Recommendation

For most subject areas analyzed (in both March and April administrations), statistically and practically significant between-mode differences were found. With the exception of March ELA, the other analyzed tests exhibited a mode effect at the 'Met the Standard' performance level.

Based on these comparisons, Pearson recommended using a separate online conversion table for March TAKS exit level retest subject areas: mathematics, science and social studies; and for April TAKS exit level retest subject areas: mathematics and science.

References

American Psychological Association Committee on Professional Standards and Committee on Psychological Tests and Assessments (APA). (1986). *Guidelines for computer-based tests and interpretations.* Washington, DC: Author.

American Educational Research Association (AERA), American Psychological Association (APA), and the National Council on Measurement in Education (NCME). (1999). S*tandards for educational and psychological testing.* Washington, DC: AERA.

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.

Table 1a. Summary of Sample Statistics and Regression Results, by Student Population and Subject Area Tested
March 2008

Subject Tested	Student Population	Intercept	Prev_ELA	Prev_Math	Prev_Scie	Prev_Sost	r-square	PaperMean	PaperN	OnlineN	OnlineMean
ELA	First timers	-90.158	0.047	0.005	0.007	0.007	0.570	49.80	3822	258	53.32
	Retesters	-123.759	0.060	0.005	0.005	0.010	0.316	38.56	4992	351	38.59
Math	First timers	-97.757	0.008	0.039	0.011	0.004	0.518	31.27	6105	380	32.98
	Retesters	-93.098	0.009	0.037	0.010	0.003	0.204	28.32	18736	1442	27.36
Science	First timers	-91.189	0.011	0.011	0.022	0.014	0.557	29.18	5647	329	30.70
	Retesters	-94.385	0.014	0.012	0.019	0.013	0.256	26.37	17783	1365	25.52
Social Studies	First timers	-84.392	0.020	0.001	0.012	0.023	0.585	36.84	3390	235	40.04
	Retesters	-81.258	0.022	0.004	0.009	0.019	0.210	27.11	3510	306	27.02

Table 1b. Summary of Sample Statistics and Regression Results, by Student Population and Subject Area Tested
April 2008

Subject Tested	Student Population	Intercept	Prev_ELA	Prev_Math	Prev_Scie	Prev_Sost	r-square	PaperMean	PaperN	OnlineN	OnlineMean
Math	First timers	-82.702	0.010	0.029	0.011	0.004	0.311	27.52	3543	231	27.61
	Retesters	-89.531	0.010	0.033	0.012	0.002	0.176	27.41	12354	1076	26.45
Science	First timers	-79.616	0.013	0.015	0.012	0.012	0.347	26.36	3196	203	26.96
	Retesters	-88.455	0.013	0.017	0.015	0.011	0.229	25.77	11376	966	24.92

Table 2a. Student Demographic and Descriptive Information for March 2008 Comparability Analyses*

Subject	Number of Campuses		Number of Students		Mean Raw Score		Estimated Raw Score, or y-hat		Demographic Information**											
									Male		White		Hispanic		African American		Other Ethnicity		Special Education[***]	
	CBT	PAP	CBT	PAP	CBT	PAP	CBT	PAP	CBT	PAP	CBT	PAP	CBT	PAP	CBT	PAP	CBT	PAP	CBT	PAP
ELA	102	1027	609	8814	44.83	43.43	45.14	43.43	54	53	25	15	56	66	14	16	5	3	4	5
Mathematics	135	1326	1822	24841	28.54	29.04	29.58	29.04	43	44	29	19	46	56	23	24	2	2	4	4
Science	135	1330	1694	23430	26.53	27.05	27.55	27.05	39	41	22	14	52	63	22	21	4	2	4	3
Social Studies	105	982	541	6900	32.67	31.89	33.13	31.89	40	38	25	17	54	63	16	18	5	2	4	3

Table 2b. Student Demographic and Descriptive Information for April 2008 Comparability Analyses*

Subject	Number of Campuses		Number of Students		Mean Raw Score		Estimated Raw Score, or y-hat[#]		Demographic Information**											
									Male		White		Hispanic		African American		Other Ethnicity		Special Education[***]	
	CBT	PAP	CBT	PAP	CBT	PAP	CBT	PAP	CBT	PAP	CBT	PAP	CBT	PAP	CBT	PAP	CBT	PAP	CBT	PAP
Mathematics	106	1177	1076	12354	26.45	27.41	27.70	27.41	41	43	22	16	55	56	21	27	2	1	4	3
Science	100	1148	966	11376	24.92	25.77	25.95	25.77	35	39	15	10	63	64	18	23	4	2	4	3

*: CBT-Online administration; PAP-Paper administration.
#: The y-hat values were not calculated for ELA or social studies since no matching was done.
**: Cell entries represent percentages rounded to the nearest integer. Due to rounding, the percentages may not add up to 100%.
***: Special education status is included in the demographics table, but was not used as a matching variable.

Table 3. Student Performance* on March 2008 Exit Level ELA Multiple Choice, Open-Ended Items, and Essay by Test Mode

	ONLINE				PAPER			
Multiple Choice	36.26				36.64			
	0	1	2	3	0	1	2	3
OE 1	8.11%	49.80%	41.96%	0.13%	8.97%	53.15%	37.76%	0.12%
OE 2	12.98%	65.13%	21.89%	0.00%	12.68%	67.26%	20.00%	0.06%
OE 3	25.30%	52.59%	22.11%	0.00%	27.24%	54.08%	18.66%	0.03%
	0	1	2	3	4	0	1	2	3	4
Essay	1.20%	25.08%	42.93%	25.48%	5.32%	1.62%	24.88%	52.90%	18.40%	2.21%

*: based on the eligible students included in the analyses.

Table 4. Summary of the March 2008 Exit Level Online TAKS Comparability Analyses

Grade /Subject	Sample Size		Raw Score Cuts*		Pass Rates (%) Comparison**			Number of Raw Score Points with Meaningful Difference***	Decision (Conclusion)
	Online	Paper	Online	Paper	Paper w/Paper	Online w/Online	Online w/Paper		
Exit Level ELA****	609	8814	43	**43**	46.5	42.7	42.7	0 (73)	Paper Conversion Table (No Mode Effect)
			63	**63**	3.1	8.5	8.5		
Exit Level Math	1822	24841	**32**	33	30.8	30.9	26.8	47(60)	Alternate Conversion Table (Mode Effect)
			53	54	0.6	1.2	0.9		
Exit Level Science	1694	23430	**29**	30	34.5	40.2	29.8	42 (55)	Alternate Conversion Table (Mode Effect)
			49	50	0.4	0.7	0.5		
Exit Level Social Studies	541	6900	**27**	28	56.5	58.1	54.4	5 (55)	Alternate Conversion Table (Mode Effect)
			49	49	6.6	10.5	10.5		

*: Raw score points corresponding to 'Met Standards' (top) and 'Commended' (bottom) levels. Final RS cuts in bold.

**: 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.

****: Note that for ELA, an essay score of 2 or above is required in addition to the scale score requirements to be considered passing.

