PARCC

Early Post-Equating Sampling

Pearson

March 24, 2016

Executive Summary

This study focuses on the psychometric considerations for obtaining robust scale scores when sampling a subset of students for post equating, such that equating and score reporting can occur earlier than would be the case if equating was based on all students testing throughout the administration window. The source data for the analyses were the eight states planning to administer PARCC in spring 2016: Colorado (CO), District of Columbia (DC), Illinois (IL), Maryland (MD), Massachusetts (MA), New Jersey (NJ), New Mexico (NM), and Rhode Island (RI).

The study had two phases of analyses: sampling analyses and post-equating analyses. The sampling analyses were conducted for all grades and assessments. The baseline data set and four sample data sets representing 25%, 30%, 40%, and 50% of the baseline were compared in terms of demographic representativeness for all assessments for evaluation. The post-equating analyses were conducted on a selection of assessments that represented the grade bands. The sample item parameter estimates were compared to the baseline item parameter estimates. The raw score to scale score conversion files were compared for meaningful differences on the reported scale score and the performance level categories.

The first analysis for sampling identified the demographic variables tending to explain more of the variability in the summative scale scores than other variables. For grades 3–8 ELA/L and math the demographic variables Students with Disabilities, ethnicity, and economically disadvantaged tended to explain more of the variability in the summative scale scores compared to gender and English language learner. For the high school assessments, the demographic variables grade level, Students with Disabilities, ethnicity, and English language learners tended to explain more of the variability in the summative scale scores.

Of the eight states scheduled to participate in the spring 2016 administrations, five states were found to be consistently represented in each of the early equating samples (25%, 30%, 40%, and 50% samples). The states were Colorado (CO), Illinois (IL), Maryland (MD), New Jersey (NJ), and New Mexico (NM). For some assessments and smaller sample sizes, the five states were not represented in the same proportion as the baseline. The proportions tended to be within 11%, with Colorado and New Jersey being overrepresented and Illinois being underrepresented. District of Columbia, Rhode Island, and Massachusetts were not represented in many of the samples due to their testing later in the school year.

For grades 3–8 ELA/L and math, sample sizes of approximately 25% of the baseline (about 85,000-90,000) were sufficient to support stable item parameter estimates, similar summative scale scores and consistent performance level agreement when compared to the baseline data sets. For grades 9–11 ELA/L, Algebra 1, Geometry, and Algebra 2, the large number of students removed from the equating due to attemptedness and filtering rules required larger sample sizes. These assessments also required larger samples for an individual item to have student responses in each of the score categories. Therefore, sample sizes of approximately 40% of the baseline (about 75,000-80,000 for Algebra 2 and grade 10 ELA/L) are recommended for stable item parameter estimates, similar summative scale scores

and consistent performance level agreement when compared to the baseline data sets. For Integrated Math I, II, III, the sample sizes for the eight states in spring 2016 were small and resulted in many items not having enough responses in all of the score categories. As a result, at least 90% of the population is recommended for these assessments to be post-equated.

Item-level and test-level analyses for the multiple samples resulted in the minimal sample size of approximately 5,000, with the average response per item around 20,000, resulting in stable statistics and robust IRT parameters. Overall, the results for each of the samples were very similar to the baseline across the assessments. The results from the study will be used to determine criteria for sample sizes and acceptable differences between the baseline demographic distributions and the sample demographic distributions. The proposed criteria reflect the study results associated with the recommended sample size.

Table of Contents

List of Tables

List of Figures

Introduction

The Partnership for Assessment of Readiness for College and Careers (PARCC) is a state-led consortium working to develop next-generation assessments that more accurately measure student progress toward college and career readiness. The PARCC assessments are aligned to the Common Core State Standards (CCSS) and include both English Language Arts/Literacy (ELA/L) and Mathematics assessments in grades 3 to 8 and high school. The PARCC assessments consist of two components: the Performance Based Assessment (PBA) and the End-of-Year assessment (EOY). Performance on both components contributes to students' summative PARCC scale score. The assessments were first administered operationally during the 2014–2015 academic year. The PBA component is administered about three quarters through the academic year and the EOY assessments is administered near the end of the school year.

The post-equating analyses for the initial (2014–2015) administration included all student data, referred to as census data, which required the scoring of all student assessments, including items that needed hand-scoring. As a result, analyses were conducted after the assessments were administered and student score reports were available several months after the administration. Pearson is investigating the possibility of early post-equating for the PARCC assessments by conducting post-equating on a representative sample of students who completed testing early in the testing window. Post-equating procedures are well-documented and have been proven effective in a wide variety of K-12 testing programs. A common approach for post-equating is to base the equating on a representative sample of students who have tested early in the administration window. This allows analyses to begin while the assessment window is still going on, and for student score reports to be available in a timelier manner. The purpose of this document is to outline analyses for evaluating an early post-equating sampling process based on operational test data from spring 2015. Multiple samples of various sizes were selected and evaluated to identify a set of criteria for pulling the early post-equating sample in spring 2016.

Section 1: Framework for Analyses

PARCC assessments are post-equated through a common-item non-equivalent groups equating design (Kolen & Brennan, 2004) in which items are estimated with the two-parameter logistic/generalized partial credit model (2PL/GPC: Muraki, 1992). An early post-equating sample allows score reporting to occur sooner; however, the representativeness of the sample and the size of the sample impacts the robustness of the calibration, scaling, and equating results. Characteristics to consider when determining representativeness include demographic variables, state representation, and performance data. The size of the sample impacts the stability of the item parameter estimates and the size of the standard errors. These psychometric considerations need to be evaluated prior to implementing an early post-equating sampling process.

The research questions of interest for the early post-equating sampling process are:

1. Which characteristics of the student population should be considered in determining a representative sample?
2. Given the differences in states' spring testing schedules, can state-representative samples be obtained early in the administration window to support robust post-equating analyses?
3. What is the minimum sample size needed to obtain stable item parameter estimates through post-equating to support PARCC scores?
4. Which analyses and statistical criteria should be used to determine the minimum sample size for early post-equating for spring 2016?

1.1 Demographic Representativeness

The representativeness of the sample impacts the fidelity of the PARCC scores when implementing a common-item non-equivalent groups equating design (Kolen & Brennan, 2004). Demographic characteristics to consider include gender, ethnicity, economically disadvantaged, English language learners, Students with Disabilities, and grade level for high school courses. The degree to which demographic information is available early in the testing window may impact the ability to include these in the analyses.

Spring 2016 testing windows may begin as early as mid-March for some states with the majority of states testing by mid-April. Due to variability in the timing of the testing windows across states, it may be challenging to have samples that are representative of the state distribution. The distribution of states within samples will be evaluated with respect to the equating analyses. This study will use spring 2015 data from states planning to administer PARCC assessments in spring 2016.

1.2 Stability of PARCC Scores

There are four kinds of PARCC student outcomes reported: summative performance levels, full summative scale scores, reading and writing claim scores (for ELA/L only), and sub-claim performance levels. Summative scale scores range from 650 to 850 and categorize students into one of five

summative performance levels with a 700 representing the threshold of Level 2, 725 representing the threshold of Level 3, and 750 representing the threshold of Level 4 which represents college and career readiness (CCR). The threshold score for Level 5 varies slightly by test and is approximately 800. The reading and writing claim scores are reported for the ELA/L assessments only and reflect student performance on the reading and writing items, respectively. The reading claim score ranges from 10-90 and the writing claim score ranges from 10-60. ELA/L and mathematics have a subset of skills, or sub-claims, in which additional information regarding student performance is provided. The sub-claim performance levels categorize students into one of three levels based on the average performance of students at the summative Performance Level 3 and Level 4: below expectations, nearly meets expectations, or meets/exceeds expectations. The sub-claim performance levels provide information regarding targeted instructional needs.

The focus of this study is the full summative score and classification of students into the summative performance levels. The distribution of the raw scores and percent of students in the performance levels contributes to the sample size selection and whether there is sufficient data at the extremes of the summative score range. The reading and writing claim scores are generated based on the item parameter estimates that determine the full summative score; therefore, the results that are robust for the full summative score should also be robust for the claim scores. The sub-claim performance levels are based on the average performance of students at Level 3 and Level 4 on the summative score. The early post-equating sample is evaluated for minimum sample size at the Level 3 and Level 4 cut scores per form to support reporting the sub-claim performance levels.

The spring 2015 summative scale scores and performance level distributions are used to evaluate the representativeness of the samples based on the distribution of scores in the baseline data. Prior PARCC assessment data were not available for this study since spring 2015 was the first operational assessment. However, the sampling in spring 2016 can use the spring 2015 data to evaluate the representation of the students' prior performance when evaluating the early equating samples.

1.3 Item Parameter Estimates

Operational items are concurrently calibrated with the two-parameter logistic/generalized partial credit model (2PL/GPC: Muraki, 1992). The number of test forms, unique items, common-items, and parameters to estimate influence the overall sample size needed for stable estimates. Operational analyses established that estimating the 2PL/GPC item parameters for various types of PARCC items requires a minimum of 500 student responses per item; however, spring 2015 analyses indicated larger numbers of student responses may be needed to have scores at each score category (e.g., very able students to support estimation of higher categories for difficult polytomous items). Frequency distributions of the items' score categories for each of the samples determines if larger samples are needed in order to have student responses in each score category.

Section 2: Study Design

To reflect best practices for future operational administrations, early return samples should reflect the total test population in terms of performance and demographics. There are two ways to approach selection of an early return sample. The first approach requires schools to test a specific sample of students early in the testing window that results in an early return sample that is representative of the testing population. This increases the testing burden on schools and may be inconsistent with school testing schedules. The second approach is to collect the data in the order in which the data are returned and evaluate the sample for representativeness of the population. This study evaluated sample sizes to obtain a representative sample using the second approach. The study design was structured to be consistent with operational administrations and operational analyses.

This study has two phases of analyses: sampling analyses and post-equating analyses. The sampling analyses were conducted for all grades and assessments. The baseline data set and four sample data sets were identified for all assessments for evaluation. The post-equating analyses were conducted on a selection of assessments that reflected the grade bands. The sample item parameter estimates were compared to the baseline item parameter estimates. The raw score to scale score conversion files were compared for meaningful differences on the reported scale score and the performance level categories.

2.1 Data Collection

This study used the spring 2015 operational online data consisting of the states participating in the spring 2016 administration as the source data which included Colorado (CO), District of Columbia (DC), Illinois (IL), Maryland (MD), Massachusetts (MA), New Jersey (NJ), New Mexico (NM), and Rhode Island (RI). Data for states not expected to participate in spring 2016 were excluded from the study; therefore, the reported spring 2015 operational results were not used as a baseline comparison to the sample results. The data sets for the participating states were filtered to include students who took both the PBA and EOY components. The resulting data sets were defined as the baseline data sets that served as the population from which the sample data sets were created and the results were compared. Table 2.1 lists the total count of students for the online baseline data sets for each PARCC assessment.

For each student, the following information was combined:

- Summative data – performance levels, summative scale scores, demographic characteristics, and state for spring 2015.
- Scored response string for PBA and EOY – scored response string for the incomplete data matrices with indicator for records that did not meet requirements for inclusion (e.g. attemptedness).
- Testing date – student records ordered by date of administration in the baseline data.

Table 2.1 Baseline Total Counts by PARCC Online Assessments

Mathematics Assessments	Baseline N Counts	ELA/L Assessments	Baseline N Counts
Algebra 2	191,763	ELA/L Grade 11	182,112
Geometry	152,044	ELA/L Grade 10	219,484
Algebra 1	298,756	ELA/L Grade 9	267,452
Integrated Math III	10,501		
Integrated Math II	12,495		
Integrated Math I	22,689		
Math Grade 8	306,693	ELA/L Grade 8	379,162
Math Grade 7	370,028	ELA/L Grade 7	383,470
Math Grade 6	380,873	ELA/L Grade 6	383,658
Math Grade 5	363,495	ELA/L Grade 5	370,295
Math Grade 4	349,198	ELA/L Grade 4	356,184
Math Grade 3	324,657	ELA/L Grade 3	323,724

The date of administration for assessments was based on the date students finished testing in spring 2015. The test date was available for online assessments only. Paper answer documents were returned after testing windows and did not indicate the actual date the student tested. Thus, paper data will not be considered in this research. However, the results of this research will be used as a basis for further consideration for how early post-equating for paper data might be structured.

For the sampling analyses, all the spring 2015 PARCC assessments were included:

- ELA/L Grades 3–11
- Mathematics Grades 3–8
- Algebra 1, Geometry, Algebra 2
- Integrated Mathematics I, II, and III

For the post-equating analyses, the following spring 2015 PARCC assessments were selected. :

- ELA/L Grade 10
- Algebra 2
- Integrated Mathematics II
- ELA/L Grade 6
- Mathematics Grade 5
- ELA/L Grade 3

2.2 Sampling Methodology

Understanding the relationship between students' demographic characteristics and students' performance on the PARCC assessments was important in creating post-equating samples that were representative of the population. The baseline data sets were analyzed to identify the demographic characteristics associated with the summative scores. Samples were drawn from the baseline data and evaluated with respect to the demographic characteristics and student performance data. This relationship was operationalized as the amount of variance in students' summative scale scores that was explained by demographic variables through a generalized linear model: analysis of variance (ANOVA). The baseline data were used as the source for the analysis and consisted of students testing in spring 2015 that took both the EOY and PBA components and received a summative scale score.

The analysis included seven categorical demographic variables for evaluating sample representativeness: gender, ethnicity, economically disadvantaged, English language learners, student with disabilities, grade level, and state. For the analyses, gender was categorized as male or female. Economically disadvantaged, English language learner, and Students with Disabilities were categorized as yes, no, or blank. Ethnicity was based on the seven federal ethnicity categories: American Indian or Alaska Native, Asian, Black or African American, Hispanic, Native Hawaiian or Other Pacific Islander, White, and Two or More Races. A student's grade level was based on the grade level when the student was administered the test. Eight states comprised the state variable: Colorado (CO), District of Columbia (DC), Illinois (IL), Maryland (MD), Massachusetts (MA), New Jersey (NJ), New Mexico (NM), and Rhode Island (RI).

Student performance was based on students' 2015 summative scale scores, which reflected student performance on both the EOY and PBA components. The summative scale scores ranged from 650 to 850 for all assessments.

The ANOVA included the demographic variables as main effects. The interactions were not included in the ANOVA analysis. The result of interest was the effect size, Eta-squared (η^2), overall and for each of the demographic variables. The overall Eta-squared (η^2) represented the variance explained in the summative scale score based on the seven demographic variables. The Eta-squared for the demographic variables represented the variance explained by a demographic variable after the effects of the other demographic variables are partialed out of the demographic variable of interest but not partialed out of the summative scale scores. The ethnicity variable, with seven categories, represents the variance explained by the summative scale scores given the differences in the ethnicity categories. The overall Eta-squared (η^2) is expressed as

$$\hat{\eta}^2 = \frac{SS}{SS_{total}}$$

where SS_{total} is the total sample (corrected) sum of squares, and SS is the observed Type III sum of squares due to the effect being tested. The ANOVA results were used to evaluate the representativeness of the samples.

For sampling, the baseline data were ordered based on the date students finished testing the online version of the PBA component. The testing date for PBA was selected since it was the first component administered. For spring 2016, PBA and EOY will be combined into a single administration with one test date. Four samples were generated from the baseline data. The sample sizes were the first 50%, 40%, 30% and 25% of students testing in spring 2015. As a result, the smaller samples were a subset of the larger samples. The simplest method for sampling and easiest method to implement operationally was to select the percent of students from the baseline data set based on the date of administration, only. This was the initial sampling approach for this study. These data sets were referred to as the *Initial Baselin*e and *Initial Sample* data sets. Each sample was further reduced to meet the post-equating inclusion rules. These data sets were referred to as the *Equating Baseline* and *Equating Sample* data sets. The sampling process included the following steps:

1. Sort the student data by the date of administration for the PBA assessment.
2. Select the student data that reflects the sample size of interest.
3. Only students who took both EOY and PBA were included in the analyses since the spring 2016 assessment will consist of a single test form with both sections. (*Initial Data Sets*)
4. Apply the filtering rules (e.g. 25% of items attempted) used for the operational analyses.
5. Compare the sample data to the baseline data in terms of demographic characteristics and score distributions (scale score and performance level). (*Equating Data Sets*)

The differences in the samples' demographic characteristics and score distributions as compared with the baseline data were used to empirically determine the criteria needed to evaluate representativeness. After sampling, there were five data sets for each assessment in the post-equating analyses: the baseline data set and four sample data sets that reflected the early student returns and represented the baseline population in terms of demographics and score distributions.

The sample data sets were evaluated with respect to the baseline data sets for each of the demographic variables and performance data for the following analyses:

- distribution of states within each sample compared to the baseline
- percent of the equating samples to the initial baseline sample
- percent difference in the summative 2015 scale score cumulative distribution functions
- percent difference in the proportion of students in each performance level in 2015

2.3 Post-Equating Methodology

The PARCC assessments are post-equated through a common-item non-equivalent groups equating design (Kolen & Brennan, 2004). Multiple test forms were constructed using unique operational items and internal sets of common items, which allowed for equating across forms (Kolen & Brennan, 2004). Test forms were administered online and on paper. The test forms were spiraled at the student-level. PARCC components (EOY and PBA) consisted of dichotomous and polytomous items. Students' EOY and PBA response data were merged and calibrated together through an incomplete data matrix (IDM).

Items not administered to students were treated as not reached. The operational items in the IDM were concurrently calibrated with the two-parameter logistic/generalized partial credit model (2PL/GPC: Muraki, 1992). The 2PL/GPC is denoted

$$p_{im}(\theta_j) = \frac{\exp\left[\sum_{k=0}^{m} Da_i(\theta_j - b_i + d_{ik})\right]}{\sum_{v=0}^{M_i-1} \exp\left[\sum_{k=0}^{v} Da_i(\theta_j - b_i + d_{iv})\right]},$$

where $a_i(\theta_j - b_i + d_{i0}) \equiv 0$; $p_{im}(\theta_j)$ is the probability of a test taker with θ_j getting score m on item i; M_i is the number of score categories of item i with possible item scores as consecutive integers from 0 to $M_i - 1$; D is the IRT scale constant (1.7).

The item parameter estimates for the paper forms were placed on the online scale through the Stocking and Lord (1983) characteristic curve transformation method. The resulting item parameter estimates were on the same underlying scale. For each EOY and PBA form combination, the operational item parameter estimates were used to generate a raw score to theta scale, which was linearly transformed to a reported scale score.

The post-equating analyses for subsequent administrations will also include a common-item non-equivalent groups equating design to transform the item parameter estimates for an administration to the base scale. For this study, the post-equating analyses for the equating baseline data and the sample data followed the spring 2015 operational process. The equating baseline item parameter estimates and the sample item parameter estimates were transformed to the 2015 operational scale through the Stocking and Lord (1983) characteristic curve transformation method. This allowed for comparisons between the baseline and samples' item parameter estimates. Since the 2015 operational scale consisted of the same test forms as the baseline and sample data files; all the operational items were used as common items.

Raw score to theta conversion tables were generated using the baseline item parameter estimates and samples' item parameter estimates. Since the item parameter estimates were on the same scale (operational 2015 scale), the PARCC operational scaling constants and performance levels were applied to create the reported summative scale score conversion tables through a linear transformation of the theta values.

Comparisons between the equating baseline result and the four equating samples' results were done at the item level and the test level. The statistics for comparison included:

Item Level

- Item analysis: average item score, item percent correct (p-value), item-to-total correlation, percent omitted, percent attempted, and distribution of the item scores for each response category.
- IRT parameters: item parameter estimates, standard errors of the estimates, item information functions, and model fit.

Summary statistics for the items were calculated for each of the four data sets by component (EOY and PBA). Item means or average item score were computed separately and compared for each data set to the baseline results. The item percent correct was computed by dividing the average item score by the maximum possible points for each item. A one-way t-test was conducted on the item mean and the item percent correct for each sample when compared to the baseline to identify if a sample was statistically significantly different from the baseline. An effect size estimate (Cohen's D) was computed to determine whether the differences were of practical significance.

For each sample, item-to-total correlations were computed for each item and compared to the items' corresponding value for the baseline data set. Pearson correlations were generated for the average item scores, item percent correct, and item parameter estimates between the baseline and sample results.

For each item, plots of the expected item characteristic curves (ICCs) given the item parameter estimates and the observed proportion correct response relative to theta were generated. Yen's Q1 item-fit statistic (Yen, 1981) was computed for each item and compared across the baseline and sample data sets.

Test Level

- Reported scale score and percent of students in each performance level
- Test characteristic curves, test information functions and conditional standard errors

Meaningful differences at the reported scale score and the percent of students in each performance level were evaluated. Differences in test characteristic curves were reflected in conversion table differences. Comparing the sample conversion tables to the baseline conversion table provided a direct evaluation of whether the equating was robust to changes in the various equating samples. In evaluating the effectiveness of equating in generating conversion tables, the "difference that matters" criterion (Dorans & Feigenbaum, 1994; Dorans, Holland, Thayer, & Tateneni, 2003) was used. This criterion was intended to reflect the magnitude of effect that would cause a student's scale score to round up or down based on sample size differences. Practical significance was evaluated based on scale score differences greater than half a scale score point (0.5).

As recommended by the PARCC Technical Advisory Committee, the evaluations of the results focused on the most critical results – the reported scale scores and performance levels. Scale scores for all students based on the baseline and sample calibrations were computed to examine impact on scores. Histograms

of the scale score distributions for the baseline and sample data sets were produced and visually compared for consistency. Decision consistency tables (5x5) reflecting the agreement in the number of students in each of the performance levels were created for the baseline and compared to each sample.

Section 3: Sampling Results

The sampling analyses were conducted for all PARCC assessments to inform the spring 2016 early post-equating process. All assessments were included to identify any differences or similarities between the grades and subjects. The results serve as a baseline when evaluating the spring 2016 samples. The initial analysis investigated demographic variables to consider when determining if samples were representative of the population. The remaining analyses evaluated the samples in comparison to the baseline data. This section contains the following results:

3.1 Demographic variables and student performance
3.2 Sample data sets
3.3 State representation of the baseline and sample data sets
3.4 Demographic representation of the baseline and sample data sets
3.5 Spring 2015 scale score distribution for the equating baseline and sample data sets
3.6 Spring 2015 performance level distribution for the equating baseline and sample data sets

The sampling results were consistent across grades and subjects with some differences when comparing grades 3–8 to the high school assessments. Since the results and interpretations were similar, Sections 3.2–3.5 included the sampling results for the six assessments that are also included in the equating analyses. The results for the other assessments are included in the Appendices.

3.1 Demographic Variables and Student Performance

Identifying the demographic variables that explain more variation in the summative scale scores than other variables may inform the criteria needed to evaluate early post-equating sample. This analysis was done through an Analysis of Variance (ANOVA). The analysis included seven categorical demographic variables for evaluating sample representativeness: gender, ethnicity, economically disadvantaged, English language learners, student with disabilities, grade level, and state. The dependent variable was the spring 2015 summative scale scores.

ANOVA required that each student had values for each of the variables in the model; otherwise, the student data was dropped from the ANOVA analysis. Consequently, many students were dropped from the analysis due to missing values for the variables: Students with Disabilities, economically disadvantaged and English language learner. In the case of the variable Students with Disabilities, as many as a third of the population were removed because states did not include this information for any of their students. As a result, ANOVAs were run both with and without the variable Students with Disabilities. The decrease in the R^2 when not including Students with Disabilities was fairly large (ranging from 0.10 – 0.12) and the rank ordering of the demographic variables by the magnitude of the Eta-squared value was consistent with the rank order of the variables when the Students with Disabilities was included. The Students with Disabilities variable is a relatively strong predictor of performance even though it is missing for a lot of the students. Therefore the ANOVA results for this study included the variable Students with Disabilities. For the sampling analyses and post-equating analyses, students

missing values for the Students with Disabilities, economically disadvantaged and English language learner variables were not removed from the analyses.

Due to large sample sizes, the statistical tests were significant for all analyses. The result of interest was the effect size, Eta-squared (η^2), overall and for each of the demographic variables. For example, the Eta-squared for gender represented the variance explained in the summative scale scores based on gender after the effects of the other demographic variables are partialed out of the gender variable but not partialed out of the summative scale scores. The ethnicity variable, with seven categories, represented the variance explained by the summative scale scores given the differences in the ethnicity categories. Tables 3.1–3.4 lists the Eta-squared values for the overall ANOVA and by demographic variable included in the model for ELA/L grades 3–8, mathematics grades 3–8, ELA/L grades 9–11, and high school math, respectively. The tables also included the total N count for the baseline data and the N count for the data included in the ANOVA.

For grades 3–8, the overall Eta-squared for ELA/L ranged from 0.31 to 0.39 (Table 3.1) and the overall Eta-squared for mathematics ranged from 0.29 to 0.35 (Table 3.2). Tables 3.1–3.2 indicated that the variable Students with Disabilities had the largest Eta-squared values for grades 3–8 ELA/L and math, respectively. For ELA/L grades 3–8, the variable economically disadvantaged had the next highest Eta-squared values with the variable ethnicity having slightly smaller values. For mathematics grades 3–8, the variable ethnicity had the next highest Eta-squared values with the variable economically disadvantaged having slightly smaller values. The remaining variables (gender, English language learner, grade level, and state) had very small Eta-squared values ranging from 0 – 0.0275.

Table 3.1 Eta-Squared Overall and by Demographic Variable for the Baseline Data for ELA/L Grades 3–8

Variable	ELA/L Grade 3	ELA/L Grade 4	ELA/L Grade 5	ELA/L Grade 6	ELA/L Grade 7	ELA/L Grade 8
Overall Eta-Squared	0.31	0.37	0.37	0.38	0.39	0.38
Gender	0.0063	0.0104	0.0113	0.0199	0.0250	0.0275
Economically Disadvantaged	0.0432	0.0419	0.0415	0.0352	0.0336	0.0297
English Language Learner	0.0199	0.0250	0.0221	0.0251	0.0264	0.0254
Student with Disabilities	0.0677	0.0855	0.0945	0.1044	0.0955	0.0946
Ethnicity	0.0311	0.0351	0.0325	0.0322	0.0372	0.0395
Grade Level	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
State	0.0092	0.0155	0.0170	0.0198	0.0208	0.0184
Included N Count	233,044	260,594	261,920	248,733	247,722	233,832
Total N Count	319,153	350,900	364,494	373,911	370,449	362,589

Table 3.2 Eta-Squared Overall and by Demographic Variable for the Baseline Data for Math Grades 3–8

Variable	Math Grade 3	Math Grade 4	Math Grade 5	Math Grade 6	Math Grade 7	Math Grade 8
Overall Eta-Squared	0.29	0.32	0.32	0.35	0.34	0.29
Gender	0.0004	0.0003	0.0000	0.0000	0.0001	0.0015
Economically Disadvantaged	0.0389	0.0373	0.0349	0.0350	0.0274	0.0177
English Language Learner	0.0117	0.0135	0.0124	0.0158	0.0176	0.0162
Student with Disabilities	0.0634	0.0690	0.0730	0.0877	0.0985	0.0870
Ethnicity	0.0471	0.0533	0.0494	0.0521	0.0441	0.0398
Grade Level	0.0001	0.0001	0.0000	0.0001	0.0005	0.0016
State	0.0122	0.0184	0.0198	0.0174	0.0193	0.0147
Included N Count	237,697	256,498	257,559	247,875	237,035	168,444
Total N Count	324,214	346,607	360,018	373,082	359,548	294,987

For high school ELA/L assessments, the overall Eta-squared for ELA/L ranged from 0.22 to 0.31 (Table 3.3). Table 3.3 indicated that the variable Students with Disabilities had the largest Eta-squared values. The variables English language learner and ethnicity had the next highest Eta-squared values. The remaining variables (gender, economically disadvantaged, grade level, and state) had very small Eta-squared values ranging from 0.0015 – 0.0250.

Table 3.3 Eta-Squared Overall and by Demographic Variable for the Baseline Data for ELA/L Grades 9–11

Variable	ELA/L Grade 9	ELA/L Grade 10	ELA/L Grade 11
Overall Eta-Squared	0.31	0.26	0.22
Gender	0.0254	0.0230	0.0250
Economically Disadvantaged	0.0192	0.0196	0.0108
English Language Learner	0.0363	0.0242	0.0262
Student with Disabilities	0.0790	0.0667	0.0669
Ethnicity	0.0354	0.0359	0.0279
Grade Level	0.0015	0.0050	0.0015
State	0.0043	0.0035	0.0102
Included N Count	161,795	178,116	111,755
Total N Count	234,515	189,899	148,966

For high school mathematics assessments, the overall Eta-squared ranged from 0.24 to 0.44 (Table 3.4). Table 3.4 indicated that the variable grade level had the largest Eta-squared values. The variable Students with Disabilities had the next highest Eta-squared values with the variable ethnicity having slightly smaller values. The remaining variables (gender, English language learner, economically disadvantaged, and state) had very small Eta-squared values ranging from 0.0001 – 0.0238.

Table 3.4 Eta-Squared Overall and by Demographic Variable for the Baseline Data for High School Math

Variable	Algebra 1	Algebra 2	Geometry	Integrated Math I	Integrated Math II	Integrated Math III
Overall Eta-Squared	0.44	0.35	0.4	0.33	0.29	0.24
Gender	0.0004	0.0002	0.0001	0.0001	0.0003	0.0007
Economically Disadvantaged	0.0095	0.0091	0.0087	0.0022	0.0097	0.0019
English Language Learner	0.0045	0.0045	0.0042	0.0238	0.0076	0.0135
Student with Disabilities	0.0254	0.0190	0.0258	0.0543	0.0217	0.0297
Ethnicity	0.0231	0.0356	0.0258	0.0143	0.0227	0.0298
Grade Level	0.1723	0.1389	0.1601	0.0934	0.1060	0.0895
State	0.0063	0.0158	0.0034	0.0048	0.0067	0.0007
Included N Count	217,379	131,394	124,579	11,587	7,375	4,860
Total N Count	268,079	163,635	132,789	19,517	9,803	7,239

The ANOVA results indicated similar relationships between the demographic variables and the summative scale scores for grades 3–8 based on the subject and similar relationships for the high school assessments by subject. Although the magnitude of the Eta-squared values varied, the order of the demographic variables with the highest Eta-squared values was consistent across grades within the grade level groups and the subject groups as listed in Table 3.5. For grades 3–8 for both ELA/L and math, the demographic variables explaining the most variance were, generally, the variables Students with Disabilities, ethnicity and economically disadvantaged. For ELA/L grades 9–11, the demographic variables explaining the most variance were, generally, the variables Students with Disabilities, ethnicity and English language learner. For high school mathematics assessments, the demographic variable explaining the most variance was grade level. The variables Students with Disabilities and ethnicity explained some variance but much less than the grade level for the high school mathematics assessments.

Table 3.5 Demographic Variables Explaining the Most Variance in Summative Scale Scores

Grade and Subject	Demographic Variables in order of Magnitude for Eta-Squared		
ELA/L Grades 3–8	Student with Disabilities	Economically disadvantaged	Ethnicity
Math Grades 3–8	Student with Disabilities	Ethnicity	Economically disadvantaged
ELA/L Grades 9–11	Student with Disabilities	English language learner	Ethnicity
High School Math	Grade Level	Student with Disabilities	Ethnicity

Table 3.5 lists three demographic variables for each grade and subject that explained the most variance in summative scores compared to the other demographic variables. This information was used to evaluate the demographic representation of the samples when compared to the baseline data set for the sampling analyses.

3.2 Sample Data Sets

For each assessment, the baseline student data was sorted by the date the student tested the PBA component. Then each sample was selected based on the required sample proportion starting with the earliest test dates. As a result the samples are not independently drawn from the baseline data set. The

smaller samples are subsets of the larger samples. Figure 3.1 illustrates four samples selected from a baseline data set. For example, the students included in the 25% sample are also included in the other three sample data sets and the baseline data set.

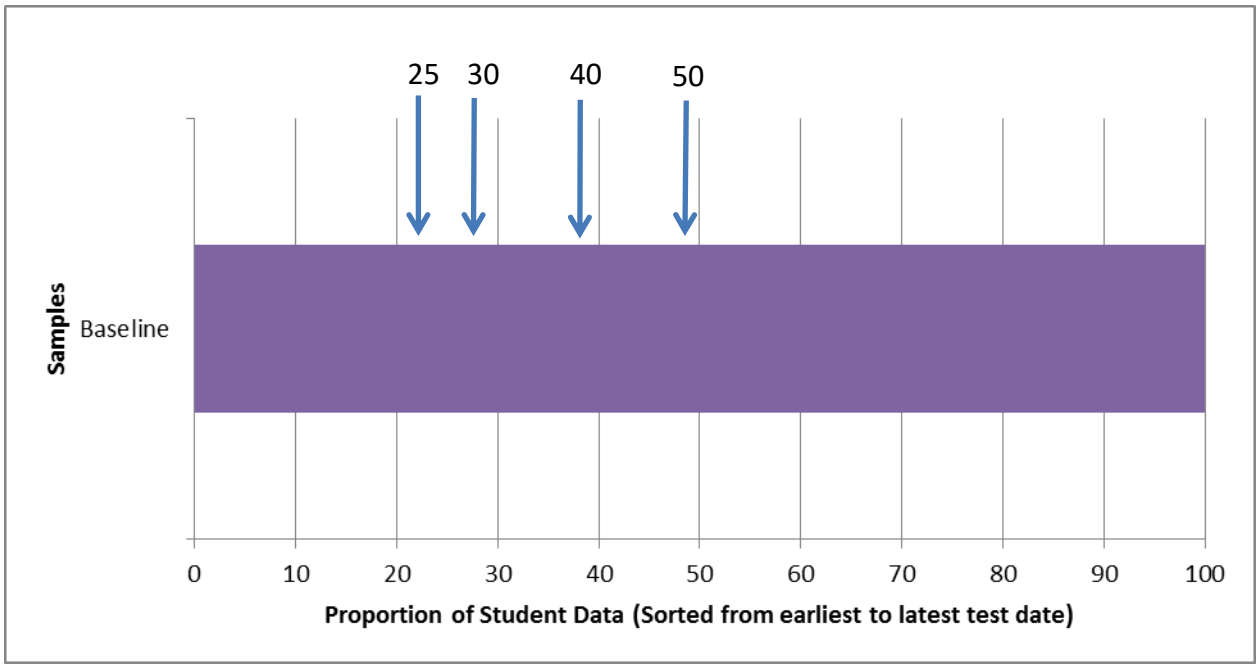

Figure 3.1 Sampling Illustration of an Initial Baseline and Sample Data Sets

The scheduled testing windows varied across the eight states testing in spring 2015. Therefore, the distribution of state representation changed across the samples. Table 3.6 and Figure 3.2 provide results for ELA/L grade 6 as an example of the state representation changing as the testing window progressed. For this illustration, the samples were generated in increments of 10% of the baseline data (100%). Table 3.6 lists the percent of students from each of the states that is represented in the samples. The overall total for each sample is also provided. Figure 3.2 illustrates the change in state representation in the data at 10% increments of the baseline data for ELA/L grade 6. The horizontal axis represents increasing the student data by 10% increments and provides the number of students included at each increment. The vertical axis represents the proportion of students from each of the states.

For example, the data set representing the first 40% of student testers in the baseline data file has a total of 153,463 students. Six of the eight states are represented in the 40% sample. However, District of Columbia is less than 1% of this sample. Figure 3.2 illustrates that for the states included in the 40% sample, their relative representation is similar to what it is in the baseline (100%) sample.

Table 3.6 Percent of State Representation of the Baseline Data for ELA/L Grade 6

State	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
CO	26.49	19.55	18.77	21.39	19.99	19.16	18.89	18.34	17.09	16.57
DC			0.01	0.14	0.24	0.32	0.39	0.53	0.66	1.06
IL	14.95	16.93	22.59	27.04	29.98	30.72	32.24	33.23	33.10	32.41
MA						0.75	1.00	1.36	2.87	5.31
MD	3.91	10.62	11.43	11.13	10.96	11.70	11.52	11.27	11.64	11.72
NJ	37.17	39.52	35.89	30.57	30.42	29.26	28.02	27.83	26.97	24.75
NM	17.48	13.38	11.30	9.74	8.41	7.85	7.43	6.91	6.51	5.97
RI					0.00	0.24	0.51	0.54	1.16	2.22
Total N Count	**38,365**	**76,731**	**115,097**	**153,463**	**191,829**	**230,194**	**268,560**	**306,926**	**345,292**	**383,658**

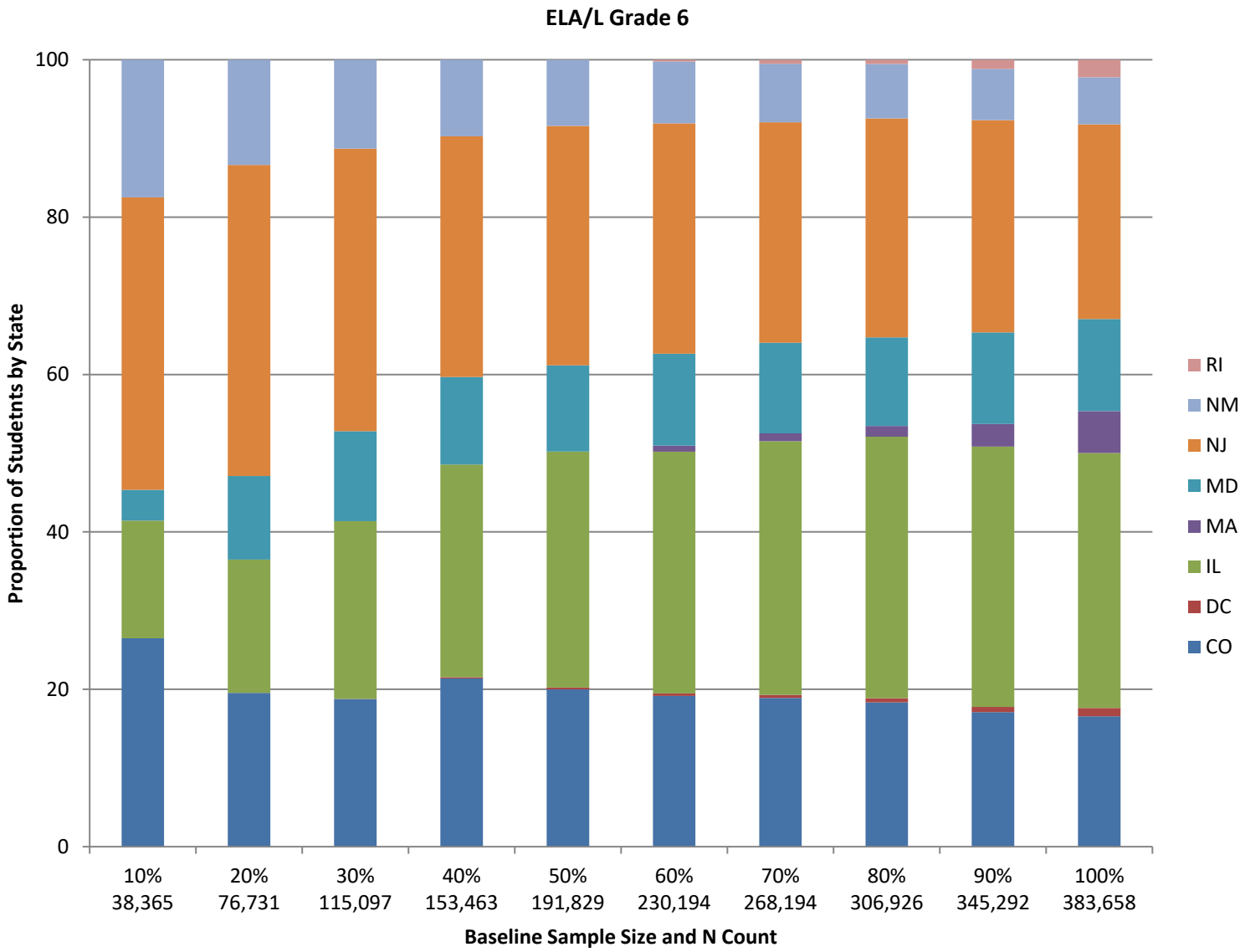

Figure 3.2 Proportion of State Representation of the Baseline Data for ELA/L Grade 6

Tables 3.7–3.12 list the number of students in the baseline data set and each of the sample data sets (25%, 30%, 40%, and 50%) for ELA/L grade 10, Algebra 2, Integrated Mathematics II, ELA/L grade 6, Mathematics grade 5, and ELA/L grade 3, respectively. The number of students in the Initial Sample column represents the proportion of the baseline data for each of the four sample data sets. The Percent of the Baseline columns is the proportion of each initial sample relative to the Baseline under the Initial Sample column. The percent reflect the desired sample size. The number of students in the Equating Sample column reflects the filtering criteria used in post-equating for the attemptedness rule. The Percent of the Baseline columns is the proportion of the equating sample relative to the Baseline under the Initial Sample column.

For Tables 3.7–3.9, the percent for the equating samples are smaller than the percent for the initial baseline samples. For ELA/L grade 10, the difference in the equating sample and the initial sample ranges between 4-6% with the total equating sample representing 86% of the total baseline sample.

Table 3.7 Percent of Students in the Baseline and Sample Data Sets for ELA/L Grade 10

	Initial Sample		Equating Sample	
Grade 10 ELA/L	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	54,871	25.00	47,746	21.79
Sample 30%	65,845	30.00	57,495	26.23
Sample 40%	87,793	40.00	77,027	35.15
Sample 50%	109,742	50.00	96,624	44.10
Baseline	219,484	100.00	189,433	86.42

For Algebra 2, the difference in the equating samples and the initial baseline samples ranges between 4-7% with the total equating sample representing 84% of the total baseline sample.