Table 5. Summary of Comparability Analysis – March Exit Level ELA

RAW	CBT_RS	RS_SD	PAP_SS	CBT_SS	SS_SD	RS_DIF	SS_DIF	SIG Final
0	0.30	0.114	1351.49	1386.11	12.830	0.30	34.62	1351
1	1.00	0.302	1465.21	1461.15	28.516	-0.00	-4.06	1465
2	1.96	0.468	1547.73	1539.01	31.221	-0.04	-8.72	1548
3	2.92	0.561	1597.64	1590.85	25.952	-0.08	-6.79	1598
4	3.89	0.604	1634.18	1628.55	21.106	-0.11	-5.63	1634
5	4.87	0.615	1663.42	1658.61	17.352	-0.13	-4.81	1663
6	5.86	0.609	1688.05	1683.86	14.521	-0.14	-4.19	1688
7	6.85	0.594	1709.49	1705.80	12.377	-0.15	-3.69	1709
8	7.85	0.577	1728.60	1725.33	10.749	-0.15	-3.27	1729
9	8.85	0.561	1745.95	1743.04	9.510	-0.15	-2.91	1746
10	9.85	0.547	1761.92	1759.33	8.554	-0.15	-2.59	1762
11	10.86	0.537	1776.76	1774.47	7.812	-0.14	-2.29	1777
12	11.86	0.529	1790.68	1788.66	7.241	-0.14	-2.02	1791
13	12.87	0.525	1803.84	1802.08	6.800	-0.13	-1.76	1804
14	13.89	0.524	1816.36	1814.85	6.455	-0.11	-1.51	1816
15	14.90	0.525	1828.32	1827.06	6.188	-0.10	-1.26	1828
16	15.92	0.527	1839.83	1838.80	5.984	-0.08	-1.03	1840
17	16.93	0.531	1850.93	1850.13	5.819	-0.07	-0.80	1851
18	17.95	0.536	1861.67	1861.09	5.691	-0.05	-0.58	1862
19	18.97	0.542	1872.11	1871.75	5.594	-0.03	-0.36	1872
20	19.99	0.548	1882.29	1882.15	5.516	-0.01	-0.14	1882
21	21.01	0.554	1892.23	1892.30	5.454	0.01	0.07	1892
22	22.03	0.560	1901.97	1902.26	5.411	0.03	0.29	1902
23	23.06	0.566	1911.55	1912.04	5.377	0.06	0.49	1912
24	24.08	0.571	1920.98	1921.67	5.347	0.08	0.69	1921
25	25.10	0.576	1930.28	1931.16	5.323	0.10	0.88	1930
26	26.12	0.580	1939.47	1940.55	5.307	0.12	1.08	1939
27	27.14	0.584	1948.58	1949.84	5.293	0.14	1.26	1949
28	28.16	0.587	1957.61	1959.06	5.286	0.16	1.45	1958
29	29.18	0.588	1966.61	1968.22	5.278	0.18	1.61	1967
30	30.20	0.589	1975.56	1977.34	5.269	0.20	1.78	1976
31	31.22	0.589	1984.49	1986.43	5.261	0.22	1.94	1984
32	32.23	0.589	1993.41	1995.50	5.254	0.23	2.09	1993
33	33.25	0.587	2002.34	2004.57	5.247	0.25	2.23	2002
34	34.26	0.584	2011.29	2013.66	5.241	0.26	2.37	2011
35	35.27	0.580	2020.28	2022.76	5.229	0.27	2.48	2020
36	36.28	0.575	2029.31	2031.90	5.219	0.28	2.59	2029
37	37.29	0.568	*2038.41*	*2041.09*	5.209	0.29	2.68	*2045*
38	38.30	0.561	2047.59	2050.36	5.199	0.30	2.77	2048
39	39.30	0.553	2056.88	2059.70	5.182	0.30	2.82	2057
40	40.30	0.544	*2066.27*	*2069.13*	5.163	0.30	2.86	*2072*
41	41.30	0.534	2075.79	2078.68	5.146	0.30	2.89	2076
42	42.29	0.523	2085.47	2088.36	5.123	0.29	2.89	2085
43	43.28	0.511	**2095.31**	**2098.17**	5.098	0.28	2.86	**2100**
44	44.27	0.498	2105.34	2108.15	5.073	0.27	2.81	2105
45	45.26	0.484	2115.59	2118.31	5.042	0.26	2.72	2116
46	46.24	0.470	2126.07	2128.68	5.012	0.24	2.61	2126
47	47.22	0.455	2136.82	2139.28	4.982	0.22	2.46	2137
48	48.20	0.439	2147.87	2150.14	4.952	0.20	2.27	2148
49	49.17	0.423	2159.26	2161.30	4.920	0.17	2.04	2159
50	50.14	0.407	2171.02	2172.78	4.891	0.14	1.76	2171
51	51.11	0.391	2183.20	2184.63	4.872	0.11	1.43	2183
52	52.07	0.374	2195.88	2196.92	4.854	0.07	1.04	2196
53	53.04	0.358	2209.08	2209.69	4.843	0.04	0.61	2209
54	54.00	0.342	2222.90	2223.00	4.844	0.00	0.10	2223
55	54.96	0.327	2237.41	2236.95	4.863	-0.04	-0.46	2237
56	55.92	0.313	2252.75	2251.65	4.903	-0.08	-1.10	2253
57	56.89	0.299	2269.00	2267.21	4.967	-0.11	-1.79	2269

Table 5. Summary of Comparability Analysis – March Exit Level ELA
(Continued)

RAW	CBT_RS	RS_SD	PAP_SS	CBT_SS	SS_SD	RS_DIF	SS_DIF	SIG Final
58	57.85	0.287	2286.35	2283.79	5.075	-0.15	-2.56	2286
59	58.82	0.276	2304.97	2301.59	5.242	-0.18	-3.38	2305
60	59.79	0.269	2325.10	2320.85	5.509	-0.21	-4.25	2325
61	60.76	0.266	2347.05	2341.89	5.943	-0.24	-5.16	2347
62	61.75	0.270	2371.16	2365.09	6.659	-0.25	-6.07	2371
63	62.74	0.287	**2397.93**	**2390.96**	7.868	-0.26	-6.97	**2400**
64	63.73	0.317	2427.97	2420.14	9.845	-0.27	-7.83	2428
65	64.74	0.362	2462.07	2453.47	12.914	-0.26	-8.60	2462
66	65.75	0.412	2501.50	2492.15	17.312	-0.25	-9.35	2501
67	66.76	0.450	2548.26	2538.23	22.958	-0.24	-10.03	2548
68	67.79	0.450	2606.59	2596.27	29.688	-0.21	-10.32	2607
69	68.85	0.399	2686.46	2677.95	37.583	-0.15	-8.51	2686
70	69.97	0.387	2807.35	2807.52	51.513	-0.03	0.17	2807
71	71.05	0.461	2952.71	2960.36	67.523	0.05	7.65	2953
72	72.02	0.364	3100.24	3102.65	52.928	0.02	2.41	3100
73	72.70	0.149	3242.87	3199.72	21.226	-0.30	-43.15	3243

Table 6. Summary of Comparability Analysis – March Exit Level Mathematics

RAW	CBT_RS	RS_SD	PAP_SS	CBT_SS	SS_SD	RS_DIF	SS_DIF	SIG	Final
0	0.33	0.008	1313.70	1360.90	1.078	0.33	47.20		1314
1	1.10	0.024	1455.77	1466.35	2.462	0.10	10.58		1466
2	2.20	0.046	1557.07	1569.31	2.807	0.20	12.24		1569
3	3.29	0.066	1617.95	1630.83	2.918	0.29	12.88		1631
4	4.37	0.084	1662.31	1675.48	2.951	0.37	13.17		1675
5	5.45	0.100	1697.62	1710.94	2.956	0.45	13.32		1711
6	6.52	0.115	1727.23	1740.61	2.946	0.52	13.38	*	1741
7	7.59	0.128	1752.91	1766.36	2.939	0.59	13.45	*	1766
8	8.65	0.141	1775.82	1789.17	2.901	0.65	13.35	*	1789
9	9.70	0.152	1796.42	1809.74	2.879	0.70	13.32	*	1810
10	10.76	0.163	1815.32	1828.60	2.852	0.76	13.28	*	1829
11	11.81	0.173	1832.86	1846.09	2.824	0.81	13.23	*	1846
12	12.85	0.182	1849.29	1862.46	2.794	0.85	13.17	*	1862
13	13.89	0.190	1864.79	1877.89	2.768	0.89	13.10	*	1878
14	14.93	0.198	1879.51	1892.54	2.743	0.93	13.03	*	1893
15	15.96	0.205	1893.56	1906.53	2.723	0.96	12.97	*	1907
16	17.00	0.211	1907.05	1919.96	2.705	1.00	12.91	*	1920
17	18.02	0.217	1920.05	1932.89	2.682	1.02	12.84	*	1933
18	19.05	0.223	1932.63	1945.34	2.668	1.05	12.71	*	1945
19	20.07	0.227	1944.77	1957.48	2.661	1.07	12.71	*	1957
20	21.09	0.232	1956.65	1969.29	2.650	1.09	12.64	*	1969
21	22.11	0.235	1968.24	1980.83	2.641	1.11	12.59	*	1981
22	23.13	0.239	1979.60	1992.13	2.633	1.13	12.53	*	1992
23	24.14	0.241	1990.75	2003.23	2.626	1.14	12.48	*	2003
24	25.15	0.244	2001.73	*2014.16*	2.618	1.15	12.43	*	*2015*
25	26.16	0.246	*2012.56*	2024.93	2.613	1.16	12.37	*	2025
26	27.16	0.247	2023.26	2035.59	2.608	1.16	12.33	*	2036
27	28.17	0.248	2033.87	2046.15	2.603	1.17	12.28	*	2046
28	29.17	0.249	2044.40	*2056.65*	2.600	1.17	12.25	*	*2058*
29	30.17	0.249	*2054.88*	2067.09	2.596	1.17	12.21	*	2067
30	31.17	0.248	2065.33	2077.50	2.594	1.17	12.17	*	2078
31	32.16	0.247	2075.77	2087.90	2.590	1.16	12.13	*	2088
32	33.15	0.246	2086.22	**2098.32**	2.589	1.15	12.10	*	**2100**
33	34.14	0.244	**2096.70**	2108.77	2.587	1.14	12.07	*	2109
34	35.13	0.242	2107.24	2119.28	2.587	1.13	12.04	*	2119
35	36.12	0.239	2117.85	2129.87	2.585	1.12	12.02	*	2130
36	37.10	0.236	2128.57	2140.55	2.585	1.10	11.98	*	2141
37	38.09	0.232	2139.41	2151.38	2.584	1.09	11.97	*	2151
38	39.07	0.228	2150.41	2162.36	2.585	1.07	11.95	*	2162
39	40.04	0.224	2161.60	2173.52	2.584	1.04	11.92	*	2174
40	41.02	0.219	2173.00	2184.90	2.584	1.02	11.90	*	2185
41	41.99	0.213	2184.66	2196.55	2.584	0.99	11.89	*	2197
42	42.96	0.207	2196.62	2208.50	2.583	0.96	11.88	*	2208
43	43.93	0.200	2208.93	2220.79	2.582	0.93	11.86	*	2221
44	44.90	0.193	2221.64	2233.50	2.580	0.90	11.86	*	2233
45	45.86	0.186	2234.82	2246.67	2.578	0.86	11.85	*	2247
46	46.82	0.178	2248.55	2260.41	2.577	0.82	11.86	*	2260
47	47.78	0.169	2262.93	2274.80	2.574	0.78	11.87	*	2275
48	48.74	0.160	2278.08	2289.96	2.572	0.74	11.88	*	2290
49	49.69	0.150	2294.13	2306.04	2.579	0.69	11.91	*	2306
50	50.65	0.140	2311.29	2323.24	2.589	0.65	11.95	*	2323
51	51.59	0.129	2329.80	2341.80	2.603	0.59	12.00	*	2342
52	52.54	0.117	2349.99	2362.08	2.626	0.54	12.09	*	2362
53	53.48	0.105	2372.34	**2384.54**	2.652	0.48	12.20		**2400**
54	54.42	0.092	**2397.53**	2409.89	2.692	0.42	12.36		2410
55	55.36	0.079	2426.63	2439.22	2.748	0.36	12.59		2439
56	56.30	0.065	2461.41	2474.39	2.839	0.30	12.98		2474
57	57.23	0.050	2505.21	2518.93	3.006	0.23	13.72		2519

Table 6. Summary of Comparability Analysis – March Exit Level Mathematics
(Continued)

RAW	CBT_RS	RS_SD	PAP_SS	CBT_SS	SS_SD	RS_DIF	SS_DIF	SIG	Final
58	58.16	0.034	2565.52	2581.15	3.433	0.16	15.63		2581
59	59.08	0.017	2666.22	2677.26	2.430	0.08	11.04		2677
60	59.72	0.005	2805.15	2766.84	0.743	-0.28	-38.31		2805

Table 7. Summary of Comparability Analysis – March Exit Level Science

RAW	CBT_RS	RS_SD	PAP_SS	CBT_SS	SS_SD	RS_DIF	SS_DIF	SIG	Final
0	0.32	0.006	1394.72	1437.38	0.745	0.32	42.66		1395
1	1.08	0.018	1525.98	1533.58	1.711	0.08	7.60		1534
2	2.16	0.035	1619.55	1628.42	1.968	0.16	8.87		1628
3	3.23	0.050	1675.98	1685.35	2.050	0.23	9.37		1685
4	4.29	0.064	1717.04	1726.70	2.086	0.29	9.66		1727
5	5.36	0.076	1749.78	1759.62	2.095	0.36	9.84		1760
6	6.42	0.088	1777.28	1787.24	2.093	0.42	9.96		1787
7	7.47	0.098	1801.17	1811.22	2.084	0.47	10.05		1811
8	8.53	0.108	1822.43	1832.56	2.072	0.53	10.13	*	1833
9	9.58	0.116	1841.70	1851.89	2.058	0.58	10.19	*	1852
10	10.62	0.124	1859.41	1869.66	2.042	0.62	10.25	*	1870
11	11.67	0.131	1875.87	1886.17	2.023	0.67	10.30	*	1886
12	12.71	0.138	1891.30	1901.60	1.995	0.71	10.30	*	1902
13	13.75	0.144	1905.80	1916.20	1.987	0.75	10.40	*	1916
14	14.79	0.149	1919.66	1930.11	1.970	0.79	10.45	*	1930
15	15.82	0.154	1932.92	1943.43	1.951	0.82	10.51	*	1943
16	16.86	0.158	1945.67	1956.23	1.934	0.86	10.56	*	1956
17	17.89	0.162	1957.98	1968.61	1.920	0.89	10.63	*	1969
18	18.92	0.165	1969.93	1980.62	1.906	0.92	10.69	*	1981
19	19.95	0.168	1981.56	1992.32	1.895	0.95	10.76	*	1992
20	20.97	0.171	1992.92	2003.76	1.883	0.97	10.84	*	2004
21	22.00	0.173	2004.06	2014.97	1.874	1.00	10.91	*	2015
22	23.02	0.174	2015.00	*2025.99*	1.867	1.02	10.99	*	*2035*
23	24.04	0.176	*2025.78*	2036.86	1.860	1.04	11.08	*	2037
24	25.06	0.176	2036.44	2047.60	1.855	1.06	11.16	*	2048
25	26.07	0.177	2047.00	*2058.26*	1.851	1.07	11.26	*	*2068*
26	27.09	0.177	2057.49	2068.84	1.847	1.09	11.35	*	2069
27	28.10	0.177	*2067.94*	2079.39	1.843	1.10	11.45	*	2079
28	29.11	0.176	2078.37	2089.92	1.842	1.11	11.55	*	2090
29	30.11	0.175	2088.80	**2100.45**	1.842	1.11	11.65	*	**2100**
30	31.12	0.174	**2099.26**	2111.03	1.840	1.12	11.77	*	2111
31	32.12	0.172	2109.79	2121.65	1.840	1.12	11.86	*	2122
32	33.12	0.170	2120.39	2132.37	1.842	1.12	11.98	*	2132
33	34.11	0.168	2131.11	2143.21	1.844	1.11	12.10	*	2143
34	35.10	0.165	2141.98	2154.19	1.847	1.10	12.21	*	2154
35	36.09	0.161	2153.02	2165.36	1.849	1.09	12.34	*	2165
36	37.08	0.158	2164.28	2176.74	1.853	1.08	12.46	*	2177
37	38.06	0.154	2175.79	2188.38	1.857	1.06	12.59	*	2188
38	39.04	0.150	2187.61	2200.32	1.861	1.04	12.71	*	2200
39	40.02	0.145	2199.78	2212.62	1.864	1.02	12.84	*	2213
40	40.99	0.140	2212.37	2225.34	1.868	0.99	12.97	*	2225
41	41.96	0.134	2225.45	2238.56	1.867	0.96	13.11	*	2239
42	42.92	0.128	2239.12	2252.36	1.868	0.92	13.24	*	2252
43	43.88	0.121	2253.47	2266.87	1.867	0.88	13.40	*	2267
44	44.84	0.114	2268.66	2282.22	1.867	0.84	13.56	*	2282
45	45.79	0.107	2284.86	2298.58	1.867	0.79	13.72	*	2299
46	46.73	0.099	2302.29	2316.20	1.879	0.73	13.91	*	2316
47	47.67	0.091	2321.26	2335.39	1.897	0.67	14.13	*	2335
48	48.61	0.082	2342.21	2356.60	1.922	0.61	14.39	*	2357
49	49.54	0.072	2365.78	**2380.47**	1.953	0.54	14.69	*	**2400**
50	50.47	0.062	**2392.93**	2408.04	2.000	0.47	15.11		2408
51	51.39	0.051	2425.32	2441.01	2.071	0.39	15.69		2441
52	52.30	0.039	2466.01	2482.76	2.205	0.30	16.75		2483
53	53.21	0.027	2522.06	2541.24	2.520	0.21	19.18		2541
54	54.11	0.014	2615.18	2629.12	1.831	0.11	13.94		2629
55	54.73	0.004	2746.23	2711.19	0.562	-0.27	-35.04		2746