Table 3.8 Percent of Students in the Baseline and Sample Data Sets for Algebra 2

	Initial Sample		Equating Sample	
Algebra 2	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	47,940	25.00	41,030	21.40
Sample 30%	57,528	30.00	49,259	25.69
Sample 40%	76,705	40.00	65,984	34.41
Sample 50%	95,881	50.00	83,210	43.39
Baseline	191,763	100.00	162,272	84.62

For Integrated Mathematics II, the difference in the equating samples and the initial baseline samples ranges between 7-12% with the total equating sample representing 78% of the total baseline sample.

Table 3.9 Percent of Students in the Baseline and Sample Data Sets for Integrated Math II

	Initial Sample		Equating Sample	
Integrated Math II	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	3,123	24.99	2,305	18.45
Sample 30%	3,748	30.00	2,828	22.63
Sample 40%	4,998	40.00	3,871	30.98
Sample 50%	6,247	50.00	4,834	38.69
Baseline	12,495	100.00	9,748	78.02

For Tables 3.10-3.12, the percent for the equating samples are very similar to the percent for the initial baseline samples. The overall equating sample represented 97–98% of the initial baseline sample; therefore, very few students were removed due to attemptedness for the grades 3–8 assessments. The difference in the equating sample percentages for the high school assessments compared to the assessments for grades 3–8 are likely due to motivation.

Table 3.10 Percent of Students in the Baseline and Sample Data Sets for ELA/L Grade 6

	Initial Sample		Equating Sample	
Grade 6 ELA/L	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	95,914	25.00	93,786	24.45
Sample 30%	115,097	30.00	112,608	29.35
Sample 40%	153,463	40.00	150,259	39.16
Sample 50%	191,829	50.00	187,931	48.98
Baseline	383,658	100.00	373,852	97.44

Table 3.11 Percent of Students in the Baseline and Sample Data Sets for Math Grade 5

	Initial Sample		Equating Sample	
Grade 5 Math	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	90,873	25.00	89,633	24.66
Sample 30%	109,048	30.00	107,574	29.59
Sample 40%	145,398	40.00	143,548	39.49
Sample 50%	181,747	50.00	179,456	49.37
Baseline	363,495	100.00	357,829	98.44

Table 3.12 Percent of Students in the Baseline and Sample Data Sets for ELA/L Grade 3

	Initial Sample		Equating Sample	
Grade 3 ELA/L	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	80,931	25.00	79,925	24.69
Sample 30%	97,117	30.00	95,952	29.64
Sample 40%	129,489	40.00	127,996	39.54
Sample 50%	161,862	50.00	160,033	49.44
Baseline	323,724	100.00	319,133	98.58

Appendix A lists tables for the remaining PARCC assessments with the number of students in the baseline and sample data sets. These assessments show similar patterns for the initial samples and the equating samples with respect to the differences found for the high school assessments compared to the assessments in grades 3–8.

3.3 State Representation of the Baseline and Sample Data Sets

The state testing windows determine the flow of the student data and the states' representativeness in each of the samples. Tables 13–18 list the percent of the states represented in each of the sample data sets, denoted Sample 25%, 30%, 40%, and 50%, and the baseline data set for ELA/L grade 10, Algebra 2, Integrated Mathematics II, ELA/L grade 6, Mathematics grade5, and ELA/L grade 3, respectively. Appendix B lists the tables of distribution of state representation in the baseline and sample data sets for the remaining PARCC assessments. The results reflect the initial data sets prior to filtering for equating. The distribution of the states in the baseline data and the total number of students in each sample and the baseline are bolded in the tables. In addition, the baseline data sets are compared to the sample data sets through an effect size (ES) which is the proportion of students in the baseline minus the proportion of students in the sample divided by the standard deviation of the baseline.

$$ES = \frac{P_b - P_s}{\sqrt{P_b * (1 - P_b)}}$$

where P_b is the proportion of students for the baseline data set for a particular state and P_s is the proportion of students for the sample data set for a particular state. Positive effect size values indicate that the baseline percent is larger than the sample percent. Negative effect size values indicate that the baseline percent is less than the sample percent.

The results indicate that Colorado, New Jersey, and New Mexico are generally overrepresented in the samples compared to the baseline data sets and District of Columbia, Illinois, Massachusetts, Maryland, and Rhode Island are generally underrepresented in the samples compared to the baseline data sets. As expected, the differences in state distributions to the baseline distribution are larger for the smaller samples compared to the larger samples. The desired effect size is a value near zero (within +/- 0.10) which indicates the difference in the proportion of students from a state for the sample and baseline are

within 0.10 standard deviations. Due to the differences in the testing windows, the effect size values generally range from 0.11 – 0.25 indicating the proportions of the states for the samples differ with respect to the baseline distributions.

Table 3.13 Distribution of State Representation in the Baseline and Samples for ELA/L Grade 10

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	35.77	33.08	32.25	30.69	**24.05**	-0.274	-0.211	-0.192	-0.155
DC	0.05	0.04	0.15	0.17	**1.72**	0.128	0.129	0.121	0.119
IL	0.55	1.01	1.37	1.95	**2.56**	0.128	0.098	0.076	0.039
MD	14.19	13.32	13.90	15.51	**20.96**	0.166	0.188	0.173	0.134
NJ	34.53	38.45	38.94	39.00	**37.36**	0.059	-0.023	-0.033	-0.034
NM	14.91	14.09	13.40	12.68	**10.46**	-0.145	-0.119	-0.096	-0.073
RI					**2.89**				
Total	**54,871**	**65,845**	**87,793**	**109,742**	**219,484**				

Table 3.14 Distribution of State Representation in the Baseline and Samples for Algebra 2

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	21.75	23.64	23.10	21.78	**16.69**	-0.136	-0.186	-0.172	-0.137
DC					**0.12**				
IL	8.76	9.17	13.42	15.02	**17.04**	0.220	0.209	0.096	0.054
MA				0.04	**2.52**				0.158
MD	19.28	21.07	21.11	21.67	**19.87**	0.015	-0.030	-0.031	-0.045
NJ	39.38	36.32	32.53	30.54	**34.03**	-0.113	-0.048	0.032	0.074
NM	10.84	9.79	9.84	10.95	**9.73**	-0.037	-0.002	-0.004	-0.041
RI					**0.00**				
Total	**47,940**	**57,528**	**76,705**	**95,881**	**191,763**				

Table 3.15 Distribution of State Representation in the Baseline and Samples for Integrated Math II

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	94.59	95.49	96.04	95.09	**89.17**	-0.174	-0.203	-0.221	-0.190
IL					**5.05**				
MA					**0.20**				
NJ					**0.07**				
NM	5.41	4.51	3.96	4.91	**5.51**	0.004	0.044	0.068	0.026
Total	**3,123**	**3,748**	**4,998**	**6,247**	**12,495**				

Table 3.16 Distribution of State Representation in the Baseline and Samples for ELA/L Grade 6

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	19.45	18.77	21.39	19.99	**16.57**	-0.077	-0.059	-0.130	-0.092
DC		0.01	0.14	0.24	**1.06**	0.000	0.103	0.090	0.080
IL	18.57	22.59	27.04	29.98	**32.41**	0.296	0.210	0.115	0.052
MA					**5.31**				
MD	11.67	11.43	11.13	10.96	**11.72**	0.001	0.009	0.018	0.023
NJ	38.41	35.89	30.57	30.42	**24.75**	-0.317	-0.258	-0.135	-0.131
NM	11.90	11.30	9.74	8.41	**5.97**	-0.250	-0.225	-0.159	-0.103
RI					**2.22**				
Total	**95,914**	**115,097**	**153,463**	**191,829**	**383,658**				

Table 3.17 Distribution of State Representation in the Baseline and Samples for Math Grade 5

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	23.10	23.66	20.90	19.90	**15.39**	-0.214	-0.229	-0.153	-0.125
DC	0.12	0.14	0.29	0.33	**1.20**	0.099	0.097	0.084	0.080
IL	24.40	25.37	27.33	27.63	**27.33**	0.066	0.044	0.000	-0.007
MA				0.21	**5.42**				0.230
MD	10.99	11.15	14.03	14.84	**15.64**	0.128	0.123	0.044	0.022
NJ	33.25	31.38	29.98	29.22	**26.34**	-0.157	-0.115	-0.083	-0.065
NM	8.15	8.29	7.47	7.60	**6.32**	-0.075	-0.081	-0.047	-0.053
RI				0.27	**2.37**				0.138
Total	**90,873**	**109,048**	**145,398**	**181,747**	**363,495**				

Table 3.18 Distribution of State Representation in the Baseline and Samples for ELA/L Grade 3

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	21.50	19.18	22.52	21.94	**17.36**	-0.109	-0.048	-0.136	-0.121
DC	0.01	0.04	0.18	0.33	**1.61**	0.127	0.124	0.113	0.102
IL	16.00	17.77	21.66	25.30	**25.76**	0.223	0.183	0.094	0.010
MA					**6.07**				
MD	10.68	12.48	12.57	12.96	**12.92**	0.067	0.013	0.010	-0.001
NJ	43.58	43.30	35.87	33.01	**29.67**	-0.304	-0.298	-0.136	-0.073
NM	8.24	7.24	7.21	6.46	**4.40**	-0.188	-0.138	-0.137	-0.101
RI					**2.21**				
Total	**80,931**	**97,117**	**129,489**	**161,862**	**323,724**				

3.4 Demographic Representation of the Baseline and Sample Data Sets

Tables 19–24 list the percent of the demographic variables represented in each of the sample data sets, denoted Sample 25%, 30%, 40%, and 50%, and the baseline data set for ELA/L grade 10, Algebra 2, Integrated Mathematics II, ELA/L grade 6, Mathematics grade 5, and ELA/L grade 3, respectively. The results reflect the initial data sets prior to filtering for equating. Appendix C lists the tables of demographic representation in the baseline and sample data sets for the remaining PARCC assessments. The distribution of the demographic variables in the baseline data and the total number of students in each sample and the baseline are bolded in the tables. In addition, the sample data sets are compared to the baseline data sets through the difference in the percent of students in the baseline minus the percent of students in the sample for each demographic variable. This is reported as an effect size. Positive effect size values indicate that the baseline percent is larger than the sample percent. Negative effect size values indicate that the baseline percent is less than the sample percent.

For the Students with Disabilities, economically disadvantaged, or English language learner variables the tables provide distributions for the values Yes, No, and Blank. Students with missing data can often be considered to be included in the No category; however, since some states chose not to respond to the variable regardless of students representation in the category this assumption cannot be made. In evaluating these variables, the focus is placed on the representation of the students in the Yes response compared to the baseline Yes.

For grades 3–11 ELA/L and grades 3–8 math, the variable Students with Disabilities explained the most variance in the summative scores compared to the other demographic variables. For high school math, it explained the second-most variance (second to grade level). For Students with Disabilities, the results indicate that the sample Yes distributions for all four samples are within approximately 1% of the distribution for the baseline for all assessments. The difference between the baseline and the samples for the Blank and No categories varied with percent differences ranging from 3-14%. The stability of the Yes percentages relative to the baseline indicate that the variability in No and Blank categories may be explained by Blank being considered No. The effect size values for the Yes category for all of the samples compared to the baselines for Student with Disabilities are within ± 0.10.

For all the assessments, the variable Ethnicity was one of the top three variables that explained the most variance in the summative scores (see Table 3.5). The distributions for the Ethnicity categories were similar to the baseline for many of the categories. The categories representing the majority of the distribution for Ethnicity, in decreasing order, were Whites, Hispanic/Latino, African American, and Asian. In general, African Americans were underrepresented in the samples when compared to the baselines for the majority of the assessments. The sample 25% tended to have the largest percent of African American underrepresented; however, the value did not exceed eight percent with many assessments having values around five percent. The corresponding Ethnicity categories that were overrepresented tended to be the Hispanic/Latino category or the White category. The effect size values for the majority of ethnicity categories for all of the samples compared to the baselines are within ±

0.10. Grades 8–11 ELA/L and the high school mathematics assessments report effect size values ranging from 0.11–0.17 for African Americans.

For grades 3–8 ELA/L and math, the variable economically disadvantaged was one of the top three variables contributing to the variance explained in the summative scale scores. For economically disadvantaged, the results indicate that the Yes proportions for the 25% sample tended to have the largest percent difference ranging from approximately 2-6%, except for the Integrated Mathematics assessments which had percent differences ranging from 8-12% compared to the baseline. As the samples sizes increased the percent difference between the sample and baseline decreased sharply. Similar results were found for ELA/L grades 9-11, Algebra 1, Geometry, and Algebra 2, although with lower magnitudes for the percent differences. The effect size values for the majority of economically disadvantaged Yes categories for all of the samples compared to the baselines are within ± 0.10 except for a sample for grade 9 ELA/L and grade 8 mathematics.

For the variable gender, the results indicate that the sample distributions for all four samples are within one percent of the distribution for the baseline for all assessments. Gender was not a strong indicator for explaining the variability in the summative scores for any of the assessments. The effect size values for all gender distributions for all of the samples compared to the baselines are within ± 0.10.

For ELA/L grade 9-11, the variable English Language Learner accounted for the second highest amount of variability in the summative scale scores. The sample representation was very similar to the baseline for all sample sizes. Although there was more fluctuation in the percent differences for category values No and Blank, the category value Yes reported approximately less than one percent difference for all assessments and sample sizes when compared to the baseline proportions. The effect size values for the English Language Learner Yes category for all of the samples compared to the baselines are within ± 0.10.

The variable grade level contributed to the variance explained for the high school mathematics tests more than any other variable. For Algebra 1, Geometry, and Algebra 2, the distribution of grade levels for the samples were all within approximately one percent of the baseline distributions with effect size values all within ± 0.10. For Integrated Mathematics I, II, and III, the distribution of grade levels for the 25 and 30% samples ranged from 5-13% and the larger samples differed by approximately two percent with effect size values larger than 0.10.

For Integrated Mathematics I, II, and III, the effect size values for the various demographic comparisons tend to be larger than the other grades with many ranging from 0.15–0.25.

Table 3.19 Demographic Representation for Baseline and Sample Data Sets for ELA/L Grade 10

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	48.80	48.83	48.71	48.76	**48.58**	-0.004	-0.005	-0.003	-0.004
Male	51.20	51.17	51.29	51.24	**51.42**	0.004	0.005	0.003	0.004
Ethnicity									
American Indian/ Alaska Native	1.77	1.59	1.47	1.48	**1.50**	-0.022	-0.007	0.002	0.002
Asian	6.05	6.38	6.45	6.53	**6.19**	0.006	-0.008	-0.011	-0.014
African American	10.25	11.19	11.58	12.33	**15.84**	0.153	0.128	0.117	0.096
Hispanic/Latino	30.02	30.09	28.57	27.55	**25.79**	-0.097	-0.098	-0.064	-0.040
Native Hawaiian/Other Pacific Islander	0.33	0.48	0.44	0.44	**0.60**	0.034	0.015	0.020	0.020
White	45.85	44.90	46.21	46.41	**45.47**	-0.008	0.011	-0.015	-0.019
Two or More Races	2.08	1.88	1.87	1.83	**1.85**	-0.017	-0.002	-0.001	0.002
Not Provided	3.66	3.48	3.41	3.42	**2.76**	-0.055	-0.044	-0.040	-0.041
Students with Disabilities									
Blank	5.31	5.58	5.84	6.46	**7.06**	0.068	0.058	0.048	0.024
No	84.49	84.03	83.91	83.33	**82.31**	-0.057	-0.045	-0.042	-0.027
Yes	10.20	10.39	10.25	10.21	**10.63**	0.014	0.008	0.012	0.013
Economically Disadvantaged									
Blank	3.62	3.57	3.46	3.41	**2.45**	-0.075	-0.072	-0.065	-0.062
No	61.47	61.16	62.31	62.64	**61.63**	0.003	0.010	-0.014	-0.021
Yes	34.91	35.27	34.23	33.94	**35.91**	0.021	0.013	0.035	0.041
English Language Learners									
Blank	39.80	38.09	37.81	37.07	**33.24**	-0.139	-0.103	-0.097	-0.081
No	59.03	60.65	61.02	61.81	**65.49**	0.136	0.102	0.094	0.077
Yes	1.17	1.26	1.17	1.12	**1.27**	0.009	0.001	0.009	0.014
Grade Level									
Grade 9	2.04	1.88	1.80	1.77	**2.51**	0.030	0.040	0.046	0.047
Grade 10	97.35	97.48	97.48	97.47	**96.31**	-0.055	-0.062	-0.062	-0.062
Grade 11	0.55	0.59	0.66	0.69	**1.01**	0.046	0.043	0.035	0.033
Total N Count	**54,871**	**65,845**	**87,793**	**109,742**	**219,484**				

Table 3.20 Demographic Representation for Baseline and Sample Data Sets for Algebra 2

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	50.10	50.19	50.21	50.26	**50.25**	0.003	0.001	0.001	0.000
Male	49.90	49.81	49.79	49.74	**49.75**	-0.003	-0.001	-0.001	0.000
Ethnicity									
American Indian/ Alaska Native	1.08	0.99	0.89	0.89	**1.22**	0.012	0.020	0.030	0.030
Asian	8.38	7.90	8.04	7.98	**7.11**	-0.049	-0.031	-0.036	-0.034
African American	11.99	12.54	12.81	13.40	**15.12**	0.087	0.072	0.065	0.048
Hispanic/Latino	25.38	24.49	23.74	23.81	**22.21**	-0.076	-0.055	-0.037	-0.039
Native Hawaiian/Other Pacific Islander	0.41	0.55	0.85	0.74	**0.62**	0.028	0.009	-0.029	-0.015
White	50.15	50.06	50.28	49.89	**50.46**	0.006	0.008	0.004	0.011
Two or More Races	1.97	2.02	1.94	1.93	**1.57**	-0.032	-0.036	-0.030	-0.029
Not Provided	0.65	1.45	1.45	1.35	**1.70**	0.081	0.019	0.019	0.027
Students with Disabilities									
Blank	10.19	11.37	15.27	16.78	**19.37**	0.232	0.202	0.104	0.065
No	83.90	82.82	79.32	77.85	**74.52**	-0.215	-0.191	-0.110	-0.076
Yes	5.91	5.81	5.41	5.37	**6.11**	0.009	0.013	0.029	0.031
Economically Disadvantaged									
Blank	0.31	1.12	1.04	0.93	**2.89**	0.154	0.106	0.111	0.117
No	68.86	68.39	67.92	67.49	**64.03**	-0.101	-0.091	-0.081	-0.072
Yes	30.83	30.49	31.04	31.58	**33.08**	0.048	0.055	0.043	0.032
English Language Learners									
Blank	36.39	38.19	41.42	41.52	**40.62**	0.086	0.050	-0.016	-0.018
No	63.01	61.26	58.00	57.81	**58.73**	-0.087	-0.051	0.015	0.019
Yes	0.61	0.55	0.57	0.67	**0.65**	0.005	0.012	0.009	-0.002
Grade Level									
Grade 9	9.37	9.28	9.28	9.38	**8.74**	-0.022	-0.019	-0.019	-0.023
Grade 10	35.76	36.04	37.35	37.33	**34.65**	-0.023	-0.029	-0.057	-0.056
Grade 11	50.68	50.00	48.40	48.32	**51.31**	0.013	0.026	0.058	0.060
Grade 12	3.58	4.09	4.35	4.30	**4.55**	0.047	0.022	0.010	0.012
Total N Count	**47,940**	**57,528**	**76,705**	**95,881**	**191,763**				

Table 3.21 Demographic Representation for Baseline and Sample Data Sets for Integrated Math II

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	48.00	48.67	48.64	48.63	**48.74**	0.015	0.001	0.002	0.002
Male	52.00	51.33	51.36	51.37	**51.26**	-0.015	-0.001	-0.002	-0.002
Ethnicity									
American Indian/ Alaska Native	2.63	2.29	2.18	2.15	**2.26**	-0.024	-0.002	0.006	0.008
Asian	1.06	2.00	2.48	2.56	**2.42**	0.089	0.027	-0.004	-0.009
African American	1.83	1.89	2.16	2.82	**5.27**	0.154	0.151	0.139	0.110
Hispanic/Latino	22.19	21.58	22.61	24.94	**28.92**	0.149	0.162	0.139	0.088
Native Hawaiian/Other Pacific Islander		0.03	0.14	0.78	**0.50**	0.000	0.067	0.051	-0.041
White	35.83	38.50	36.25	37.75	**41.77**	0.120	0.066	0.112	0.082
Two or More Races	0.93	0.99	0.98	1.17	**1.34**	0.036	0.031	0.032	0.015
Not Provided	35.54	32.71	33.19	27.84	**17.52**	-0.474	-0.400	-0.412	-0.271
Students with Disabilities									
Blank	33.49	31.00	30.91	26.11	**21.49**	-0.292	-0.232	-0.229	-0.112
No	55.68	59.12	59.46	64.30	**69.27**	0.294	0.220	0.212	0.108
Yes	10.82	9.87	9.62	9.59	**9.24**	-0.055	-0.022	-0.013	-0.012
Economically Disadvantaged									
Blank	35.61	32.79	33.27	27.90	**17.70**	-0.469	-0.395	-0.408	-0.267
No	41.31	45.12	44.38	47.25	**50.28**	0.179	0.103	0.118	0.060
Yes	23.09	22.09	22.35	24.84	**32.02**	0.191	0.213	0.207	0.154
English Language Learners									
Blank	94.97	95.81	96.30	95.33	**94.49**	-0.021	-0.058	-0.079	-0.036
No	4.29	3.58	3.16	4.21	**5.11**	0.037	0.070	0.089	0.041
Yes	0.74	0.61	0.54	0.46	**0.39**	-0.055	-0.035	-0.024	-0.012
Grade Level									
Grade 9	28.63	23.93	19.31	19.93	**18.87**	-0.249	-0.129	-0.011	-0.027
Grade 10	49.57	57.39	64.45	63.77	**63.18**	0.282	0.120	-0.026	-0.012
Grade 11	21.33	18.28	15.47	15.51	**16.61**	-0.127	-0.045	0.031	0.030
Total N Count	**3,123**	**3,748**	**4,998**	**6,247**	**12,495**				

Table 3.22 Demographic Representation for Baseline and Sample Data Sets for ELA/L Grade 6

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	48.71	48.92	48.87	48.95	**48.92**	0.004	0.000	0.001	-0.001
Male	51.29	51.08	51.13	51.05	**51.08**	-0.004	0.000	-0.001	0.001
Ethnicity									
American Indian/ Alaska Native	1.27	1.44	1.19	1.07	**0.91**	-0.037	-0.055	-0.029	-0.017
Asian	6.62	6.34	6.23	6.15	**5.89**	-0.031	-0.019	-0.014	-0.011
African American	12.42	12.59	12.82	13.45	**15.24**	0.078	0.074	0.067	0.050
Hispanic/Latino	27.92	28.51	29.33	29.18	**27.09**	-0.019	-0.032	-0.050	-0.047
Native Hawaiian/Other Pacific Islander	0.19	0.20	0.19	0.18	**0.17**	-0.005	-0.008	-0.005	-0.002
White	48.73	48.00	47.25	47.01	**47.50**	-0.025	-0.010	0.005	0.010
Two or More Races	1.76	1.67	1.66	1.57	**1.50**	-0.021	-0.014	-0.013	-0.006
Not Provided	1.10	1.26	1.34	1.38	**1.70**	0.046	0.034	0.028	0.024
Students with Disabilities									
Blank	19.75	23.91	28.30	31.05	**34.09**	0.302	0.215	0.122	0.064
No	69.81	66.43	62.77	60.46	**57.34**	-0.252	-0.184	-0.110	-0.063
Yes	10.44	9.67	8.93	8.49	**8.57**	-0.067	-0.039	-0.013	0.003
Economically Disadvantaged									
Blank	0.50	0.56	0.60	0.53	**4.36**	0.189	0.186	0.184	0.188
No	58.42	56.93	55.50	54.89	**49.92**	-0.170	-0.140	-0.112	-0.099
Yes	41.08	42.51	43.90	44.59	**45.72**	0.093	0.064	0.037	0.023
English Language Learners									
Blank	27.86	32.04	38.11	40.62	**42.02**	0.287	0.202	0.079	0.028
No	70.89	66.57	60.62	58.19	**56.91**	-0.282	-0.195	-0.075	-0.026
Yes	1.26	1.38	1.28	1.19	**1.06**	-0.019	-0.031	-0.021	-0.012
Total N Count	**95,914**	**115,097**	**153,463**	**191,829**	**383,658**				

Table 3.23 Demographic Representation for Baseline and Sample Data Sets for Math Grade 5

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	48.61	48.63	48.80	48.71	**48.84**	0.236	0.210	0.044	0.131
Male	51.39	51.37	51.20	51.29	**51.16**	-0.236	-0.210	-0.044	-0.131
Ethnicity									
American Indian/ Alaska Native	0.75	0.75	0.77	0.80	**0.96**	0.210	0.208	0.189	0.164
Asian	6.70	6.55	6.85	6.79	**6.21**	-0.494	-0.345	-0.645	-0.584
African American	11.87	11.88	12.52	12.66	**14.77**	2.903	2.893	2.255	2.111
Hispanic/Latino	27.75	27.82	27.14	26.53	**25.34**	-2.410	-2.488	-1.806	-1.198
Native Hawaiian/Other Pacific Islander	0.17	0.16	0.15	0.17	**0.17**	-0.006	0.003	0.012	0.002
White	49.54	49.64	49.25	49.73	**49.24**	-0.303	-0.405	-0.013	-0.491
Two or More Races	1.67	1.66	1.73	1.72	**1.62**	-0.055	-0.038	-0.116	-0.104
Not Provided	1.55	1.53	1.58	1.60	**1.70**	0.155	0.173	0.124	0.101
Students with Disabilities									
Blank	25.50	26.44	28.39	28.84	**28.98**	3.48	2.54	0.59	0.15
No	64.70	63.87	62.59	62.15	**61.54**	-3.16	-2.33	-1.05	-0.61
Yes	9.80	9.69	9.02	9.01	**9.48**	-0.32	-0.22	0.46	0.47
Economically Disadvantaged									
Blank	0.58	0.58	0.53	0.66	**4.16**	3.58	3.58	3.63	3.50
No	57.46	57.01	57.02	57.04	**51.74**	-5.72	-5.27	-5.28	-5.30
Yes	41.96	42.41	42.45	42.30	**44.10**	2.14	1.69	1.65	1.80
English Language Learners									
Blank	36.37	37.70	38.32	38.41	**36.65**	0.28	-1.05	-1.67	-1.76
No	62.17	60.89	60.41	60.32	**62.04**	-0.13	1.15	1.62	1.71
Yes	1.46	1.41	1.27	1.26	**1.32**	-0.14	-0.10	0.05	0.05
Total N Count	**90,873**	**109,048**	**145,398**	**181,747**	**363,495**				

Table 3.24 Demographic Representation for Baseline and Sample Data Sets for ELA/L Grade 3

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	49.10	49.01	49.07	49.10	**48.95**	-0.003	-0.001	-0.002	-0.003
Male	50.90	50.99	50.93	50.90	**51.05**	0.003	0.001	0.002	0.003
Ethnicity									
American Indian/ Alaska Native	1.47	1.31	1.23	1.12	**0.96**	-0.053	-0.036	-0.027	-0.017
Asian	6.77	7.31	7.08	6.93	**6.27**	-0.020	-0.043	-0.033	-0.027
African American	14.00	13.98	13.47	13.45	**13.87**	-0.004	-0.003	0.012	0.012
Hispanic/Latino	28.33	27.75	27.29	26.89	**26.33**	-0.045	-0.032	-0.022	-0.013
Native Hawaiian/Other Pacific Islander	0.23	0.21	0.22	0.21	**0.19**	-0.011	-0.006	-0.007	-0.005
White	45.74	46.04	47.23	47.90	**48.75**	0.060	0.054	0.030	0.017
Two or More Races	2.16	2.16	2.16	2.13	**1.92**	-0.017	-0.017	-0.018	-0.015
Not Provided	1.30	1.24	1.33	1.36	**1.72**	0.032	0.036	0.030	0.027
Students with Disabilities									
Blank	17.76	19.32	23.01	26.50	**27.57**	0.220	0.185	0.102	0.024
No	72.73	71.23	68.21	65.14	**63.82**	-0.186	-0.154	-0.091	-0.028
Yes	9.50	9.45	8.78	8.35	**8.61**	-0.032	-0.030	-0.006	0.009
Economically Disadvantaged									
Blank	0.66	0.57	0.50	0.47	**4.40**	0.182	0.187	0.190	0.191
No	53.33	54.00	54.71	54.87	**50.92**	-0.048	-0.062	-0.076	-0.079
Yes	46.01	45.43	44.79	44.66	**44.68**	-0.027	-0.015	-0.002	0.000
English Language Learners									
Blank	25.50	26.15	32.76	36.15	**36.22**	0.223	0.209	0.072	0.001
No	72.81	72.38	65.67	62.33	**62.30**	-0.217	-0.208	-0.070	-0.001
Yes	1.68	1.47	1.57	1.52	**1.48**	-0.016	0.001	-0.007	-0.003
Total N Count	**80,931**	**97,117**	**129,489**	**161,862**	**323,724**				

3.5 Spring 2015 Scale Score Distribution for the Equating Baseline and Sample Data Sets

In addition to demographic representation, the representativeness of student performance within the samples was evaluated. The current year summative scale scores were used in this study due to spring 2015 being the first administration for the PARCC assessments and prior assessment scores were not available. For spring 2016, the 2015 summative scale score for students' prior grade assessment in the same content area (where available) will be used to evaluate the distribution of student scores. For example, when evaluating the early post-equating sample for ELA/L grade 6, the students included in the early post-equating sample in spring 2016 will be evaluated based on their prior year, ELA/L grade 5 performance, with respect to the consortiums population distribution for last year. As a result, the 2015 conversion files were used for each sample and the same 2015 scale score value was assigned to a student regardless of the sample.

For this analysis, the initial baseline data sets and initial sample data sets were filtered for attemptedness to meet the post-equating criteria for inclusion in calibrations. The spring 2015 summative scale scores and performance level designations that were reported for the students was merged with the equating baseline and sample data sets to evaluate student performance. The cumulative distributions (CDF_{base}) for the equating baseline data were compared to the cumulative distributions for each of the sample data sets ($CDF_{25\%}$, $CDF_{30\%}$, $CDF_{40\%}$, and $CDF_{50\%}$).

For each scale score, a *cumulative percent difference* (CPD) value was computed by subtracting the cumulative percent for the sample data set cumulative distribution from the cumulative percent for the baseline data set cumulative distribution at each summative scale score value. For each scale score,

$$CPD_{ss} = \mathrm{CDF}_{baseSS} - \mathrm{CDF}_{25\%SS}$$

where *SS* represents the scale score values from 650 – 850, CPD_{ss} is the cumulative percent difference for *SS*, CDF_{baseSS} is the cumulative percent for the baseline cumulative distribution at *SS*, $\mathrm{CDF}_{25\%SS}$ is the cumulative percent for Sample 25% cumulative distribution at *SS*.

Tables 3.25–3.30 provide descriptive statistics, mean, standard deviation (STD), minimum (min) and maximum (max), for the cumulative percent difference for the samples when compared to the baseline across the summative scale score for ELA/L grade 10, Algebra 2, Integrated Mathematics II, ELA/L grade 6, Mathematics grade 5, and ELA/L grade 3, respectively. The number of students in each of the samples is also provided. The number of students should be consistent with the equating baseline and sample counts, rather than the initial baseline and sample counts. Positive cumulative percent differences represent the baseline cumulative percent being larger than the sample cumulative percent at the same summative scale score. Negative cumulative percent differences represent the baseline cumulative percent being smaller than the sample cumulative percent at the same summative scale score. Appendix D provides descriptive statistics for the cumulative percent difference for the remaining PARCC assessments.

Figures 3.3–3.8 shows the cumulative distribution of the summative scale scores for the equating baseline data and sampling data sets for ELA/L grade 10, Algebra 2, Integrated Mathematics II, ELA/L grade 6, Mathematics grade 5, and ELA/L grade 3, respectively. The horizontal axis is the summative scale score values and the vertical axis is the cumulative percent for the students in each of the data sets. Appendix D provides the cumulative distributions for the remaining PARCC assessments.

For high school grades 9–11 ELA/L, the average cumulative percent differences between the baseline cumulative percent and the sample cumulative percent across all sample data sets were below 3% with many of the average cumulative percent differences being less than 2%. The maximum cumulative percent differences ranged from 4–8%. Algebra 1, Algebra 2, and Geometry had similar results for the comparison between the baseline cumulative percent distributions for the summative scale score. The average cumulative percent difference was less than 1.7% for all sample data sets and the maximum differences did not exceed 4%. The Integrated Mathematics assessments listed less stable cumulative distributions for the sample sets when compared to the baseline data sets. The cumulative distributions showed more variation graphically across the samples and when compared to the baseline. The average cumulative percent difference was approximately 3% for many of the sample data sets and the maximum differences were as high as 6-9%.

For grades 3–8 ELA/L and math, the average cumulative percent differences between the baseline cumulative percent and the sample cumulative percent across all sample data sets were below 1.0%, except for grades 5 and 8 ELA/L which have average cumulative percent differences around 1.5% for some sample sizes. The maximum cumulative percent differences were below 3.7% for all sample sizes. In general, the cumulative distributions for grades 3–8 were very similar for all sample data sets along the entire summative scale score range. For example, in Table 3.28 for grade 6 ELA/L the average cumulative percent difference for the 25% sample was 0.945% with a minimum percent difference of 0.0%. The largest cumulative percent difference between the baseline data and the 25% sample data was 2.438%.

Table 3.25 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for ELA/L Grade 10

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	47,745	2.384	1.577	0.000	4.736
Sample 30%	57,494	2.224	1.493	0.000	4.474
Sample 40%	77,026	2.474	1.552	0.000	4.723
Sample 50%	96,623	2.007	1.249	0.000	3.794

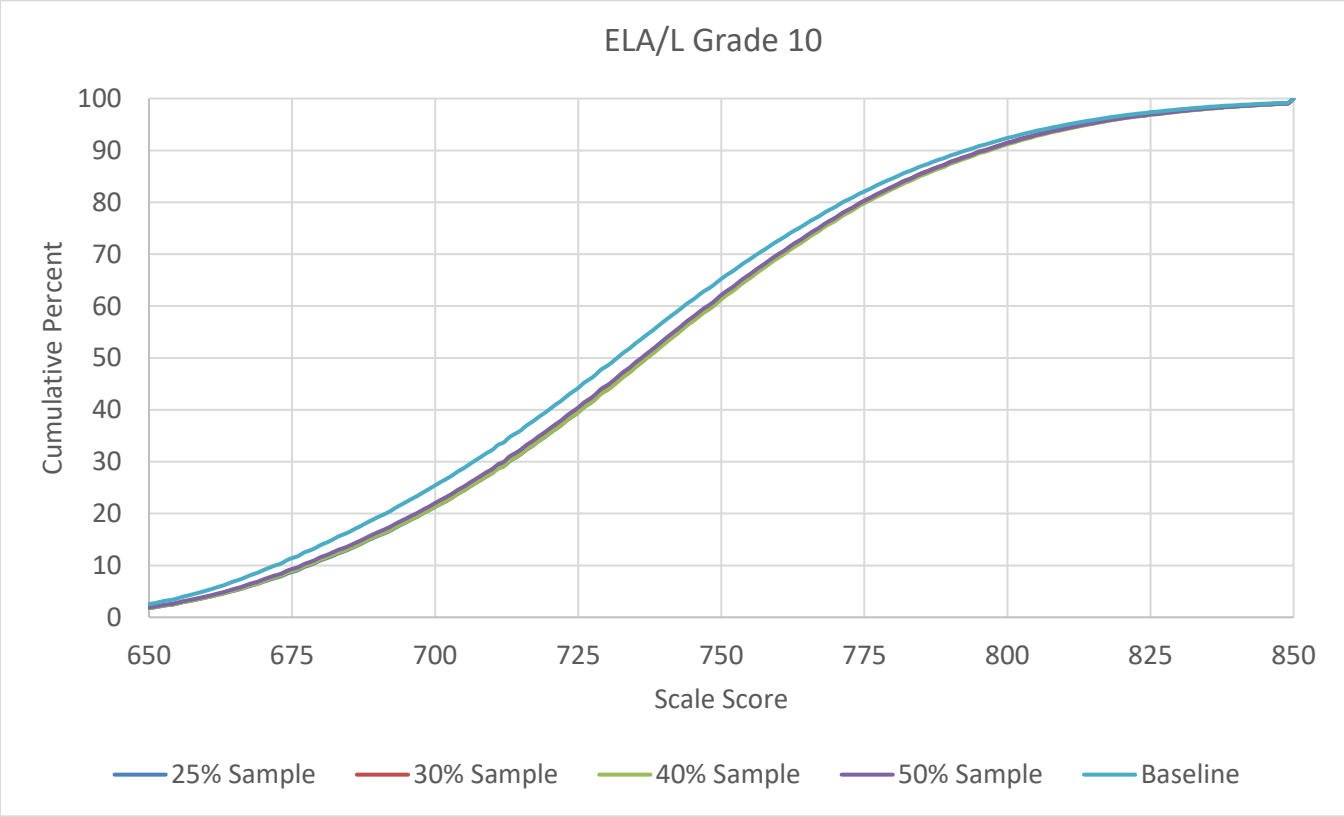

Figure 3.3 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for ELA/L Grade 10

Table 3.26 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for Algebra 2

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	**N**	**Mean**	**STD**	**Min**	**Max**
Sample 25%	41,030	1.749	1.572	-0.051	4.455
Sample 30%	49,259	1.360	1.303	-0.076	3.617
Sample 40%	65,984	1.281	1.223	-0.092	3.375
Sample 50%	83,210	1.212	1.101	-0.056	3.017

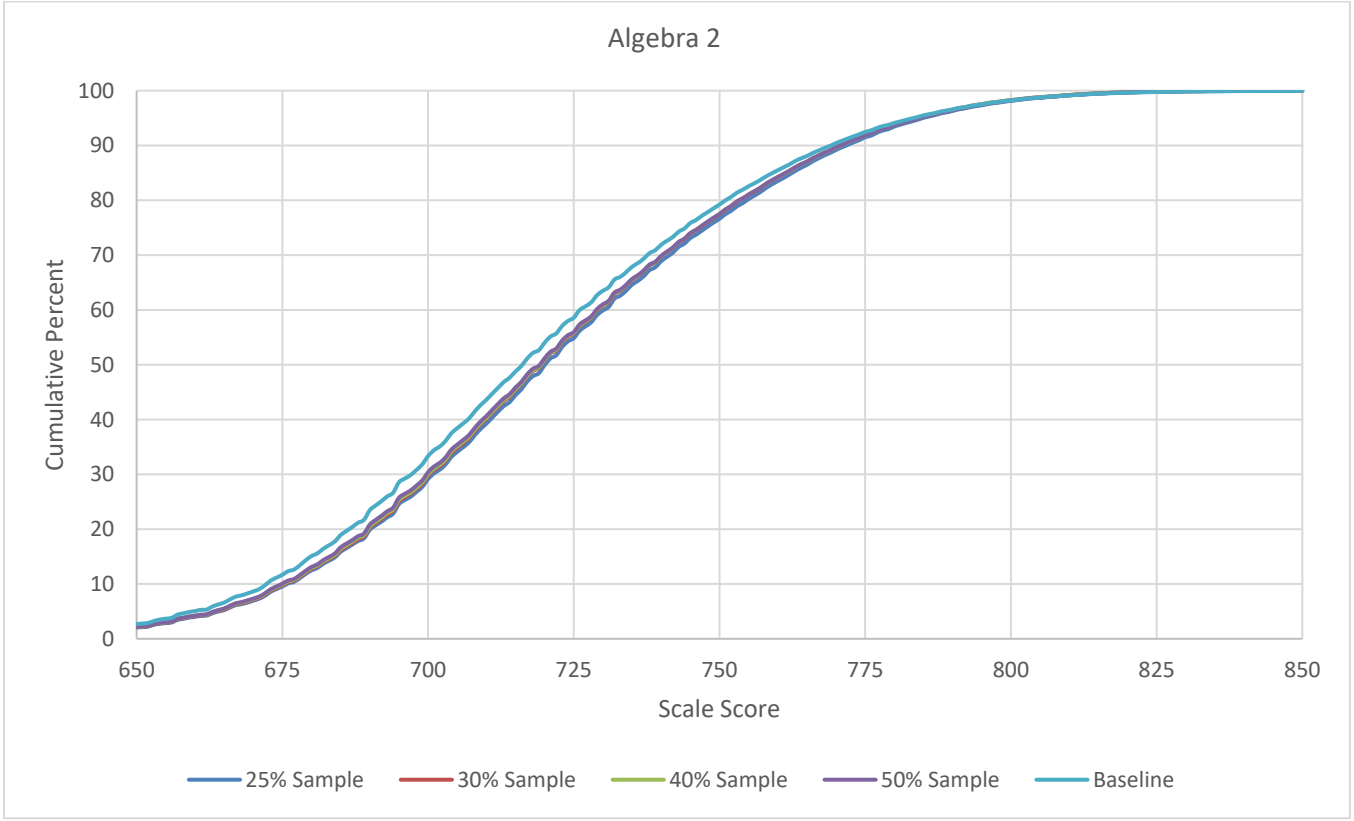

Figure 3.4 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for Algebra 2

Table 3.27 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for Integrated Math II

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	2,305	-3.178	2.650	-8.153	-0.069
Sample 30%	2,828	-0.941	1.023	-3.290	0.029
Sample 40%	3,871	-0.479	0.889	-2.661	0.434
Sample 50%	4,834	0.636	0.639	-0.219	1.915