Table 8. Summary of Comparability Analysis – March Exit Level Social Studies

RAW	CBT_RS	RS_SD	PAP_SS	CBT_SS	SS_SD	RS_DIF	SS_DIF	SIG	Final
0	0.32	0.016	1429.65	1470.65	2.062	0.32	41.00		1430
1	1.06	0.051	1557.66	1563.45	4.799	0.06	5.79		1563
2	2.12	0.094	1648.85	1655.34	5.530	0.12	6.49		1655
3	3.17	0.133	1703.70	1710.51	5.540	0.17	6.81		1711
4	4.22	0.169	1743.71	1750.65	5.521	0.22	6.94		1751
5	5.26	0.201	1775.61	1782.62	5.488	0.26	7.01		1783
6	6.31	0.231	1802.40	1809.49	5.469	0.31	7.09		1809
7	7.34	0.258	1825.77	1832.87	5.409	0.34	7.10		1833
8	8.38	0.283	1846.48	1853.62	5.368	0.38	7.14		1854
9	9.42	0.306	1865.26	1872.45	5.337	0.42	7.19		1872
10	10.45	0.327	1882.55	1889.77	5.295	0.45	7.22		1890
11	11.48	0.347	1898.62	1905.89	5.259	0.48	7.27		1906
12	12.51	0.364	1913.72	1921.04	5.222	0.51	7.32		1921
13	13.54	0.380	1928.01	1935.38	5.188	0.54	7.37		1935
14	14.57	0.395	1941.63	1949.05	5.154	0.57	7.42		1949
15	15.60	0.408	1954.68	1962.15	5.117	0.60	7.47		1962
16	16.62	0.419	1967.24	1974.76	5.089	0.62	7.52		1975
17	17.64	0.429	1979.40	1986.97	5.056	0.64	7.57		1987
18	18.66	0.438	1991.20	1998.83	5.027	0.66	7.63		1999
19	19.68	0.445	2002.71	2010.39	5.001	0.68	7.68		2010
20	20.70	0.451	2013.97	2021.71	4.971	0.70	7.74		2022
21	21.72	0.455	*2025.02*	*2032.81*	4.942	0.72	7.79		*2033*
22	22.73	0.459	2035.89	2043.73	4.915	0.73	7.84		2044
23	23.74	0.461	2046.62	2054.51	4.886	0.74	7.89		2055
24	24.76	0.462	*2057.23*	*2065.18*	4.861	0.76	7.95		*2067*
25	25.76	0.462	2067.76	2075.75	4.833	0.76	7.99		2076
26	26.77	0.461	2078.22	2086.27	4.810	0.77	8.05		2086
27	27.78	0.459	2088.66	**2096.76**	4.781	0.78	8.10		**2100**
28	28.78	0.455	**2099.08**	2107.23	4.756	0.78	8.15		2107
29	29.78	0.451	2109.52	2117.71	4.728	0.78	8.19		2118
30	30.78	0.445	2119.99	2128.24	4.704	0.78	8.25		2128
31	31.78	0.439	2130.53	2138.83	4.681	0.78	8.30		2139
32	32.77	0.431	2141.17	2149.51	4.652	0.77	8.34		2150
33	33.77	0.423	2151.92	2160.30	4.628	0.77	8.38		2160
34	34.76	0.414	2162.82	2171.25	4.602	0.76	8.43		2171
35	35.75	0.403	2173.90	2182.37	4.577	0.75	8.47		2182
36	36.74	0.392	2185.20	2193.72	4.553	0.74	8.52		2194
37	37.72	0.380	2196.76	2205.32	4.527	0.72	8.56		2205
38	38.70	0.367	2208.62	2217.22	4.503	0.70	8.60		2217
39	39.68	0.353	2220.84	2229.49	4.478	0.68	8.65		2229
40	40.66	0.338	2233.48	2242.17	4.452	0.66	8.69		2242
41	41.64	0.322	2246.61	2255.34	4.429	0.64	8.73	*	2255
42	42.61	0.306	2260.32	2269.10	4.407	0.61	8.78	*	2269
43	43.58	0.288	2274.72	2283.51	4.368	0.58	8.79	*	2284
44	44.55	0.270	2289.86	2298.76	4.382	0.55	8.90	*	2299
45	45.51	0.251	2306.10	2315.06	4.373	0.51	8.96	*	2315
46	46.48	0.231	2323.55	2332.58	4.375	0.48	9.03		2333
47	47.44	0.210	2342.53	2351.64	4.381	0.44	9.11		2352
48	48.39	0.188	2363.44	2372.65	4.403	0.39	9.21		2373
49	49.35	0.165	**2386.91**	**2396.25**	4.442	0.35	9.34		**2400**
50	50.30	0.141	2413.89	2423.41	4.512	0.30	9.52		2423
51	51.24	0.116	2445.98	2455.80	4.641	0.24	9.82		2456
52	52.19	0.089	2486.19	2496.56	4.897	0.19	10.37		2497
53	53.13	0.061	2541.25	2553.08	5.587	0.13	11.83		2553
54	54.07	0.032	2632.85	2641.36	4.040	0.07	8.51		2641
55	54.72	0.010	2761.11	2725.23	1.243	-0.28	-35.88		2761

Table 9. Summary of the April 2008 Exit Level Online TAKS Comparability Analyses

Grade /Subject	Sample Size		Raw Score Cuts*		Pass Rates (%) Comparison**			Number of Raw Score Points with Meaningful Difference***	Decision (Conclusion)
	Online	Paper	Online	Paper	Paper w/Paper	Online w/Online	Online w/Paper		
Exit Level Math	1076	12354	**32**	33	24.6	24.1	21.1	50(60)	Alternate Conversion Table (Mode Effect)
			53	53	0.1	0.2	0.2		
Exit Level Science	966	11376	**29**	30	29.0	29.7	24.8	41(55)	Alternate Conversion Table (Mode Effect)
			50	50	0	0	0		

*: Raw score points corresponding to 'Met Standards' (top) and 'Commended' (bottom) levels. Final RS cuts in bold.

**: 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 10. Summary of Comparability Analysis – April Exit Level Mathematics

RAW	CBT_RS	RS_SD	PAP_SS	CBT_SS	SS_SD	RS_DIF	SS_DIF	SIG	Final
0	0.34	0.009	1307.21	1354.58	1.208	0.34	47.37		1307
1	1.11	0.028	1448.61	1460.17	2.793	0.11	11.56		1460
2	2.22	0.053	1549.94	1563.40	3.210	0.22	13.46		1563
3	3.32	0.076	1610.86	1625.12	3.359	0.32	14.26		1625
4	4.42	0.097	1655.25	1669.93	3.417	0.42	14.68		1670
5	5.50	0.116	1690.61	1705.54	3.437	0.50	14.93	*	1706
6	6.59	0.133	1720.26	1735.35	3.436	0.59	15.09	*	1735
7	7.66	0.150	1746.00	1761.19	3.418	0.66	15.19	*	1761
8	8.74	0.164	1768.87	1784.13	3.391	0.74	15.25	*	1784
9	9.81	0.178	1789.58	1804.87	3.352	0.81	15.29	*	1805
10	10.87	0.191	1808.59	1823.89	3.307	0.87	15.30	*	1824
11	11.93	0.202	1826.23	1841.53	3.267	0.93	15.30	*	1842
12	12.98	0.212	1842.76	1858.05	3.233	0.98	15.30	*	1858
13	14.04	0.222	1858.36	1873.66	3.212	1.04	15.30	*	1874
14	15.08	0.230	1873.19	1888.53	3.179	1.08	15.34	*	1889
15	16.13	0.238	1887.43	1902.67	3.152	1.13	15.24	*	1903
16	17.17	0.245	1901.00	1916.24	3.134	1.17	15.24	*	1916
17	18.21	0.251	1914.10	1929.34	3.119	1.21	15.24	*	1929
18	19.24	0.257	1926.80	1942.04	3.105	1.24	15.24	*	1942
19	20.27	0.262	1939.14	1954.38	3.092	1.27	15.24	*	1954
20	21.30	0.266	1951.17	1966.41	3.080	1.30	15.24	*	1966
21	22.33	0.270	1962.94	1978.18	3.067	1.33	15.24	*	1978
22	23.35	0.273	1974.48	1989.72	3.055	1.35	15.24	*	1990
23	24.37	0.275	1985.83	2001.07	3.044	1.37	15.24	*	2001
24	25.38	0.277	1997.01	*2012.24*	3.028	1.38	15.23	*	*2015*
25	26.40	0.279	*2008.06*	2023.27	3.022	1.40	15.21	*	2023
26	27.41	0.280	2018.97	2034.20	3.014	1.41	15.23	*	2034
27	28.42	0.280	2029.81	2045.04	3.005	1.42	15.24	*	2045
28	29.42	0.280	2040.58	*2055.82*	2.997	1.42	15.24	*	*2058*
29	30.42	0.279	*2051.31*	2066.56	2.990	1.42	15.25	*	2067
30	31.42	0.278	2062.03	2077.28	2.985	1.42	15.26	*	2077
31	32.42	0.277	2072.74	2088.00	2.980	1.42	15.26	*	2088
32	33.41	0.274	2083.47	**2098.74**	2.975	1.41	15.27	*	**2100**
33	34.41	0.272	**2094.25**	2109.53	2.971	1.41	15.28	*	2110
34	35.39	0.269	2105.09	2120.38	2.969	1.39	15.29	*	2120
35	36.38	0.265	2116.02	2131.33	2.966	1.38	15.30	*	2131
36	37.36	0.262	2127.07	2142.38	2.965	1.36	15.32	*	2142
37	38.34	0.257	2138.25	2153.58	2.966	1.34	15.33	*	2154
38	39.32	0.252	2149.60	2164.94	2.967	1.32	15.34	*	2165
39	40.30	0.247	2161.14	2176.50	2.969	1.30	15.36	*	2176
40	41.27	0.241	2172.92	2188.29	2.972	1.27	15.37	*	2188
41	42.24	0.235	2184.96	2200.34	2.975	1.24	15.38	*	2200
42	43.20	0.228	2197.31	2212.71	2.981	1.20	15.40	*	2213
43	44.17	0.221	2210.02	2225.43	2.985	1.17	15.41	*	2225
44	45.13	0.213	2223.15	2238.57	2.989	1.13	15.42	*	2239
45	46.08	0.205	2236.76	2252.19	2.992	1.08	15.43	*	2252
46	47.04	0.196	2250.93	2266.37	2.994	1.04	15.44	*	2266
47	47.99	0.186	2265.76	2281.21	2.992	0.99	15.45	*	2281
48	48.93	0.176	2281.36	2296.82	2.989	0.93	15.47	*	2297
49	49.88	0.166	2297.88	2313.37	2.981	0.88	15.49	*	2313
50	50.82	0.154	2315.52	2331.04	2.975	0.82	15.52	*	2331
51	51.75	0.143	2334.52	2350.10	2.972	0.75	15.58	*	2350
52	52.68	0.130	2355.21	2370.87	2.980	0.68	15.66	*	2371
53	53.61	0.117	**2378.07**	**2393.84**	3.000	0.61	15.77	*	**2400**
54	54.54	0.103	2403.78	2419.72	3.039	0.54	15.95	*	2420
55	55.46	0.088	2433.40	2449.61	3.100	0.46	16.22		2450
56	56.38	0.072	2468.71	2485.38	3.196	0.38	16.67		2485
57	57.29	0.056	2513.04	2530.65	3.389	0.29	17.61		2531

Table 10. Summary of Comparability Analysis – April Exit Level Mathematics
(Continued)

RAW	CBT_RS	RS_SD	PAP_SS	CBT_SS	SS_SD	RS_DIF	SS_DIF	SIG	Final
58	58.20	0.038	2574.06	2594.01	3.853	0.20	19.94		2594
59	59.10	0.020	2675.35	2689.70	2.785	0.10	14.34		2690
60	59.73	0.006	2817.73	2779.39	0.852	-0.27	-38.34		2818

Table 11. Summary of Comparability Analysis – April Exit Level Science

RAW	CBT_RS	RS_SD	PAP_SS	CBT_SS	SS_SD	RS_DIF	SS_DIF	SIG	Final
0	0.33	0.008	1397.89	1440.89	0.998	0.33	43.00		1398
1	1.09	0.025	1528.64	1537.43	2.313	0.09	8.78		1537
2	2.18	0.048	1621.95	1632.20	2.683	0.18	10.24		1632
3	3.27	0.069	1678.03	1688.89	2.830	0.27	10.86		1689
4	4.34	0.089	1718.88	1730.07	2.900	0.34	11.20		1730
5	5.42	0.108	1751.39	1762.79	2.937	0.42	11.39		1763
6	6.49	0.125	1778.66	1790.17	2.956	0.49	11.51		1790
7	7.55	0.141	1802.31	1813.90	2.962	0.55	11.59	*	1814
8	8.61	0.156	1823.32	1834.96	2.964	0.61	11.64	*	1835
9	9.67	0.170	1842.34	1854.00	2.955	0.67	11.66	*	1854
10	10.72	0.182	1859.80	1871.47	2.941	0.72	11.67	*	1871
11	11.77	0.194	1876.00	1887.71	2.932	0.77	11.71	*	1888
12	12.81	0.205	1891.25	1902.89	2.901	0.81	11.64	*	1903
13	13.86	0.214	1905.57	1917.19	2.884	0.86	11.63	*	1917
14	14.89	0.223	1919.18	1930.79	2.871	0.89	11.61	*	1931
15	15.93	0.231	1932.21	1943.80	2.857	0.93	11.59	*	1944
16	16.96	0.238	1944.73	1956.30	2.845	0.96	11.57	*	1956
17	17.99	0.245	1956.83	1968.38	2.835	0.99	11.54	*	1968
18	19.01	0.250	1968.57	1980.08	2.825	1.01	11.52	*	1980
19	20.03	0.255	1979.99	1991.48	2.816	1.03	11.49	*	1991
20	21.05	0.259	1991.15	2002.61	2.808	1.05	11.46	*	2003
21	22.07	0.263	2002.09	2013.52	2.798	1.07	11.43	*	2014
22	23.08	0.265	2012.84	2024.24	2.791	1.08	11.40	*	2024
23	24.09	0.267	2023.44	*2034.81*	2.783	1.09	11.37	*	*2035*
24	25.09	0.268	*2033.92*	2045.25	2.772	1.09	11.33	*	2045
25	26.10	0.269	2044.30	2055.59	2.766	1.10	11.29	*	2056
26	27.10	0.269	2054.60	*2065.87*	2.760	1.10	11.27	*	*2068*
27	28.09	0.268	*2064.88*	2076.11	2.753	1.09	11.23	*	2076
28	29.09	0.267	2075.14	2086.34	2.747	1.09	11.20	*	2086
29	30.08	0.265	2085.42	**2096.58**	2.740	1.08	11.16	*	**2100**
30	31.07	0.262	**2095.73**	2106.86	2.735	1.07	11.13	*	2107
31	32.06	0.259	2106.10	2117.19	2.728	1.06	11.09	*	2117
32	33.04	0.255	2116.57	2127.62	2.723	1.04	11.06	*	2128
33	34.02	0.251	2127.15	2138.17	2.717	1.02	11.02	*	2138
34	35.00	0.246	2137.89	2148.87	2.712	1.00	10.98	*	2149
35	35.98	0.240	2148.81	2159.75	2.705	0.98	10.94	*	2160
36	36.95	0.234	2159.95	2170.85	2.699	0.95	10.90	*	2171
37	37.93	0.227	2171.36	2182.21	2.693	0.93	10.86	*	2182
38	38.89	0.220	2183.07	2193.89	2.685	0.89	10.82	*	2194
39	39.86	0.212	2195.15	2205.92	2.678	0.86	10.78	*	2206
40	40.82	0.203	2207.65	2218.39	2.671	0.82	10.74	*	2218
41	41.79	0.194	2220.66	2231.35	2.662	0.79	10.70	*	2231
42	42.75	0.185	2234.25	2244.92	2.656	0.75	10.66	*	2245
43	43.70	0.175	2248.55	2259.18	2.649	0.70	10.63	*	2259
44	44.66	0.164	2263.69	2274.29	2.646	0.66	10.60	*	2274
45	45.61	0.152	2279.84	2290.43	2.651	0.61	10.59	*	2290
46	46.56	0.140	2297.24	2307.82	2.661	0.56	10.58	*	2308
47	47.51	0.128	2316.19	2326.79	2.676	0.51	10.59	*	2327
48	48.45	0.115	2337.14	2347.76	2.699	0.45	10.62		2348
49	49.39	0.101	2360.71	2371.40	2.732	0.39	10.69		2371
50	50.33	0.086	**2387.88**	**2398.71**	2.785	0.33	10.82		**2400**
51	51.27	0.071	2420.30	2431.38	2.873	0.27	11.08		2431
52	52.21	0.054	2461.04	2472.66	3.038	0.21	11.62		2473
53	53.14	0.037	2517.00	2530.13	3.467	0.14	13.13		2530
54	54.07	0.019	2610.17	2619.47	2.482	0.07	9.29		2619
55	54.72	0.006	2739.75	2703.70	0.761	-0.28	-36.05		2740

Table 12. Summary of March Exit Level Item-Level Analyses and Subgroup Analyses

Grade/ Subject	Number of Items with Significant Mode Differences*	Mean RS Difference (Effect Size) between Online and Paper over Replications**	Subgroup Analysis--Mean RS Difference (Effect Size) between Online and Paper over Replications**				
			Gender		Ethnicity		
			Male	Female	African American	Hispanic	White
Exit Level ELA	3 (52)	-0.34 (-0.03)	-0.56 (-0.05)	-0.08 (-0.01)	-0.01 (<-0.01)	-1.05 (-0.09)	0.95 (0.10)
Exit Level Math	21 (60)	**-1.07 (-0.13)**	**-0.95 (-0.12)**	**-1.15 (-0.14)**	-0.77 (-0.11)	**-0.87 (-0.11)**	**-1.38 (-0.16)**
Exit Level Science	15 (55)	**-1.03 (-0.14)**	-0.63 (-0.08)	**-1.28 (-0.18)**	**-0.86 (-0.14)**	**-1.06 (-0.15)**	**-1.03 (-0.12)**
Exit Level Social Studies	5 (55)	-0.60 (-0.06)	-0.71 (-0.07)	-0.52 (-0.05)	-0.57 (-0.07)	-0.72 (-0.08)	-0.21 (-0.02)

*: Items with significant mean differences (p<.05) over replications. Total number of items in parentheses.

**: RS Significant Differences (p<.05) are in bold-face. Negative values indicate a lower mean RS for the online group.

Table 13. Summary of April Exit Level Item-Level Analyses and Subgroup Analyses

Grade/ Subject	Number of Items with Significant Mode Differences*	Mean RS Difference (Effect Size) between Online and Paper over Replications**	Subgroup Analysis--Mean RS Difference (Effect Size) between Online and Paper over Replications**				
			Gender		Ethnicity		
			Male	Female	African American	Hispanic	White
Exit Level Math	16 (60)	**-1.37** **(-0.19)**	**-0.98** **(-0.13)**	**-1.63** **(-0.23)**	-0.67 (-0.09)	**-1.38** **(-0.19)**	**-1.92** **(-0.28)**
Exit Level Science	6 (55)	**-1.04** **(-0.16)**	-0.32 (-0.05)	**-1.43** **(-0.23)**	-0.58 (-0.09)	**-1.27** **(-0.20)**	-0.79 (-0.13)

*: Items with significant mean differences (p<.05) over replications. Total number of items in parentheses.