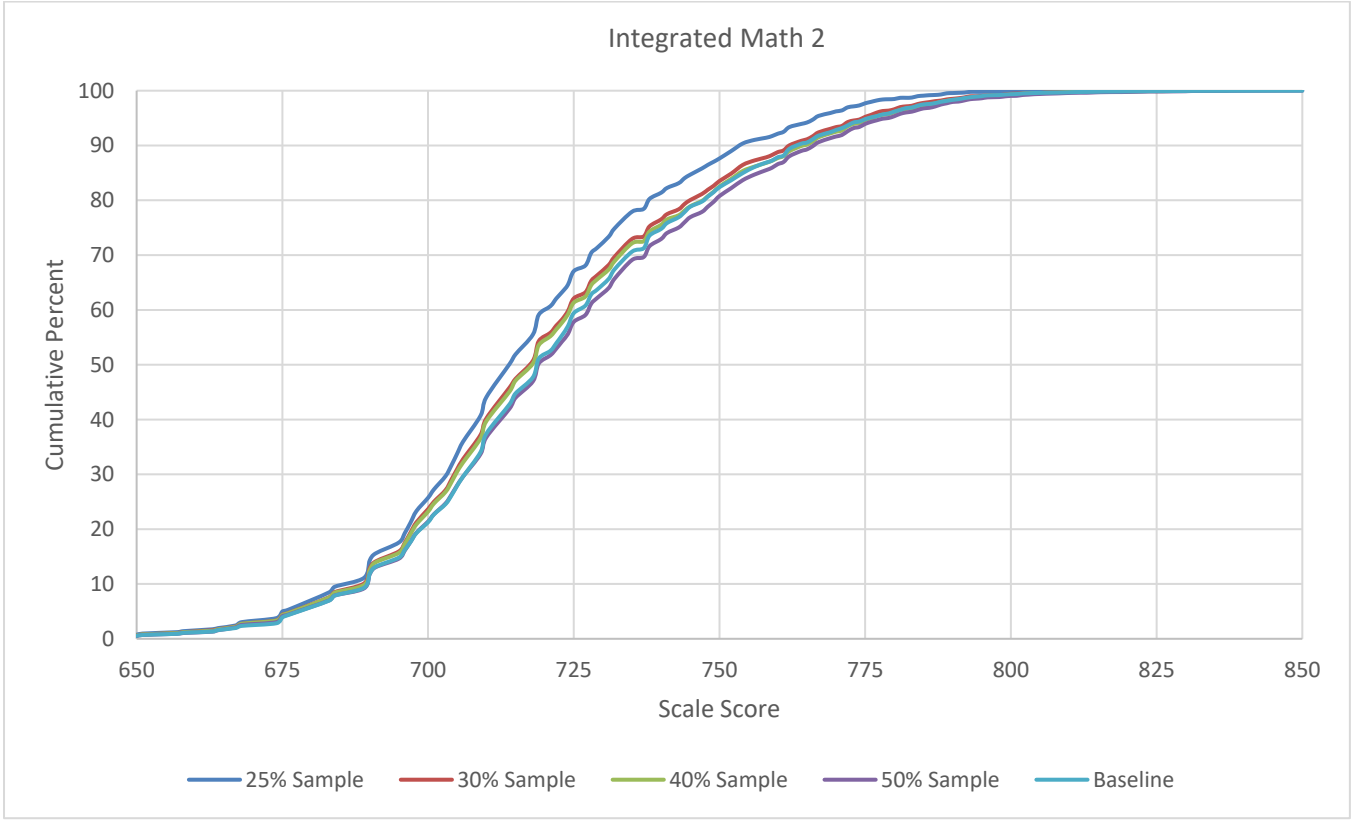

Figure 3.5 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for Integrated Math II

Table 3.28 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for ELA/L Grade 6

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	93,786	0.945	0.868	0.000	2.438
Sample 30%	112,608	0.787	0.729	-0.002	2.033
Sample 40%	150,259	0.522	0.493	-0.001	1.396
Sample 50%	187,931	0.533	0.501	0.000	1.380

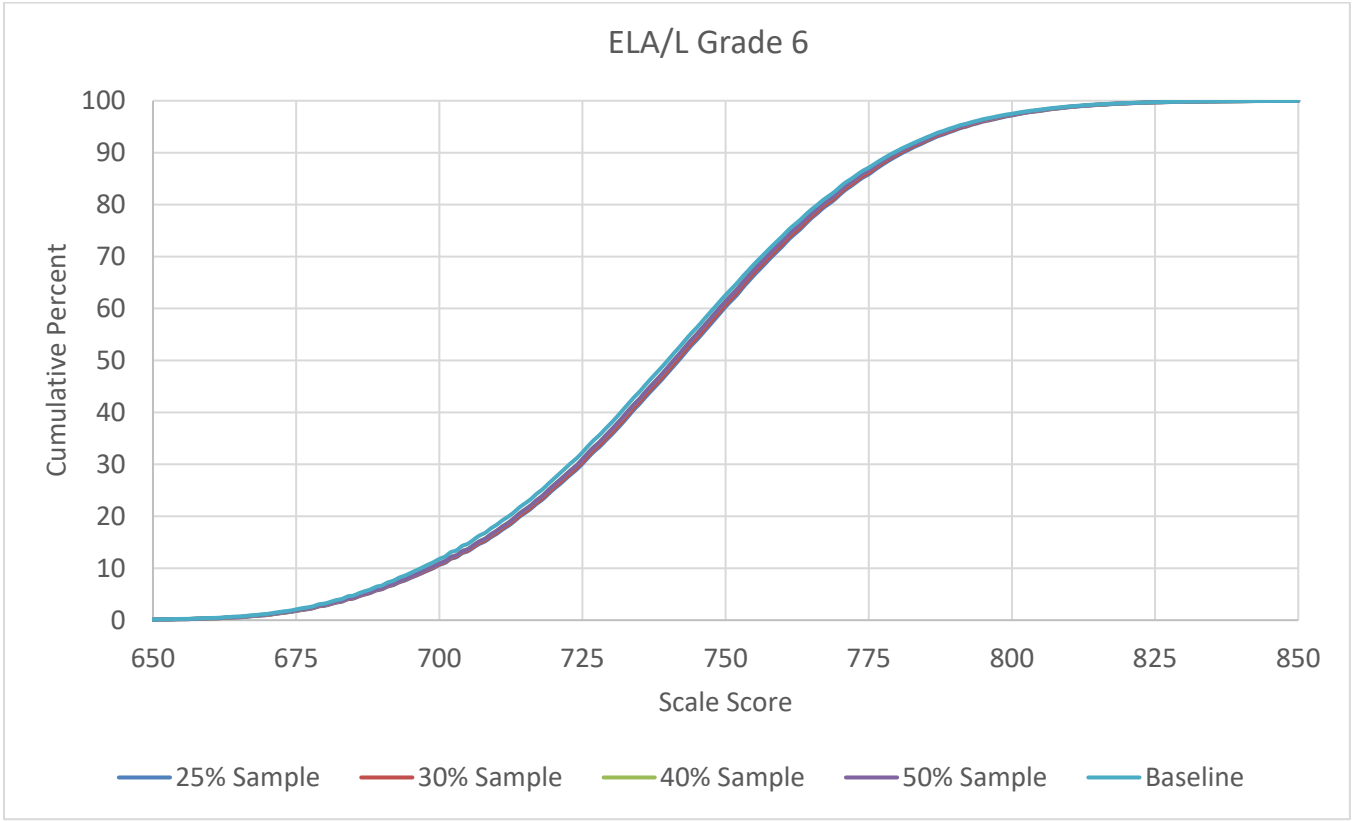

Figure 3.6 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for ELA/L Grade 6

Table 3.29 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for Math Grade 5

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	89,633	0.281	0.350	-0.018	1.116
Sample 30%	107,574	0.078	0.180	-0.162	0.520
Sample 40%	143,548	0.319	0.289	0.000	0.857
Sample 50%	179,456	0.283	0.260	0.000	0.787

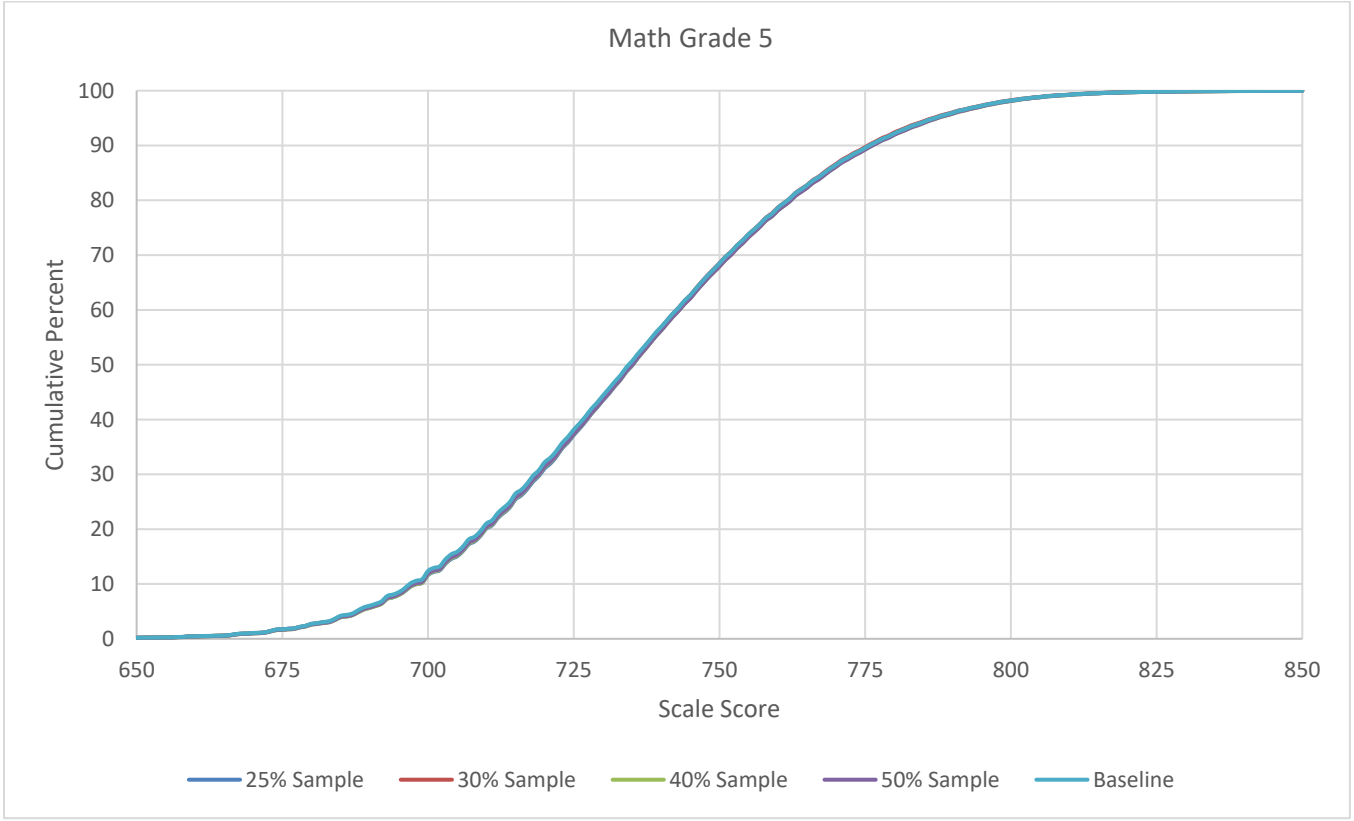

Figure 3.7 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for Math Grade 5

Table 3.30 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for ELA/L Grade 3

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	79,925	-0.214	0.220	-0.631	0.055
Sample 30%	95,952	-0.018	0.119	-0.292	0.180
Sample 40%	127,996	0.018	0.078	-0.183	0.149
Sample 50%	160,033	-0.125	0.161	-0.445	0.045

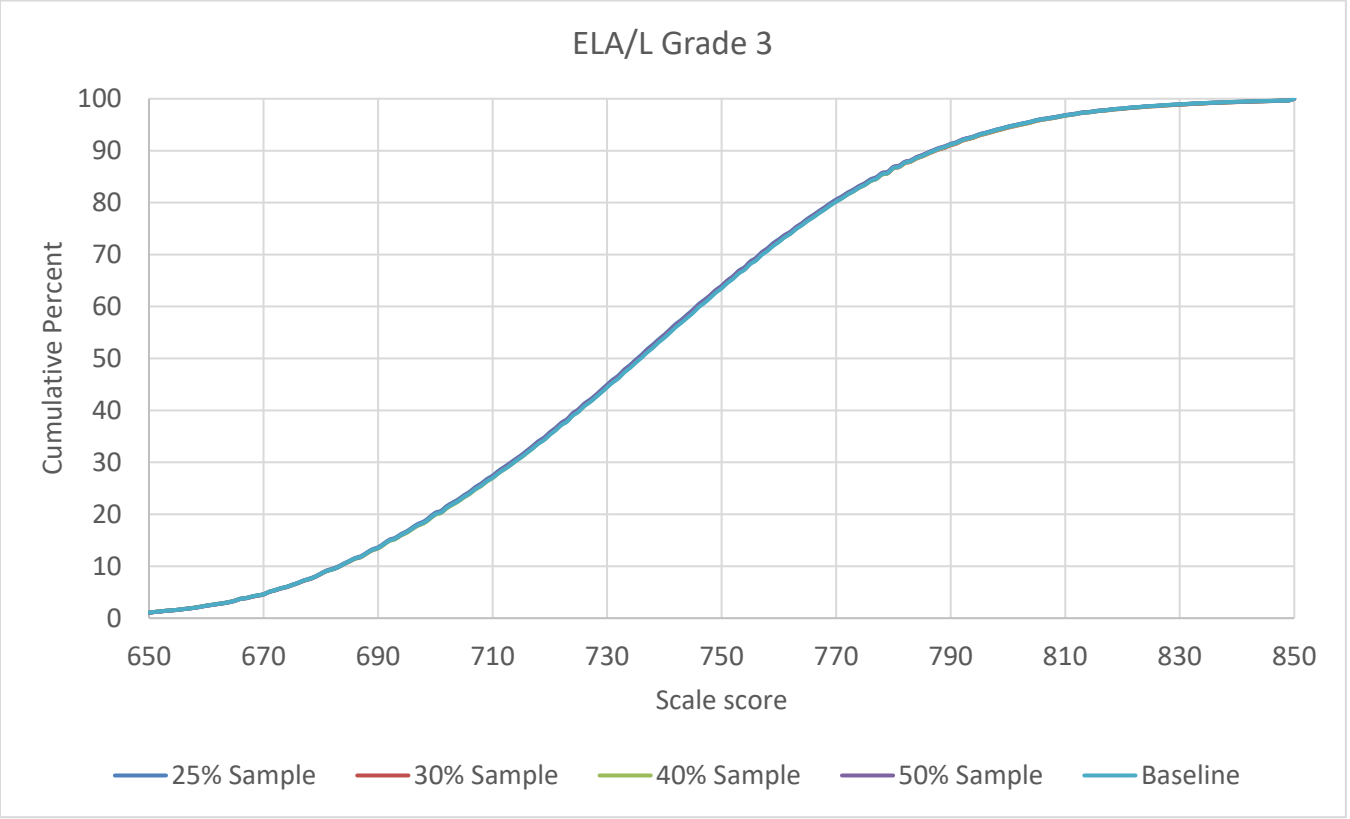

Figure 3.8 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for ELA/L Grade 3

3.6 Spring 2015 Performance Level Distribution for the Equating Baseline and Sample Data Sets

In addition to evaluating the samples based on the 2015 scale score values, the students' 2015 performance level classifications were compared. For PARCC assessments, students are classified into one of five performance levels based on their summative scale score. Level 1 represents lower performing students and Level 5 represents higher performing students. Tables 31–36 provide the performance level distributions for each of the equating sample data sets, denoted Sample 25%, 30%, 40%, and 50%, and the baseline data set for ELA/L grade 10, Algebra 2, Integrated Mathematics II, ELA/L grade 6, Mathematics grade 5, and ELA/L grade 3, respectively. The results reflect the equating data sets after filtering for calibrations. Appendix E provides tables of performance level distributions for the equating baseline and sample data sets for the remaining PARCC assessments. The total N count for the number of students in each sample and the baseline data sets are provided. In addition, the sample data sets are compared to the baseline data sets through an effect size comparing the difference in the percent of students in the sample minus the percent of students in the baseline divided by the standard deviation of the baseline for each performance level. Positive effect size values indicate that the baseline percent is larger than the sample percent. Negative effect size values indicate that the baseline percent is less than the sample percent.

For ELA/L grades 5–11 and Algebra 1 and Algebra 2, the performance levels for the samples when compared to the baseline data sets indicated a similar pattern in which the percent of students in Level 1 is smaller than the baseline and the percent of students in Level 2 is larger than the baseline data sets. The Level 2, 3, and 5 tend to be similar or less than 1-1.5% compared to the baseline. For Mathematics grades 3 and 4 and Integrated Mathematics I, II, and III, the distributions of the performance levels compared to the baseline show the samples to have slightly larger proportions in Level 1 and Level 2 and smaller proportions in Level 4. The following assessments list similar distributions across the samples when compared to the baseline samples: ELA/L grades 3–4, Mathematics grades 6–8, and Geometry. Except for grades 9 and 11 ELA/L and Integrated Mathematics I and II, the effect size values for all samples for each performance level are within ± 0.10 for all assessments. Grades 9 and 11 ELA/L and Integrated Mathematics I and II report effect size values ranging from 0.11 – 0.127.

Table 3.31 Spring 2015 Performance Level Distribution for the Equating Baseline and Sample Data Sets for ELA/L Grade 10

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	20.66	20.84	20.66	21.38	**24.77**	0.095	0.091	0.095	0.079
Level 2	18.05	18.17	18.06	18.30	**18.67**	0.016	0.013	0.016	0.009
Level 3	21.88	21.89	21.60	21.47	**20.92**	-0.024	-0.024	-0.017	-0.014
Level 4	28.22	28.01	28.25	27.72	**25.72**	-0.057	-0.052	-0.058	-0.046
Level 5	11.19	11.10	11.43	11.13	**9.92**	-0.042	-0.039	-0.050	-0.040
Total N Count	**47,746**	**57,495**	**77,027**	**96,624**	**189,433**				

Table 3.32 Spring 2015 Performance Level Distribution for the Equating Baseline and Sample Data Sets for Algebra 2

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	27.70	28.25	28.51	28.88	**31.81**	0.088	0.076	0.071	0.063
Level 2	26.57	26.83	26.74	26.50	**26.17**	-0.009	-0.015	-0.013	-0.008
Level 3	21.69	21.67	21.57	21.44	**20.57**	-0.028	-0.027	-0.025	-0.022
Level 4	22.95	22.20	22.15	22.09	**20.37**	-0.064	-0.045	-0.044	-0.043
Level 5	1.10	1.05	1.02	1.08	**1.08**	-0.002	0.003	0.006	0.001
Total N Count	**41,030**	**49,259**	**65,984**	**83,210**	**162,272**				

Table 3.33 Spring 2015 Performance Level Distribution for the Equating Baseline and Sample Data Sets for Integrated Math II

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	23.30	21.25	20.87	19.34	**19.29**	-0.102	-0.050	-0.040	-0.001
Level 2	41.30	38.54	38.31	36.26	**37.76**	-0.073	-0.016	-0.011	0.031
Level 3	22.43	22.84	22.40	24.12	**24.44**	0.047	0.037	0.047	0.007
Level 4	11.97	14.92	15.53	16.76	**15.79**	0.105	0.024	0.007	-0.027
Level 5	1.00	2.44	2.89	3.52	**2.73**	0.106	0.018	-0.010	-0.048
Total N Count	**2,305**	**2,828**	**3,871**	**4,834**	**9,748**				

Table 3.34 Spring 2015 Performance Level Distribution for the Equating Baseline and Sample Data Sets for ELA/L Grade 6

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	10.03	10.15	10.35	10.34	**11.17**	0.036	0.032	0.026	0.026
Level 2	18.98	19.17	19.51	19.49	**20.02**	0.026	0.021	0.013	0.013
Level 3	30.04	30.16	30.32	30.32	**30.18**	0.003	0.001	-0.003	-0.003
Level 4	34.88	34.61	34.14	34.16	**33.21**	-0.035	-0.030	-0.020	-0.020
Level 5	6.07	5.91	5.67	5.68	**5.42**	-0.029	-0.022	-0.011	-0.012
Total N Count	**93,786**	**112,608**	**150,259**	**187,931**	**373,852**				

Table 3.35 Spring 2015 Performance Level Distribution for the Equating Baseline and Sample Data Sets for Math Grade 5

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	10.23	10.54	10.25	10.41	**10.87**	0.021	0.011	0.020	0.015
Level 2	25.59	25.87	25.74	25.66	**25.95**	0.008	0.002	0.005	0.007
Level 3	31.22	31.03	30.87	30.81	**30.58**	-0.014	-0.010	-0.006	-0.005
Level 4	28.63	28.31	28.63	28.64	**28.28**	-0.008	-0.001	-0.008	-0.008
Level 5	4.33	4.24	4.51	4.48	**4.32**	-0.001	0.004	-0.009	-0.008
Total N Count	**89,633**	**107,574**	**143,548**	**179,456**	**357,829**				

Table 3.36 Spring 2015 Performance Level Distribution for the Equating Baseline and Sample Data Sets for ELA/L Grade 3

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	19.44	19.19	19.04	19.22	**19.17**	-0.007	0.000	0.003	-0.001
Level 2	20.02	19.95	19.98	20.08	**19.81**	-0.005	-0.003	-0.004	-0.007
Level 3	23.74	23.73	23.78	23.79	**23.70**	-0.001	-0.001	-0.002	-0.002
Level 4	33.44	33.66	33.72	33.50	**33.90**	0.010	0.005	0.004	0.008
Level 5	3.36	3.48	3.47	3.40	**3.42**	0.003	-0.004	-0.003	0.001
Total N Count	**79,925**	**95,952**	**127,996**	**160,033**	**319,133**				

Section 4: Post-Equating Results

For the post-equating analyses, the following spring 2015 PARCC assessments were selected. These assessments were selected because minimal operational items were dropped during the operational calibrations:

- ELA/L Grade 10
- Algebra 2
- Integrated Mathematics II
- ELA/L Grade 6
- Mathematics Grade 5
- ELA/L Grade 3

For each of the six assessments, the equating baseline data set and the four equating sample data sets, representing 25%, 30%, 40%, and 50% of the baseline data set, were analyzed through classical item analysis and IRT calibrations. The resulting IRT parameter estimates for each data set were used to calculate the summative scale score conversion files. This resulted in five sets of conversion files for each of the assessments except for Algebra 2 and Integrated Mathematics II, which had at least one item in a sample that could not be calibrated.

All operational items in the sample data sets were required to have data in all of the response categories and the item had to converge during IRT calibrations in order for the IRT calibrations to be consistent with the spring 2015 operational procedures. The same sets of operational items were required for each sample in order to generate conversion files and allow for comparisons across samples. For Algebra 2, the sample representing 25% of the baseline data set had one item with no students attaining the highest score category. As a result the score category for the item could not be calibrated. Therefore, the 25% sample was determined to be too small for Algebra 2 and the results for the IRT calibrations and conversion files were not able to be generated. For Integrated Mathematics II, the equating baseline data set and all four of the sample data sets had items missing responses for one or more score categories. All samples for Integrated Mathematics II were determined to be too small. Although the calibrations could not be obtained, the classical item analyses were conducted for all items and all data sets for Algebra 2 and Integrated Mathematics II.

This section contains the following results:

4.1 Classical Item Analysis for the Equating Baseline and Sample Data Sets

For each of the six assessments, classical item statistics were generated: item mean, item percent correct, response category distributions, and item-to-total correlations. The item mean is the average item score across all students in the data set that responded to the item and ranges from 0 to the maximum possible score points for an item. The item percent correct is the item mean value divided by the maximum possible points for an item and ranges from zero to one. The response category distribution is the proportion of students in each response category. The item-to-total correlation is an indicator of item discrimination and is computed by correlating the item score with the total raw score. It ranges from -1 to 1. As with the spring 2015 operational analyses, the item analyses were done by component: EOY and PBA. The results for the item analysis were similar across the six assessments; therefore, only the ELA/L grade 10 results are shown in this section. The results for the other five assessments are presented in Appendix F.

Tables 4.1 and 4.2 list descriptive statistics for the item percent correct value and the item mean value, respectively, for the equating baseline data set and each of the equating sample data sets for ELA/L EOY grade 10. A Pearson correlation for each sample with the baseline data set is listed. The mean of the item percent correct values and the item mean for each sample were compared to the mean of the baseline data set through a t-test (T) and Cohen's D effect size. The desired effect size is a value near zero (within +/- 0.10) which indicates the difference in the means for the sample and baseline are within 0.10 standard deviations.

The results shown for ELA/L grade 10 are similar to the results for the other five assessments for EOY which are listed in Appendix F. The number of items, mean, standard deviation (STD), minimum (Min), and Maximum (Max) for the item percent correct and item mean are very similar across all the data sets for each assessment. The correlations of the sample data sets with the equating baseline data set are all larger than 0.999, except Integrated Mathematics II which has all values larger than 0.996. The t-test is below the critical value for all comparisons (t-critical = 1.962). Cohen's D is below 0.05 for the majority of the comparisons. For the smaller sample size for Integrated Mathematics II, Cohen's D was -0.12. For ELA/L grade 10, Cohen's D ranged from 0.05 – 0.10 for all of the samples. While the effect size estimates are larger for Integrated Mathematics II and ELA/L grade 10, the t-test indicates that differences are not statistically significant.

Table 4.1 EOY Item Percent Correct for Equating Baseline and Sample Data Sets for ELA/L Grade 10

EOY Item Percent Correct								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	87	0.397	0.131	0.102	0.702	0.99949	0.487	0.074
Sample 30%	87	0.396	0.131	0.101	0.698	0.99963	0.424	0.064
Sample 40%	87	0.399	0.132	0.102	0.701	0.99969	0.563	0.085
Sample 50%	87	0.396	0.130	0.101	0.699	0.99983	0.405	0.061
Baseline	87	0.388	0.128	0.096	0.693			

Table 4.2 EOY Item Mean for Equating Baseline and Sample Data Sets for ELA/L Grade 10

EOY Item Mean								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	87	0.795	0.263	0.203	1.404	0.99952	0.485	0.074
Sample 30%	87	0.792	0.262	0.203	1.396	0.99963	0.425	0.064
Sample 40%	87	0.798	0.263	0.205	1.402	0.99969	0.563	0.085
Sample 50%	87	0.792	0.261	0.203	1.398	0.99983	0.409	0.062
Baseline	87	0.776	0.256	0.192	1.386			

For ELA/L grades 3, 6, 10 and mathematics grade 5, the EOY items had a maximum score point value of two resulting in three possible response categories (0, 1, 2). For Algebra 2 and Integrated Mathematics II the EOY items had a maximum score point value of four resulting in five possible response categories (0, 1, 2, 3, 4). Table 4.3 provides descriptive statistics for the response category distributions for the equating baseline data set and each of the equating sample data sets for ELA/L EOY grade 10 and Appendix F provides the tables for the other assessments.

For each item's response category value, the difference between the proportion of students for the baseline data set and the proportion of students for the sample data set was calculated. Descriptive statistics showing the mean, standard deviation (STD), minimum (Min), and maximum (Max) of the baseline and sample proportion differences are also listed in Table 4.3. Positive difference values indicated the equating baseline proportion was larger than the sample proportion for the response category. Negative difference values indicate the equating baseline proportion was smaller than the sample proportion for the reporting category.

For ELA/L grades 3, 6, 10, math grade 5 and Algebra 2, the proportion of students in a response category is within one percent, on average, to the item's proportion of students in the response category for the equating baseline data set. The largest difference in proportion of students in any response category between a sample data set and the equating baseline data set is 2.7% for ELA/L grade 10, 2.2% for ELA/L grade 6, 1.4% for ELA/L grade 3, 1.9% for mathematics grade 5, and 3.5% for Algebra 2.

For Integrated Mathematics II, large mean differences between the sample and baseline proportions were found for response categories 0-3 which ranged from 1.5 – 2.45%. The largest difference for an item's response category was a magnitude of 6.8%.

Table 4.3 EOY Response Category Distributions for Equating Baseline and Sample Data Sets for ELA/L Grade 10

	Score Category 0					Difference (Baseline - Sample) for Score 0			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	87	48.02	14.37	16.50	81.00	1.02	0.61	-0.70	2.10
Sample 30%	87	48.16	14.34	16.60	81.10	0.88	0.54	-0.70	2.10
Sample 40%	87	47.88	14.33	16.60	80.90	1.16	0.51	-0.10	2.10
Sample 50%	87	48.18	14.26	16.60	81.10	0.86	0.38	-0.10	1.60
Baseline	87	49.04	14.04	17.40	82.10				

	Score Category 1					Difference (Baseline - Sample) for Score 1			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	87	24.47	13.43	6.80	66.90	-0.11	0.62	-1.60	1.10
Sample 30%	87	24.43	13.37	6.80	66.80	-0.08	0.53	-1.20	1.00
Sample 40%	87	24.45	13.36	7.00	66.50	-0.10	0.50	-1.40	0.90
Sample 50%	87	24.47	13.26	7.20	66.60	-0.11	0.35	-1.00	0.60
Baseline	87	24.35	13.03	7.40	66.30				

	Score Category 2					Difference (Baseline - Sample) for Score 2			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	87	27.52	15.12	1.30	63.30	-0.91	0.60	-2.70	0.20
Sample 30%	87	27.41	15.09	1.40	63.00	-0.80	0.53	-2.20	0.20
Sample 40%	87	27.67	15.15	1.40	63.30	-1.06	0.55	-2.10	0.10
Sample 50%	87	27.35	15.00	1.40	63.00	-0.74	0.39	-1.60	-0.10
Baseline	87	26.61	14.70	1.30	62.40				

Tables 4.4 and 4.5 list descriptive statistics for the item percent correct value and the item mean value, respectively, for the equating baseline data set and each of the equating sample data sets for ELA/L PBA grade 10. A Pearson correlation for each sample with the baseline data set is listed. The mean of the item percent correct and item mean values for each sample was compared to the mean of the baseline data set through a t-test (T) and Cohen's D effect size.

The results shown for ELA/L grade 10 are similar to the results for the other five assessments for PBA. The number of items, mean, standard deviation (STD), minimum (Min), and Maximum (Max) for the item percent correct and item mean are very similar across all the data sets. The correlations of the sample data sets with the equating baseline data set are all larger than 0.999, except Integrated Mathematics II which has all values larger than 0.994. The t-test is below the critical value for all comparisons (t-critical = 1.962). Cohen's D is below 0.05 for the majority of the comparisons. For the smaller sample size for Integrated Mathematics II, Cohen's D was -0.16. For ELA/L grade 10 and Algebra 2, Cohen's D ranged from 0.05 – 0.14 for most of the samples. The larger deviations indicate greater differences; however, these are not significant given the t-tests.

Table 4.4 PBA Item Percent Correct for the Equating Data Sets for ELA/L Grade 10

PBA Item Percent Correct								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	86	0.424	0.136	0.152	0.801	0.99741	0.847	0.129
Sample 30%	86	0.423	0.136	0.152	0.801	0.99768	0.814	0.124
Sample 40%	86	0.426	0.136	0.156	0.804	0.99784	0.940	0.143
Sample 50%	86	0.423	0.136	0.157	0.802	0.99872	0.798	0.122
Baseline	86	0.406	0.132	0.153	0.783			

Table 4.5 PBA Item Mean for the Equating Data Sets for ELA/L Grade 10

PBA Item Mean								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	86	1.663	2.027	0.304	7.661	0.99970	0.376	0.057
Sample 30%	86	1.657	2.015	0.304	7.622	0.99977	0.356	0.054
Sample 40%	86	1.666	2.023	0.312	7.653	0.99981	0.386	0.059
Sample 50%	86	1.645	1.985	0.314	7.561	0.99988	0.321	0.049
Baseline	86	1.552	1.825	0.307	7.073			

For ELA/L grades 3, 6, and 10, the majority of the non-PCR PBA items have a maximum score point of two. However, the PBA PCR items are scored on 2-3 traits ranging in scores from 0-12, depending on the trait and weighting. For the item analysis, the PCR scores is reflected as the sum of the scored traits; therefore, for ELA/L grades 6 and 10, the PBA PCR items had a maximum score point value of 19 resulting in 20 possible response categories and the grade 3 PCRs had a maximum score point value of 15 resulting in 16 possible response categories. For Algebra 2, Integrated Mathematics II, and mathematics grade 5, the PBA items had a maximum score point value of six resulting in seven possible response categories. Table 4.6 provides descriptive statistics for the response category distributions for the equating baseline data set and each of the equating sample data sets for ELA/L PBA grade 10 and Appendix F provides the tables for the other assessments.

For each item's response category value, the difference between the proportion of students for the baseline data set and the proportion of students for the sample data set was calculated. Descriptive statistics showing the mean, standard deviation (STD), minimum (Min), and maximum (Max) of the baseline and sample proportion differences are also listed in Table 4.6. Positive difference values indicated the equating baseline proportion was larger than the sample proportion for the response category. Negative difference values indicate the equating baseline proportion was smaller than the sample proportion for the response category.

For ELA/L PBA grades 3, 6, and 10, the response category results are similar to the results for EOY. The proportion of students in a response category is within one percent, on average, to the item's proportion of students in the response category for the equating baseline data set. For ELA/L grade 10, the PBA maximum difference values were larger for PBA then EOY for the score categories 0, 1, and 2. The magnitude of the value ranged from 2.2 – 6.3. For ELA/L grades 3 and 6, the magnitudes of the difference values for PBA were similar to EOY.

For PBA mathematics grade 5 items, the proportion of students in a response category is less than one percent when compared to the item's proportion of students in the response category for the equating baseline data set. The magnitudes of the difference values were slightly higher than those values for EOY for score categories zero and one. For Algebra 2, the PBA results were slightly higher than the EOY results for the mean and the magnitude of the difference. The samples' differences had magnitudes as high as 3.5-5.2% compared to the baseline proportions for response category zero and as high as 3.4 for response category one.

For Integrated Mathematics II, large mean differences between the sample and baseline proportions were found for most of the response categories which ranged from 1.74-2.45%. The largest difference for an item's response category was a magnitude of 7.7%.

Table 4.6 PBA Response Category Distributions for the Equating Data Sets for ELA/L Grade 10

	Score Category 0					Difference (Baseline - Sample) for Score 0			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	86	40.60	16.43	1.40	74.10	2.16	1.59	-0.40	6.30
Sample 30%	86	40.71	16.38	1.50	74.20	2.05	1.50	-0.40	5.90
Sample 40%	86	40.50	16.27	1.50	73.60	2.25	1.36	0.20	5.60
Sample 50%	86	40.85	16.09	1.60	73.40	1.90	1.07	0.20	4.60
Baseline	86	42.76	15.47	3.00	73.90				

	Score Category 1					Difference (Baseline - Sample) for Score 1			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	86	23.81	17.21	1.30	73.40	-0.30	0.77	-2.20	1.00
Sample 30%	86	23.81	17.20	1.30	73.40	-0.30	0.74	-2.20	0.90
Sample 40%	86	23.75	17.20	1.30	74.00	-0.23	0.76	-2.20	1.00
Sample 50%	86	23.74	17.12	1.40	73.50	-0.23	0.63	-2.00	0.80
Baseline	86	23.51	16.72	1.40	72.00				

	Score Category 2					Difference (Baseline - Sample) for Score 2			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	86	25.63	18.75	0.20	74.00	-1.10	0.94	-3.50	0.30
Sample 30%	86	25.57	18.70	0.20	74.00	-1.04	0.89	-3.60	0.30
Sample 40%	86	25.85	18.82	0.20	74.50	-1.31	0.98	-4.00	0.10
Sample 50%	86	25.65	18.68	0.30	74.20	-1.11	0.81	-3.10	0.10
Baseline	86	24.53	17.94	0.30	72.10				

	Score Category 3					Difference (Baseline - Sample) for Score 3			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	12	1.33	1.63	0.00	4.30	0.03	0.15	-0.30	0.30
Sample 30%	12	1.32	1.63	0.00	4.30	0.03	0.13	-0.20	0.30
Sample 40%	12	1.30	1.59	0.10	4.20	0.05	0.13	-0.10	0.40
Sample 50%	12	1.30	1.59	0.10	4.20	0.05	0.12	0.00	0.40
Baseline	12	1.35	1.67	0.10	4.30				

	Score Category 4					Difference (Baseline - Sample) for Score 4			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	12	5.93	5.01	1.90	14.70	-0.23	0.35	-0.80	0.20
Sample 30%	12	5.89	4.97	1.90	14.40	-0.20	0.32	-0.90	0.20
Sample 40%	12	5.78	4.89	1.90	14.00	-0.09	0.29	-0.80	0.40
Sample 50%	12	5.78	4.82	1.90	13.90	-0.08	0.22	-0.60	0.30
Baseline	12	5.69	4.74	1.90	13.90				

	Score Category 5	Difference (Baseline - Sample) for Score 5

PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	12	14.98	8.24	3.70	25.40	-0.22	0.34	-0.80	0.30
Sample 30%	12	14.95	8.22	3.60	25.60	-0.19	0.35	-0.70	0.50
Sample 40%	12	14.82	8.16	3.50	25.50	-0.06	0.29	-0.50	0.50
Sample 50%	12	14.78	8.19	3.40	25.40	-0.03	0.30	-0.40	0.50
Baseline	12	14.76	8.25	3.50	25.40				

		Score Category 6				Difference (Baseline - Sample) for Score 6			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	12	3.32	1.65	0.60	5.10	-0.23	0.26	-0.60	0.10
Sample 30%	12	3.30	1.66	0.60	5.10	-0.22	0.26	-0.60	0.10
Sample 40%	12	3.28	1.61	0.60	5.00	-0.19	0.19	-0.50	0.10
Sample 50%	12	3.26	1.58	0.70	4.90	-0.18	0.15	-0.50	0.00
Baseline	12	3.08	1.49	0.70	4.60				

		Score Category 7				Difference (Baseline - Sample) for Score 7			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	12	2.38	2.26	0.70	7.90	-0.23	0.29	-0.90	0.20
Sample 30%	12	2.40	2.27	0.70	7.90	-0.24	0.29	-0.90	0.10
Sample 40%	12	2.33	2.15	0.70	7.50	-0.17	0.18	-0.50	0.10
Sample 50%	12	2.28	2.12	0.70	7.40	-0.13	0.14	-0.40	0.00
Baseline	12	2.16	2.00	0.70	7.00				

		Score Category 8				Difference (Baseline - Sample) for Score 8			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	12	5.21	7.03	0.20	16.20	-0.47	0.69	-1.80	0.00
Sample 30%	12	5.18	6.97	0.20	16.20	-0.43	0.61	-1.60	0.00
Sample 40%	12	5.14	6.97	0.20	16.00	-0.40	0.56	-1.40	0.00
Sample 50%	12	5.06	6.85	0.20	15.60	-0.32	0.42	-1.20	0.00
Baseline	12	4.74	6.48	0.10	14.90				

		Score Category 9				Difference (Baseline - Sample) for Score 9			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	12	3.06	0.77	2.00	4.60	-0.28	0.27	-0.90	0.20
Sample 30%	12	3.05	0.76	1.90	4.50	-0.27	0.26	-0.80	0.20
Sample 40%	12	3.01	0.72	2.00	4.40	-0.23	0.22	-0.70	0.20
Sample 50%	12	2.99	0.72	1.90	4.20	-0.21	0.14	-0.50	0.00
Baseline	12	2.78	0.67	1.70	3.80				

		Score Category 10				Difference (Baseline - Sample) for Score 10			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	12	12.80	8.94	0.50	20.60	-0.95	0.93	-2.30	0.10
Sample 30%	12	12.77	8.92	0.40	20.40	-0.92	0.86	-2.10	0.10
Sample 40%	12	12.82	8.93	0.40	20.50	-0.97	0.82	-2.10	0.10
Sample 50%	12	12.65	8.81	0.40	20.20	-0.80	0.67	-1.70	0.10
Baseline	12	11.85	8.25	0.30	19.40				

Score Category 11						Difference (Baseline - Sample) for Score 11			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	12	4.15	2.13	2.20	9.90	-0.54	0.37	-1.50	-0.10
Sample 30%	12	4.16	2.10	2.20	9.80	-0.55	0.33	-1.40	-0.20
Sample 40%	12	4.15	2.08	2.30	9.70	-0.54	0.26	-1.30	-0.30
Sample 50%	12	4.05	2.04	2.10	9.40	-0.44	0.23	-1.00	-0.10
Baseline	12	3.61	1.86	2.00	8.40				

Score Category 12						Difference (Baseline - Sample) for Score 12			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	12	3.97	4.93	0.30	11.90	-0.38	0.42	-1.20	0.00
Sample 30%	12	3.93	4.93	0.30	11.90	-0.33	0.42	-1.20	0.00
Sample 40%	12	4.03	5.05	0.30	12.20	-0.44	0.53	-1.30	0.00
Sample 50%	12	3.94	4.95	0.30	12.00	-0.35	0.43	-1.10	0.00
Baseline	12	3.59	4.53	0.30	11.10				

Score Category 13						Difference (Baseline - Sample) for Score 13			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	12	0.25	0.25	0.00	0.70	-0.03	0.07	-0.10	0.10
Sample 30%	12	0.27	0.25	0.00	0.70	-0.05	0.05	-0.10	0.00
Sample 40%	12	0.24	0.25	0.00	0.70	-0.03	0.05	-0.10	0.00
Sample 50%	12	0.24	0.25	0.00	0.70	-0.03	0.05	-0.10	0.00
Baseline	12	0.22	0.21	0.00	0.60				

Score Category 14						Difference (Baseline - Sample) for Score 14			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	12	1.82	1.07	0.50	3.90	-0.28	0.24	-0.70	0.00
Sample 30%	12	1.78	1.06	0.50	3.90	-0.23	0.23	-0.70	0.10
Sample 40%	12	1.78	1.03	0.50	3.80	-0.23	0.18	-0.60	0.00
Sample 50%	12	1.75	1.00	0.50	3.70	-0.21	0.16	-0.50	0.00
Baseline	12	1.54	0.88	0.40	3.20				

Score Category 15						Difference (Baseline - Sample) for Score 15			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	12	8.64	1.34	6.90	10.50	-1.03	0.42	-1.70	-0.50
Sample 30%	12	8.60	1.32	6.90	10.40	-0.99	0.29	-1.40	-0.60
Sample 40%	12	8.77	1.38	7.20	10.70	-1.16	0.21	-1.40	-0.70
Sample 50%	12	8.54	1.35	7.00	10.40	-0.93	0.14	-1.20	-0.60
Baseline	12	7.61	1.35	6.00	9.50				

Score Category 16						Difference (Baseline - Sample) for Score 16			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	8	0.61	0.26	0.30	1.20	-0.11	0.08	-0.30	0.00
Sample 30%	8	0.60	0.23	0.30	1.10	-0.10	0.05	-0.20	0.00
Sample 40%	8	0.63	0.23	0.30	1.10	-0.13	0.05	-0.20	-0.10
Sample 50%	8	0.61	0.23	0.30	1.10	-0.11	0.06	-0.20	0.00

Baseline	8	0.50	0.19	0.20	0.90

		Score Category 17*				Difference (Baseline - Sample) for Score 17			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sample 30%	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sample 40%	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sample 50%	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Baseline	8	0.00	0.00	0.00	0.00				

		Score Category 18				Difference (Baseline - Sample) for Score 18			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	8	0.39	0.22	0.10	0.70	-0.08	0.07	-0.20	0.00
Sample 30%	8	0.36	0.24	0.10	0.70	-0.05	0.05	-0.10	0.00
Sample 40%	8	0.40	0.23	0.10	0.70	-0.09	0.06	-0.20	0.00
Sample 50%	8	0.38	0.23	0.10	0.70	-0.06	0.07	-0.20	0.00
Baseline	8	0.31	0.19	0.10	0.60				

		Score Category 19				Difference (Baseline - Sample) for Score 19			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	8	4.28	1.26	2.90	6.50	-0.70	0.25	-1.00	-0.40
Sample 30%	8	4.18	1.19	2.90	6.30	-0.60	0.17	-0.80	-0.40
Sample 40%	8	4.31	1.27	3.00	6.60	-0.74	0.21	-1.10	-0.50
Sample 50%	8	4.16	1.26	2.90	6.40	-0.59	0.19	-0.90	-0.40
Baseline	8	3.58	1.11	2.50	5.50				

*The score point of 17 has zero students due to 17 being difficult to attain when combining the traits. It requires students to get almost zero points on one trait and full credit on other traits.