**: RS Significant Differences (p<.05) are in bold-face. Negative values indicate a lower mean RS for the online group.

Table 14. Summary of Item-Level Analysis – March Exit Level ELA

ITEM	CBT_PVAL	PAP_PVAL	DIF_PVAL	DIF_STD	Z_DIF	SIG	EFFECT_SIZE
1	0.59	0.65	-0.06	0.03	-2.27	*	-0.13
2	0.88	0.92	-0.03	0.02	-1.92		-0.10
3	0.87	0.87	0.00	0.02	-0.15		-0.01
4	0.90	0.88	0.02	0.02	1.38		0.07
5	0.92	0.91	0.00	0.02	0.27		0.02
6	0.54	0.57	-0.03	0.03	-1.04		-0.06
7	0.74	0.72	0.02	0.02	0.89		0.05
8	0.76	0.77	-0.01	0.02	-0.49		-0.03
9	0.75	0.75	-0.01	0.02	-0.40		-0.02
10	0.66	0.69	-0.03	0.02	-1.30		-0.07
11	0.78	0.80	-0.02	0.02	-1.15		-0.06
12	0.67	0.68	-0.01	0.02	-0.21		-0.01
13	0.74	0.76	-0.01	0.02	-0.62		-0.03
14	0.84	0.85	-0.01	0.02	-0.28		-0.02
15	0.78	0.77	0.01	0.02	0.37		0.02
16	0.65	0.66	0.00	0.03	-0.14		-0.01
17	0.64	0.65	-0.01	0.03	-0.52		-0.03
18	0.89	0.88	0.01	0.02	0.59		0.03
19	0.50	0.52	-0.02	0.03	-0.88		-0.05
20	0.81	0.84	-0.03	0.02	-1.61		-0.08
21	0.58	0.60	-0.02	0.03	-0.75		-0.04
22	0.91	0.87	0.04	0.02	2.17	*	0.12
23	0.72	0.73	-0.01	0.02	-0.35		-0.02
24	0.42	0.48	-0.06	0.03	-2.40	*	-0.12
25	0.46	0.48	-0.02	0.03	-0.78		-0.04
26	0.70	0.73	-0.03	0.02	-1.30		-0.07
27	0.77	0.75	0.02	0.02	0.94		0.05
28	0.78	0.79	-0.01	0.02	-0.60		-0.03
29	1.01	0.97	0.04	0.03	1.14		0.06
30	1.18	1.19	-0.01	0.04	-0.38		-0.02
31	1.03	0.98	0.05	0.03	1.39		0.07
32	0.49	0.48	0.01	0.03	0.56		0.03
33	0.70	0.67	0.03	0.03	1.01		0.06
34	0.63	0.61	0.02	0.03	0.77		0.04
35	0.78	0.77	0.01	0.02	0.35		0.02
36	0.61	0.64	-0.03	0.03	-1.25		-0.07
37	0.55	0.56	-0.01	0.03	-0.28		-0.01

Table 14. Summary of Item-Level Analysis – March Exit Level ELA
(Continued)

ITEM	CBT_PVAL	PAP_PVAL	DIF_PVAL	DIF_STD	Z_DIF	SIG	EFFECT_SIZE
38	0.83	0.79	0.04	0.02	1.89		0.10
39	0.90	0.90	0.00	0.02	0.07		0.00
40	0.67	0.67	0.00	0.02	0.06		0.00
41	0.64	0.66	-0.02	0.02	-0.87		-0.04
42	0.52	0.55	-0.02	0.03	-0.95		-0.05
43	0.84	0.86	-0.02	0.02	-1.02		-0.05
44	0.67	0.67	0.00	0.02	-0.07		0.00
45	0.55	0.55	0.00	0.03	0.14		0.01
46	0.78	0.76	0.01	0.02	0.56		0.03
47	0.76	0.76	0.00	0.02	0.02		0.00
48	0.73	0.73	0.00	0.02	-0.20		-0.01
49	0.66	0.63	0.03	0.02	1.07		0.05
50	0.65	0.64	0.01	0.03	0.51		0.03
51	0.77	0.80	-0.03	0.02	-1.48		-0.08
52	7.66	7.76	-0.10	0.14	-0.75		-0.03

Table 15. Summary of Item-Level Analysis – March Exit Level Mathematics

ITEM	CBT_PVAL	PAP_PVAL	DIF_PVAL	DIF_STD	Z_DIF	SIG	EFFECT_SIZE
1	0.77	0.77	-0.01	0.01	-0.61		-0.02
2	0.85	0.86	-0.01	0.01	-1.08		-0.03
3	0.57	0.57	0.00	0.01	-0.07		0.00
4	0.77	0.79	-0.01	0.01	-0.92		-0.03
5	0.55	0.58	-0.04	0.02	-2.42	*	-0.08
6	0.56	0.56	0.00	0.02	-0.08		0.00
7	0.42	0.41	0.01	0.01	1.00		0.03
8	0.28	0.27	0.00	0.01	0.33		0.01
9	0.43	0.45	-0.02	0.02	-1.04		-0.03
10	0.36	0.36	0.00	0.02	-0.19		-0.01
11	0.54	0.55	-0.02	0.02	-1.04		-0.04
12	0.56	0.55	0.01	0.02	0.50		0.02
13	0.42	0.46	-0.05	0.02	-2.62	*	-0.09
14	0.41	0.42	-0.01	0.02	-0.75		-0.03
15	0.55	0.56	-0.02	0.02	-0.92		-0.03
16	0.60	0.62	-0.02	0.01	-1.43		-0.04
17	0.67	0.69	-0.03	0.02	-1.70		-0.06
18	0.54	0.53	0.01	0.02	0.53		0.02
19	0.45	0.57	-0.12	0.01	-8.51	*	-0.24
20	0.57	0.52	0.05	0.02	2.45	*	0.09
21	0.46	0.48	-0.02	0.02	-1.49		-0.05
22	0.22	0.22	0.00	0.01	-0.16		-0.01
23	0.45	0.44	0.01	0.02	0.86		0.03
24	0.42	0.44	-0.02	0.02	-1.47		-0.05
25	0.30	0.31	-0.01	0.01	-0.99		-0.03
26	0.30	0.31	-0.01	0.01	-0.54		-0.02
27	0.78	0.80	-0.03	0.01	-2.00	*	-0.07
28	0.46	0.48	-0.02	0.02	-1.37		-0.05
29	0.34	0.34	0.00	0.01	-0.24		-0.01
30	0.22	0.27	-0.05	0.01	-3.10	*	-0.11
31	0.31	0.37	-0.06	0.02	-3.46	*	-0.13
32	0.45	0.41	0.04	0.02	2.12	*	0.07
33	0.41	0.42	0.00	0.02	-0.22		-0.01
34	0.33	0.37	-0.04	0.02	-2.27	*	-0.08
35	0.68	0.65	0.03	0.02	2.26	*	0.07
36	0.41	0.43	-0.01	0.02	-0.86		-0.03
37	0.47	0.46	0.01	0.02	0.71		0.02

Table 15. Summary of Item-Level Analysis – March Exit Level Mathematics
(Continued)

ITEM	CBT_PVAL	PAP_PVAL	DIF_PVAL	DIF_STD	Z_DIF	SIG	EFFECT_SIZE
38	0.44	0.45	-0.01	0.02	-0.54		-0.02
39	0.47	0.53	-0.05	0.02	-3.26	*	-0.11
40	0.25	0.25	0.01	0.01	0.42		0.01
41	0.25	0.26	-0.01	0.01	-1.19		-0.03
42	0.38	0.42	-0.05	0.01	-3.24	*	-0.10
43	0.52	0.48	0.04	0.02	2.67	*	0.08
44	0.74	0.75	-0.01	0.01	-0.94		-0.03
45	0.42	0.39	0.03	0.02	1.69		0.06
46	0.33	0.35	-0.02	0.02	-1.21		-0.04
47	0.36	0.36	-0.01	0.02	-0.46		-0.02
48	0.35	0.37	-0.02	0.02	-1.22		-0.04
49	0.74	0.75	0.00	0.01	-0.12		0.00
50	0.53	0.56	-0.04	0.02	-2.17	*	-0.07
51	0.20	0.23	-0.03	0.01	-2.52	*	-0.07
52	0.50	0.49	0.00	0.02	0.28		0.01
53	0.37	0.47	-0.10	0.02	-5.93	*	-0.20
54	0.63	0.73	-0.10	0.01	-6.91	*	-0.22
55	0.62	0.65	-0.03	0.02	-2.04	*	-0.07
56	0.40	0.41	-0.01	0.02	-0.67		-0.02
57	0.41	0.46	-0.05	0.02	-3.27	*	-0.10
58	0.44	0.47	-0.03	0.02	-1.77		-0.06
59	0.65	0.72	-0.06	0.02	-4.17	*	-0.14
60	0.68	0.73	-0.05	0.01	-3.05	*	-0.10