Tables 4.7 and 4.8 provide descriptive statistics for the item-to-total correlations for the equating baseline data set and each of the equating sample data sets for ELA/L grade 10 EOY and PBA, respectively. Appendix F provides the corresponding tables for the other assessments. For each item, the difference between the baseline data set item-to-total correlation and the sample data set item-to-total correlation was calculated. Descriptive statistics showing the mean, standard deviation (STD), minimum (Min), and maximum (Max) of the baseline and sample item-to-total correlation differences are also listed in Table 4.7. Positive difference values indicated the equating baseline item-to-total correlation was larger than the sample item-to-total correlation. Negative difference values indicate the equating baseline item-to-total correlation was smaller than the sample item-to-total correlation.

For EOY and PBA, the item-to-total correlation results across all samples and assessments were very similar to the equating baseline item-to-total correlation. The mean differences are approximately zero with very small standard deviations of the differences.

Table 4.7 EOY Item-to-Total Correlations for the Equating Data Sets for ELA/L Grade 10

		Item-to-Total Correlation				Difference (Baseline - Sample)			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	87	0.506	0.096	0.272	0.698	0.005	0.007	-0.016	0.024
Sample 30%	87	0.507	0.096	0.272	0.694	0.004	0.006	-0.016	0.021
Sample 40%	87	0.509	0.096	0.270	0.693	0.002	0.006	-0.023	0.015
Sample 50%	87	0.509	0.097	0.265	0.693	0.002	0.004	-0.017	0.013
Baseline	87	0.511	0.098	0.267	0.692				

Table 4.8 PBA Item-to-Total Correlations for the Equating Data Sets for ELA/L Grade 10

		Item-to-Total Correlation				Difference (Baseline - Sample)			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	86	0.464	0.177	0.189	0.875	0.009	0.010	-0.009	0.029
Sample 30%	86	0.464	0.177	0.186	0.875	0.009	0.009	-0.009	0.030
Sample 40%	86	0.466	0.176	0.186	0.876	0.008	0.008	-0.008	0.027
Sample 50%	86	0.466	0.176	0.185	0.875	0.007	0.007	-0.009	0.025
Baseline	86	0.474	0.175	0.186	0.871				

4.2 Stocking and Lord Scaling Transformation Constants for the Equating Baseline and Sample Data Sets

Each equating baseline and equating sample data set for each of the assessments was calibrated separately. In order to compare results across samples and evaluate student performance on the summative scale score and performance levels, the estimated IRT parameters for each data set was transformed through Stocking and Lord test characteristic curve method to be on the same IRT scale as the spring 2015 operational items for an assessment. Table 4.9 lists the Stocking and Lord slope and intercept for each data set that were applied to the IRT estimates to place all the item parameters for an assessment on the same scale. Also, Table 4.9 provides the difference between the baseline and each sample for the slope and the intercept. For an assessment, the slope and intercept for the equating samples were very similar to the baseline slope and intercept. Generally, larger differences were found for the smallest sample size and smaller differences for the larger sample sizes.

Table 4.9 Stocking and Lord Scaling Transformation Constants for the Equating Data Sets

Assessment	Sample	Slope	Intercept	Difference in the Slope	Difference in the Intercept
ELA/L Grade 10	Sample 25%	0.975222	0.061505	0.018085	-0.093794
	Sample 30%	0.975333	0.054745	0.017974	-0.087034
	Sample 40%	0.982186	0.067430	0.011121	-0.099719
	Sample 50%	0.984012	0.047051	0.009295	-0.079340
	Baseline	0.993307	-0.032289		
Algebra 2	Sample 30%	0.961327	0.113767	0.025183	-0.078181
	Sample 40%	0.963707	0.108784	0.022803	-0.073198
	Sample 50%	0.969678	0.103716	0.016832	-0.068130
	Baseline	0.986510	0.035586		
ELA/L Grade 6	Sample 25%	1.010839	0.073577	0.005357	-0.056016
	Sample 30%	1.009909	0.063015	0.006287	-0.045454
	Sample 40%	1.008543	0.046398	0.007653	-0.028837
	Sample 50%	1.008817	0.046592	0.007379	-0.029031
	Baseline	1.016196	0.017561		
Math Grade 5	Sample 25%	0.998362	-0.009894	0.008007	-0.017383
	Sample 30%	0.998655	-0.023493	0.007714	-0.003784
	Sample 40%	1.004054	-0.006915	0.002315	-0.020362
	Sample 50%	1.005654	-0.008912	0.000715	-0.018365
	Baseline	1.006369	-0.027277		
ELA/L Grade 3	Sample 25%	1.012533	0.019577	0.002301	0.014389
	Sample 30%	1.015642	0.029289	-0.000808	0.004677
	Sample 40%	1.014719	0.033305	0.000115	0.000661
	Sample 50%	1.012327	0.026317	0.002507	0.007649
	Baseline	1.014834	0.033966		

4.3 IRT Descriptive Statistics for the Baseline and Sample Data Sets

The EOY and PBA items were concurrently calibrated in an incomplete data matrix (IDM) with the two-parameter logistic/generalized partial credit model (2PL/GPC: Muraki, 1992). The IRT parameter estimates were transformed through the Stocking and Lord slope and intercept to be on the 2015 IRT scale. The samples' IRT parameter estimates and associated standard errors were then compared to the equating baseline IRT parameters and standard errors. Descriptive statistics, including mean, standard deviation (STD), minimum (Min), and maximum (Max), for the number student responses for the items were provided for the equating baseline and the sample data sets. Descriptive statistics for the IRT parameters and standard errors were computed for the equating baseline and the sample data sets.

Table 4.10 summarizes the average number of student responses for the items in each of the samples and the equating baseline data set for ELA/L grade 10. The average number of student responses for the equating baseline data set was almost 44,000 with a range of 15,000 to almost 94,000 responses. The sample data sets are proportionally smaller based on the percent of the baseline with the 25% sample having the fewest student responses of 3,855.

Table 4.10 Descriptive Statistics for the Number of Student Responses per Item for ELA/L Grade 10

Item N count					
	N	Mean	STD	Min	Max
Sample 25%	193	11,066.04	4,003.73	3,855	23,762
Sample 30%	193	13,321.79	4,825.48	4,646	28,589
Sample 40%	193	17,850.16	6,463.50	6,230	38,310
Sample 50%	193	22,391.05	8,112.26	7,836	48,117
Baseline	193	43,745.73	15,858.11	15,308	93,878

Table 4.11 provides descriptive statistics for the IRT parameters (*a*, *b*, and *d2–d5*) and standard errors of the parameters for ELA/L grade 10. The *a*-parameter is the slope; *b*-parameter is the difficulty/location parameter; and the *d2–d5* parameters are the thresholds for the score categories (1-4). For the 2PL/GPC model, *d1* (score category 0) is set to zero and therefore is not included in the table. The IRT parameters were highly correlated with the baseline parameters for all the sample data sets. For the *a*- and *b*-parameters the correlations ranged from 0.997-0.999. The standard errors decreased as the sample sizes increased and the correlation of the standard errors with the baseline data sets were high ranging from 0.997-0.999. Cohen's D is below 0.05 for the *a*, *b*, and *d2-d4* parameters. The *d5* parameter, which is the highest score category, has effect size values of 0.15-0.23.

Table 4.11 Descriptive Statistics for the IRT Parameters and Standard Errors for ELA/L Grade 10

IRT A Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	193	0.476	0.178	0.148	0.852	0.997	0.178	0.018
Sample 30%	193	0.476	0.178	0.145	0.861	0.998	0.171	0.017
Sample 40%	193	0.475	0.177	0.146	0.865	0.998	0.093	0.009
Sample 50%	193	0.474	0.175	0.144	0.847	0.999	0.034	0.003
Baseline	193	0.473	0.175	0.142	0.837			

IRT A Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	193	0.021	0.009	0.009	0.045	0.998	15.065	1.534
Sample 30%	193	0.019	0.008	0.009	0.041	0.999	13.282	1.352
Sample 40%	193	0.017	0.007	0.007	0.036	0.999	10.347	1.053
Sample 50%	193	0.015	0.006	0.007	0.032	0.999	7.876	0.802
Baseline	193	0.011	0.004	0.005	0.023			

IRT B Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	193	0.667	0.884	-2.104	3.585	0.998	0.069	0.007
Sample 30%	193	0.668	0.887	-2.080	3.641	0.999	0.079	0.008
Sample 40%	193	0.667	0.887	-2.068	3.658	0.999	0.070	0.007
Sample 50%	193	0.665	0.887	-2.090	3.701	0.999	0.048	0.005
Baseline	193	0.661	0.893	-2.029	3.798			

IRT B Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	193	0.031	0.025	0.012	0.198	0.997	7.415	0.755
Sample 30%	193	0.028	0.023	0.011	0.187	0.997	6.509	0.663
Sample 40%	193	0.024	0.019	0.009	0.159	0.998	4.892	0.498
Sample 50%	193	0.021	0.018	0.008	0.146	0.998	3.672	0.374
Baseline	193	0.016	0.013	0.006	0.110			

IRT D2 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	193	-0.222	1.516	-5.721	2.647	0.999	-0.080	-0.008
Sample 30%	193	-0.222	1.514	-5.683	2.656	0.999	-0.081	-0.008
Sample 40%	193	-0.227	1.512	-5.639	2.607	0.999	-0.116	-0.012
Sample 50%	193	-0.221	1.495	-5.536	2.606	1.000	-0.078	-0.008
Baseline	193	-0.210	1.459	-5.276	2.701			

IRT D2 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	193	0.054	0.038	0.015	0.235	0.997	9.114	0.928
Sample 30%	193	0.049	0.035	0.013	0.220	0.997	8.084	0.823
Sample 40%	193	0.042	0.030	0.011	0.189	0.998	6.351	0.646
Sample 50%	193	0.037	0.026	0.010	0.163	0.999	4.895	0.498
Baseline	193	0.026	0.018	0.007	0.113			

IRT D3 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	193	0.460	1.408	-2.647	5.721	0.999	0.132	0.013
Sample 30%	193	0.460	1.407	-2.656	5.683	0.999	0.132	0.013
Sample 40%	193	0.464	1.407	-2.607	5.639	0.999	0.158	0.016
Sample 50%	193	0.457	1.392	-2.606	5.536	0.999	0.108	0.011
Baseline	193	0.442	1.361	-2.701	5.276			

IRT D3 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	193	0.055	0.038	0.015	0.235	0.996	9.238	0.940
Sample 30%	193	0.050	0.035	0.013	0.220	0.997	8.195	0.834
Sample 40%	193	0.043	0.030	0.011	0.189	0.998	6.430	0.655
Sample 50%	193	0.038	0.026	0.010	0.163	0.999	4.961	0.505
Baseline	193	0.027	0.018	0.007	0.113			

IRT D4 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	32	-0.637	0.217	-1.071	-0.269	0.979	-0.322	-0.080
Sample 30%	32	-0.636	0.218	-1.080	-0.258	0.987	-0.293	-0.073
Sample 40%	32	-0.631	0.217	-1.077	-0.237	0.992	-0.217	-0.054
Sample 50%	32	-0.627	0.220	-1.074	-0.222	0.995	-0.134	-0.034
Baseline	32	-0.620	0.220	-1.059	-0.209			

IRT D4 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	32	0.044	0.010	0.026	0.058	0.988	10.353	2.588
Sample 30%	32	0.040	0.009	0.024	0.053	0.991	8.988	2.247
Sample 40%	32	0.034	0.008	0.021	0.047	0.993	6.711	1.678
Sample 50%	32	0.031	0.007	0.019	0.042	0.996	5.126	1.281
Baseline	32	0.023	0.005	0.014	0.032			

IRT D5 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	20	-1.283	0.160	-1.611	-1.021	0.922	-0.703	-0.222
Sample 30%	20	-1.284	0.155	-1.565	-0.987	0.946	-0.737	-0.233
Sample 40%	20	-1.275	0.149	-1.545	-0.978	0.959	-0.569	-0.180
Sample 50%	20	-1.271	0.154	-1.510	-0.957	0.971	-0.469	-0.148
Baseline	20	-1.249	0.142	-1.457	-0.958			

IRT D5 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	20	0.062	0.010	0.041	0.079	0.978	11.337	3.585
Sample 30%	20	0.057	0.009	0.038	0.074	0.979	9.893	3.129
Sample 40%	20	0.049	0.008	0.033	0.064	0.986	7.319	2.315
Sample 50%	20	0.044	0.007	0.030	0.058	0.993	5.480	1.733
Baseline	20	0.033	0.006	0.022	0.044			

Table 4.12 summarizes the average number of student responses for the items in each of the samples and the equating baseline data set for Algebra 2. The average number of student responses for the equating baseline data set was almost 43,000 with a range of 23,370 to almost 105,000 responses. The sample data sets are proportionally smaller based on the percent of the baseline with the 30% sample having the fewest student responses of 7,000. The 25% sample is not included due to an item missing responses at the highest score category.

Table 4.12 Descriptive Statistics for the Number of Student Responses per Item for Algebra 2

Item N count					
	N	Mean	STD	Min	Max
Sample 30%	200	13,040.75	5,700.42	7,000	31,708
Sample 40%	200	17,465.15	7,662.89	9,332	42,532
Sample 50%	200	22,023.55	9,603.44	11,923	53,724
Baseline	200	42,890.51	18,538.42	23,370	104,718

Table 4.13 provides descriptive statistics for the IRT parameters (*a*, *b*, and *d2-d7*) and standard errors of the parameters for Algebra 2. The *a*-parameter is the slope; *b*-parameter is the difficulty/location parameter; and the *d2–d7* parameters are the thresholds for the score categories (1-6). For the 2PL/GPC model, *d1* (score category 0) is set to zero and therefore is not included in the table. The IRT *a*-parameters were highly correlated with the baseline parameters for all the sample data sets with

correlations ranging from 0.995-0.998. For the *b*-parameters the correlations ranged from 0.924-0.987. The *b*-parameters had items with very large values which inflated the mean *b*-parameter and mean standard error. These items also impacted the *d2-d7* parameters. Except for the *b*-parameter, the standard errors decreased as the sample sizes increased and the correlation of the standard errors with the baseline data sets were high ranging from 0.965-0.999. Cohen's D is below 0.10 for all parameters.

Table 4.13 Descriptive Statistics for the IRT Parameters and Standard Errors for Algebra 2

IRT A Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 30%	200	0.690	0.331	0.029	1.701	0.995	0.095	0.010
Sample 40%	200	0.689	0.333	0.028	1.819	0.997	0.068	0.007
Sample 50%	200	0.690	0.333	0.026	1.795	0.998	0.086	0.009
Baseline	200	0.687	0.330	0.022	1.767			

IRT A Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 30%	200	0.037	0.027	0.009	0.180	0.997	7.311	0.731
Sample 40%	200	0.032	0.024	0.007	0.164	0.998	5.555	0.556
Sample 50%	200	0.028	0.022	0.007	0.142	0.999	4.187	0.419
Baseline	200	0.021	0.015	0.005	0.102			

IRT B Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 30%	200	2.054	2.012	-1.223	17.241	0.987	-0.081	-0.008
Sample 40%	200	2.091	2.344	-1.219	24.553	0.983	0.083	0.008
Sample 50%	200	2.170	3.271	-1.224	41.123	0.924	0.349	0.035
Baseline	200	2.072	2.281	-1.244	20.866			

IRT B Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 30%	200	0.165	0.748	0.011	8.763	0.987	0.900	0.090
Sample 40%	200	0.185	1.183	0.009	15.893	0.947	0.865	0.087
Sample 50%	200	0.291	2.809	0.008	39.461	0.873	0.919	0.092
Baseline	200	0.104	0.585	0.006	6.685			

IRT D2 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 30%	88	0.779	1.678	-2.317	8.256	0.984	-0.088	-0.013
Sample 40%	88	0.786	1.734	-2.271	8.899	0.987	-0.059	-0.009
Sample 50%	88	0.787	1.744	-2.277	9.365	0.992	-0.054	-0.008
Baseline	88	0.802	1.867	-2.401	11.400			

IRT D2 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 30%	88	0.141	0.392	0.014	3.240	0.969	0.978	0.147
Sample 40%	88	0.127	0.379	0.012	3.104	0.965	0.737	0.111
Sample 50%	88	0.112	0.361	0.011	3.072	0.982	0.484	0.073
Baseline	88	0.087	0.336	0.008	3.049			

IRT D3 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 30%	88	-0.228	1.820	-8.256	4.590	0.987	0.138	0.021
Sample 40%	88	-0.238	1.877	-8.899	4.600	0.989	0.102	0.015
Sample 50%	88	-0.243	1.888	-9.365	4.548	0.993	0.085	0.013
Baseline	88	-0.268	2.005	-11.400	4.528			

IRT D3 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 30%	88	0.148	0.392	0.014	3.240	0.969	1.038	0.157
Sample 40%	88	0.133	0.379	0.012	3.104	0.965	0.784	0.118
Sample 50%	88	0.118	0.361	0.011	3.072	0.982	0.518	0.078
Baseline	88	0.091	0.336	0.008	3.049			

IRT D4 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 30%	47	-0.120	0.983	-2.522	1.960	0.992	0.265	0.055
Sample 40%	47	-0.138	1.004	-2.658	1.913	0.997	0.175	0.036
Sample 50%	47	-0.166	0.977	-2.527	1.863	0.998	0.037	0.008
Baseline	47	-0.174	0.975	-2.697	1.804			

IRT D4 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 30%	47	0.135	0.127	0.027	0.545	0.988	3.436	0.709
Sample 40%	47	0.117	0.117	0.023	0.627	0.972	2.714	0.560
Sample 50%	47	0.097	0.088	0.021	0.450	0.986	2.107	0.435
Baseline	47	0.066	0.053	0.014	0.234			

IRT D5 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 30%	34	-0.557	1.092	-2.901	2.339	0.982	-0.345	-0.084
Sample 40%	34	-0.535	1.078	-3.046	2.376	0.995	-0.262	-0.064
Sample 50%	34	-0.503	1.055	-3.114	2.350	0.991	-0.138	-0.034
Baseline	34	-0.468	1.027	-2.912	2.335			

IRT D5 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 30%	34	0.209	0.249	0.030	0.955	0.965	2.614	0.634
Sample 40%	34	0.169	0.187	0.026	0.673	0.987	2.224	0.539
Sample 50%	34	0.136	0.134	0.023	0.500	0.986	1.645	0.399
Baseline	34	0.091	0.083	0.016	0.297			

IRT D6 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 30%	12	-0.487	1.259	-3.458	1.106	0.994	-0.032	-0.013
Sample 40%	12	-0.422	1.267	-3.539	1.122	0.998	0.093	0.038
Sample 50%	12	-0.454	1.275	-3.616	1.101	0.999	0.030	0.012
Baseline	12	-0.470	1.277	-3.636	1.104			

IRT D6 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 30%	12	0.276	0.288	0.088	1.120	0.997	1.424	0.581
Sample 40%	12	0.243	0.286	0.074	1.111	0.998	1.082	0.442
Sample 50%	12	0.210	0.237	0.065	0.921	0.999	0.832	0.340
Baseline	12	0.142	0.149	0.046	0.587			

IRT D7 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 30%	11	-1.641	2.147	-5.577	1.181	0.998	-0.075	-0.032
Sample 40%	11	-1.682	2.078	-5.640	1.102	0.997	-0.123	-0.052
Sample 50%	11	-1.595	2.068	-5.553	1.017	0.998	-0.025	-0.011
Baseline	11	-1.574	2.038	-5.482	1.089			

IRT D7 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 30%	11	0.547	0.819	0.098	2.868	0.998	1.117	0.476
Sample 40%	11	0.522	0.926	0.083	3.266	0.992	0.921	0.393
Sample 50%	11	0.415	0.675	0.073	2.394	0.997	0.730	0.311
Baseline	11	0.248	0.346	0.054	1.243			

Table 4.14 summarizes the average number of student responses for the items in each of the samples and the equating baseline data set for ELA/L grade 6. The average number of student responses for the equating baseline data set was almost 107,000 with a range of 20,000 to almost 187,000 responses. The sample data sets are proportionally smaller based on the percent of the baseline with the 25% sample having the fewest student responses of 5,006.

Table 4.14 Descriptive Statistics for the Number of Student Responses per Item for ELA/L Grade 6

Item N count					
	N	Mean	STD	Min	Max
Sample 25%	157	26,862.12	9,092.35	5,006	46,861
Sample 30%	157	32,253.73	10,912.19	6,007	56,317
Sample 40%	157	43,034.31	14,544.79	8,069	75,113
Sample 50%	157	53,822.62	18,183.21	10,178	93,912
Baseline	157	106,991.97	36,138.30	20,130	186,455

Table 4.15 provides descriptive statistics for the IRT parameters (a, b, and d2-d5) and standard errors of the parameters for ELA/L grade 6. The *a*-parameter is the slope; *b*-parameter is the difficulty/location parameter; and the *d2–d5* parameters are the thresholds for the score categories (1-4). For the 2PL/GPC model, *d1* (score category 0) is set to zero and therefore is not included in the table. The IRT parameters were highly correlated with the baseline parameters for all the sample data sets. For the *a*- and *b*-parameters the correlations ranged from 0.999-1.000. The standard errors decreased as the sample sizes increased and the correlation of the standard errors with the baseline data sets were high ranging from 0.999-1.000. Cohen's D is below 0.05 for the *a*, *b*, and *d2-d5* parameters except for the *d4* and *d5* parameter for the Sample 25% which has effect size values of 0.12.

Table 4.15 Descriptive Statistics for the IRT Parameters and Standard Errors for ELA/L Grade 6

IRT A Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	157	0.441	0.191	0.107	0.910	0.999	-0.152	-0.017
Sample 30%	157	0.442	0.191	0.107	0.910	0.999	-0.126	-0.014
Sample 40%	157	0.443	0.191	0.109	0.897	1.000	-0.067	-0.008
Sample 50%	157	0.443	0.191	0.109	0.889	1.000	-0.071	-0.008
Baseline	157	0.445	0.192	0.108	0.902			

IRT A Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	157	0.014	0.006	0.007	0.042	1.000	11.768	1.328
Sample 30%	157	0.012	0.006	0.006	0.038	1.000	10.421	1.176
Sample 40%	157	0.011	0.005	0.005	0.033	1.000	8.102	0.914
Sample 50%	157	0.010	0.005	0.005	0.030	1.000	6.192	0.699
Baseline	157	0.007	0.003	0.003	0.021			

IRT B Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	157	0.577	1.078	-2.201	6.232	0.999	0.047	0.005
Sample 30%	157	0.579	1.079	-2.181	6.230	1.000	0.057	0.006
Sample 40%	157	0.576	1.076	-2.168	6.268	1.000	0.035	0.004
Sample 50%	157	0.577	1.078	-2.168	6.387	1.000	0.043	0.005
Baseline	157	0.572	1.068	-2.155	6.130			

IRT B Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	157	0.023	0.025	0.008	0.258	0.999	5.010	0.565
Sample 30%	157	0.021	0.023	0.007	0.237	0.999	4.464	0.504
Sample 40%	157	0.018	0.020	0.006	0.209	0.999	3.517	0.397
Sample 50%	157	0.016	0.019	0.005	0.193	0.999	2.699	0.305
Baseline	157	0.011	0.012	0.004	0.126			

IRT D2 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	157	-0.299	1.998	-9.500	3.877	0.999	-0.112	-0.013
Sample 30%	157	-0.293	1.988	-9.223	3.905	1.000	-0.084	-0.009
Sample 40%	157	-0.282	1.975	-9.311	3.948	1.000	-0.034	-0.004
Sample 50%	157	-0.280	1.977	-9.260	4.037	1.000	-0.026	-0.003
Baseline	157	-0.274	1.959	-8.822	3.882			

IRT D2 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	157	0.042	0.044	0.009	0.382	0.999	5.494	0.620
Sample 30%	157	0.038	0.040	0.008	0.349	0.999	4.909	0.554
Sample 40%	157	0.033	0.034	0.007	0.289	0.999	3.873	0.437
Sample 50%	157	0.029	0.030	0.006	0.257	0.999	3.018	0.341
Baseline	157	0.021	0.021	0.004	0.181			

IRT D3 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	157	0.553	1.866	-3.877	9.500	0.999	0.092	0.010
Sample 30%	157	0.549	1.854	-3.905	9.223	1.000	0.071	0.008
Sample 40%	157	0.539	1.840	-3.948	9.311	1.000	0.022	0.003
Sample 50%	157	0.537	1.842	-4.037	9.260	1.000	0.017	0.002
Baseline	157	0.534	1.822	-3.882	8.822			

IRT D3 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	157	0.042	0.044	0.009	0.382	0.999	5.529	0.624
Sample 30%	157	0.038	0.040	0.008	0.349	0.999	4.938	0.557
Sample 40%	157	0.033	0.034	0.007	0.289	0.999	3.897	0.440
Sample 50%	157	0.030	0.030	0.006	0.257	0.999	3.034	0.342
Baseline	157	0.021	0.021	0.004	0.181			

IRT D4 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	24	-0.878	0.239	-1.234	-0.397	0.988	0.414	0.120
Sample 30%	24	-0.885	0.245	-1.236	-0.374	0.988	0.313	0.090
Sample 40%	24	-0.894	0.244	-1.230	-0.400	0.993	0.189	0.055
Sample 50%	24	-0.897	0.242	-1.229	-0.381	0.996	0.145	0.042
Baseline	24	-0.907	0.238	-1.243	-0.388			

IRT D4 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	24	0.031	0.012	0.015	0.054	0.999	5.479	1.582
Sample 30%	24	0.029	0.011	0.014	0.049	1.000	4.883	1.410
Sample 40%	24	0.025	0.010	0.012	0.043	0.999	3.857	1.113
Sample 50%	24	0.022	0.009	0.011	0.038	0.999	2.935	0.847
Baseline	24	0.016	0.006	0.008	0.027			

IRT D5 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	15	-1.253	0.139	-1.440	-0.951	0.940	0.321	0.117
Sample 30%	15	-1.262	0.138	-1.467	-0.970	0.944	0.135	0.049
Sample 40%	15	-1.258	0.129	-1.428	-1.010	0.950	0.213	0.078
Sample 50%	15	-1.261	0.129	-1.432	-1.007	0.975	0.160	0.059
Baseline	15	-1.268	0.127	-1.425	-1.003			

IRT D5 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	15	0.047	0.015	0.025	0.068	0.998	4.912	1.793
Sample 30%	15	0.043	0.014	0.023	0.063	0.998	4.384	1.601
Sample 40%	15	0.038	0.013	0.020	0.056	0.998	3.477	1.269
Sample 50%	15	0.034	0.011	0.018	0.049	0.999	2.654	0.969
Baseline	15	0.025	0.008	0.013	0.036			

Table 4.16 summarizes the average number of student responses for the items in each of the samples and the equating baseline data set for mathematics grade 5. The average number of student responses for the equating baseline data set was around 90,000 with a range of 46,500 to almost 170,000 responses. The sample data sets are proportionally smaller based on the percent of the baseline with the 25% sample having the fewest student responses of 11,972.

Table 4.16 Descriptive Statistics for the Number of Student Responses per Item for Math Grade 5

Item N count					
	N	Mean	STD	Min	Max
Sample 25%	206	22,605.42	10,455.83	11,972	41,226
Sample 30%	206	27,128.07	12,647.64	14,291	49,756
Sample 40%	206	36,200.00	17,019.17	18,970	66,785
Sample 50%	206	45,253.79	21,528.67	23,572	84,265
Baseline	206	90,202.33	43,649.79	46,545	169,791

Table 4.17 provides descriptive statistics for the IRT parameters (*a*, *b*, and *d2-d7*) and standard errors of the parameters for mathematics grade 5. The *a*-parameter is the slope; *b*-parameter is the difficulty/location parameter; and the *d2–d7* parameters are the thresholds for the score categories (1-6). For the 2PL/GPC model, *d1* (score category 0) is set to zero and therefore is not included in the table. The IRT parameters were highly correlated with the baseline parameters for all the sample data sets. For the *a*- and *b*-parameters the correlations ranged from 0.998-1.000. The standard errors decreased as

the sample sizes increased and the correlation of the standard errors with the baseline data sets were high ranging from 0.998-0.999. Cohen's D is below 0.05 for the *a*, *b*, and *d2-d7* parameters.

Table 4.17 Descriptive Statistics for the IRT Parameters and Standard Errors for Math Grade 5

IRT A Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	206	0.689	0.243	0.188	1.556	0.998	0.100	0.010
Sample 30%	206	0.689	0.243	0.193	1.559	0.998	0.091	0.009
Sample 40%	206	0.688	0.241	0.204	1.528	0.999	0.069	0.007
Sample 50%	206	0.688	0.241	0.202	1.516	0.999	0.044	0.004
Baseline	206	0.687	0.242	0.200	1.517			

IRT A Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	206	0.023	0.009	0.008	0.065	0.998	16.697	1.645
Sample 30%	206	0.021	0.008	0.007	0.059	0.998	14.778	1.456
Sample 40%	206	0.019	0.007	0.006	0.052	0.999	11.555	1.139
Sample 50%	206	0.017	0.006	0.006	0.047	0.999	8.872	0.874
Baseline	206	0.012	0.005	0.004	0.032			

IRT B Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	206	0.390	1.279	-2.755	5.070	0.999	0.009	0.001
Sample 30%	206	0.391	1.281	-2.789	5.171	0.999	0.014	0.001
Sample 40%	206	0.390	1.277	-2.858	4.940	1.000	0.007	0.001
Sample 50%	206	0.390	1.277	-2.852	4.891	1.000	0.003	0.000
Baseline	206	0.389	1.280	-2.847	4.976			

IRT B Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	206	0.025	0.029	0.007	0.354	0.998	5.591	0.551
Sample 30%	206	0.023	0.027	0.007	0.341	0.998	4.889	0.482
Sample 40%	206	0.020	0.022	0.006	0.266	0.999	3.871	0.381
Sample 50%	206	0.018	0.019	0.005	0.234	0.999	2.970	0.293
Baseline	206	0.013	0.014	0.004	0.174			

IRT D2 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	67	0.285	0.845	-2.118	2.181	0.999	-0.061	-0.010
Sample 30%	67	0.288	0.842	-2.138	2.170	0.999	-0.038	-0.007
Sample 40%	67	0.288	0.843	-2.113	2.174	0.999	-0.040	-0.007
Sample 50%	67	0.290	0.841	-2.096	2.176	1.000	-0.026	-0.004
Baseline	67	0.293	0.834	-2.099	2.197			

IRT D2 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	67	0.025	0.017	0.008	0.091	0.999	5.541	0.957
Sample 30%	67	0.023	0.015	0.007	0.081	0.999	4.959	0.857
Sample 40%	67	0.020	0.013	0.006	0.072	0.999	3.802	0.657

Sample 50%	67	0.018	0.012	0.006	0.064	0.999	2.926	0.505
Baseline	67	0.012	0.008	0.004	0.045			

IRT D3 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	67	0.028	0.769	-1.177	2.118	0.999	0.077	0.013
Sample 30%	67	0.026	0.769	-1.192	2.138	0.999	0.055	0.009
Sample 40%	67	0.024	0.763	-1.180	2.113	1.000	0.041	0.007
Sample 50%	67	0.023	0.763	-1.188	2.096	1.000	0.037	0.006
Baseline	67	0.018	0.761	-1.191	2.099			

IRT D3 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	67	0.027	0.017	0.008	0.091	0.999	5.662	0.978
Sample 30%	67	0.025	0.016	0.007	0.081	0.999	5.060	0.874
Sample 40%	67	0.021	0.014	0.006	0.072	0.999	3.879	0.670
Sample 50%	67	0.019	0.012	0.006	0.064	0.999	2.979	0.515
Baseline	67	0.013	0.009	0.004	0.045			

IRT D4 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	25	-0.162	0.641	-1.271	1.394	0.997	-0.021	-0.006
Sample 30%	25	-0.165	0.642	-1.273	1.390	0.997	-0.035	-0.010
Sample 40%	25	-0.160	0.646	-1.272	1.412	0.999	-0.009	-0.003
Sample 50%	25	-0.162	0.638	-1.265	1.414	0.999	-0.020	-0.006
Baseline	25	-0.158	0.640	-1.257	1.398			

IRT D4 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	25	0.039	0.018	0.019	0.084	0.997	4.945	1.399
Sample 30%	25	0.036	0.017	0.017	0.077	0.998	4.383	1.240
Sample 40%	25	0.031	0.014	0.015	0.068	0.998	3.340	0.945
Sample 50%	25	0.027	0.013	0.013	0.060	0.998	2.536	0.717
Baseline	25	0.019	0.009	0.009	0.041			

IRT D5 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	11	-0.730	0.625	-1.692	0.077	0.998	-0.019	-0.008
Sample 30%	11	-0.724	0.615	-1.680	0.069	0.999	0.002	0.001
Sample 40%	11	-0.719	0.619	-1.661	0.083	0.999	0.020	0.008
Sample 50%	11	-0.720	0.615	-1.646	0.052	0.999	0.016	0.007
Baseline	11	-0.725	0.605	-1.645	0.071			

IRT D5 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	11	0.046	0.017	0.026	0.084	0.998	3.987	1.700
Sample 30%	11	0.042	0.016	0.024	0.079	0.999	3.523	1.502
Sample 40%	11	0.036	0.013	0.020	0.066	0.999	2.707	1.154
Sample 50%	11	0.032	0.012	0.018	0.058	0.999	2.062	0.879
Baseline	11	0.023	0.009	0.013	0.042			

IRT D6 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	4	-0.942	0.926	-2.070	0.051	0.998	-0.018	-0.013
Sample 30%	4	-0.943	0.910	-2.081	0.002	1.000	-0.019	-0.013
Sample 40%	4	-0.940	0.899	-2.066	0.000	1.000	-0.014	-0.010
Sample 50%	4	-0.947	0.890	-2.057	-0.009	1.000	-0.026	-0.018
Baseline	4	-0.931	0.892	-2.063	-0.012			

IRT D6 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	4	0.054	0.005	0.048	0.059	0.980	10.532	7.447
Sample 30%	4	0.049	0.004	0.044	0.054	0.990	8.962	6.337
Sample 40%	4	0.042	0.004	0.037	0.046	0.994	6.917	4.891
Sample 50%	4	0.037	0.003	0.033	0.041	0.996	5.144	3.637
Baseline	4	0.027	0.003	0.024	0.030			

IRT D7 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	4	-1.280	0.676	-2.178	-0.550	1.000	0.055	0.039
Sample 30%	4	-1.287	0.694	-2.219	-0.548	1.000	0.041	0.029
Sample 40%	4	-1.294	0.671	-2.184	-0.563	0.999	0.026	0.019
Sample 50%	4	-1.299	0.686	-2.217	-0.562	0.999	0.017	0.012
Baseline	4	-1.307	0.688	-2.231	-0.581			

IRT D7 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	4	0.064	0.020	0.039	0.087	0.998	2.956	2.090
Sample 30%	4	0.059	0.018	0.037	0.081	0.999	2.635	1.863
Sample 40%	4	0.050	0.015	0.031	0.067	0.999	2.008	1.420
Sample 50%	4	0.045	0.014	0.028	0.061	0.999	1.515	1.071
Baseline	4	0.032	0.010	0.020	0.044			

Table 4.18 summarizes the average number of student responses for the items in each of the samples and the equating baseline data set for ELA/L grade 3. The average number of student responses for the equating baseline data set was almost 92,000 with a range of 34,000 to over 132,000 responses. The sample data sets are proportionally smaller based on the percent of the baseline with the 25% sample having the fewest student responses of 8,561.

Table 4.18 Descriptive Statistics for the Number of Student Responses per Item for ELA/L Grade 3

Item N count					
	N	Mean	STD	Min	Max
Sample 25%	111	22,962.14	7,185.04	8,561	32,987
Sample 30%	111	27,570.95	8,636.21	10,247	39,685
Sample 40%	111	36,772.97	11,509.36	13,661	52,805
Sample 50%	111	45,974.63	14,392.62	17,036	66,081
Baseline	111	91,663.14	28,737.31	34,223	132,187

Table 4.19 provides descriptive statistics for the IRT parameters (*a*, *b*, and *d2-d4*) and standard errors of the parameters for ELA/L grade 3. The *a*-parameter is the slope; *b*-parameter is the difficulty/location parameter; and the *d2–d4* parameters are the thresholds for the score categories (1-3). For the 2PL/GPC model, *d1* (score category 0) is set to zero and therefore is not included in the table. The IRT parameters were highly correlated with the baseline parameters for all the sample data sets. For the *a*- and *b*-parameters the correlations ranged from 0.998-1.000. The standard errors decreased as the sample sizes increased and the correlation of the standard errors with the baseline data sets were high ranging from 0.996-1.000. Cohen's D is below 0.05 for the *a*, *b*, and *d2-d4* parameters. The *d5* parameter, which is the highest score category, has effect size values of 0.06-0.10.

Table 4.19 Descriptive Statistics for the IRT Parameters and Standard Errors for ELA/L Grade 3

IRT A Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	111	0.500	0.200	0.094	1.026	0.999	-0.133	-0.018
Sample 30%	111	0.501	0.200	0.097	1.026	0.999	-0.107	-0.014
Sample 40%	111	0.501	0.199	0.099	1.021	0.999	-0.083	-0.011
Sample 50%	111	0.503	0.200	0.097	1.021	1.000	-0.038	-0.005
Baseline	111	0.504	0.200	0.095	1.021			

IRT A Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	111	0.016	0.007	0.007	0.034	0.999	10.718	1.439
Sample 30%	111	0.015	0.006	0.007	0.031	0.999	9.488	1.274
Sample 40%	111	0.013	0.006	0.006	0.027	0.999	7.438	0.998
Sample 50%	111	0.011	0.005	0.005	0.024	1.000	5.703	0.766
Baseline	111	0.008	0.004	0.004	0.017			

IRT B Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	111	0.846	1.455	-1.436	8.102	0.998	0.017	0.002
Sample 30%	111	0.846	1.460	-1.438	8.348	0.999	0.016	0.002
Sample 40%	111	0.841	1.449	-1.440	8.303	0.999	-0.008	-0.001
Sample 50%	111	0.839	1.438	-1.438	8.077	0.999	-0.019	-0.003
Baseline	111	0.843	1.459	-1.434	8.651			

IRT B Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	111	0.034	0.067	0.007	0.520	0.990	2.405	0.323
Sample 30%	111	0.031	0.062	0.007	0.502	0.996	2.102	0.282
Sample 40%	111	0.027	0.052	0.006	0.432	0.998	1.621	0.218
Sample 50%	111	0.024	0.046	0.005	0.368	0.996	1.253	0.168
Baseline	111	0.017	0.034	0.004	0.298			

IRT D2 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	111	-0.059	1.749	-7.378	5.259	1.000	-0.020	-0.003
Sample 30%	111	-0.064	1.759	-7.537	5.258	1.000	-0.043	-0.006
Sample 40%	111	-0.061	1.750	-7.468	5.149	1.000	-0.031	-0.004
Sample 50%	111	-0.056	1.751	-7.424	5.153	1.000	-0.007	-0.001
Baseline	111	-0.054	1.748	-7.220	5.200			

IRT D2 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	111	0.043	0.048	0.008	0.330	0.997	4.342	0.583
Sample 30%	111	0.040	0.044	0.008	0.309	0.999	3.827	0.514
Sample 40%	111	0.034	0.037	0.007	0.264	0.999	3.010	0.404
Sample 50%	111	0.031	0.034	0.006	0.235	0.998	2.333	0.313
Baseline	111	0.022	0.023	0.004	0.157			

IRT D3 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	111	0.369	1.523	-5.259	7.378	1.000	-0.013	-0.002
Sample 30%	111	0.375	1.534	-5.258	7.537	1.000	0.015	0.002
Sample 40%	111	0.373	1.522	-5.149	7.468	1.000	0.008	0.001
Sample 50%	111	0.368	1.523	-5.153	7.424	1.000	-0.016	-0.002
Baseline	111	0.372	1.513	-5.200	7.220			

IRT D3 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	111	0.044	0.048	0.008	0.330	0.997	4.453	0.598
Sample 30%	111	0.040	0.044	0.008	0.309	0.999	3.920	0.526
Sample 40%	111	0.035	0.037	0.007	0.264	0.999	3.087	0.414
Sample 50%	111	0.031	0.034	0.006	0.235	0.998	2.393	0.321
Baseline	111	0.022	0.023	0.004	0.157			

IRT D4 Parameter								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	24	-1.435	0.344	-2.255	-0.825	0.994	0.344	0.099
Sample 30%	24	-1.435	0.343	-2.285	-0.815	0.995	0.340	0.098
Sample 40%	24	-1.442	0.335	-2.260	-0.845	0.997	0.275	0.079
Sample 50%	24	-1.445	0.337	-2.278	-0.862	0.997	0.241	0.069
Baseline	24	-1.468	0.325	-2.315	-0.935			

IRT D4 Parameter Standard Error								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	24	0.056	0.024	0.025	0.121	0.997	4.850	1.400
Sample 30%	24	0.050	0.022	0.022	0.107	0.998	4.190	1.209
Sample 40%	24	0.044	0.019	0.019	0.090	0.998	3.356	0.969

Sample 50%	24	0.040	0.017	0.018	0.085	0.998	2.604	0.752
Baseline	24	0.028	0.012	0.012	0.059			

Yen's Q1 item fit statistic (Yen, 1981) was computed for each of the item in the sample data sets and baseline data sets. Table 4.20 lists the correlation of the Q1 statistic for each sample with the baseline items for the equating assessments. In general, the results show that the item fit for the samples were similar to the item fit for the estimates in the baseline data sets.

Table 4.20 Correlation of Yen's Q1 Statistic between the Sample and Baseline Data Sets

Assessment	Sample	Q1 Correlation with Baseline
ELA/L Grade 10	Sample 25%	0.990792
	Sample 30%	0.992408
	Sample 40%	0.993108
	Sample 50%	0.993657
	Baseline	
Algebra 2	Sample 30%	0.795107
	Sample 40%	0.960262
	Sample 50%	0.980667
	Baseline	
ELA/L Grade 6	Sample 25%	0.993952
	Sample 30%	0.992143
	Sample 40%	0.991870
	Sample 50%	0.992174
	Baseline	
Math Grade 5	Sample 25%	0.986743
	Sample 30%	0.984161
	Sample 40%	0.987221
	Sample 50%	0.988233
	Baseline	
ELA/L Grade 3	Sample 25%	0.991956
	Sample 30%	0.993620
	Sample 40%	0.992996
	Sample 50%	0.992448
	Baseline	

4.4 Summative Scale Scores for the Equating Baseline and Sample Data Sets

Summative scale score conversion files were generated for each EOY and PBA form combination for the equating baseline and the sample data sets based on the respective IRT parameter estimates. The conversion files were applied to all the students in the *initial baseline* data set in order to evaluate the

results. Tables 4.21–4.25 provide descriptive statistics for the summative scale scores after applying the conversion files from each sample, the baseline, and the spring 2015 operational files for ELA/L grade 10, Algebra 2, ELA/L grade 6, mathematics grade 5, and ELA/L grade 3, respectively. The summative scale scores from the sample conversion files were correlated with the scale scores from the baseline conversion file and the actual summative scale score for spring 2015 operational conversion files.

For each assessment, Tables 4.21-4.25 show similar means and standard deviations for the various conversion files. The correlation of the summative scale score for the samples' conversion file with the summative scale score for the baseline conversion file is above 0.9999. The correlation of the summative scale score for the samples' conversion file with the summative scale score for the spring 2015 conversion file is above 0.9987.

Table 4.21 Descriptive Statistics for the Summative Scale Score by Conversion Files for ELA/L Grade 10

Samples, Baseline, and Spring 2015 Conversion Files Applied to the Initial Baseline Data							
Conversion Files	N	Mean	STD	Min	Max	Correlation Baseline	Correlation Spring 2015
Sample 25%	189,367	732.500	45.177	650	850	0.99993	0.99993
Sample 30%	189,367	732.533	45.191	650	850	0.99994	0.99994
Sample 40%	189,367	732.592	45.211	650	850	0.99994	0.99993
Sample 50%	189,367	732.654	45.217	650	850	0.99994	0.99993
Baseline	189,367	733.053	45.164	650	850		0.99996
Spring 2015	189,367	733.175	45.139	650	850		

Table 4.22 Descriptive Statistics for the Summative Scale Score by Conversion Files for Algebra 2

Samples, Baseline, and Spring 2015 Conversion Files Applied to the Initial Baseline Data							
Conversion Files	N	Mean	STD	Min	Max	Correlation Baseline	Correlation Spring 2015
Sample 30%	162,209	720.029	36.708	650	850	0.99991	0.99943
Sample 40%	162,209	719.992	36.720	650	850	0.99992	0.99946
Sample 50%	162,209	719.950	36.718	650	850	0.99994	0.99947
Baseline	162,209	719.890	36.796	650	850		0.99954
Spring 2015	162,209	719.292	36.876	650	850		

Table 4.23 Descriptive Statistics for the Summative Scale Score by Conversion Files for ELA/L Grade 6

Samples, Baseline, and Spring 2015 Conversion Files Applied to the Initial Baseline Data							
Conversion Files	N	Mean	STD	Min	Max	Correlation Baseline	Correlation Spring 2015
Sample 25%	373,634	739.141	31.513	650	850	0.99992	0.99988
Sample 30%	373,634	739.139	31.519	650	850	0.99992	0.99988
Sample 40%	373,634	739.188	31.518	650	850	0.99993	0.99988
Sample 50%	373,634	739.171	31.543	650	850	0.99993	0.99988
Baseline	373,634	739.330	31.549	650	850		0.99989
Spring 2015	373,634	739.682	31.630	650	850		

Table 4.24 Descriptive Statistics for the Summative Scale Score by Conversion Files for Math Grade 5

Samples, Baseline, and Spring 2015 Conversion Files Applied to the Initial Baseline Data							
Conversion Files	N	Mean	STD	Min	Max	Correlation Baseline	Correlation Spring 2015
Sample 25%	357,674	736.998	30.934	650	850	0.99994	0.99884
Sample 30%	357,674	737.006	30.931	650	850	0.99995	0.99885
Sample 40%	357,674	737.008	30.939	650	850	0.99995	0.99884
Sample 50%	357,674	737.017	30.947	650	850	0.99996	0.99882
Baseline	357,674	737.038	30.952	650	850		0.99878
Spring 2015	357,674	735.998	30.349	650	850		

Table 4.25 Descriptive Statistics for the Summative Scale Score by Conversion Files for ELA/L Grade 3

Samples, Baseline, and Spring 2015 Conversion Files Applied to the Initial Baseline Data							
Conversion Files	N	Mean	STD	Min	Max	Correlation Baseline	Correlation Spring 2015
Sample 25%	318,869	736.007	39.649	650	850	0.99993	0.99990
Sample 30%	318,869	735.988	39.634	650	850	0.99993	0.99990
Sample 40%	318,869	736.064	39.675	650	850	0.99995	0.99992
Sample 50%	318,869	736.112	39.676	650	850	0.99996	0.99993
Baseline	318,869	736.131	39.716	650	850		0.99994
Spring 2015	318,869	736.325	39.699	650	850		

The “difference that matters” criteria is based on the magnitude of the effect of the sample sizes on the conversion files resulting in scale score differences greater than half a scale score point (0.5) which would round to a different scale score value. This value is determined by calculating the difference between the summative scale score based on the baseline conversion files minus the summative scale score based on the sample conversion files for each student. Positive values indicate the baseline scale score value was larger than the summative scale score value resulting in a drop in the summative scale score for students meeting this criterion in the sample data set. Negative values indicate the baseline scale score value was smaller than the summative scale score value resulting in an increase in the summative scale score for students meeting this criterion in the sample data set.

Tables 4.26– 4.30 list the summative scale score difference values and the corresponding proportion of students with the difference value for each sample’s conversion file for ELA/L grade 10, Algebra 2, ELA/L grade 6, mathematics grade 5, and ELA/L grade 3, respectively. For ELA/L grade 10, less than 50% of students maintain the same summative scale score when implementing the 25% and 30% samples. The majority of students’ summative scale scores would decrease by one scale score point with the 25% and 30% samples compared to the baseline samples. For the 40% and 50% samples, this trends switches such that over 50% of the students summative scale scores would be the same as the baseline conversion file assigned summative scale score. Approximately 46-40% of students’ summative scale scores would decrease by one scale score point.

Table 4.26 Summative Scale Score Difference by Conversion Files for ELA/L Grade 10

Summative Scale Score Difference (Baseline - Sample)	Sample 25% Conversion File	Sample 30% Conversion File	Sample 40% Conversion File	Sample 50% Conversion File
-1	0.81	0.23	0.09	0.14
0	44.69	47.75	53.76	59.79
1	52.86	51.79	46.10	40.07
2	1.64	0.23	0.05	0.00
Total	189,367	189,367	189,367	189,367

For Algebra 2, the majority of students' summative scale scores (82-77%) would remain the same scale score value as the summative scale score from the baseline conversion file across all the sample sizes. For the remaining students (12-20%), the summative scale scores would increase by one scale score point compared to the baseline samples. Approximately 6% of students' summative scale scores would decrease by one scale score point.

Table 4.27 Summative Scale Score Difference by Conversion Files for Algebra 2

Summative Scale Score Difference (Baseline - Sample)	Sample 30% Conversion File	Sample 40% Conversion File	Sample 50% Conversion File
-3	0.005	0.004	
-2	0.18	0.08	0.01
-1	20.46	16.55	12.15
0	72.36	76.83	81.67
1	6.99	6.53	6.17
Total	162,209	162,209	162,209

For ELA/L grade 6, the majority of students' summative scale scores (80-85%) would remain the same scale score value as the summative scale score from the baseline conversion file across all the sample sizes. For the remaining students (16-19%), the summative scale scores would decrease by one scale score point compared to the baseline samples.

Table 4.28 Summative Scale Score Difference by Conversion Files for ELA/L Grade 6

Summative Scale Score Difference (Baseline - Sample)	Sample 25% Conversion File	Sample 30% Conversion File	Sample 40% Conversion File	Sample 50% Conversion File
-1	0.37	0.37	0.45	0.03
0	80.38	80.19	84.87	84.06
1	19.25	19.44	14.68	15.91
Total	373,634	373,634	373,634	373,634

For mathematics grade 5, the majority of students' summative scale scores (88-93%) would remain the same scale score value as the summative scale score from the baseline conversion file across all the

sample sizes. For the remaining students (5-8%), the summative scale scores would decrease by one scale score point compared to the baseline samples. Approximately 2-4% of the summative scale scores for the samples would increase by one scale score point compared to the baseline conversion file.

Table 4.29 Summative Scale Score Difference by Conversion Files for Math Grade 5

Summative Scale Score Difference (Baseline - Sample)	Sample 25% Conversion File	Sample 30% Conversion File	Sample 40% Conversion File	Sample 50% Conversion File
-1	4.16	3.65	2.86	2.38
0	87.65	89.50	91.26	93.09
1	8.19	6.85	5.88	4.53
Total	357,674	357,674	357,674	357,674

For ELA/L grade 3, the majority of students' summative scale scores (74-86%) would remain the same scale score value as the summative scale score from the baseline conversion file across all the sample sizes. For the remaining students (8-18%), the summative scale scores would decrease by one scale score point compared to the baseline samples. Approximately 5% of the summative scale scores for the samples would increase by one scale score point compared to the baseline conversion file.

Table 4.30 Summative Scale Score Difference by Conversion Files for ELA/L Grade 3

Summative Scale Score Difference (Baseline - Sample)	Sample 25% Conversion File	Sample 30% Conversion File	Sample 40% Conversion File	Sample 50% Conversion File
-1	5.68	5.75	5.44	5.63
0	76.28	74.36	82.47	86.81
1	17.93	19.68	12.04	7.55
2	0.11	0.21	0.05	0.00
Total	318,869	318,869	318,869	318,869

4.5 Performance Level Agreement for the Equating Baseline and Sample Data Sets

The PARCC summative scale scores are categorized into one of five performance levels. Tables 4.31-4.35 indicate the proportion of students in each performance level based on the sample conversion file compared to the baseline conversion file for ELA/L grade 10, Algebra 2, ELA/L grade 6, mathematics grade 5, and ELA/L grade 3, respectively. The percent of students in each performance level for the baseline conversion files are represented in the last column of the table. The percent of students in each performance level for the sample conversion files are represented in the last row of the table. The proportion of students with same performance level for the sample and the baseline are provided in the center diagonal cells in the table. The off-diagonal cells represent the percent of students in different performance level categorized for the sample and baseline conversion files.

For ELA/L grade 10, Table 4.31 indicates that in general the sample conversion files result in the same performance level category for the majority of students as the baseline conversion files. In general, less than half a percent of students are classified differently for a given performance level with the overall percent of students classified in different performance levels ranging from 1.2-1.52 percent.

Table 4.31 Performance Level Percent Agreement Between the Baseline and Samples for ELA/L Grade 10

Performance Level	Sample 25%					
Baseline	1	2	3	4	5	Grand Total
1	24.83					24.83
2	0.42	18.20				18.62
3		0.56	20.40			20.96
4			0.42	25.36		25.78
5				0.12	9.68	9.80
Grand Total	**25.25**	**18.77**	**20.81**	**25.49**	**9.68**	**100.00**

Performance Level	Sample 30%					
Baseline	1	2	3	4	5	Grand Total
1	24.83					24.83
2	0.49	18.14				18.62
3		0.54	20.42			20.96
4			0.42	25.36		25.78
5				0.09	9.72	9.80
Grand Total	**25.32**	**18.67**	**20.84**	**25.45**	**9.72**	**100.00**

Performance Level	Sample 40%					
Baseline	1	2	3	4	5	Grand Total
1	24.83					24.83
2	0.45	18.17				18.62
3		0.50	20.46			20.96
4			0.38	25.39	0.01	25.78
5				0.05	9.75	9.80
Grand Total	**25.28**	**18.68**	**20.84**	**25.43**	**9.77**	**100.00**

Performance Level	Sample 50%					
Baseline	1	2	3	4	5	Grand Total
1	24.83					24.83
2	0.37	18.25				18.62
3		0.45	20.51			20.96
4			0.32	25.46		25.78
5				0.06	9.74	9.80
Grand Total	**25.20**	**18.70**	**20.83**	**25.52**	**9.74**	**100.00**

For Algebra 2, Table 4.32 indicates that in general the sample conversion files result in the same performance level category for the majority of students as the baseline conversion files. In general, less than 0.15 percent of students are classified differently for a given performance level with the overall percent of students classified in different performance levels ranging from 0.15-0.26 percent.

Table 4.32 Performance Level Percent Agreement Between the Baseline and Samples for Algebra 2

Performance Level	Sample 30%					
Baseline	1	2	3	4	5	Grand Total
1	31.44					31.44
2		25.91				25.91
3			20.93	0.03		20.95
4			0.09	20.43	0.02	20.54
5				0.01	1.14	1.15
Grand Total	**31.44**	**25.91**	**21.02**	**20.47**	**1.16**	**100.00**

Performance Level	Sample 40%					
Baseline	1	2	3	4	5	Grand Total
1	31.44					31.44
2		25.91				25.91
3		0.15	20.77	0.03		20.95
4			0.06	20.47	0.01	20.54
5				0.01	1.15	1.15
Grand Total	**31.44**	**26.06**	**20.83**	**20.51**	**1.16**	**100.00**

Performance Level	Sample 50%					
Baseline	1	2	3	4	5	Grand Total
1	31.44					31.44
2		25.91				25.91
3		0.15	20.77	0.03		20.95
4			0.06	20.47	0.01	20.54
5				0.00	1.15	1.15
Grand Total	**31.44**	**26.06**	**20.83**	**20.50**	**1.16**	**100.00**

For ELA/L grade 6, Table 4.33 indicates that in general the sample conversion files result in the same performance level category for the majority of students as the baseline conversion files. In general, less than 0.18 percent of students are classified differently for a given performance level with the overall percent of students classified in different performance levels ranging from 0.41-0.53 percent.

Table 4.33 Performance Level Percent Agreement Between the Baseline and Samples for ELA/L Grade 6

Performance Level	Sample 25%					
Baseline	1	2	3	4	5	Grand Total
1	11.30					11.30
2	0.09	20.24				20.33
3		0.07	30.14			30.21
4			0.18	32.74		32.92
5				0.14	5.09	5.24
Grand Total	**11.39**	**20.31**	**30.32**	**32.88**	**5.09**	**100.00**

Performance Level	Sample 30%					
Baseline	1	2	3	4	5	Grand Total
1	11.30					11.30
2	0.09	20.24				20.33
3		0.12	30.09			30.21
4			0.18	32.74		32.92
5				0.14	5.09	5.24
Grand Total	**11.39**	**20.36**	**30.27**	**32.88**	**5.09**	**100.00**

Performance Level	Sample 40%					
Baseline	1	2	3	4	5	Grand Total
1	11.30					11.30
2	0.09	20.20	0.04			20.33
3		0.08	30.13			30.21
4			0.07	32.85		32.92
5				0.13	5.10	5.24
Grand Total	**11.39**	**20.28**	**30.24**	**32.98**	**5.10**	**100.00**

Performance Level	Sample 50%					
Baseline	1	2	3	4	5	Grand Total
1	11.30					11.30
2	0.05	20.27				20.33
3		0.12	30.10			30.21
4			0.15	32.77		32.92
5				0.13	5.10	5.24
Grand Total	**11.36**	**20.39**	**30.25**	**32.90**	**5.10**	**100.00**

For mathematics grade 5, Table 4.34 indicates that in general the sample conversion files result in the same performance level category for the majority of students as the baseline conversion files. In general, less than 0.28 percent of students are classified differently for a given performance level with the overall percent of students classified in different performance levels ranging from 0.37-0.60 percent.

Table 4.34 Performance Level Percent Agreement Between the Baseline and Samples for Grade 5 Math

Performance Level	Sample 25%					
Baseline	1	2	3	4	5	Grand Total
1	10.64					10.64
2	0.04	25.17				25.21
3		0.28	29.95	0.04		30.28
4			0.17	28.75	0.03	28.95
5				0.04	4.90	4.93
Grand Total	**10.68**	**25.45**	**30.12**	**28.83**	**4.93**	**100.00**

Performance Level	Sample 30%					
Baseline	1	2	3	4	5	Grand Total
1	10.64					10.64
2	0.04	25.17				25.21
3		0.28	29.90	0.09		30.28
4			0.05	28.89	0.01	28.95
5				0.06	4.87	4.93
Grand Total	**10.68**	**25.45**	**29.95**	**29.04**	**4.89**	**100.00**

Performance Level	Sample 40%					
Baseline	1	2	3	4	5	Grand Total
1	10.64					10.64
2	0.04	25.17				25.21
3		0.22	29.97	0.09		30.28
4			0.05	28.90		28.95
5				0.02	4.91	4.93
Grand Total	**10.68**	**25.38**	**30.02**	**29.01**	**4.91**	**100.00**

Performance Level	Sample 50%					
Baseline	1	2	3	4	5	Grand Total
1	10.64					10.64
2	0.04	25.17				25.21
3		0.22	30.01	0.05		30.28
4			0.05	28.90		28.95
5				0.01	4.92	4.93
Grand Total	**10.68**	**25.38**	**30.06**	**28.96**	**4.92**	**100.00**

For ELA/L grade 3, Table 4.35 indicates that in general the sample conversion files result in the same performance level category for the majority of students as the baseline conversion files. In general, less than 0.11 percent of students are classified differently for a given performance level with the overall percent of students classified in different performance levels ranging from 0.15-0.54 percent.

Table 4.35 Performance Level Percent Agreement Between the Baseline and Samples for ELA/L Grade 3

Performance Level	Sample 25%					
Baseline	1	2	3	4	5	Grand Total
1	19.31					19.31
2	0.10	19.53	0.08			19.70
3		0.06	23.62	0.08		23.76
4			0.05	33.82		33.87
5				0.09	3.27	3.36
Grand Total	**19.41**	**19.59**	**23.74**	**33.99**	**3.27**	**100.00**

Performance Level	Sample 30%					
Baseline	1	2	3	4	5	Grand Total
1	19.31					19.31
2	0.10	19.53	0.08			19.70
3		0.06	23.62	0.08		23.76
4			0.11	33.76		33.87
5				0.11	3.24	3.36
Grand Total	**19.41**	**19.59**	**23.81**	**33.95**	**3.24**	**100.00**

Performance Level	Sample 40%					
Baseline	1	2	3	4	5	Grand Total
1	19.31					19.31
2	0.04	19.63	0.03			19.70
3		0.06	23.59	0.11		23.76
4				33.87		33.87
5				0.07	3.28	3.36
Grand Total	**19.35**	**19.69**	**23.62**	**34.05**	**3.28**	**100.00**

Performance Level	Sample 50%					
Baseline	1	2	3	4	5	Grand Total
1	19.31					19.31
2		19.70				19.70
3			23.65	0.11		23.76
4				33.87		33.87
5				0.04	3.31	3.36
Grand Total	**19.31**	**19.70**	**23.65**	**34.02**	**3.31**	**100.00**

Section 5: Study Summary

This research study investigated analyses for selecting an early post-equating sample for the PARCC assessments during operational administration in spring 2016. The process included evaluating the representativeness of multiple convenience samples that were selected based on student administration date. The samples were analyzed through classical item analyses and IRT calibrations. The results were compared to a baseline sample.

The results are summarized in the following sections:

5.1 Sample Size and Representativeness

The sample sizes for the study were selected to balance the need for adequate data to produce valid and reliable test scores and the desire for early reporting of student performance. The samples sizes investigated were 25% 30%, 40%, and 50% of the population or initial baseline data. For the high school assessments, the sample sizes were reduced 5-7% after applying the filtering requirements for equating analyses, specifically the attemptedness rules filtered out the majority of the students that were dropped. As a result, sampling for high school assessments needs to consider the attainable sample that will be available for calibrations. On the other hand, grades 3–8 for ELA/L and mathematics did not result in drops in the overall sample size proportion when the data were filtered for the equating requirements.

The representativeness of a sample impacts the fidelity of the PARCC scores when implementing a common-item non-equivalent groups equating design (Kolen & Brennan, 2004). In this study, demographic characteristics evaluated included gender, ethnicity, economically disadvantaged, English language learners, Students with Disabilities, and grade level for high school courses. An ANOVA estimated the amount of variance in the 2015 summative scale scores attributable to the demographic variables. Although the demographic variables did not explain much of the variability in the summative scale scores, for each assessment, three demographic variables with the largest Eta-squared values were selected for evaluating the representativeness of the various samples.

In general, the distribution of the demographic variables became more similar to the baseline as the sample sizes increased. Except for Integrated Mathematics I, II, and III, the differences in the proportions for the demographic variables fell within 10 percentage points with many differences below 5 percentage points.

The 2015 summative scale scores were used to evaluate the ability distribution of the samples and the representativeness of the score distributions. For the majority of the assessments, the distribution of the summative scale scores and the proportion of students in each of the performance levels were very similar to the baseline.

Eight states are participating in the spring 2016 administration: Colorado (CO), District of Columbia (DC), Illinois (IL), Maryland (MD), Massachusetts (MA), New Jersey (NJ), New Mexico (NM), and Rhode Island (RI). Based on the testing windows for spring 2015, sample sizes of approximately 25-30% of the baseline resulted in the inclusion of five of the eight states in the early equating samples: CO, IL, MD, NJ, and NM. DC, MA, and RI often required 50% or more of the baseline sample to be represented due to their testing later in the school year. Although, the five states were not represented in the same proportion as the baseline for some assessments, the proportions tended to be within 11 percentage points with CO and NJ overrepresented and IL underrepresented.

5.2 Item-Level Analyses

Classical item analyses and IRT calibrations were conducted to evaluate the sample results compared to the baseline results. Both the classical and IRT analyses for all samples were very similar to the baseline across all assessments, with larger differences found for the Integrated Mathematics I, II, and III assessments.

5.3 Test-Level Analyses

The summative scale score and the performance level categories were evaluated to compare the samples to the baseline data sets. Conversion files based on the IRT parameter estimates from the samples and the baseline were generated and applied to the population. In addition, the students' actual summative scale score from the spring 2015 administration were compared to the sample results. As with the item-level results, the test-level results were similar when comparing the samples to the baseline data sets and the spring 2015 results. However, there were distinctive patterns for the test-level analyses such that the early samples had larger differences in the summative scale score conversion files compared to the baseline, then the larger samples. Small differences (1-2 scale score points) for the summative scale score were found for some students especially in the smaller samples; however, the majority (98%) of students maintained the same performance level designation regardless of the sample when compared to the baseline performance level designation.

5.4 Research Questions

The following are answers to the research questions of interest for the early post-equating sampling process based on the results from this study:

1. Which characteristics of the student population should be considered in determining a representative sample?

In this study, gender, economically disadvantaged, English language learner, Students with Disabilities, ethnicity, grade level, and state were evaluated for each of the samples and compared to the baseline data set. Each of these demographic variables was considered in evaluating the representativeness of the samples. However, some variables tended to explain more of the variability in the summative scale scores which led to additional investigation. The demographic variables Students with Disabilities, ethnicity, and economically disadvantaged tended to explain more of the variability in the summative scale scores compared to gender and English language learner for grades 3–8 ELA/L and math. For the high school assessments, the variables of interest were identified as grade level, Students with Disabilities, ethnicity, and English language learners. The 2015 summative scale score was also identified as a critical variable to consider.

2. Given the differences in states' spring testing schedules, can state-representative samples be obtained early in the administration window to support robust post-equating analyses?

   Of the eight states scheduled to participate in the spring 2016 administrations, five states were found to be consistently represented in each of the early equating samples (25%, 30%, 40%, and 50% samples). The states were Colorado (CO), Illinois (IL), Maryland (MD), New Jersey (NJ), and New Mexico (NM). For some assessments and smaller sample sizes, the five states were not represented in the same proportion as the baseline. The proportions tended to be within 11 percentage points with Colorado and New Jersey being overrepresented and Illinois being underrepresented. District of Columbia, Rhode Island, and Massachusetts were not represented in many of the samples due to their testing later in the school year. Even though the state representation did not include all the participant states, the IRT parameter estimates, summative scale scores, and performance level results were similar to the baseline which included all states. The five states consistently represented in the samples due to earlier testing windows result in a robust post-equating process.

3. What is the minimum sample size needed to obtain stable item parameter estimates through post-equating to support PARCC scores?

   For ELA/L and mathematics assessments at grades 3–8, total sample sizes of approximately 25% of the baseline (about 85,000-90,000) were sufficient to support stable item parameter estimates, similar summative scale scores and consistent performance level agreement when compared to the baseline data sets. For grades 9–11 ELA/L, Algebra 1, Geometry, and Algebra 2, the large number of students removed from the equating due to attemptedness and filtering rules required larger sample sizes. These assessments also required larger samples for an individual item to have student responses in each of the score categories. Therefore, sample sizes of approximately 40% of the baseline (about 75,000-80,000 for Algebra 2 and ELA/L grade 10) are recommended for stable item parameter estimates, similar summative scale scores and consistent performance level agreement when compared to the baseline data sets. For Integrated Mathematics I, II, III, the sample sizes for the eight states in spring 2016 were small

and resulted in many items not having enough responses in all of the score categories. As a result, at least 90% of the population is recommended for these assessments to be post-equated.

4. Which analyses and statistical criteria should be used to determine the minimum sample size for early post-equating for spring 2016?

   Item-level and test-level analyses for the multiple samples resulted in the minimal sample size of approximately 5,000 with the average response total around 20,000 having stable statistics and robust IRT parameters. Overall, the results for each of the samples were very similar to the baseline across the assessments. Summarizing these results can provide guidelines for the early equating sampling criteria.

5.5 Proposed Early Post-Equating Sampling Criteria

The spring 2015 PARCC operational data for the eight states participating in PARCC in spring 2016 was the data source for the early sampling analyses. Table 5.1 lists the PARCC assessments and number of testers which served as the population for this study. The four samples were compared to analyses using this data as the baseline.

Table 5.1 Spring 2015 PARCC Administrations and N Counts

Assessments	Spring 2015 N Counts	Assessments	Spring 2015 N Counts
Algebra 2	191,763	ELA/L Grade 11	182,112
Geometry	152,044	ELA/L Grade 10	219,484
Algebra 1	298,756	ELA/L Grade 9	267,452
Integrated Math III	10,501		
Integrated Math II	12,495		
Integrated Math I	22,689		
Math Grade 8	306,693	ELA/L Grade 8	379,162
Math Grade 7	370,028	ELA/L Grade 7	383,470
Math Grade 6	380,873	ELA/L Grade 6	383,658
Math Grade 5	363,495	ELA/L Grade 5	370,295
Math Grade 4	349,198	ELA/L Grade 4	356,184
Math Grade 3	324,657	ELA/L Grade 3	323,724

The results from the study were used to determine criteria for sample size and acceptable differences between the baseline demographic distributions and the sample demographic distributions. Table 5.2 lists the proposed criteria for the early post-equating sampling for spring 2016. The proposed criteria reflect the study results associated with the 2015 test design. For spring 2016, the PBA and EOY components will be combined into a single test administration. The change in test design will result in

fewer items being administered. In addition, the number of test forms will be reduced compared to the spring 2015 administration. These changes will need to be considered when determining the spring 2016 early equating criteria. The proposed criteria may need to be adjusted for larger counts per form and item to match the overall count proposed or it may need to be adjusted for lower overall counts to reflect the counts per form and item. For the demographic variables, the sampling will require representation in each category in addition to the criteria listed in Table 5.2.

Table 5.2 Proposed Early Post-Equating Sampling Criteria

Assessments	Proposed Early Post-Equating Sampling Criteria	
ELA/L Grades 3–8	**Demographic Variables**	**Online Criteria**
	Student with Disabilities (Yes)	Proportions within 3% of the 2015 distributions
	Economically disadvantaged (Yes)	Proportions within 6% of the 2015 distributions
	Ethnicity Category	Effect size less than 0.15 of the 2015 distributions for each ethnicity category
	Sample Size	
	Overall N Count	Minimum 90,000 (25%)
	Per Form	Average 22,000
	Per Item	Minimum 5,000
	State	
	Five of the 8: CO, IL, MD, NJ, NM	Proportions within 12% of the 2015 distributions
	Prior PARCC Scores	
	Summative Scale Score	Percent difference in CDF within 3% of 2015
	Performance Level	Proportions within 3% of the 2015 distribution
Math Grades 3–8	**Demographic Variables**	**Online Criteria**
	Student with Disabilities (Yes)	Proportions within 3% of the 2015 distributions
	Ethnicity Category	Effect size less than 0.15 of the 2015 distributions for each ethnicity category
	Economically disadvantaged (Yes)	Proportions within 5% of the 2015 distributions
	Sample Size	
	Overall	Minimum 90,000 (25%)
	Per Form	Average 22,000
	Per Item	Minimum 5,000
	State	
	Five of the 8: CO, IL, MD, NJ, NM	Proportions within 12% of the 2015 distributions
	Prior PARCC Scores	
	Summative Scale Score	Percent difference in CDF within 3% of 2015
	Performance Level	Proportions within 3% of the 2015 distribution

Assessments	Proposed Early Post-Equating Sampling Criteria	
ELA/L Grades 9-11	**Demographic Variables**	**Online Criteria**
	Student with Disabilities (Yes)	Proportions within 3% of the 2015 distributions
	English Language Learner (Yes)	Proportions within 6% of the 2015 distributions
	Ethnicity Category	Effect size less than 0.15 of the 2015 distributions for each ethnicity category
	Sample Size	
	Overall	Minimum 80,000 (30-40%)
	Per Form	Average 22,000
	Per Item	Minimum 5,000
	State	
	Five of the 8: CO, IL, MD, NJ, NM	Proportions within 12% of the 2015 distributions
	Prior PARCC Scores	
	Summative Scale Score	Percent difference in CDF within 3% of 2015
	Performance Level	Proportions within 3% of the 2015 distribution
Algebra 1, Geometry, Algebra 2	**Demographic Variables**	**Online Criteria**
	Grade Level	Proportions within 3% of the 2015 distributions
	Student with Disabilities (Yes)	Proportions within 6% of the 2015 distributions
	Ethnicity Category	Effect size less than 0.15 of the 2015 distributions for each ethnicity category
	Sample Size	
	Overall	Minimum 75,000-90,000 (30-40%)
	Per Form	Average 22,000
	Per Item	Minimum 5,000
	State	
	Five of the 8: CO, IL, MD, NJ, NM	Proportions within 12% of the 2015 distributions
	Prior PARCC Scores	
	Summative Scale Score	Percent difference in CDF within 3% of 2015
	Performance Level	Proportions within 3% of the 2015 distribution
Integrated Math I, II, III	**Demographic Variables**	**Online Criteria**
	Grade Level	All testers (90%)
	Student with Disabilities	All testers (90%)
	Ethnicity Category	All testers (90%)
	Sample Size	
	Overall	All testers (90%)
	Per Form	Average 5,000
	Per Item	Minimum 1,000
	State	
	States	All state participants
	Prior PARCC Scores	
	Summative Scale Score	All testers (90%)
	Performance Level	All testers (90%)

Section 6: Next Steps

This study focuses on the online test forms. For paper forms, the answer documents were returned at the end of the testing windows in spring 2015; therefore, the actual date a student tested is not known. Early post-equating for the paper test forms will be evaluated in a follow-up study in which additional analyses for evaluating student performance across modes will be considered. These results are provided in the *Addendum: Early Post-Equating Sampling for Paper Report*.

An early post-equating sample enables student reports to be provided in a timely manner to help students and schools determine next steps for instruction. This study focused on the psychometric considerations for obtaining robust scale scores through a sampling process prior to post equating. Additional considerations such as the time required to hand score items, state testing windows, and paper test form processing need to be considered.

References

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.

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

Muraki, E. (1992). A generalized partial credit model: Applications of an EM algorithm. *Applied Psychological Measurement, 16,* 159-176.

Stocking, M. L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement, 7,* 201-210.

Yen, W. (1981). Using simulation results to choose a latent trait model. *Applied Psychological Measurement*, 5, 245-262.

Appendix A. Number of Students in the Baseline and Sample Data Sets

ELA/L

Table A.1 Percent of Students in the Baseline and Sample Data Sets for ELA/L Grade 11

	Initial Sample		Equating Sample	
Grade 11 ELA/L	N Count	Percent of Baseline	N Count2	Percent of Baseline3
Sample 25%	45,528	25.00	37,219	20.44
Sample 30%	54,633	30.00	45,190	24.81
Sample 40%	72,844	40.00	61,138	33.57
Sample 50%	91,056	50.00	77,035	42.30
Baseline	182,112	100.00	148,041	81.29

Table A.2 Percent of Students in the Baseline and Sample Data Sets for ELA/L Grade 9

	Initial Sample		Equating Sample	
Grade 9 ELA/L	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	66,863	25.00	59,745	22.34
Sample 30%	80,235	30.00	71,998	26.92
Sample 40%	106,980	40.00	96,324	36.02
Sample 50%	133,726	50.00	120,639	45.11
Baseline	267,452	100.00	233,074	87.15

Table A.3 Percent of Students in the Baseline and Sample Data Sets for ELA/L Grade 8

	Initial Sample		Equating Sample	
Grade 8 ELA/L	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	94,790	25.00	91,335	24.09
Sample 30%	113,748	30.00	109,799	28.96
Sample 40%	151,664	40.00	146,410	38.61
Sample 50%	189,581	50.00	182,623	48.16
Baseline	379,162	100.00	362,467	95.60

Table A.4 Percent of Students in the Baseline and Sample Data Sets for ELA/L Grade 7

	Initial Sample		Equating Sample	
Grade 7 ELA/L	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	95,867	25.00	93,205	24.31
Sample 30%	115,041	30.00	111,908	29.18
Sample 40%	153,388	40.00	149,160	38.90
Sample 50%	191,735	50.00	186,700	48.69
Baseline	383,470	100.00	370,386	96.59

Table A.5 Percent of Students in the Baseline and Sample Data Sets for ELA/L Grade 5

	Initial Sample		Equating Sample	
Grade 5 ELA/L	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	92,573	25.00	91,179	24.62
Sample 30%	111,088	30.00	109,519	29.58
Sample 40%	148,118	40.00	146,148	39.47
Sample 50%	185,147	50.00	182,735	49.35
Baseline	370,295	100.00	364,281	98.38

Table A.6 Percent of Students in the Baseline and Sample Data Sets for ELA/L Grade 4

	Initial Sample		Equating Sample	
Grade 4 ELA/L	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	89,046	25.00	87,792	24.65
Sample 30%	106,855	30.00	105,414	29.60
Sample 40%	142,473	40.00	140,654	39.49
Sample 50%	178,092	50.00	175,931	49.39
Baseline	356,184	100.00	350,808	98.49

Math

Table A.7 Percent of Students in the Baseline and Sample Data Sets for Algebra 1

	Initial Sample		Equating Sample	
Algebra 1	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	74,689	25.00	67,133	22.47
Sample 30%	89,626	30.00	80,874	27.07
Sample 40%	119,502	40.00	108,338	36.26
Sample 50%	149,378	50.00	135,766	45.44
Baseline	298,756	100.00	264,367	88.49

Table A.8 Percent of Students in the Baseline and Sample Data Sets for Geometry

	Initial Sample		Equating Sample	
Geometry	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	38,011	25.00	33,030	21.72
Sample 30%	45,613	30.00	39,903	26.24
Sample 40%	60,817	40.00	53,162	34.96
Sample 50%	76,022	50.00	66,729	43.89
Baseline	152,044	100.00	131,160	86.26

Table A.9 Percent of Students in the Baseline and Sample Data Sets for Integrated Math III

	Initial Sample		Equating Sample	
Integrated Math III	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	2,625	25.00	1,705	16.24
Sample 30%	3,150	30.00	2,132	20.30
Sample 40%	4,200	40.00	3,012	28.68
Sample 50%	5,250	50.00	3,836	36.53
Baseline	10,501	100.00	7,210	68.66

Table A.10 Percent of Students in the Baseline and Sample Data Sets for Integrated Math I

	Initial Sample		Equating Sample	
Integrated Math I	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	5,672	25.00	4,776	21.05
Sample 30%	6,806	30.00	5,686	25.06
Sample 40%	9,075	40.00	7,765	34.22
Sample 50%	11,344	50.00	9,781	43.11
Baseline	22,689	100.00	19,296	85.05

Table A.11 Percent of Students in the Baseline and Sample Data Sets for Math Grade 8

	Initial Sample		Equating Sample	
Grade 8 Math	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	76,673	25.00	73,654	24.02
Sample 30%	92,007	30.00	88,372	28.81
Sample 40%	122,677	40.00	118,109	38.51
Sample 50%	153,346	50.00	147,820	48.20
Baseline	306,693	100.00	292,854	95.49

Table A.12 Percent of Students in the Baseline and Sample Data Sets for Math Grade 7

	Initial Sample		Equating Sample	
Grade 7 Math	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	92,507	25.00	89,786	24.26
Sample 30%	111,008	30.00	107,654	29.09
Sample 40%	148,011	40.00	143,964	38.91
Sample 50%	185,014	50.00	179,923	48.62
Baseline	370,028	100.00	357,488	96.61

Table A.13 Percent of Students in the Baseline and Sample Data Sets for Math Grade 6

	Initial Sample		Equating Sample	
Grade 6 Math	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	95,218	25.00	93,174	24.46
Sample 30%	114,261	30.00	111,849	29.37
Sample 40%	152,349	40.00	149,247	39.19
Sample 50%	190,436	50.00	186,636	49.00
Baseline	380,873	100.00	371,256	97.48

Table A.14 Percent of Students in the Baseline and Sample Data Sets for Math Grade 4

	Initial Sample		Equating Sample	
Grade 4 Math	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	87,299	25.00	86,193	24.68
Sample 30%	104,759	30.00	103,419	29.62
Sample 40%	139,679	40.00	137,948	39.50
Sample 50%	174,599	50.00	172,479	49.39
Baseline	349,198	100.00	344,191	98.57

Table A.15 Percent of Students in the Baseline and Sample Data Sets for Math Grade 3

	Initial Sample		Equating Sample	
Grade 3 Math	N Count	Percent of Baseline	N Count	Percent of Baseline
Sample 25%	81,164	25.00	80,169	24.69
Sample 30%	97,397	30.00	96,228	29.64
Sample 40%	129,862	40.00	128,368	39.54
Sample 50%	162,328	50.00	160,488	49.43
Baseline	324,657	100.00	320,146	98.61

Appendix B. State Representation for the Baseline and Sample Data Sets

ELA/L

Table B.1 State Representation in the Baseline and Sample Data Sets for ELA/L Grade 11