Table 16. Summary of Item-Level Analysis – March Exit Level Science

ITEM	CBT_PVAL	PAP_PVAL	DIF_PVAL	DIF_STD	Z_DIF	SIG	EFFECT_SIZE
1	0.75	0.78	-0.04	0.02	-2.36	*	-0.09
2	0.70	0.71	-0.01	0.02	-0.57		-0.02
3	0.42	0.45	-0.03	0.02	-2.05	*	-0.07
4	0.71	0.73	-0.01	0.02	-0.91		-0.03
5	0.40	0.42	-0.03	0.02	-1.71		-0.06
6	0.56	0.56	0.00	0.02	-0.16		-0.01
7	0.71	0.71	0.00	0.02	-0.32		-0.01
8	0.51	0.51	-0.01	0.02	-0.37		-0.01
9	0.64	0.62	0.02	0.02	1.46		0.05
10	0.48	0.47	0.01	0.02	0.32		0.01
11	0.53	0.58	-0.04	0.02	-2.43	*	-0.09
12	0.23	0.29	-0.06	0.02	-3.76	*	-0.14
13	0.32	0.34	-0.02	0.02	-1.34		-0.04
14	0.57	0.54	0.04	0.02	2.25	*	0.08
15	0.28	0.28	-0.01	0.02	-0.46		-0.02
16	0.58	0.61	-0.03	0.02	-1.68		-0.06
17	0.34	0.37	-0.03	0.02	-1.88		-0.06
18	0.43	0.44	-0.02	0.02	-0.90		-0.03
19	0.25	0.31	-0.06	0.01	-3.84	*	-0.12
20	0.23	0.23	-0.01	0.01	-0.43		-0.01
21	0.38	0.41	-0.03	0.02	-1.60		-0.05
22	0.38	0.45	-0.07	0.02	-4.02	*	-0.13
23	0.64	0.68	-0.03	0.01	-2.25	*	-0.07
24	0.38	0.42	-0.04	0.02	-2.38	*	-0.08
25	0.26	0.33	-0.06	0.02	-4.26	*	-0.14
26	0.44	0.45	-0.01	0.02	-0.69		-0.02
27	0.65	0.68	-0.03	0.01	-2.16	*	-0.07
28	0.76	0.78	-0.02	0.01	-1.54		-0.05
29	0.65	0.65	0.00	0.02	-0.17		-0.01
30	0.57	0.57	-0.01	0.02	-0.42		-0.02
31	0.30	0.31	-0.01	0.02	-0.71		-0.02
32	0.46	0.47	-0.01	0.02	-0.55		-0.02
33	0.49	0.53	-0.04	0.02	-2.38	*	-0.09
34	0.28	0.30	-0.02	0.02	-1.55		-0.05
35	0.55	0.55	0.00	0.02	0.15		0.00
36	0.46	0.50	-0.03	0.02	-1.96		-0.06
37	0.45	0.48	-0.03	0.02	-1.84		-0.06

Table 16. Summary of Item-Level Analysis – March Exit Level Science
(Continued)

ITEM	CBT_PVAL	PAP_PVAL	DIF_PVAL	DIF_STD	Z_DIF	SIG	EFFECT_SIZE
38	0.42	0.44	-0.02	0.02	-1.15		-0.04
39	0.59	0.57	0.02	0.02	1.30		0.04
40	0.48	0.52	-0.04	0.02	-2.35	*	-0.08
41	0.18	0.20	-0.02	0.01	-1.34		-0.05
42	0.33	0.36	-0.03	0.02	-1.71		-0.06
43	0.62	0.62	0.00	0.02	0.13		0.00
44	0.62	0.63	-0.01	0.01	-0.62		-0.02
45	0.48	0.50	-0.03	0.02	-1.59		-0.05
46	0.40	0.40	0.00	0.02	-0.26		-0.01
47	0.28	0.32	-0.03	0.02	-2.08	*	-0.08
48	0.50	0.49	0.01	0.02	0.36		0.01
49	0.36	0.36	0.00	0.01	0.13		0.00
50	0.37	0.39	-0.02	0.02	-1.55		-0.05
51	0.63	0.65	-0.02	0.02	-1.61		-0.05
52	0.40	0.41	-0.01	0.01	-0.80		-0.02
53	0.77	0.76	0.01	0.01	0.64		0.02
54	0.66	0.66	0.00	0.02	0.05		0.00
55	0.75	0.79	-0.04	0.01	-2.48	*	-0.09

Table 17. Summary of Item-Level Analysis – March Exit Level Social Studies

ITEM	CBT_PVAL	PAP_PVAL	DIF_PVAL	DIF_STD	Z_DIF	SIG	EFFECT_SIZE
1	0.90	0.92	-0.02	0.02	-0.96		-0.06
2	0.55	0.56	-0.01	0.03	-0.45		-0.02
3	0.80	0.79	0.01	0.02	0.24		0.01
4	0.76	0.79	-0.03	0.02	-1.22		-0.07
5	0.62	0.62	0.00	0.02	-0.19		-0.01
6	0.78	0.80	-0.02	0.02	-0.78		-0.04
7	0.70	0.70	0.00	0.03	-0.10		-0.01
8	0.58	0.56	0.01	0.03	0.48		0.03
9	0.51	0.51	0.00	0.03	-0.17		-0.01
10	0.30	0.29	0.01	0.03	0.37		0.02
11	0.63	0.60	0.04	0.02	1.47		0.08
12	0.51	0.52	-0.01	0.03	-0.49		-0.03
13	0.51	0.53	-0.02	0.03	-0.83		-0.05
14	0.79	0.82	-0.03	0.03	-1.25		-0.08
15	0.53	0.59	-0.07	0.03	-2.13	*	-0.13
16	0.52	0.53	-0.01	0.03	-0.26		-0.01
17	0.52	0.52	0.00	0.03	0.07		0.00
18	0.54	0.52	0.01	0.03	0.42		0.02
19	0.66	0.63	0.04	0.03	1.40		0.08
20	0.53	0.55	-0.02	0.03	-0.65		-0.04
21	0.54	0.57	-0.03	0.03	-1.24		-0.07
22	0.59	0.57	0.02	0.03	0.93		0.05
23	0.69	0.69	0.00	0.03	0.01		0.00
24	0.54	0.55	-0.02	0.03	-0.52		-0.03
25	0.72	0.73	-0.02	0.03	-0.58		-0.03
26	0.31	0.30	0.01	0.02	0.29		0.02
27	0.39	0.40	-0.01	0.03	-0.26		-0.01
28	0.40	0.40	-0.01	0.03	-0.25		-0.01
29	0.30	0.37	-0.07	0.03	-2.48	*	-0.15
30	0.62	0.62	0.00	0.03	0.00		0.00
31	0.62	0.70	-0.08	0.03	-2.69	*	-0.17
32	0.36	0.41	-0.05	0.03	-1.70		-0.10
33	0.81	0.78	0.04	0.02	1.65		0.09
34	0.54	0.59	-0.05	0.03	-1.63		-0.09
35	0.67	0.73	-0.06	0.03	-2.23	*	-0.14
36	0.43	0.44	-0.01	0.03	-0.36		-0.02
37	0.58	0.58	0.00	0.03	-0.02		0.00

Table 17. Summary of Item-Level Analysis – March Exit Level Social Studies
(Continued)

ITEM	CBT_PVAL	PAP_PVAL	DIF_PVAL	DIF_STD	Z_DIF	SIG	EFFECT_SIZE
38	0.53	0.46	0.07	0.03	2.73	*	0.14
39	0.75	0.75	-0.01	0.02	-0.26		-0.01
40	0.72	0.70	0.02	0.02	0.86		0.05
41	0.68	0.69	-0.01	0.02	-0.41		-0.02
42	0.52	0.57	-0.05	0.03	-1.74		-0.10
43	0.57	0.59	-0.03	0.02	-1.05		-0.05
44	0.59	0.64	-0.04	0.03	-1.61		-0.09
45	0.46	0.49	-0.04	0.03	-1.31		-0.07
46	0.31	0.29	0.02	0.02	0.67		0.03
47	0.66	0.69	-0.03	0.02	-1.39		-0.07
48	0.80	0.77	0.03	0.02	1.41		0.08
49	0.65	0.67	-0.01	0.03	-0.53		-0.03
50	0.77	0.73	0.04	0.02	1.71		0.10
51	0.74	0.76	-0.01	0.03	-0.51		-0.03
52	0.50	0.53	-0.03	0.03	-1.05		-0.06
53	0.72	0.72	0.00	0.02	-0.05		0.00
54	0.68	0.71	-0.03	0.02	-1.43		-0.08
55	0.74	0.76	-0.02	0.02	-0.96		-0.05