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	37.54	35.92	34.83	32.28	**25.37**	-0.280	-0.242	-0.217	-0.159
DC					**0.03**				
IL	6.56	8.76	12.91	14.37	**20.48**	0.345	0.290	0.187	0.151
MA				0.07	**3.33**				0.182
MD				0.00	**0.01**				0.009
NJ	36.86	36.88	36.49	38.60	**39.59**	0.056	0.055	0.063	0.020
NM	19.05	18.45	15.78	14.68	**11.19**	-0.249	-0.230	-0.146	-0.111
RI					**0.00**				
Total	**45,528**	**54,633**	**72,844**	**91,056**	**182,112**				

Table B.2 State Representation in the Baseline and Sample Data Sets for ELA/L Grade 9

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	27.71	28.83	28.62	27.70	**21.77**	-0.144	-0.171	-0.166	-0.144
DC		0.01	0.01	0.29	**1.32**		0.115	0.114	0.090
IL	21.53	21.47	23.15	25.54	**28.88**	0.162	0.164	0.126	0.074
MA				0.00	**2.65**				0.165
MD			0.00	0.00	**0.01**			0.008	0.008
NJ	41.07	38.73	37.30	36.22	**33.57**	-0.159	-0.109	-0.079	-0.056
NM	9.69	10.96	10.92	10.24	**9.04**	-0.023	-0.067	-0.065	-0.042
RI				0.01	**2.76**				0.168
Total	**66,863**	**80,235**	**106,980**	**133,726**	**267,452**				

Table B.3 State Representation in the Baseline and Sample Data Sets for ELA/L Grade 8

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	21.11	20.30	21.83	21.30	**15.81**	-0.145	-0.123	-0.165	-0.150
DC	0.01	0.05	0.14	0.25	**1.10**	0.105	0.100	0.092	0.081
IL	24.66	27.43	30.16	32.40	**33.23**	0.182	0.123	0.065	0.018
MA				0.08	**5.37**				0.235
MD	11.99	11.75	11.18	11.65	**12.17**	0.005	0.013	0.030	0.016
NJ	32.38	30.84	28.42	26.44	**24.29**	-0.189	-0.153	-0.096	-0.050
NM	9.85	9.63	8.27	7.84	**5.96**	-0.164	-0.155	-0.098	-0.080
RI				0.03	**2.06**				0.143
Total	**94,790**	**113,748**	**151,664**	**189,581**	**379,162**				

Table B.4 State Representation in the Baseline and Sample Data Sets for ELA/L Grade 7

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	22.39	22.33	22.76	20.67	**16.05**	-0.173	-0.171	-0.183	-0.126
DC	0.04	0.04	0.09	0.12	**1.05**	0.099	0.100	0.094	0.091
IL	20.12	20.52	22.80	26.68	**31.57**	0.246	0.238	0.189	0.105
MA				0.00	**5.50**				0.241
MD	11.39	11.81	12.10	11.80	**13.75**	0.068	0.056	0.048	0.057
NJ	32.05	32.51	31.82	31.70	**24.25**	-0.182	-0.193	-0.177	-0.174
NM	14.01	12.80	10.43	9.04	**6.00**	-0.337	-0.286	-0.187	-0.128
RI					**1.83**				
Total	**95,867**	**115,041**	**153,388**	**191,735**	**383,470**				

Table B.5 State Representation in the Baseline and Sample Data Sets for ELA/L Grade 5

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	21.06	19.79	22.23	21.66	**17.16**	-0.104	-0.070	-0.135	-0.120
DC	0.05	0.07	0.19	0.25	**1.18**	0.105	0.103	0.092	0.086
IL	15.86	18.71	21.76	24.95	**26.91**	0.249	0.185	0.116	0.044
MA				0.00	**5.30**				0.237
MD	15.29	16.85	17.47	16.90	**15.03**	-0.007	-0.051	-0.068	-0.052
NJ	37.32	35.19	29.93	28.50	**25.89**	-0.261	-0.212	-0.092	-0.060
NM	10.43	9.39	8.41	7.74	**6.21**	-0.175	-0.132	-0.091	-0.063
RI					**2.32**				
Total	**92,573**	**111,088**	**148,118**	**185,147**	**370,295**				

Table B.6 State Representation in the Baseline and Sample Data Sets for ELA/L Grade 4

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	22.54	23.00	24.22	23.62	**17.90**	-0.121	-0.133	-0.165	-0.149
DC	0.01	0.05	0.13	0.27	**1.35**	0.116	0.112	0.105	0.093
IL	15.90	17.82	19.93	22.34	**24.47**	0.199	0.155	0.106	0.050
MA					**5.46**				
MD	13.19	12.85	13.44	13.26	**15.70**	0.069	0.078	0.062	0.067
NJ	34.70	31.76	30.52	30.22	**26.62**	-0.183	-0.116	-0.088	-0.082
NM	13.66	14.52	11.74	10.28	**6.13**	-0.314	-0.350	-0.234	-0.173
RI					**2.38**				
Total	**89,046**	**106,855**	**142,473**	**178,092**	**356,184**				

Math

Table B.7 State Representation in the Baseline and Sample Data Sets for Algebra 1

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	24.95	24.94	23.75	23.18	**15.79**	-0.251	-0.251	-0.219	-0.203
DC	0.07	0.16	0.40	0.52	**1.29**	0.108	0.100	0.079	0.068
IL	11.86	12.59	13.80	14.63	**17.93**	0.158	0.139	0.108	0.086
MA				0.02	**2.86**				0.170
MD	14.65	15.86	16.65	16.94	**18.21**	0.092	0.061	0.040	0.033
NJ	41.34	39.04	37.09	34.68	**32.72**	-0.184	-0.135	-0.093	-0.042
NM	7.13	7.40	8.27	9.58	**8.33**	0.043	0.033	0.002	-0.045
RI			0.04	0.46	**2.89**			0.170	0.145
Total	**74,689**	**89,626**	**119,502**	**149,378**	**298,756**				

Table B.8 State Representation in the Baseline and Sample Data Sets for Geometry

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	33.63	32.50	31.78	30.87	**23.48**	-0.240	-0.213	-0.196	-0.174
DC	0.03	0.23	0.29	0.54	**2.40**	0.155	0.142	0.138	0.122
IL	2.49	2.62	3.14	4.02	**3.16**	0.039	0.031	0.001	-0.049
MA					**0.77**				
MD					**0.00**				
NJ	50.15	51.08	51.90	49.55	**51.44**	0.026	0.007	-0.009	0.038
NM	13.70	13.58	12.90	14.79	**13.76**	0.002	0.005	0.025	-0.030
RI				0.24	**5.00**				0.219
Total	**38,011**	**45,613**	**60,817**	**76,022**	**152,044**				

Table B.9 State Representation in the Baseline and Sample Data Sets for Integrated Math III

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	96.04	96.57	96.10	92.44	**83.97**	-0.329	-0.343	-0.330	-0.231
IL		0.06	1.26	4.15	**9.60**		0.324	0.283	0.185
MA					**1.80**				
NJ					**0.04**				
NM	3.96	3.37	2.64	3.41	**4.59**	0.030	0.059	0.093	0.056
Total	**2,625**	**3,150**	**4,200**	**5,250**	**10,501**				

Table B.10 State Representation in the Baseline and Sample Data Sets for Integrated Math I

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	63.93	67.46	68.56	64.39	**57.62**	-0.128	-0.199	-0.221	-0.137
IL	32.07	28.61	27.87	32.47	**37.47**	0.112	0.183	0.198	0.103
MA					**1.73**				
NJ			0.03	0.08	**0.16**			0.032	0.020
NM	4.00	3.94	3.54	3.07	**2.99**	-0.059	-0.055	-0.032	-0.004
RI					**0.04**				
Total	**5,672**	**6,806**	**9,075**	**11,344**	**22,689**				

Table B.11 State Representation in the Baseline and Sample Data Sets for Math Grade 8

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	21.24	20.82	18.86	17.54	**13.35**	-0.232	-0.220	-0.162	-0.123
DC	0.09	0.10	0.21	0.34	**1.03**	0.093	0.092	0.082	0.069
IL	36.21	38.00	42.95	43.50	**40.96**	0.097	0.060	-0.040	-0.052
MA			0.00	0.29	**5.71**			0.246	0.234
MD	10.93	11.11	10.31	9.97	**11.61**	0.021	0.016	0.041	0.051
NJ	24.22	22.70	20.02	20.78	**19.47**	-0.120	-0.081	-0.014	-0.033
NM	7.30	7.28	7.55	7.13	**6.03**	-0.053	-0.052	-0.064	-0.046
RI			0.10	0.45	**1.82**			0.129	0.103
Total	**76,673**	**92,007**	**122,677**	**153,346**	**306,693**				

Table B.12 State Representation in the Baseline and Sample Data Sets for Math Grade 7

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	24.76	24.79	20.58	19.55	**14.64**	-0.286	-0.287	-0.168	-0.139
DC	0.09	0.16	0.12	0.28	**1.03**	0.093	0.086	0.090	0.074
IL	28.71	29.25	30.90	30.79	**32.68**	0.085	0.073	0.038	0.040
MA				0.05	**5.71**				0.244
MD	8.62	9.41	10.20	10.92	**13.89**	0.152	0.130	0.107	0.086
NJ	27.00	25.57	28.90	29.66	**24.02**	-0.070	-0.036	-0.114	-0.132
NM	10.82	10.81	9.30	8.37	**6.14**	-0.195	-0.195	-0.132	-0.093
RI				0.38	**1.89**				0.111
Total	**92,507**	**111,008**	**148,011**	**185,014**	**370,028**				

Table B.13 State Representation in the Baseline and Sample Data Sets for Math Grade 6

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	20.70	20.83	19.26	19.62	**15.10**	-0.156	-0.160	-0.116	-0.126
DC	0.21	0.18	0.20	0.37	**1.07**	0.083	0.086	0.085	0.068
IL	26.05	29.57	31.04	32.92	**32.62**	0.140	0.065	0.034	-0.007
MA				0.35	**5.36**				0.222
MD	11.18	10.50	12.10	11.74	**12.78**	0.048	0.068	0.020	0.031
NJ	33.00	30.40	29.44	27.50	**24.83**	-0.189	-0.129	-0.107	-0.062
NM	8.86	8.53	7.96	7.25	**6.00**	-0.120	-0.107	-0.083	-0.053
RI				0.25	**2.25**				0.135
Total	**95,218**	**114,261**	**152,349**	**190,436**	**380,873**				

Table B.14 State Representation in the Baseline and Sample Data Sets for Math Grade 4

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	24.99	22.60	20.96	19.88	**16.04**	-0.244	-0.179	-0.134	-0.105
DC	0.18	0.15	0.14	0.33	**1.38**	0.103	0.106	0.107	0.090
IL	24.50	22.17	24.58	24.32	**24.86**	0.008	0.062	0.007	0.012
MA				0.07	**5.59**				0.241
MD	12.79	12.66	12.97	13.37	**16.35**	0.096	0.100	0.091	0.081
NJ	25.71	31.98	30.84	32.07	**27.10**	0.031	-0.110	-0.084	-0.112
NM	11.82	10.45	10.52	9.75	**6.25**	-0.230	-0.173	-0.176	-0.144
RI				0.21	**2.43**				0.144
Total	**87,299**	**104,759**	**139,679**	**174,599**	**349,198**				

Table B.15 State Representation in the Baseline and Sample Data Sets for Math Grade 3

State	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
CO	20.56	21.92	20.99	22.61	**17.21**	-0.089	-0.125	-0.100	-0.143
DC	0.04	0.07	0.10	0.29	**1.61**	0.124	0.122	0.120	0.105
IL	21.50	23.62	26.05	26.10	**25.43**	0.090	0.042	-0.014	-0.015
MA				0.09	**6.08**				0.251
MD	9.19	9.15	10.64	11.66	**13.37**	0.123	0.124	0.080	0.050
NJ	40.16	36.90	34.79	32.38	**29.44**	-0.235	-0.164	-0.118	-0.065
NM	8.54	8.34	7.43	6.74	**4.65**	-0.185	-0.175	-0.132	-0.099
RI				0.13	**2.21**				0.141
Total	**81,164**	**97,397**	**129,862**	**162,328**	**324,657**				

Appendix C. Demographic Representation for the Baseline and Sample Data Sets

ELA/L

Table C.1 Demographic Representation in the Baseline and Sample Data Sets for ELA/L Grade 11

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	48.48	48.48	48.61	48.38	**48.64**	0.003	0.003	0.000	0.005
Male	51.52	51.52	51.39	51.62	**51.36**	-0.003	-0.003	0.000	-0.005
Ethnicity									
American Indian/ Alaska Native	1.75	1.59	1.36	1.32	**1.49**	-0.022	-0.008	0.011	0.014
Asian	6.33	6.06	6.11	5.95	**5.51**	-0.036	-0.024	-0.027	-0.020
African American	5.97	6.59	6.95	7.78	**10.55**	0.149	0.129	0.117	0.090
Hispanic/Latino	31.77	30.68	29.65	28.44	**26.71**	-0.114	-0.090	-0.067	-0.039
Native Hawaiian/Other Pacific Islander	0.33	0.32	0.37	0.41	**0.42**	0.013	0.014	0.008	0.001
White	48.14	49.51	50.72	51.24	**51.31**	0.063	0.036	0.012	0.001
Two or More Races	1.47	1.42	1.38	1.26	**1.01**	-0.046	-0.042	-0.037	-0.025
Not Provided	4.24	3.84	3.47	3.61	**3.02**	-0.071	-0.048	-0.026	-0.034
Students with Disabilities									
Blank	11.25	13.13	16.97	18.65	**24.05**	0.300	0.256	0.166	0.126
No	79.28	77.41	73.88	71.94	**66.67**	-0.267	-0.228	-0.153	-0.112
Yes	9.48	9.47	9.16	9.41	**9.28**	-0.007	-0.006	0.004	-0.004
Economically Disadvantaged									
Blank	4.02	3.59	3.01	3.18	**4.66**	0.030	0.051	0.078	0.070
No	62.06	62.81	63.69	63.44	**59.63**	-0.049	-0.065	-0.083	-0.078
Yes	33.92	33.60	33.29	33.38	**35.71**	0.037	0.044	0.050	0.049
English Language Learners									
Blank	48.29	48.76	51.34	50.55	**52.28**	0.080	0.071	0.019	0.035
No	50.42	50.09	47.68	48.55	**46.78**	-0.073	-0.066	-0.018	-0.036
Yes	1.28	1.15	0.98	0.90	**0.93**	-0.036	-0.022	-0.005	0.004
Grade Level									
Grade 10	2.07	1.95	1.78	1.94	**2.18**	0.008	0.016	0.028	0.017
Grade 11	97.54	97.62	97.75	97.47	**96.84**	-0.040	-0.045	-0.052	-0.036
Grade 12	0.32	0.35	0.39	0.44	**0.75**	0.050	0.046	0.042	0.036
Total N Count	**45,528**	**54,633**	**72,844**	**91,056**	**182,112**				

Table C.2 Demographic Representation in the Baseline and Sample Data Sets for ELA/L Grade 9

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	49.07	48.98	48.87	48.82	**48.65**	-0.008	-0.007	-0.004	-0.003
Male	50.93	51.02	51.13	51.18	**51.35**	0.008	0.007	0.004	0.003
Ethnicity									
American Indian/ Alaska Native	1.04	1.18	1.39	1.50	**1.40**	0.031	0.019	0.001	-0.008
Asian	6.85	6.43	6.07	6.10	**5.59**	-0.055	-0.037	-0.021	-0.022
African American	8.56	8.47	9.09	10.08	**14.46**	0.168	0.170	0.153	0.125
Hispanic/Latino	28.92	29.20	29.91	29.36	**30.40**	0.032	0.026	0.011	0.023
Native Hawaiian/Other Pacific Islander	0.30	0.45	0.44	0.45	**0.54**	0.034	0.013	0.014	0.012
White	49.12	48.70	47.86	47.35	**43.33**	-0.117	-0.108	-0.091	-0.081
Two or More Races	1.24	1.24	1.21	1.17	**1.04**	-0.020	-0.020	-0.016	-0.013
Not Provided	3.98	4.33	4.03	3.99	**3.23**	-0.042	-0.062	-0.046	-0.043
Students with Disabilities									
Blank	26.00	26.20	26.01	28.46	**31.84**	0.125	0.121	0.125	0.072
No	64.77	64.70	64.97	62.60	**59.21**	-0.113	-0.112	-0.117	-0.069
Yes	9.23	9.10	9.02	8.94	**8.95**	-0.010	-0.005	-0.002	0.000
Economically Disadvantaged									
Blank	3.25	3.61	3.35	3.29	**4.05**	0.041	0.022	0.035	0.038
No	61.03	60.43	59.46	59.35	**54.49**	-0.131	-0.119	-0.100	-0.098
Yes	35.72	35.96	37.19	37.36	**41.47**	0.117	0.112	0.087	0.083
English Language Learners									
Blank	52.44	53.28	54.70	56.22	**54.81**	0.048	0.031	0.002	-0.028
No	46.13	45.18	43.85	42.34	**43.70**	-0.049	-0.030	-0.003	0.027
Yes	1.43	1.53	1.45	1.44	**1.49**	0.005	-0.003	0.004	0.005
Grade Level									
Grade 8	99.23	99.24	99.19	99.08	**98.47**	-0.062	-0.063	-0.059	-0.050
Grade 9	0.60	0.59	0.62	0.68	**1.02**	0.041	0.043	0.040	0.033
Total N Count	**66,863**	**80,235**	**106,980**	**133,726**	**267,452**				

Table C.3 Demographic Representation in the Baseline and Sample Data Sets for ELA/L Grade 8

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	48.51	48.52	48.67	48.70	**48.81**	0.006	0.006	0.003	0.002
Male	51.49	51.48	51.33	51.30	**51.19**	-0.006	-0.006	-0.003	-0.002
Ethnicity									
American Indian/ Alaska Native	1.15	1.16	1.01	0.95	**0.93**	-0.023	-0.024	-0.008	-0.002
Asian	6.51	6.54	6.46	6.30	**5.62**	-0.039	-0.040	-0.036	-0.029
African American	11.97	12.07	12.96	13.84	**16.57**	0.124	0.121	0.097	0.073
Hispanic/Latino	26.04	26.29	26.49	26.23	**25.71**	-0.008	-0.013	-0.018	-0.012
Native Hawaiian/Other Pacific Islander	0.28	0.31	0.33	0.38	**0.41**	0.021	0.016	0.012	0.005
White	49.73	49.49	48.47	48.26	**47.25**	-0.050	-0.045	-0.024	-0.020
Two or More Races	1.39	1.33	1.33	1.36	**1.20**	-0.017	-0.012	-0.012	-0.014
Not Provided	2.93	2.81	2.95	2.68	**2.29**	-0.043	-0.035	-0.044	-0.026
Students with Disabilities									
Blank	27.10	29.91	32.73	34.71	**35.80**	0.181	0.123	0.064	0.023
No	63.90	61.44	59.17	57.42	**55.78**	-0.164	-0.114	-0.068	-0.033
Yes	9.00	8.65	8.10	7.87	**8.42**	-0.021	-0.008	0.012	0.020
Economically Disadvantaged									
Blank	2.11	1.90	2.01	1.74	**5.18**	0.139	0.148	0.143	0.155
No	59.54	58.59	57.39	56.77	**50.45**	-0.182	-0.163	-0.139	-0.126
Yes	38.34	39.51	40.60	41.49	**44.37**	0.121	0.098	0.076	0.058
English Language Learners									
Blank	35.57	37.79	40.49	42.58	**40.95**	0.109	0.064	0.009	-0.033
No	63.25	60.88	57.91	55.78	**57.51**	-0.116	-0.068	-0.008	0.035
Yes	1.18	1.33	1.60	1.64	**1.54**	0.029	0.017	-0.005	-0.009
Total N Count	**94,790**	**113,748**	**151,664**	**189,581**	**379,162**				

Table C.4 Demographic Representation in the Baseline and Sample Data Sets for ELA/L Grade 7

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	48.73	48.83	48.87	48.91	**48.84**	0.002	0.000	-0.001	-0.001
Male	51.27	51.17	51.13	51.09	**51.16**	-0.002	0.000	0.001	0.001
Ethnicity									
American Indian/ Alaska Native	1.47	1.36	1.14	1.02	**0.97**	-0.051	-0.040	-0.017	-0.005
Asian	6.10	6.29	6.40	6.59	**5.66**	-0.019	-0.027	-0.032	-0.040
African American	11.71	11.83	12.44	13.36	**16.34**	0.125	0.122	0.106	0.081
Hispanic/Latino	30.09	29.32	29.18	27.79	**26.17**	-0.089	-0.072	-0.069	-0.037
Native Hawaiian/Other Pacific Islander	0.20	0.23	0.29	0.27	**0.23**	0.006	0.001	-0.012	-0.008
White	47.45	47.97	47.52	47.83	**47.50**	0.001	-0.009	0.000	-0.007
Two or More Races	1.55	1.60	1.57	1.51	**1.37**	-0.015	-0.019	-0.017	-0.012
Not Provided	1.44	1.40	1.46	1.64	**1.76**	0.024	0.027	0.023	0.009
Students with Disabilities									
Blank	21.90	22.39	24.53	28.41	**33.75**	0.251	0.240	0.195	0.113
No	68.30	67.99	66.23	62.75	**57.64**	-0.216	-0.209	-0.174	-0.103
Yes	9.81	9.63	9.24	8.84	**8.61**	-0.043	-0.036	-0.022	-0.008
Economically Disadvantaged									
Blank	0.84	0.79	0.78	0.82	**4.53**	0.177	0.180	0.180	0.178
No	56.51	57.02	56.56	56.88	**50.24**	-0.125	-0.136	-0.126	-0.133
Yes	42.65	42.19	42.67	42.30	**45.23**	0.052	0.061	0.052	0.059
English Language Learners									
Blank	31.99	32.53	34.56	37.52	**40.43**	0.172	0.161	0.120	0.059
No	66.41	65.85	63.83	61.04	**58.36**	-0.163	-0.152	-0.111	-0.054
Yes	1.60	1.61	1.61	1.44	**1.21**	-0.036	-0.037	-0.037	-0.021
Total N Count	**95,867**	**115,041**	**153,388**	**191,735**	**383,470**				

Table C.5 Demographic Representation in the Baseline and Sample Data Sets for ELA/L Grade 5

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	48.91	48.99	49.05	48.95	**48.86**	-0.001	-0.003	-0.004	-0.002
Male	51.09	51.01	50.95	51.05	**51.14**	0.001	0.003	0.004	0.002
Ethnicity									
American Indian/ Alaska Native	0.86	0.81	0.85	0.80	**0.95**	0.009	0.015	0.010	0.015
Asian	7.90	7.77	7.19	6.91	**6.09**	-0.076	-0.070	-0.046	-0.035
African American	12.08	12.34	12.58	12.67	**14.45**	0.067	0.060	0.053	0.051
Hispanic/Latino	26.90	26.34	26.19	26.36	**25.63**	-0.029	-0.016	-0.013	-0.017
Native Hawaiian/Other Pacific Islander	0.17	0.17	0.17	0.17	**0.17**	0.000	0.000	0.000	0.000
White	48.96	49.22	49.49	49.51	**49.27**	0.006	0.001	-0.004	-0.005
Two or More Races	1.99	2.01	2.01	1.95	**1.66**	-0.026	-0.028	-0.028	-0.023
Not Provided	1.14	1.34	1.51	1.62	**1.79**	0.049	0.034	0.021	0.013
Students with Disabilities									
Blank	17.02	19.94	23.16	26.26	**28.65**	0.257	0.193	0.121	0.053
No	72.13	69.77	67.31	64.62	**61.91**	-0.211	-0.162	-0.111	-0.056
Yes	10.84	10.29	9.53	9.12	**9.44**	-0.048	-0.029	-0.003	0.011
Economically Disadvantaged									
Blank	0.50	0.57	0.65	0.66	**4.19**	0.184	0.181	0.177	0.176
No	59.62	59.25	58.64	57.58	**51.83**	-0.156	-0.149	-0.136	-0.115
Yes	39.88	40.17	40.71	41.75	**43.97**	0.082	0.077	0.066	0.045
English Language Learners									
Blank	24.89	27.32	32.31	35.68	**36.72**	0.245	0.195	0.091	0.022
No	73.64	71.34	66.41	63.10	**61.95**	-0.241	-0.193	-0.092	-0.024
Yes	1.47	1.34	1.28	1.23	**1.33**	-0.012	-0.001	0.004	0.009
Total N Count	**92,573**	**111,088**	**148,118**	**185,147**	**370,295**				

Table C.6 Demographic Representation in the Baseline and Sample Data Sets for ELA/L Grade 4

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	49.08	49.16	49.14	49.06	**48.90**	-0.004	-0.005	-0.005	-0.003
Male	50.92	50.84	50.86	50.94	**51.10**	0.004	0.005	0.005	0.003
Ethnicity									
American Indian/ Alaska Native	1.22	1.71	1.47	1.33	**0.95**	-0.028	-0.078	-0.054	-0.039
Asian	7.04	6.66	6.91	6.66	**6.02**	-0.043	-0.027	-0.037	-0.027
African American	13.31	13.32	13.74	13.68	**14.72**	0.040	0.040	0.028	0.029
Hispanic/Latino	30.38	30.42	28.96	28.64	**26.41**	-0.090	-0.091	-0.058	-0.051
Native Hawaiian/Other Pacific Islander	0.18	0.18	0.20	0.20	**0.18**	0.000	0.000	-0.004	-0.004
White	44.65	44.48	45.38	46.03	**48.14**	0.070	0.073	0.055	0.042
Two or More Races	1.89	1.86	1.86	1.88	**1.83**	-0.005	-0.003	-0.003	-0.004
Not Provided	1.32	1.38	1.48	1.58	**1.75**	0.033	0.029	0.021	0.013
Students with Disabilities									
Blank	17.47	19.34	21.34	23.67	**26.30**	0.201	0.158	0.113	0.060
No	72.20	70.84	69.22	67.12	**64.30**	-0.165	-0.137	-0.103	-0.059
Yes	10.32	9.82	9.44	9.21	**9.40**	-0.032	-0.014	-0.001	0.007
Economically Disadvantaged									
Blank	0.62	0.62	0.66	0.69	**4.25**	0.180	0.180	0.178	0.176
No	53.11	52.08	53.46	53.91	**50.74**	-0.047	-0.027	-0.054	-0.063
Yes	46.27	47.29	45.89	45.40	**45.01**	-0.025	-0.046	-0.018	-0.008
English Language Learners									
Blank	25.30	28.04	31.26	34.01	**34.90**	0.201	0.144	0.076	0.019
No	72.46	69.48	66.48	63.89	**63.50**	-0.186	-0.124	-0.062	-0.008
Yes	2.24	2.49	2.25	2.11	**1.60**	-0.051	-0.070	-0.052	-0.040
Total N Count	**89,046**	**106,855**	**142,473**	**178,092**	**356,184**				

Table C.7 Demographic State Representation in the Baseline and Sample Data Sets for Algebra 1

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	48.05	48.09	48.36	48.24	**48.06**	0.000	-0.001	-0.006	-0.003
Male	51.95	51.91	51.64	51.76	**51.94**	0.000	0.001	0.006	0.003
Ethnicity									
American Indian/ Alaska Native	0.82	0.90	1.26	1.33	**1.24**	0.038	0.031	-0.001	-0.008
Asian	6.70	6.48	6.60	6.36	**5.85**	-0.036	-0.027	-0.032	-0.022
African American	15.97	16.41	16.63	17.17	**19.93**	0.099	0.088	0.082	0.069
Hispanic/Latino	29.32	29.24	28.26	29.19	**27.85**	-0.033	-0.031	-0.009	-0.030
Native Hawaiian/Other Pacific Islander	0.33	0.37	0.44	0.47	**0.37**	0.006	-0.001	-0.012	-0.018
White	43.98	43.59	43.86	42.25	**41.60**	-0.048	-0.041	-0.046	-0.013
Two or More Races	1.80	1.82	1.80	1.74	**1.55**	-0.020	-0.022	-0.020	-0.015
Not Provided	1.08	1.19	1.15	1.48	**1.61**	0.042	0.034	0.037	0.010
Students with Disabilities									
Blank	13.76	14.50	15.49	16.71	**20.59**	0.169	0.151	0.126	0.096
No	75.62	75.27	74.53	73.41	**69.44**	-0.134	-0.127	-0.111	-0.086
Yes	10.62	10.23	9.98	9.87	**9.97**	-0.022	-0.009	0.000	0.003
Economically Disadvantaged									
Blank	0.62	0.76	0.79	0.96	**2.94**	0.137	0.129	0.127	0.117
No	60.12	59.46	59.65	57.92	**53.89**	-0.125	-0.112	-0.115	-0.081
Yes	39.26	39.79	39.57	41.12	**43.17**	0.079	0.068	0.073	0.041
English Language Learners									
Blank	41.02	41.50	41.29	41.44	**38.68**	-0.048	-0.058	-0.054	-0.057
No	57.96	57.51	57.47	57.09	**59.96**	0.041	0.050	0.051	0.059
Yes	1.01	1.00	1.24	1.47	**1.36**	0.029	0.031	0.010	-0.010
Grade Level									
Grade 7	3.13	3.57	4.52	4.13	**3.59**	0.025	0.001	-0.050	-0.029
Grade 8	20.65	20.63	21.07	20.87	**21.16**	0.012	0.013	0.002	0.007
Grade 9	65.88	65.52	64.85	65.52	**65.91**	0.001	0.008	0.022	0.008
Grade 10	8.86	8.78	8.10	7.95	**7.54**	-0.050	-0.047	-0.021	-0.015
Grade 11	1.32	1.33	1.29	1.32	**1.43**	0.009	0.008	0.012	0.009
Total N Count	**74,689**	**89,626**	**119,502**	**149,378**	**298,756**				

Table C.8 Demographic State Representation in the Baseline and Sample Data Sets for Geometry

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	48.76	48.87	48.96	49.00	**48.79**	0.001	-0.001	-0.003	-0.004
Male	51.24	51.13	51.04	51.00	**51.21**	-0.001	0.001	0.003	0.004
Ethnicity									
American Indian/ Alaska Native	1.38	1.57	1.60	1.54	**1.82**	0.033	0.019	0.016	0.021
Asian	6.39	6.32	6.62	6.63	**6.55**	0.006	0.009	-0.003	-0.003
African American	8.07	7.72	8.52	8.49	**11.71**	0.113	0.124	0.099	0.100
Hispanic/Latino	31.74	31.11	31.37	31.19	**28.67**	-0.068	-0.054	-0.060	-0.056
Native Hawaiian/Other Pacific Islander	0.48	0.75	1.07	0.98	**0.77**	0.034	0.003	-0.034	-0.024
White	49.45	49.80	48.05	48.48	**47.14**	-0.046	-0.053	-0.018	-0.027
Two or More Races	1.44	1.38	1.41	1.42	**1.35**	-0.008	-0.002	-0.005	-0.007
Not Provided	1.05	1.35	1.36	1.26	**2.00**	0.068	0.046	0.046	0.053
Students with Disabilities									
Blank	5.15	5.83	5.89	6.40	**7.27**	0.082	0.056	0.053	0.033
No	84.99	84.46	84.06	83.44	**81.64**	-0.086	-0.073	-0.062	-0.047
Yes	9.86	9.71	10.05	10.15	**11.09**	0.039	0.044	0.033	0.030
Economically Disadvantaged									
Blank	0.85	1.26	1.25	1.13	**2.18**	0.091	0.063	0.063	0.072
No	64.24	64.48	63.47	64.09	**61.90**	-0.048	-0.053	-0.032	-0.045
Yes	34.91	34.26	35.28	34.79	**35.92**	0.021	0.035	0.013	0.024
English Language Learners									
Blank	41.17	40.05	39.72	39.51	**33.68**	-0.159	-0.135	-0.128	-0.123
No	57.67	58.76	59.02	59.11	**64.73**	0.148	0.125	0.119	0.118
Yes	1.15	1.19	1.25	1.38	**1.59**	0.034	0.032	0.027	0.017
Grade Level									
Grade 8	4.39	4.56	3.99	4.20	**3.61**	-0.042	-0.051	-0.020	-0.031
Grade 9	28.60	28.62	28.68	29.33	**27.82**	-0.017	-0.018	-0.019	-0.034
Grade 10	54.94	55.13	56.25	56.08	**57.89**	0.060	0.056	0.033	0.037
Grade 11	12.00	11.60	10.99	10.29	**10.48**	-0.049	-0.037	-0.017	0.006
Total N Count	**38,011**	**45,613**	**60,817**	**76,022**	**152,044**				

Table C.9 Demographic Representation in the Baseline and Sample Data Sets for Integrated Math III

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	48.11	47.90	48.29	47.89	**49.01**	0.018	0.022	0.015	0.023
Male	51.89	52.10	51.71	52.11	**50.99**	-0.018	-0.022	-0.015	-0.023
Ethnicity									
American Indian/ Alaska Native	2.06	1.78	1.64	1.50	**1.85**	-0.016	0.005	0.015	0.025
Asian	1.22	1.40	2.19	2.48	**2.74**	0.093	0.082	0.034	0.016
African American	1.49	1.33	2.76	2.48	**3.33**	0.103	0.111	0.032	0.048
Hispanic/Latino	15.96	16.29	22.24	22.55	**23.52**	0.178	0.171	0.030	0.023
Native Hawaiian/Other Pacific Islander	0.08	0.06	1.07	0.91	**0.57**	0.066	0.067	-0.066	-0.045
White	24.84	25.02	28.38	36.21	**45.22**	0.410	0.406	0.338	0.181
Two or More Races	1.10	1.08	1.10	1.03	**1.23**	0.011	0.014	0.012	0.018
Not Provided	53.26	53.05	40.62	32.84	**21.53**	-0.772	-0.767	-0.464	-0.275
Students with Disabilities									
Blank	50.40	50.76	40.00	35.49	**29.79**	-0.451	-0.459	-0.223	-0.125
No	42.97	43.24	53.90	59.09	**64.99**	0.462	0.456	0.232	0.124
Yes	6.63	6.00	6.10	5.43	**5.22**	-0.063	-0.035	-0.039	-0.009
Economically Disadvantaged									
Blank	53.41	53.17	40.71	32.88	**22.87**	-0.727	-0.721	-0.425	-0.238
No	31.39	31.84	39.52	45.26	**50.09**	0.374	0.365	0.211	0.097
Yes	15.20	14.98	19.76	21.87	**27.04**	0.266	0.271	0.164	0.116
English Language Learners									
Blank	96.11	96.70	97.36	96.63	**94.80**	-0.059	-0.085	-0.115	-0.082
No	3.58	3.05	2.45	3.22	**4.91**	0.062	0.086	0.114	0.078
Yes	0.30	0.25	0.19	0.15	**0.29**	-0.004	0.006	0.018	0.025
Grade Level									
Grade 9	0.76	0.63	0.76	2.13	**2.25**	0.100	0.109	0.100	0.008
Grade 10	28.65	27.21	21.67	22.40	**23.41**	-0.124	-0.090	0.041	0.024
Grade 11	66.55	68.19	73.64	71.50	**68.73**	0.047	0.012	-0.106	-0.060
Grade 12	4.04	3.97	3.93	3.96	**5.62**	0.069	0.072	0.073	0.072
Total N Count	**2,625**	**3,150**	**4,200**	**5,250**	**10,501**				

Table C.10 Demographic Representation in the Baseline and Sample Data Sets for Integrated Math I

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	47.57	47.91	48.28	48.33	**48.14**	0.012	0.005	-0.003	-0.004
Male	52.43	52.09	51.72	51.67	**51.86**	-0.012	-0.005	0.003	0.004
Ethnicity									
American Indian/ Alaska Native	2.05	2.17	1.87	1.60	**1.37**	-0.058	-0.070	-0.044	-0.021
Asian	1.83	1.73	2.25	2.15	**2.32**	0.032	0.039	0.005	0.011
African American	9.31	8.04	7.36	7.15	**8.06**	-0.046	0.001	0.026	0.033
Hispanic/Latino	27.03	26.01	29.29	32.92	**36.99**	0.206	0.227	0.159	0.084
Native Hawaiian/Other Pacific Islander	0.12	0.13	0.79	0.74	**0.60**	0.062	0.061	-0.025	-0.018
White	34.29	35.23	37.05	36.23	**37.98**	0.076	0.057	0.019	0.036
Two or More Races	0.74	0.71	0.91	0.91	**0.89**	0.016	0.020	-0.002	-0.001
Not Provided	24.63	25.98	20.47	18.29	**11.80**	-0.398	-0.440	-0.269	-0.201
Students with Disabilities									
Blank	52.96	51.04	45.48	41.56	**38.66**	-0.294	-0.254	-0.140	-0.059
No	37.80	39.74	45.44	48.84	**52.26**	0.290	0.251	0.137	0.069
Yes	9.24	9.21	9.08	9.61	**9.07**	-0.006	-0.005	0.000	-0.019
Economically Disadvantaged									
Blank	23.57	25.07	19.61	17.52	**12.27**	-0.344	-0.390	-0.224	-0.160
No	36.74	37.03	41.13	40.80	**43.01**	0.127	0.121	0.038	0.045
Yes	39.69	37.91	39.25	41.69	**44.72**	0.101	0.137	0.110	0.061
English Language Learners									
Blank	96.14	96.27	96.63	97.03	**95.45**	-0.033	-0.039	-0.057	-0.076
No	3.49	3.32	3.00	2.67	**4.27**	0.038	0.047	0.063	0.079
Yes	0.37	0.41	0.37	0.30	**0.29**	-0.016	-0.023	-0.016	-0.002
Grade Level									
Grade 8	4.14	6.16	7.79	8.16	**6.24**	0.087	0.003	-0.064	-0.079
Grade 9	86.72	84.85	82.48	82.35	**80.93**	-0.148	-0.100	-0.039	-0.036
Grade 10	8.06	7.76	7.83	7.50	**9.19**	0.039	0.050	0.047	0.058
Grade 11	0.76	0.93	1.36	1.33	**2.92**	0.128	0.118	0.093	0.094
Total N Count	**5,672**	**6,806**	**9,075**	**11,344**	**22,689**				

Table C.11 Demographic Representation in the Baseline and Sample Data Sets for Math Grade 8

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	48.00	48.10	48.20	48.16	**48.28**	0.006	0.004	0.002	0.002
Male	52.00	51.90	51.80	51.84	**51.72**	-0.006	-0.004	-0.002	-0.002
Ethnicity									
American Indian/ Alaska Native	1.29	1.40	1.22	1.17	**1.06**	-0.022	-0.033	-0.016	-0.011
Asian	4.89	4.89	4.77	4.72	**4.16**	-0.036	-0.036	-0.031	-0.028
African American	13.59	14.11	14.61	14.67	**18.50**	0.126	0.113	0.100	0.098
Hispanic/Latino	25.57	25.08	26.53	27.11	**27.49**	0.043	0.054	0.022	0.009
Native Hawaiian/Other Pacific Islander	0.22	0.21	0.22	0.26	**0.42**	0.032	0.032	0.031	0.024
White	49.68	49.61	48.44	48.06	**45.05**	-0.093	-0.092	-0.068	-0.060
Two or More Races	1.15	1.11	1.04	1.00	**0.99**	-0.016	-0.012	-0.005	-0.002
Not Provided	3.63	3.59	3.17	3.00	**2.32**	-0.086	-0.084	-0.056	-0.045
Students with Disabilities									
Blank	38.92	40.54	45.11	45.81	**43.09**	0.084	0.051	-0.041	-0.055
No	50.87	49.55	45.88	45.07	**46.96**	-0.078	-0.052	0.022	0.038
Yes	10.21	9.91	9.02	9.12	**9.96**	-0.008	0.001	0.031	0.028
Economically Disadvantaged									
Blank	2.33	2.23	1.75	1.78	**5.16**	0.128	0.132	0.154	0.153
No	54.59	54.43	52.89	52.39	**45.35**	-0.186	-0.182	-0.152	-0.142
Yes	43.08	43.34	45.35	45.83	**49.49**	0.128	0.123	0.083	0.073
English Language Learners									
Blank	46.93	49.08	52.72	52.36	**47.42**	0.010	-0.033	-0.106	-0.099
No	51.69	49.52	45.76	45.98	**50.73**	-0.019	0.024	0.099	0.095
Yes	1.38	1.40	1.52	1.66	**1.85**	0.035	0.033	0.025	0.014
Total N Count	**76,673**	**92,007**	**122,677**	**153,346**	**306,693**				

Table C.12 Demographic Representation in the Baseline and Sample Data Sets for Math Grade 7

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	48.86	48.75	49.02	48.91	**48.87**	0.000	0.002	-0.003	-0.001
Male	51.14	51.25	50.98	51.09	**51.13**	0.000	-0.002	0.003	0.001
Ethnicity									
American Indian/ Alaska Native	1.36	1.46	1.20	1.08	**0.99**	-0.037	-0.047	-0.021	-0.009
Asian	5.06	5.13	5.53	5.61	**5.16**	0.005	0.001	-0.017	-0.020
African American	12.20	12.17	13.30	13.94	**16.98**	0.127	0.128	0.098	0.081
Hispanic/Latino	29.68	29.16	28.36	27.63	**26.28**	-0.077	-0.065	-0.047	-0.031
Native Hawaiian/Other Pacific Islander	0.18	0.19	0.27	0.26	**0.23**	0.011	0.009	-0.007	-0.006
White	48.14	48.61	48.32	48.32	**47.36**	-0.016	-0.025	-0.019	-0.019
Two or More Races	1.33	1.37	1.29	1.34	**1.31**	-0.002	-0.005	0.002	-0.003
Not Provided	2.06	1.91	1.73	1.81	**1.69**	-0.029	-0.017	-0.003	-0.009
Students with Disabilities									
Blank	30.69	31.04	32.41	32.50	**34.76**	0.086	0.078	0.049	0.048
No	60.55	60.16	59.03	58.82	**56.37**	-0.084	-0.077	-0.054	-0.049
Yes	8.77	8.79	8.56	8.68	**8.87**	0.004	0.003	0.011	0.007
Economically Disadvantaged									
Blank	0.95	0.89	0.70	0.64	**4.57**	0.173	0.176	0.185	0.188
No	53.70	54.52	54.74	55.59	**49.14**	-0.091	-0.108	-0.112	-0.129
Yes	45.35	44.58	44.56	43.77	**46.30**	0.019	0.034	0.035	0.051
English Language Learners									
Blank	39.98	41.76	41.64	40.62	**40.66**	0.014	-0.022	-0.020	0.001
No	58.44	56.69	57.07	58.11	**58.13**	-0.006	0.029	0.021	0.000
Yes	1.58	1.55	1.29	1.27	**1.21**	-0.034	-0.031	-0.007	-0.006
Total N Count	**92,507**	**111,008**	**148,011**	**185,014**	**370,028**				

Table C.13 Demographic Representation in the Baseline and Sample Data Sets for Math Grade 6

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	48.91	48.87	48.92	48.94	**48.92**	0.000	0.001	0.000	0.000
Male	51.09	51.13	51.08	51.06	**51.08**	0.000	-0.001	0.000	0.000
Ethnicity									
American Indian/ Alaska Native	1.06	1.30	1.22	1.11	**0.92**	-0.015	-0.040	-0.032	-0.020
Asian	6.51	6.22	6.10	6.10	**5.98**	-0.022	-0.010	-0.005	-0.005
African American	12.58	13.17	13.74	13.86	**15.65**	0.085	0.068	0.053	0.049
Hispanic/Latino	29.28	28.68	28.49	28.21	**26.66**	-0.059	-0.046	-0.041	-0.035
Native Hawaiian/Other Pacific Islander	0.18	0.18	0.18	0.17	**0.17**	-0.002	-0.003	-0.003	-0.001
White	47.41	47.40	47.20	47.30	**47.51**	0.002	0.002	0.006	0.004
Two or More Races	1.63	1.58	1.58	1.57	**1.51**	-0.010	-0.006	-0.006	-0.005
Not Provided	1.36	1.47	1.49	1.69	**1.60**	0.019	0.011	0.009	-0.007
Students with Disabilities									
Blank	27.38	30.73	32.20	34.36	**34.21**	0.144	0.073	0.042	-0.003
No	63.54	60.52	59.41	57.44	**57.17**	-0.129	-0.068	-0.045	-0.006
Yes	9.08	8.75	8.39	8.19	**8.62**	-0.016	-0.005	0.008	0.015
Economically Disadvantaged									
Blank	0.47	0.42	0.47	0.74	**4.29**	0.189	0.191	0.188	0.175
No	55.98	55.17	54.60	54.28	**49.85**	-0.123	-0.107	-0.095	-0.089
Yes	43.55	44.41	44.93	44.98	**45.86**	0.046	0.029	0.019	0.018
English Language Learners									
Blank	35.06	38.76	40.17	42.80	**41.26**	0.126	0.051	0.022	-0.031
No	63.78	59.93	58.60	56.02	**57.67**	-0.124	-0.046	-0.019	0.034
Yes	1.16	1.32	1.23	1.18	**1.07**	-0.009	-0.024	-0.015	-0.011
Total N Count	**95,218**	**114,261**	**152,349**	**190,436**	**380,873**				

Table C.14 Demographic Representation in the Baseline and Sample Data Sets for Math Grade 4

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	49.04	48.76	48.97	48.93	**48.88**	-0.003	0.002	-0.002	-0.001
Male	50.96	51.24	51.03	51.07	**51.12**	0.003	-0.002	0.002	0.001
Ethnicity									
American Indian/ Alaska Native	0.86	0.88	1.31	1.26	**0.96**	0.010	0.008	-0.036	-0.031
Asian	5.55	6.24	6.44	6.53	**6.16**	0.025	-0.003	-0.012	-0.015
African American	13.04	13.57	13.92	13.97	**15.03**	0.056	0.041	0.031	0.030
Hispanic/Latino	30.93	29.84	29.23	28.57	**26.08**	-0.110	-0.086	-0.072	-0.057
Native Hawaiian/Other Pacific Islander	0.16	0.16	0.17	0.18	**0.18**	0.004	0.004	0.003	0.000
White	46.16	46.05	45.79	46.28	**48.16**	0.040	0.042	0.047	0.038
Two or More Races	1.68	1.70	1.66	1.69	**1.78**	0.008	0.006	0.009	0.007
Not Provided	1.63	1.55	1.48	1.52	**1.65**	0.001	0.008	0.013	0.010
Students with Disabilities									
Blank	25.97	23.70	25.92	25.63	**26.59**	0.014	0.065	0.015	0.022
No	64.89	66.60	64.86	65.06	**63.95**	-0.020	-0.055	-0.019	-0.023
Yes	9.14	9.69	9.22	9.31	**9.47**	0.011	-0.008	0.009	0.005
Economically Disadvantaged									
Blank	0.62	0.64	0.51	0.48	**4.20**	0.178	0.177	0.184	0.185
No	51.51	52.55	52.48	53.28	**50.68**	-0.017	-0.037	-0.036	-0.052
Yes	47.87	46.81	47.01	46.24	**45.12**	-0.055	-0.034	-0.038	-0.022
English Language Learners									
Blank	36.65	33.37	35.33	34.99	**34.68**	-0.041	0.027	-0.014	-0.006
No	61.29	64.76	62.71	63.14	**63.70**	0.050	-0.022	0.021	0.012
Yes	2.06	1.87	1.96	1.87	**1.62**	-0.035	-0.020	-0.027	-0.020
Total N Count	**87,299**	**104,759**	**139,679**	**174,599**	**349,198**				

Table C.15 Demographic Representation in the Baseline and Sample Data Sets for Math Grade 3

Gender	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Female	49.12	49.13	49.17	49.05	**48.92**	-0.004	-0.004	-0.005	-0.003
Male	50.88	50.87	50.83	50.95	**51.08**	0.004	0.004	0.005	0.003
Ethnicity									
American Indian/ Alaska Native	1.51	1.44	1.21	1.11	**0.96**	-0.056	-0.049	-0.026	-0.016
Asian	6.60	6.38	6.86	6.56	**6.34**	-0.011	-0.002	-0.021	-0.009
African American	14.02	13.36	12.88	12.69	**13.99**	-0.001	0.018	0.032	0.038
Hispanic/Latino	28.62	28.70	28.27	28.03	**26.17**	-0.056	-0.058	-0.048	-0.042
Native Hawaiian/Other Pacific Islander	0.22	0.22	0.21	0.21	**0.19**	-0.008	-0.007	-0.006	-0.006
White	45.66	46.52	47.14	47.87	**48.68**	0.060	0.043	0.031	0.016
Two or More Races	1.87	1.86	1.90	1.97	**1.92**	0.004	0.005	0.002	-0.003
Not Provided	1.50	1.53	1.52	1.55	**1.76**	0.020	0.017	0.018	0.015
Students with Disabilities									
Blank	23.22	25.20	27.38	27.41	**27.28**	0.091	0.047	-0.002	-0.003
No	67.51	65.86	64.05	64.15	**64.04**	-0.072	-0.038	0.000	-0.002
Yes	9.27	8.94	8.57	8.44	**8.68**	-0.021	-0.009	0.004	0.008
Economically Disadvantaged									
Blank	0.65	0.61	0.55	0.57	**4.44**	0.184	0.186	0.189	0.188
No	52.74	52.62	53.37	53.85	**50.72**	-0.040	-0.038	-0.053	-0.063
Yes	46.62	46.77	46.07	45.58	**44.84**	-0.036	-0.039	-0.025	-0.015
English Language Learners									
Blank	31.22	33.77	36.03	38.20	**35.98**	0.099	0.046	-0.001	-0.046
No	66.92	64.34	62.20	60.06	**62.41**	-0.093	-0.040	0.004	0.048
Yes	1.87	1.89	1.77	1.74	**1.61**	-0.021	-0.022	-0.013	-0.011
Total N Count	**81,164**	**97,397**	**129,862**	**162,328**	**324,657**				

Appendix D. Spring 2015 Summative Scale Score Cumulative Percent for the Baseline and Sample Data Sets

ELA/L

Table D.1 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for ELA/L Grade 11

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	37,219	3.253	2.340	0.000	6.851
Sample 30%	45,190	3.015	2.185	0.000	6.409
Sample 40%	61,138	2.846	2.009	0.000	5.929
Sample 50%	77,035	2.323	1.629	0.000	4.883

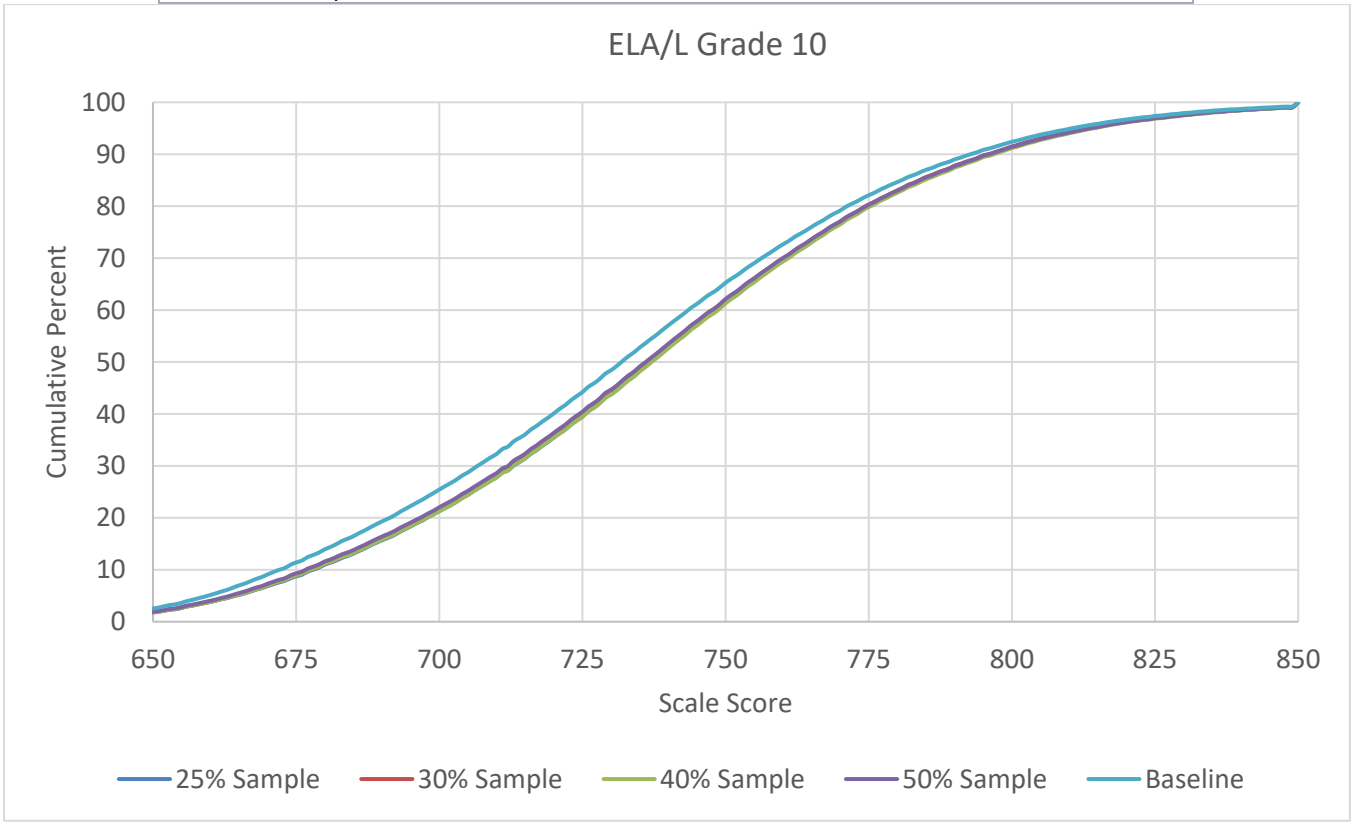

Figure D.1 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for ELA/L Grade 11

Table D.2 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for ELA/L Grade 9

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	59,745	3.306	2.786	0.000	7.807
Sample 30%	71,998	3.160	2.616	0.000	7.335
Sample 40%	96,324	2.717	2.285	0.000	6.394
Sample 50%	120,639	2.416	1.998	0.000	5.554

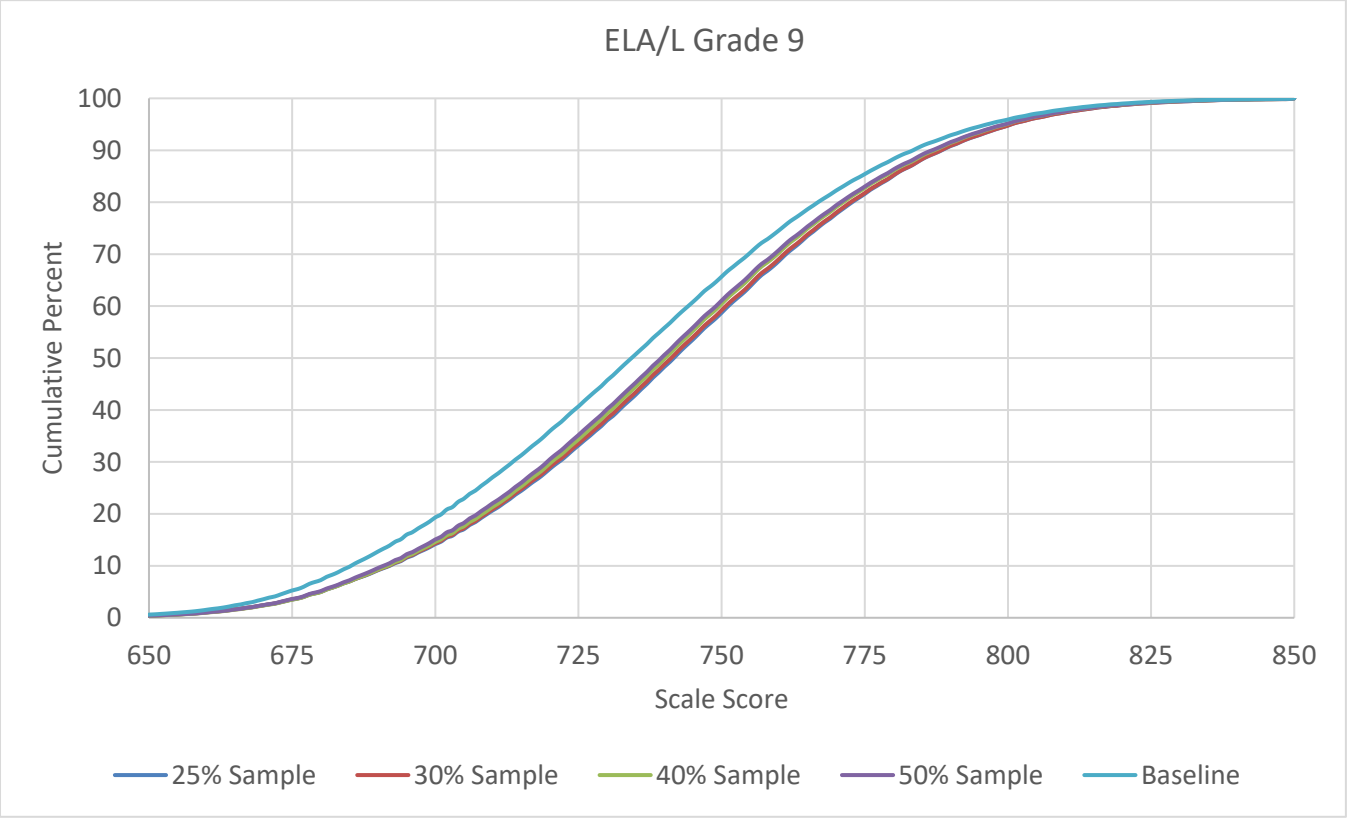

Figure D.2 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for ELA/L Grade 9

Table D.3 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for ELA/L Grade 8

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	91,335	1.443	1.024	0.000	3.005
Sample 30%	109,799	1.301	0.933	0.000	2.733
Sample 40%	146,410	1.284	0.943	0.000	2.737
Sample 50%	182,623	0.928	0.707	0.000	2.023

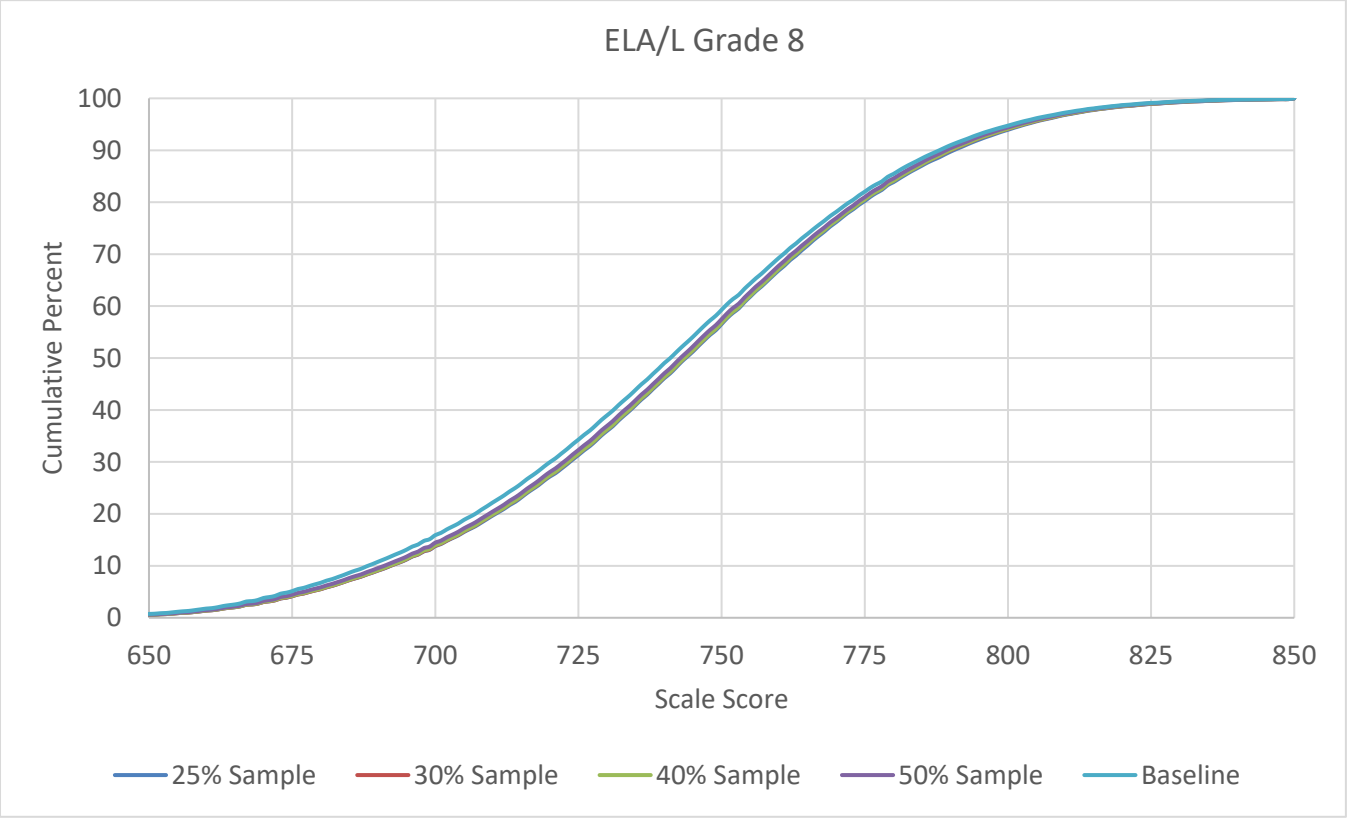

Figure D.3 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for ELA/L Grade 8

Table D.4 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for ELA/L Grade 7

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	93,205	0.407	0.228	0.000	0.814
Sample 30%	111,908	0.650	0.396	0.000	1.274
Sample 40%	149,160	0.925	0.619	0.000	1.892
Sample 50%	186,700	1.015	0.724	0.000	2.127

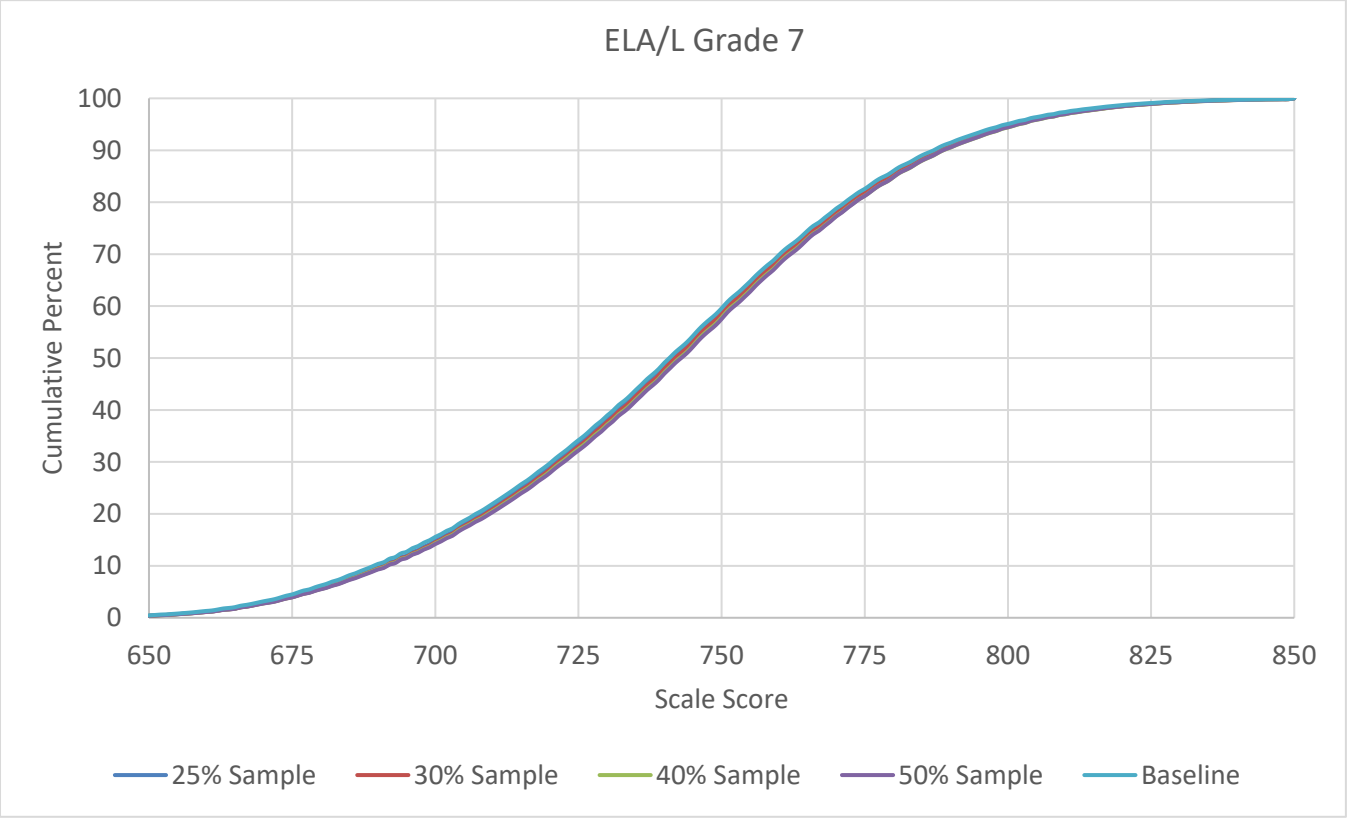

Figure D.4 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for ELA/L Grade 7

Table D.5 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for ELA/L Grade 5

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	91,179	1.451	1.341	0.000	3.799
Sample 30%	109,519	1.455	1.319	0.000	3.691
Sample 40%	146,148	1.185	1.067	0.000	2.982
Sample 50%	182,735	0.831	0.736	0.000	2.037

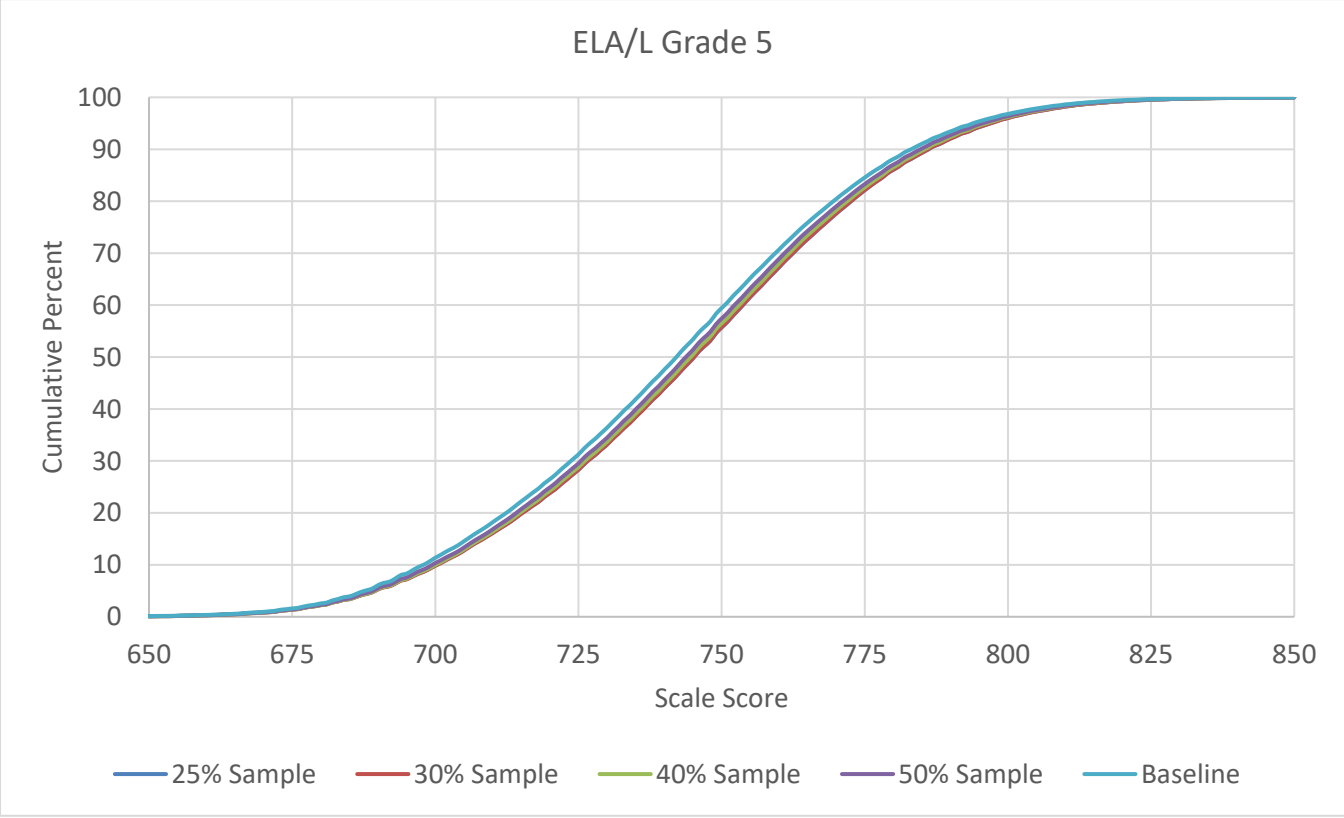

Figure D.5 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for ELA/L Grade 5

Table D.6 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for ELA/L Grade 4

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	87,792	-0.066	0.199	-0.554	0.232
Sample 30%	105,414	-0.232	0.330	-0.943	0.068
Sample 40%	140,654	0.156	0.098	0.000	0.375
Sample 50%	175,931	0.222	0.153	0.000	0.481

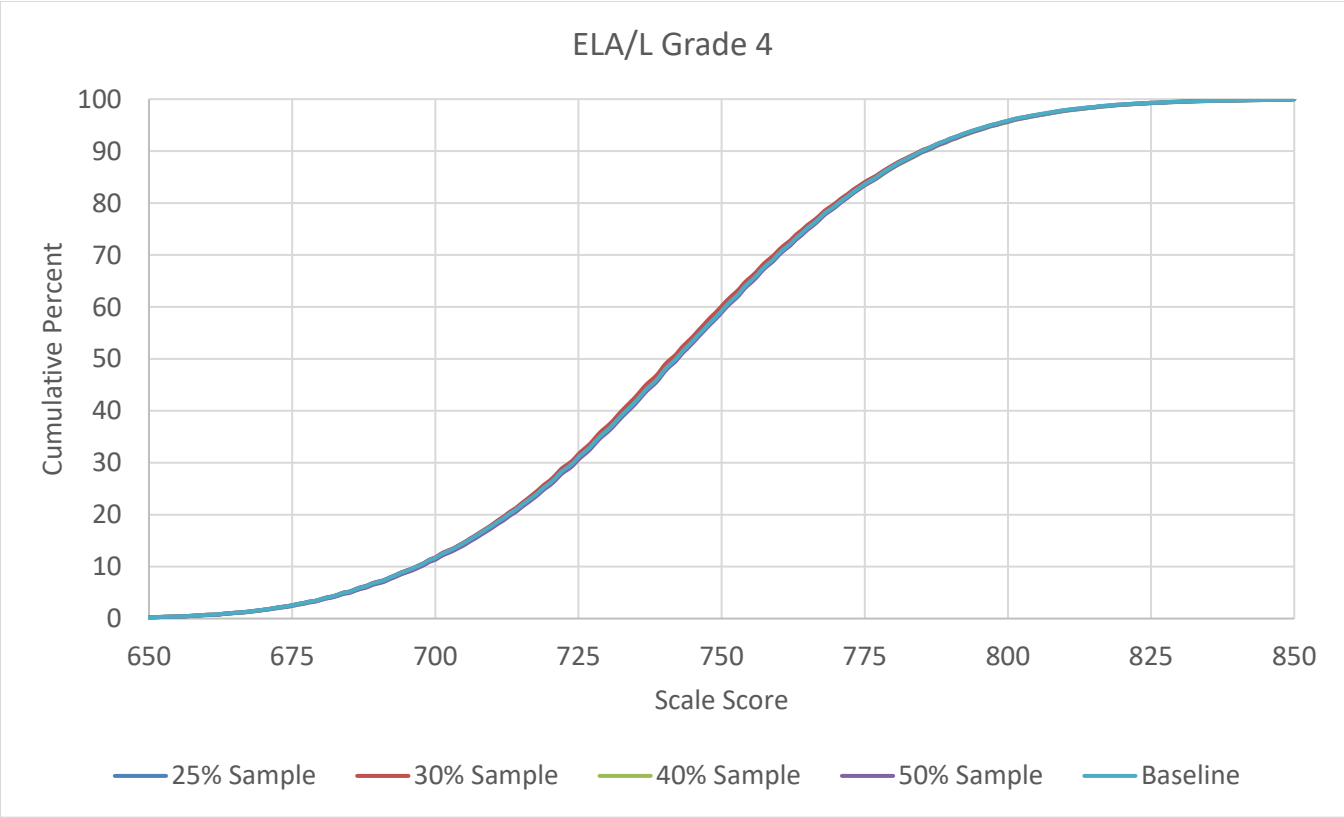

Figure D.6 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for ELA/L Grade 4

Math

Table D.7 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for Algebra 1

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	67,133	0.871	0.963	-0.038	2.871
Sample 30%	80,874	0.778	0.885	-0.048	2.593
Sample 40%	108,338	1.120	1.130	0.000	3.409
Sample 50%	135,766	0.629	0.697	-0.019	2.130

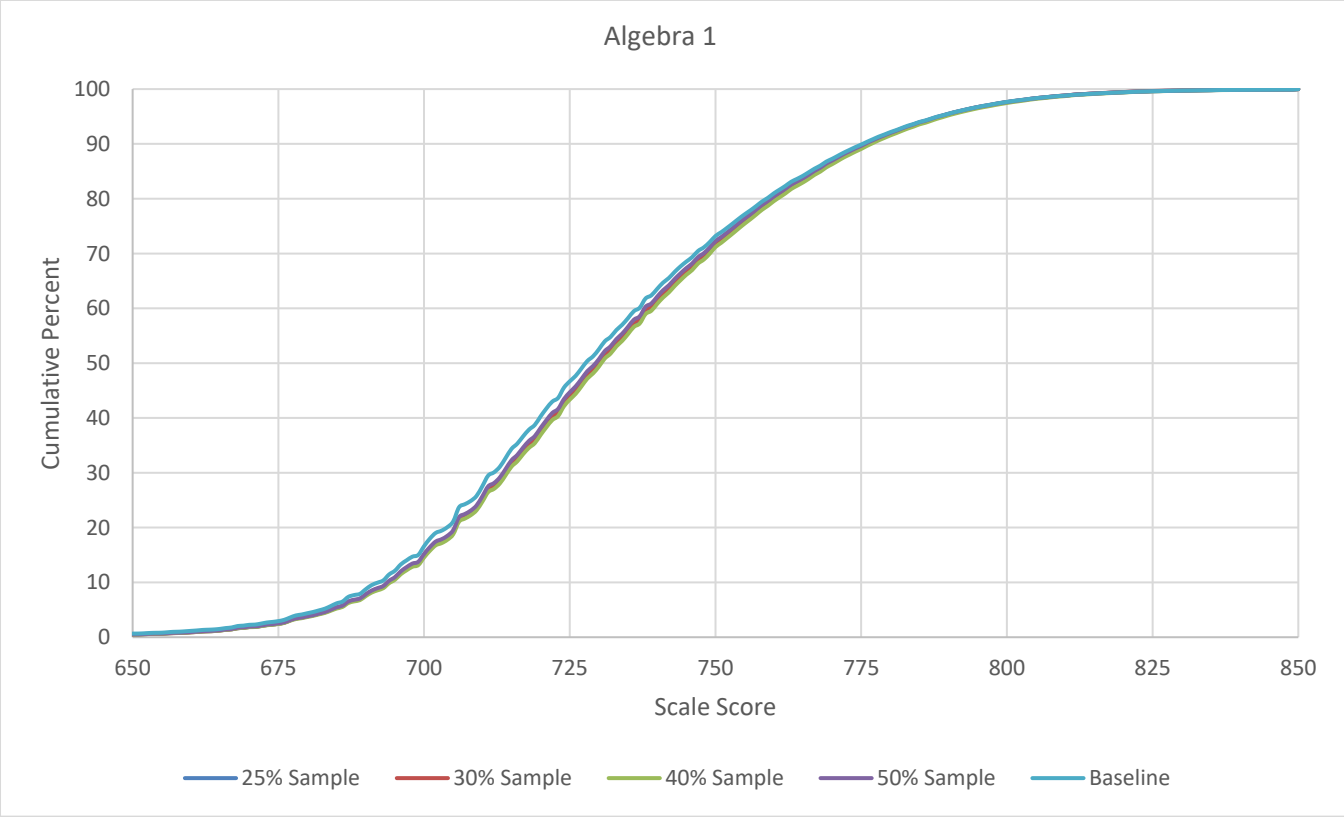

Figure D.7 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for Algebra 1

Table D.8 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for Geometry

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	33,030	0.309	0.427	-0.144	1.279
Sample 30%	39,903	0.551	0.687	-0.049	2.032
Sample 40%	53,162	0.322	0.430	-0.087	1.266
Sample 50%	66,729	0.635	0.688	-0.007	1.940

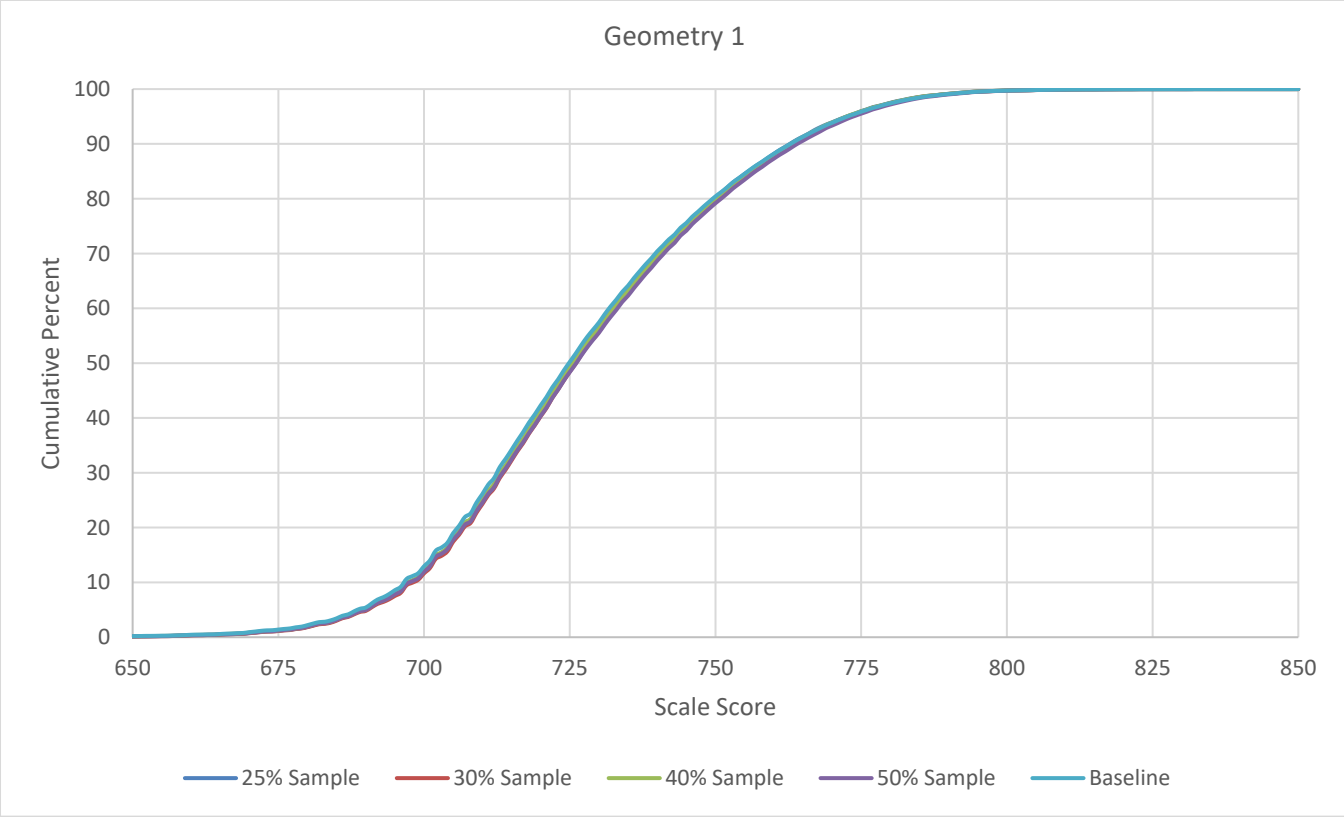

Figure D.8 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for Geometry

Table D.9 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for Integrated Math III

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	1,705	-3.133	1.821	-6.093	-0.205
Sample 30%	2,132	-3.211	1.944	-6.371	-0.217
Sample 40%	3,012	-1.595	1.086	-3.736	-0.020
Sample 50%	3,836	-0.775	0.410	-1.459	-0.036

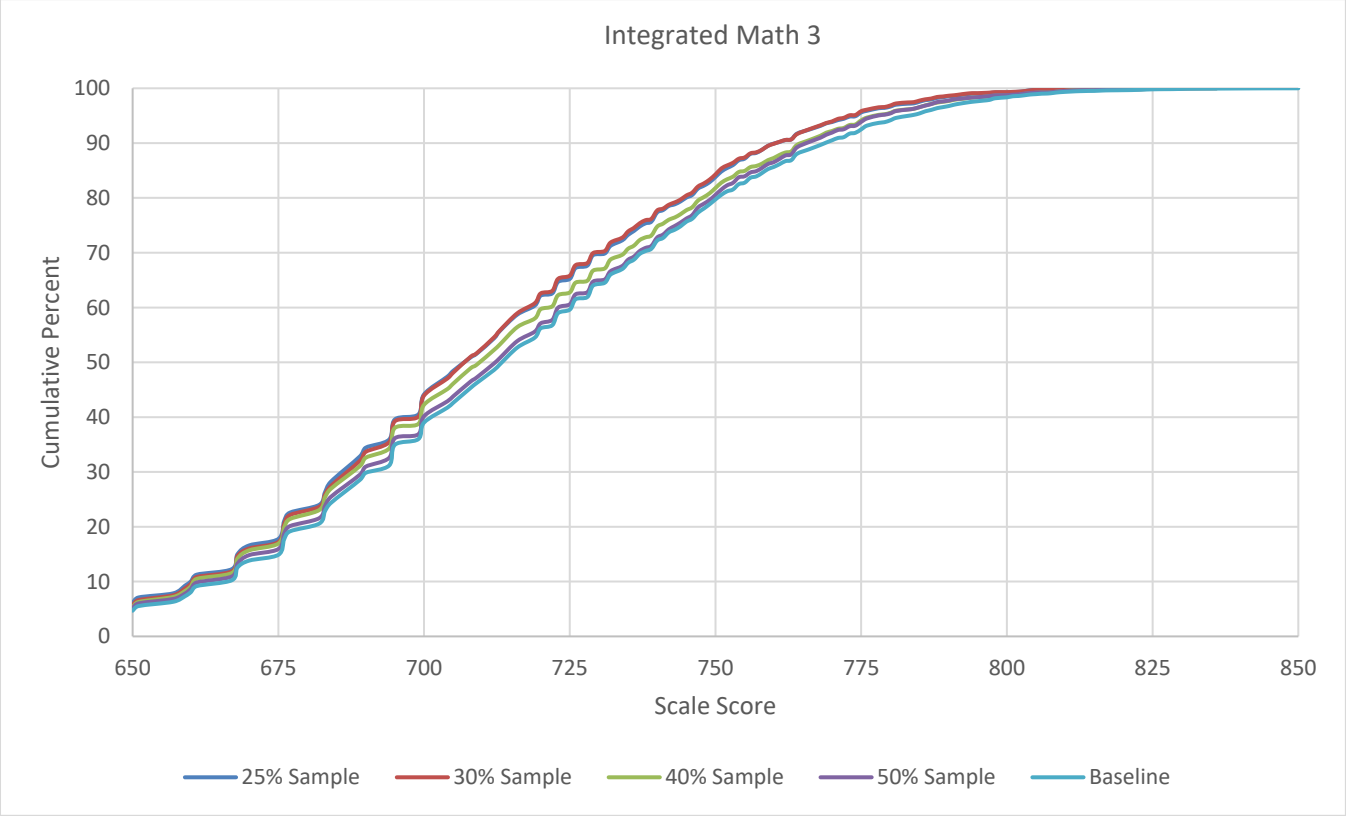

Figure D.9 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for Integrated Math III