Table 18. Summary of Item-Level Analysis – April Exit Level Mathematics

ITEM	CBT_PVAL	PAP_PVAL	DIF_PVAL	DIF_STD	Z_DIF	SIG	EFFECT_SIZE
1	0.75	0.74	0.01	0.02	0.65		0.03
2	0.73	0.73	-0.01	0.02	-0.45		-0.02
3	0.71	0.72	-0.01	0.02	-0.69		-0.03
4	0.35	0.36	-0.02	0.02	-0.89		-0.04
5	0.76	0.78	-0.03	0.02	-1.62		-0.06
6	0.56	0.59	-0.02	0.02	-1.22		-0.05
7	0.56	0.57	-0.01	0.02	-0.59		-0.03
8	0.19	0.19	0.00	0.02	-0.14		-0.01
9	0.56	0.62	-0.06	0.02	-2.67	*	-0.12
10	0.64	0.65	-0.01	0.02	-0.68		-0.03
11	0.48	0.48	0.00	0.02	-0.09		0.00
12	0.63	0.65	-0.01	0.02	-0.67		-0.02
13	0.31	0.36	-0.05	0.02	-2.53	*	-0.11
14	0.18	0.21	-0.03	0.02	-1.49		-0.06
15	0.27	0.28	-0.01	0.02	-0.62		-0.02
16	0.49	0.46	0.04	0.02	1.86		0.07
17	0.36	0.40	-0.04	0.02	-2.00	*	-0.08
18	0.49	0.51	-0.01	0.02	-0.61		-0.02
19	0.73	0.73	-0.01	0.02	-0.50		-0.02
20	0.33	0.35	-0.02	0.02	-1.32		-0.05
21	0.52	0.51	0.00	0.02	0.09		0.00
22	0.29	0.31	-0.02	0.02	-1.10		-0.04
23	0.48	0.53	-0.05	0.02	-2.34	*	-0.10
24	0.27	0.29	-0.02	0.02	-0.77		-0.03
25	0.48	0.56	-0.08	0.02	-3.52	*	-0.16
26	0.62	0.62	0.00	0.02	0.20		0.01
27	0.60	0.63	-0.02	0.02	-1.35		-0.05
28	0.29	0.29	0.00	0.02	-0.12		0.00
29	0.45	0.45	0.01	0.02	0.30		0.01
30	0.35	0.34	0.01	0.02	0.51		0.02
31	0.59	0.55	0.04	0.02	1.70		0.07
32	0.41	0.41	0.01	0.02	0.25		0.01
33	0.46	0.51	-0.05	0.02	-2.08	*	-0.10
34	0.46	0.59	-0.13	0.02	-6.71	*	-0.26
35	0.30	0.32	-0.02	0.02	-1.41		-0.05
36	0.24	0.24	0.00	0.02	-0.18		-0.01
37	0.28	0.31	-0.03	0.02	-1.47		-0.06

Table 18. Summary of Item-Level Analysis – April Exit Level Mathematics
(Continued)

ITEM	CBT_PVAL	PAP_PVAL	DIF_PVAL	DIF_STD	Z_DIF	SIG	EFFECT_SIZE
38	0.16	0.18	-0.02	0.02	-1.04		-0.04
39	0.28	0.31	-0.02	0.02	-1.25		-0.05
40	0.33	0.34	-0.01	0.02	-0.50		-0.02
41	0.20	0.24	-0.04	0.02	-2.20	*	-0.09
42	0.26	0.32	-0.06	0.02	-3.29	*	-0.14
43	0.22	0.21	0.01	0.02	0.66		0.03
44	0.21	0.25	-0.04	0.02	-2.41	*	-0.09
45	0.42	0.45	-0.03	0.02	-1.44		-0.06
46	0.48	0.54	-0.05	0.02	-2.52	*	-0.11
47	0.46	0.48	-0.02	0.02	-0.83		-0.03
48	0.72	0.74	-0.02	0.02	-1.39		-0.06
49	0.31	0.37	-0.06	0.02	-2.99	*	-0.12
50	0.45	0.43	0.02	0.02	1.04		0.04
51	0.32	0.34	-0.02	0.02	-0.83		-0.04
52	0.30	0.36	-0.06	0.02	-2.92	*	-0.12
53	0.58	0.62	-0.04	0.02	-1.74		-0.07
54	0.21	0.22	-0.01	0.02	-0.85		-0.04
55	0.54	0.54	-0.01	0.02	-0.26		-0.01
56	0.51	0.53	-0.02	0.02	-0.85		-0.04
57	0.45	0.47	-0.02	0.02	-0.75		-0.04
58	0.59	0.63	-0.05	0.02	-2.48	*	-0.09
59	0.60	0.65	-0.05	0.02	-2.36	*	-0.10
60	0.69	0.77	-0.08	0.02	-3.89	*	-0.18

Table 19. Summary of Item-Level Analysis – April Exit Level Science

ITEM	CBT_PVAL	PAP_PVAL	DIF_PVAL	DIF_STD	Z_DIF	SIG	EFFECT_SIZE
1	0.67	0.72	-0.05	0.02	-2.04	*	-0.10
2	0.51	0.54	-0.03	0.02	-1.23		-0.06
3	0.28	0.27	0.01	0.02	0.28		0.01
4	0.70	0.69	0.01	0.02	0.48		0.02
5	0.41	0.45	-0.04	0.02	-1.81		-0.08
6	0.45	0.42	0.03	0.02	1.16		0.05
7	0.59	0.63	-0.04	0.02	-1.49		-0.07
8	0.70	0.71	-0.02	0.02	-0.96		-0.04
9	0.38	0.40	-0.02	0.02	-0.87		-0.04
10	0.33	0.39	-0.06	0.02	-2.58	*	-0.13
11	0.52	0.49	0.03	0.02	1.35		0.06
12	0.42	0.52	-0.10	0.02	-4.90	*	-0.20
13	0.42	0.45	-0.02	0.02	-1.02		-0.05
14	0.49	0.52	-0.03	0.02	-1.21		-0.06
15	0.68	0.70	-0.02	0.02	-1.03		-0.04
16	0.50	0.52	-0.02	0.02	-1.04		-0.04
17	0.46	0.48	-0.02	0.02	-0.82		-0.04
18	0.42	0.42	0.00	0.02	-0.15		-0.01
19	0.31	0.30	0.01	0.02	0.54		0.03
20	0.40	0.39	0.00	0.02	0.11		0.00
21	0.42	0.41	0.01	0.02	0.66		0.03
22	0.46	0.55	-0.09	0.03	-3.49	*	-0.18
23	0.35	0.37	-0.02	0.02	-1.06		-0.04
24	0.45	0.46	-0.01	0.02	-0.39		-0.02
25	0.33	0.36	-0.03	0.02	-1.27		-0.06
26	0.51	0.49	0.01	0.02	0.49		0.02
27	0.29	0.32	-0.03	0.02	-1.36		-0.07
28	0.27	0.29	-0.02	0.02	-1.24		-0.06
29	0.39	0.43	-0.04	0.02	-2.09	*	-0.09
30	0.57	0.55	0.02	0.02	0.91		0.04
31	0.17	0.18	-0.01	0.02	-0.65		-0.03
32	0.31	0.31	0.00	0.02	0.07		0.00
33	0.58	0.62	-0.04	0.02	-1.88		-0.09
34	0.54	0.57	-0.03	0.02	-1.57		-0.06
35	0.38	0.42	-0.04	0.02	-1.66		-0.07
36	0.47	0.48	-0.02	0.02	-0.86		-0.03
37	0.42	0.44	-0.01	0.02	-0.76		-0.03

Table 19. Summary of Item-Level Analysis – April Exit Level Science
(Continued)

ITEM	CBT_PVAL	PAP_PVAL	DIF_PVAL	DIF_STD	Z_DIF	SIG	EFFECT_SIZE
38	0.38	0.40	-0.02	0.02	-0.83		-0.04
39	0.45	0.46	-0.01	0.02	-0.24		-0.01
40	0.58	0.59	-0.01	0.02	-0.53		-0.02
41	0.52	0.56	-0.04	0.02	-1.80		-0.07
42	0.36	0.41	-0.04	0.02	-1.92		-0.09
43	0.44	0.43	0.01	0.02	0.42		0.02
44	0.48	0.50	-0.02	0.02	-1.15		-0.05
45	0.47	0.51	-0.03	0.02	-1.39		-0.06
46	0.57	0.61	-0.04	0.02	-1.63		-0.07
47	0.26	0.26	0.00	0.02	0.18		0.01
48	0.60	0.59	0.02	0.02	0.82		0.04
49	0.31	0.29	0.02	0.02	0.81		0.04
50	0.35	0.36	-0.01	0.02	-0.53		-0.02
51	0.48	0.54	-0.06	0.02	-2.83	*	-0.12
52	0.32	0.34	-0.02	0.02	-0.96		-0.04
53	0.68	0.72	-0.04	0.02	-1.95		-0.09
54	0.44	0.45	-0.01	0.02	-0.44		-0.02
55	0.68	0.69	-0.01	0.02	-0.65		-0.03