Table D.10 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for Integrated Math I

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	4,776	-3.423	2.856	-8.452	-0.046
Sample 30%	5,686	-3.373	2.813	-8.377	-0.059
Sample 40%	7,765	-1.157	1.116	-3.366	0.038
Sample 50%	9,781	-0.724	0.826	-2.521	0.079

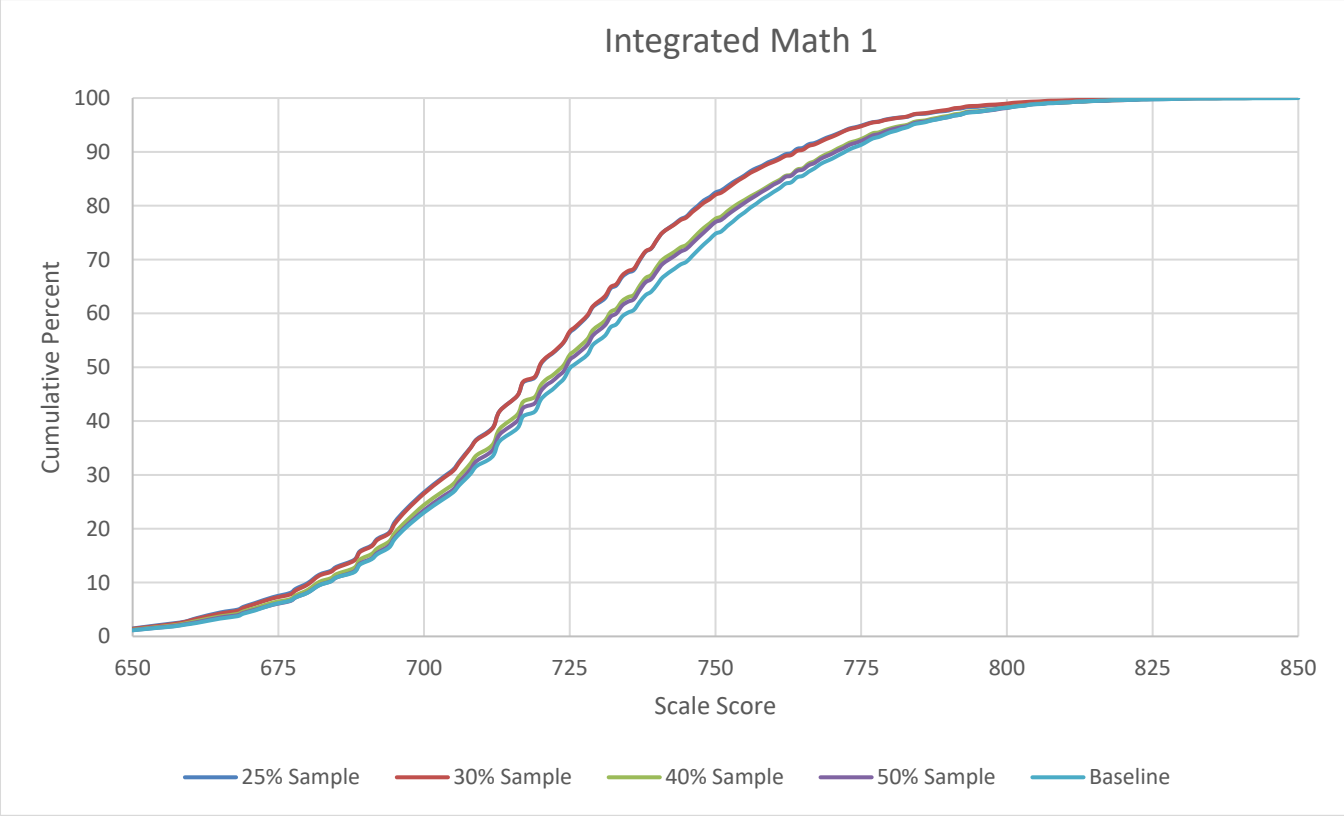

Figure D.10 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for Integrated Math I

Table D.11 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for Math Grade 8

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	73,654	0.496	0.450	0.000	1.505
Sample 30%	88,372	0.639	0.514	0.000	1.698
Sample 40%	118,109	0.617	0.490	-0.003	1.586
Sample 50%	147,820	0.551	0.464	-0.009	1.471

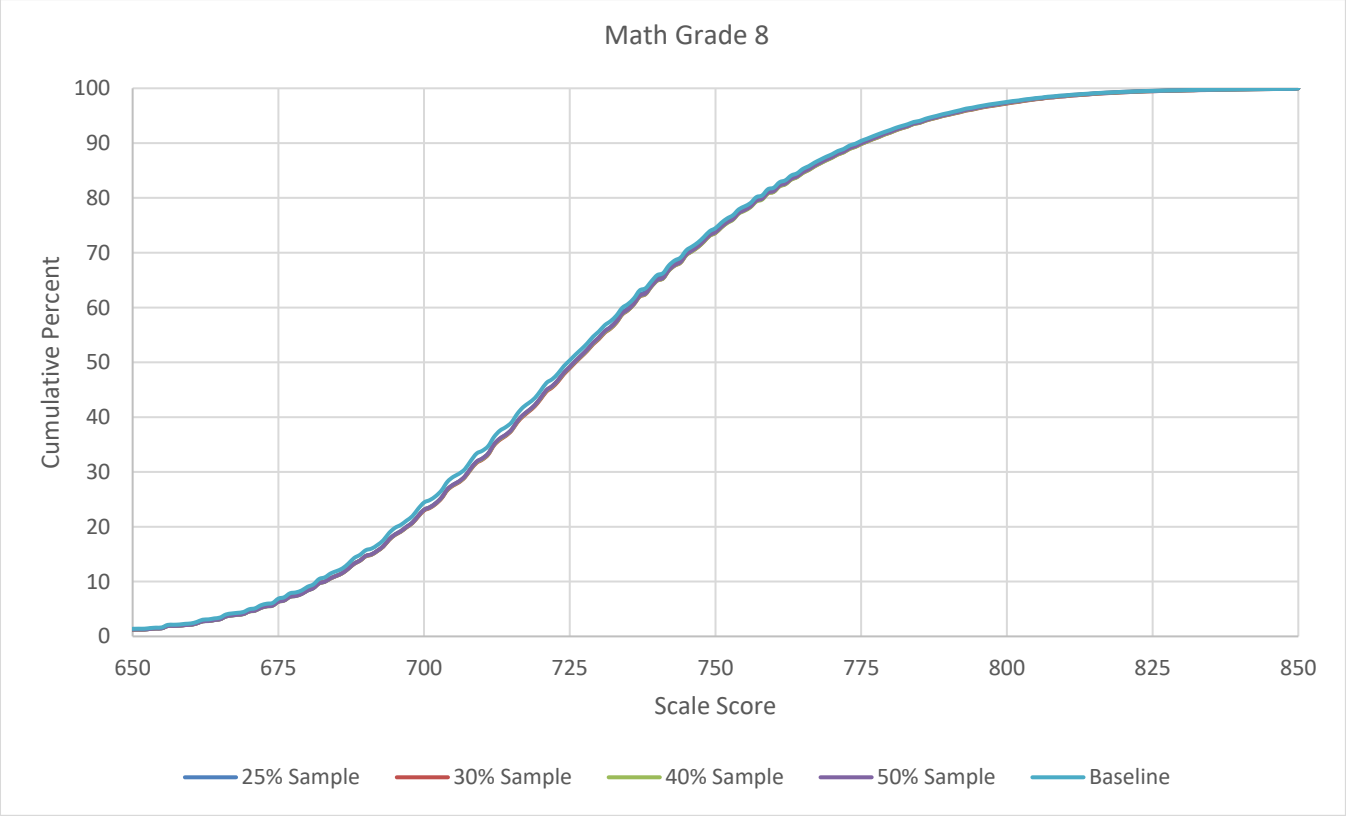

Figure D.11 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for Math Grade 8

Table D.12 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for Math Grade 7

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	89,786	-0.118	0.319	-0.842	0.395
Sample 30%	107,654	-0.038	0.250	-0.567	0.447
Sample 40%	143,964	0.314	0.357	-0.009	1.093
Sample 50%	179,923	0.379	0.412	-0.008	1.202

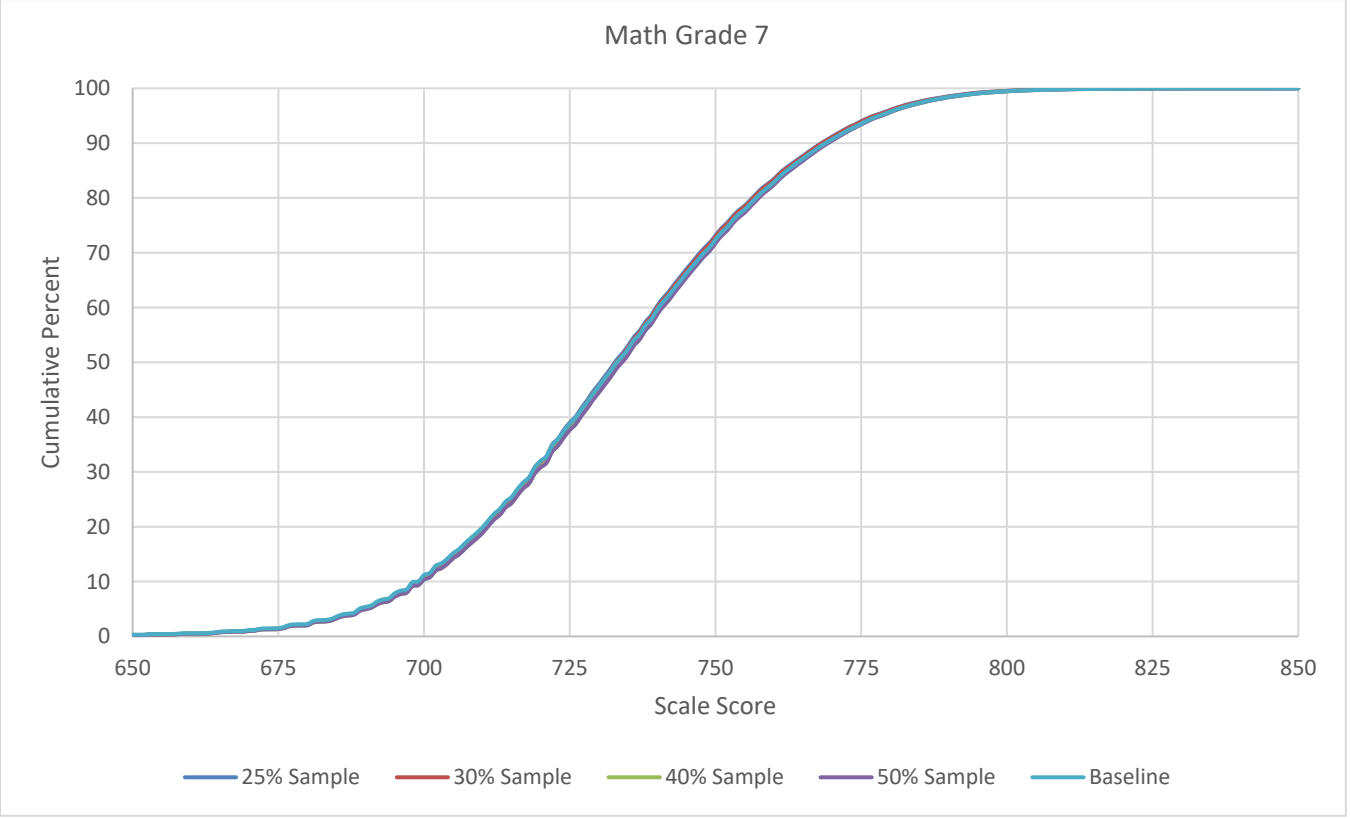

Figure D.12 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for Math Grade 7

Table D.13 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for Math Grade 6

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	93,174	0.487	0.514	-0.046	1.486
Sample 30%	111,849	0.230	0.271	-0.106	0.694
Sample 40%	149,247	0.129	0.228	-0.194	0.606
Sample 50%	186,636	0.014	0.158	-0.223	0.368

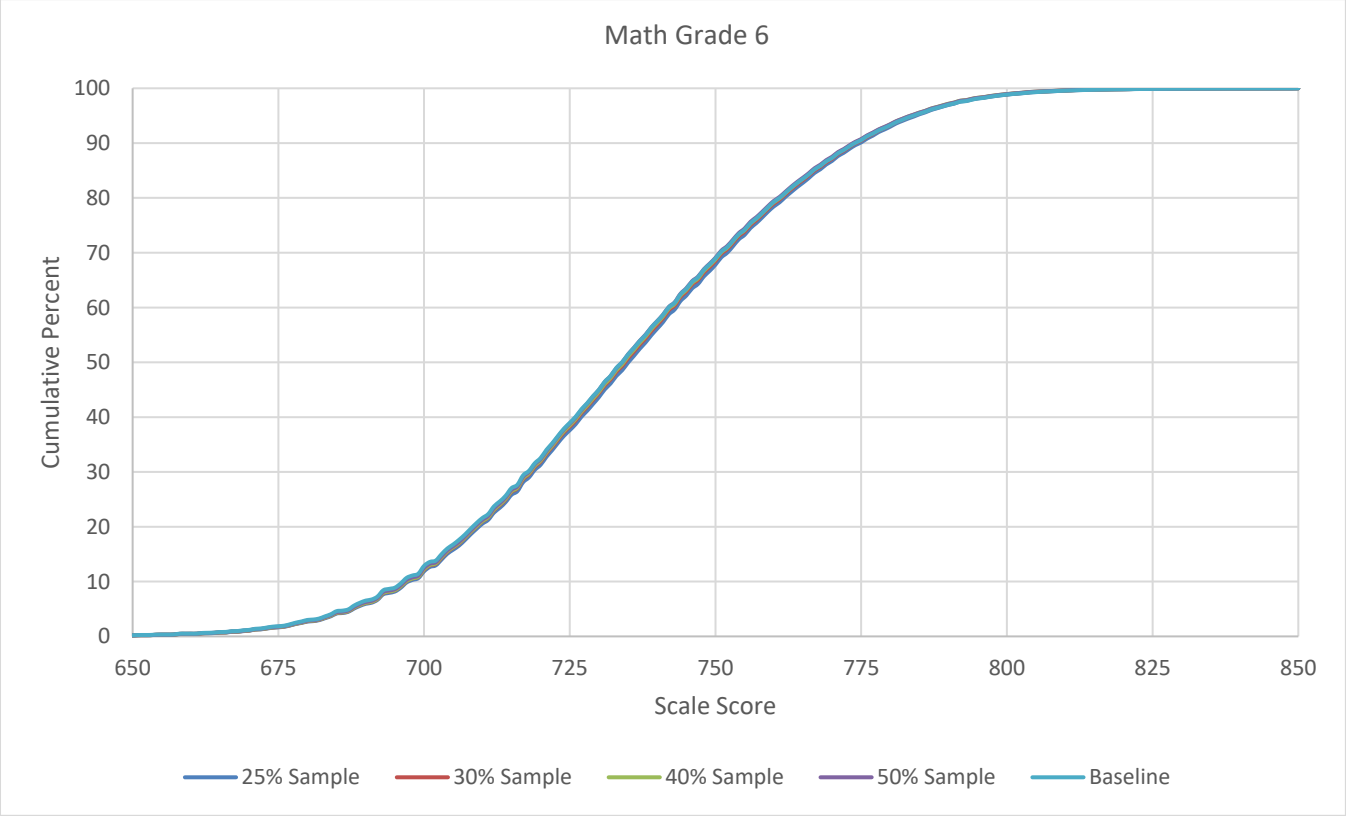

Figure D.13 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for Math Grade 6

Table D.14 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for Math Grade 4

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	86,193	-0.995	1.004	-2.842	0.022
Sample 30%	103,419	-0.503	0.540	-1.493	0.018
Sample 40%	137,948	-0.383	0.455	-1.257	0.030
Sample 50%	172,479	-0.209	0.269	-0.776	0.048

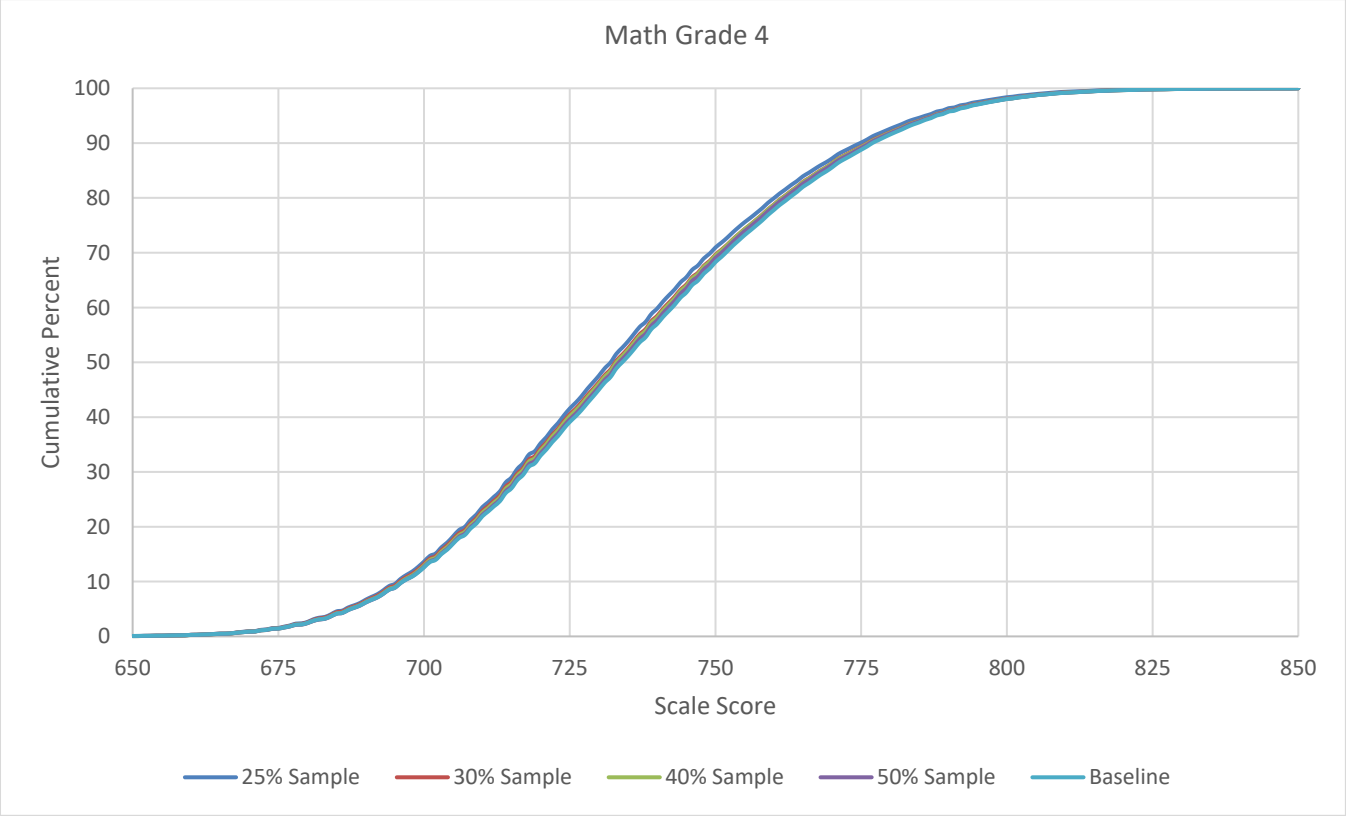

Figure D.14 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for Math Grade 4

Table D.15 Descriptive Statistics for Cumulative Percent Difference Between Equating Baseline and Sample Data for Math Grade 3

Cumulative Percent Difference Between the Baseline and Sample Distributions					
Sample Data Sets	N	Mean	STD	Min	Max
Sample 25%	80,169	-0.680	0.651	-1.860	0.057
Sample 30%	96,228	-0.665	0.638	-1.843	0.020
Sample 40%	128,368	-0.446	0.482	-1.334	0.041
Sample 50%	160,488	-0.490	0.489	-1.351	0.001

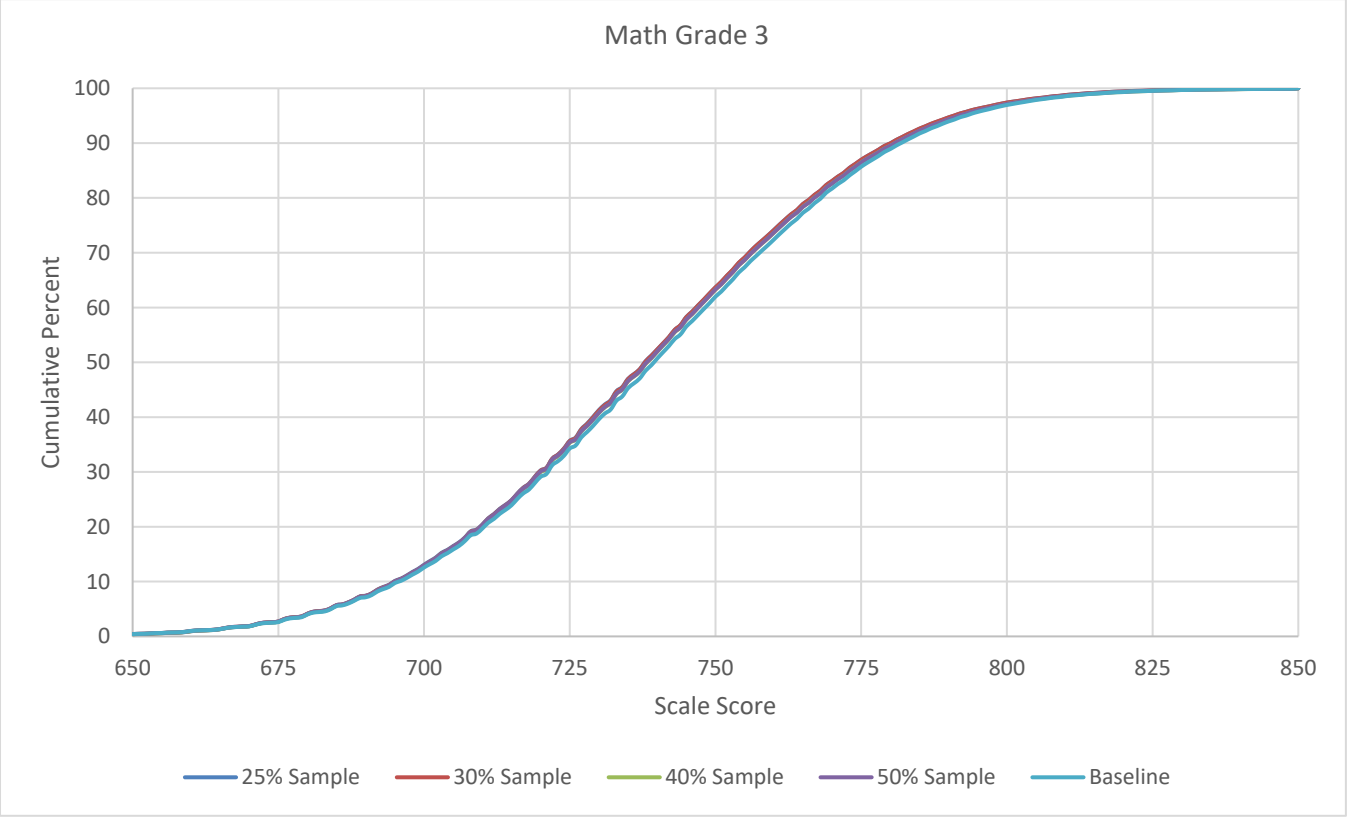

Figure D.15 Spring 2015 Summative Scale Score Cumulative Percent for the Equating Baseline and Sample Data Sets for Math Grade 3

Appendix E. Spring 2015 Performance Level Distribution for the Equating

Baseline and Sample Data Sets

ELA/L

Table E.1 Spring 2015 Performance Level Distribution for the Equating Data Sets for ELA/L Grade 11

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	12.97	13.08	13.33	14.00	**17.24**	0.113	0.110	0.103	0.086
Level 2	17.29	17.61	17.83	18.17	**19.58**	0.058	0.050	0.044	0.036
Level 3	24.57	24.59	24.57	24.64	**24.30**	-0.006	-0.007	-0.006	-0.008
Level 4	33.07	32.97	32.51	31.86	**29.50**	-0.078	-0.076	-0.066	-0.052
Level 5	12.10	11.76	11.76	11.32	**9.38**	-0.093	-0.082	-0.082	-0.067
Total N Count	**37,219**	**45,190**	**61,138**	**77,035**	**148,041**				

Table E.2 Spring 2015 Performance Level Distribution for the Equating Data Sets for ELA/L Grade 9

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	13.58	13.73	13.95	14.40	**18.51**	0.127	0.123	0.118	0.106
Level 2	18.68	18.90	19.46	19.83	**21.25**	0.063	0.057	0.044	0.035
Level 3	25.40	25.50	25.93	25.75	**24.89**	-0.012	-0.014	-0.024	-0.020
Level 4	33.16	32.76	32.10	31.60	**28.25**	-0.109	-0.100	-0.086	-0.074
Level 5	9.18	9.11	8.56	8.42	**7.10**	-0.081	-0.078	-0.057	-0.051
Total N Count	**59,745**	**71,998**	**96,324**	**120,639**	**233,074**				

Table E.3 Spring 2015 Performance Level Distribution for the Equating Data Sets for ELA/L Grade 8

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	13.08	13.13	13.24	13.71	**15.12**	0.057	0.055	0.053	0.039
Level 2	17.39	17.60	17.53	17.73	**18.30**	0.024	0.018	0.020	0.015
Level 3	24.91	24.95	24.87	24.91	**24.74**	-0.004	-0.005	-0.003	-0.004
Level 4	35.81	35.72	35.80	35.41	**34.09**	-0.036	-0.035	-0.036	-0.028
Level 5	8.82	8.59	8.56	8.24	**7.75**	-0.040	-0.032	-0.030	-0.018
Total N Count	**91,335**	**109,799**	**146,410**	**182,623**	**362,467**				

Table E.4 Spring 2015 Performance Level Distribution for the Equating Data Sets for ELA/L Grade 7

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	14.16	13.88	13.68	13.56	**14.92**	0.021	0.029	0.035	0.038
Level 2	18.42	18.23	18.00	17.80	**18.39**	-0.001	0.004	0.010	0.015
Level 3	25.42	25.30	24.98	24.99	**25.14**	-0.006	-0.004	0.004	0.004
Level 4	30.04	30.39	30.72	31.06	**29.99**	-0.001	-0.009	-0.016	-0.023
Level 5	11.95	12.20	12.62	12.59	**11.56**	-0.012	-0.020	-0.033	-0.032
Total N Count	**93,205**	**111,908**	**149,160**	**186,700**	**370,386**				

Table E.5 Spring 2015 Performance Level Distribution for the Equating Data Sets for ELA/L Grade 5

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	9.28	9.21	9.38	9.70	**10.65**	0.045	0.047	0.041	0.031
Level 2	18.12	18.20	18.54	18.90	**19.66**	0.039	0.037	0.028	0.019
Level 3	27.14	27.22	27.43	27.68	**27.98**	0.019	0.017	0.012	0.007
Level 4	40.83	40.73	40.23	39.49	**37.97**	-0.059	-0.057	-0.047	-0.031
Level 5	4.63	4.64	4.42	4.23	**3.74**	-0.047	-0.048	-0.036	-0.026
Total N Count	**91,179**	**109,519**	**146,148**	**182,735**	**364,281**				

Table E.6 Spring 2015 Performance Level Distribution for the Equating Data Sets for ELA/L Grade 4

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	11.28	11.36	11.01	10.92	**11.27**	-0.001	-0.003	0.008	0.011
Level 2	18.97	19.11	18.71	18.50	**18.61**	-0.009	-0.013	-0.002	0.003
Level 3	28.26	28.50	28.21	28.28	**28.19**	-0.001	-0.007	0.000	-0.002
Level 4	33.20	32.97	33.63	33.95	**33.86**	0.014	0.019	0.005	-0.002
Level 5	8.29	8.05	8.44	8.34	**8.06**	-0.009	0.000	-0.014	-0.010
Total N Count	**87,792**	**105,414**	**140,654**	**175,931**	**350,808**				

Math

Table E.7 Spring 2015 Performance Level Distribution for the Equating Data Sets for Algebra 1

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	13.42	13.39	13.15	13.76	**15.01**	0.044	0.045	0.052	0.035
Level 2	29.24	29.57	28.98	29.69	**30.53**	0.028	0.021	0.034	0.018
Level 3	27.69	27.80	27.82	27.65	**26.54**	-0.026	-0.029	-0.029	-0.025
Level 4	27.85	27.44	28.12	27.12	**26.16**	-0.038	-0.029	-0.045	-0.022
Level 5	1.80	1.79	1.92	1.79	**1.77**	-0.002	-0.002	-0.011	-0.001
Total N Count	**67,133**	**80,874**	**108,338**	**135,766**	**264,367**				

Table E.8 Spring 2015 Performance Level Distribution for the Equating Data Sets for Integrated Geometry

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	10.82	10.48	10.95	10.75	**11.68**	0.027	0.037	0.023	0.029
Level 2	36.88	36.39	36.64	36.09	**37.02**	0.003	0.013	0.008	0.019
Level 3	31.66	31.99	31.55	31.40	**30.81**	-0.018	-0.026	-0.016	-0.013
Level 4	18.60	19.03	18.80	19.46	**18.38**	-0.006	-0.017	-0.011	-0.028
Level 5	2.03	2.11	2.05	2.29	**2.11**	0.006	0.000	0.004	-0.013
Total N Count	**33,030**	**39,903**	**53,162**	**66,729**	**131,160**				

Table E.9 Spring 2015 Performance Level Distribution for the Equating Data Sets for Integrated Math III

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	40.47	40.06	38.75	36.86	**35.98**	-0.094	-0.085	-0.058	-0.018
Level 2	24.22	25.19	23.51	23.12	**23.01**	-0.029	-0.052	-0.012	-0.003
Level 3	18.18	18.11	18.59	19.63	**19.86**	0.042	0.044	0.032	0.006
Level 4	16.60	16.14	18.13	19.42	**19.88**	0.082	0.094	0.044	0.011
Level 5	0.53	0.52	1.03	0.96	**1.28**	0.067	0.068	0.022	0.028
Total N Count	**1,705**	**2,132**	**3,012**	**3,836**	**7,210**				

Table E.10 Spring 2015 Performance Level Distribution for the Equating Data Sets for Integrated Math I

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	23.72	23.44	21.47	20.59	**20.32**	-0.085	-0.078	-0.029	-0.007
Level 2	30.95	31.29	28.96	28.73	**27.56**	-0.076	-0.083	-0.031	-0.026
Level 3	26.93	26.45	26.27	26.70	**25.89**	-0.024	-0.013	-0.009	-0.019
Level 4	17.04	17.60	21.22	21.81	**24.14**	0.166	0.153	0.068	0.055
Level 5	1.36	1.21	2.07	2.17	**2.09**	0.051	0.061	0.001	-0.006
Total N Count	**4,776**	**5,686**	**7,765**	**9,781**	**19,296**				

Table E.11 Spring 2015 Performance Level Distribution for the Equating Data Sets for Math Grade 8

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	22.00	21.86	21.94	21.98	**23.30**	0.031	0.034	0.032	0.031
Level 2	26.10	25.99	26.09	26.13	**26.02**	-0.002	0.001	-0.002	-0.003
Level 3	25.46	25.26	25.05	25.11	**24.60**	-0.020	-0.015	-0.010	-0.012
Level 4	23.71	24.10	24.22	24.14	**23.59**	-0.003	-0.012	-0.015	-0.013
Level 5	2.72	2.78	2.70	2.64	**2.49**	-0.015	-0.019	-0.014	-0.010
Total N Count	**73,654**	**88,372**	**118,109**	**147,820**	**292,854**				

Table E.12 Spring 2015 Performance Level Distribution for the Equating Data Sets for Math Grade 7

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	9.73	9.71	9.38	9.33	**10.04**	0.010	0.011	0.022	0.023
Level 2	27.95	27.71	27.08	26.98	**27.43**	-0.011	-0.006	0.008	0.010
Level 3	34.13	34.12	34.13	34.02	**33.51**	-0.013	-0.013	-0.013	-0.011
Level 4	25.75	25.95	26.69	26.94	**26.35**	0.014	0.009	-0.008	-0.013
Level 5	2.44	2.51	2.72	2.73	**2.66**	0.014	0.010	-0.003	-0.004
Total N Count	**89,786**	**107,654**	**143,964**	**179,923**	**357,488**				

Table E.13 Spring 2015 Performance Level Distribution for the Equating Data Sets for Math Grade 6

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	10.62	10.87	10.88	11.10	**11.42**	0.025	0.017	0.017	0.010
Level 2	25.97	26.37	26.57	26.66	**26.47**	0.011	0.002	-0.002	-0.004
Level 3	30.10	30.17	30.25	30.23	**29.98**	-0.002	-0.004	-0.006	-0.005
Level 4	29.35	28.79	28.60	28.29	**28.25**	-0.025	-0.012	-0.008	-0.001
Level 5	3.96	3.81	3.70	3.72	**3.87**	-0.005	0.003	0.009	0.008
Total N Count	**93,174**	**111,849**	**149,247**	**186,636**	**371,256**				

Table E.14 Spring 2015 Performance Level Distribution for the Equating Data Sets for Math Grade 4

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	12.70	12.24	11.97	11.76	**11.68**	-0.032	-0.017	-0.009	-0.002
Level 2	27.62	27.06	27.00	26.63	**26.24**	-0.032	-0.019	-0.017	-0.009
Level 3	29.51	29.19	29.35	29.45	**29.17**	-0.007	0.000	-0.004	-0.006
Level 4	27.68	28.70	28.82	29.26	**29.95**	0.050	0.027	0.025	0.015
Level 5	2.49	2.81	2.86	2.90	**2.96**	0.028	0.009	0.006	0.004
Total N Count	**86,193**	**103,419**	**137,948**	**172,479**	**344,191**				

Table E.15 Spring 2015 Performance Level Distribution for the Equating Data Sets for Math Grade 3

Actual 2015 Performance Level	Sample 25% Percent	Sample 30% Percent	Sample 40% Percent	Sample 50% Percent	Baseline Percent	Sample 25% Effect Size	Sample 30% Effect Size	Sample 40% Effect Size	Sample 50% Effect Size
Level 1	12.33	12.23	12.11	12.20	**11.84**	-0.015	-0.012	-0.008	-0.011
Level 2	21.95	22.02	21.84	21.85	**21.12**	-0.020	-0.022	-0.018	-0.018
Level 3	28.29	28.34	28.21	28.12	**27.86**	-0.010	-0.011	-0.008	-0.006
Level 4	31.70	31.72	31.75	31.80	**32.72**	0.022	0.021	0.021	0.020
Level 5	5.72	5.70	6.09	6.02	**6.45**	0.030	0.031	0.015	0.017
Total N Count	**80,169**	**96,228**	**128,368**	**160,488**	**320,146**				

Appendix F. Item Analysis Descriptive Statistics for the Equating Data Sets

Table F.1 EOY Item Percent Correct for the Equating Data Sets for Algebra 2

EOY Item Percent Correct								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	122	0.245	0.183	0.003	0.743	0.99938	0.368	0.047
Sample 30%	122	0.242	0.182	0.004	0.741	0.99956	0.232	0.030
Sample 40%	122	0.242	0.182	0.004	0.741	0.99967	0.223	0.029
Sample 50%	122	0.242	0.182	0.004	0.742	0.99974	0.243	0.031
Baseline	122	0.237	0.179	0.004	0.734			

Table F.2 EOY Item Mean for the Equating Data Sets for Algebra 2

EOY Item Mean								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	122	0.415	0.491	0.003	2.970	0.99964	0.234	0.030
Sample 30%	122	0.410	0.488	0.004	2.963	0.99975	0.153	0.020
Sample 40%	122	0.409	0.487	0.004	2.964	0.99982	0.150	0.019
Sample 50%	122	0.410	0.488	0.004	2.968	0.99986	0.162	0.021
Baseline	122	0.400	0.477	0.004	2.935			

Table F.3 EOY Response Category Distributions for the Equating Data Sets for Algebra 2

	Score Category 0					Difference (Baseline - Sample) for Score 0			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	122	68.98	22.58	3.70	99.70	0.94	0.77	-0.80	3.50
Sample 30%	122	69.32	22.53	3.70	99.60	0.60	0.64	-0.60	2.70
Sample 40%	122	69.33	22.52	3.80	99.60	0.59	0.57	-0.50	2.30
Sample 50%	122	69.28	22.54	3.70	99.60	0.64	0.53	-0.40	2.20
Baseline	122	69.92	22.28	4.00	99.60				

	Score Category 1					Difference (Baseline - Sample) for Score 1			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	122	24.15	16.32	0.30	71.30	-0.61	0.79	-3.50	2.20
Sample 30%	122	23.92	16.26	0.40	71.20	-0.38	0.61	-2.70	1.80
Sample 40%	122	23.91	16.26	0.40	71.00	-0.37	0.54	-2.20	1.40
Sample 50%	122	23.94	16.25	0.40	71.10	-0.40	0.51	-1.90	1.30
Baseline	122	23.54	16.10	0.40	70.40				

	Score Category 2					Difference (Baseline - Sample) for Score 2			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	51	10.70	10.34	0.20	47.80	-0.53	0.81	-3.00	0.80
Sample 30%	51	10.50	10.23	0.20	47.30	-0.33	0.65	-2.10	0.70
Sample 40%	51	10.52	10.20	0.20	47.50	-0.35	0.57	-2.00	0.70
Sample 50%	51	10.56	10.20	0.30	47.50	-0.39	0.53	-1.90	0.50
Baseline	51	10.17	9.79	0.30	45.60				

Score Category 3						Difference (Baseline - Sample) for Score 3			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	10	15.10	16.81	2.70	55.90	-0.58	0.59	-1.80	0.20
Sample 30%	10	14.94	16.86	2.60	55.80	-0.42	0.45	-1.50	0.10
Sample 40%	10	14.93	16.76	2.70	55.40	-0.41	0.45	-1.50	0.10
Sample 50%	10	14.95	16.73	2.70	55.50	-0.43	0.35	-1.20	-0.10
Baseline	10	14.52	16.65	2.50	55.20				

Score Category 4						Difference (Baseline - Sample) for Score 4			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	10	14.24	17.13	0.90	49.40	-0.84	1.11	-3.10	0.10
Sample 30%	10	14.01	16.83	0.90	48.40	-0.61	0.82	-2.40	0.10
Sample 40%	10	13.92	16.76	0.90	48.10	-0.52	0.70	-1.80	0.20
Sample 50%	10	13.97	16.79	1.00	48.30	-0.57	0.68	-1.70	0.00
Baseline	10	13.40	16.20	0.90	47.10				

Table F.4 PBA Item Percent Correct for the Equating Data Sets for Algebra 2

PBA Item Percent Correct								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	78	0.192	0.152	0.003	0.696	0.99893	0.553	0.088
Sample 30%	78	0.188	0.151	0.002	0.691	0.99929	0.390	0.062
Sample 40%	78	0.187	0.150	0.002	0.688	0.99940	0.357	0.057
Sample 50%	78	0.186	0.149	0.003	0.685	0.99954	0.310	0.050
Baseline	78	0.178	0.146	0.003	0.672			

Table F.5 PBA Item Mean for the Equating Data Sets for Algebra 2

PBA Item Mean								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	78	0.393	0.361	0.003	1.634	0.99895	0.577	0.092
Sample 30%	78	0.385	0.356	0.002	1.608	0.99937	0.430	0.069
Sample 40%	78	0.384	0.353	0.002	1.583	0.99944	0.404	0.065
Sample 50%	78	0.381	0.351	0.003	1.575	0.99949	0.365	0.058
Baseline	78	0.361	0.328	0.003	1.467			

Table F.6 PBA Response Category Distributions for the Equating Data Sets for Algebra 2

Score Category 0						Difference (Baseline - Sample) for Score 0			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	78	72.76	18.47	24.60	99.70	1.96	1.33	-0.30	5.20
Sample 30%	78	73.26	18.40	25.00	99.80	1.46	1.17	-0.40	4.40
Sample 40%	78	73.38	18.28	25.20	99.80	1.34	1.05	-0.20	3.80
Sample 50%	78	73.55	18.18	25.50	99.70	1.16	0.93	-0.10	3.50
Baseline	78	74.72	17.56	28.00	99.70				

	Score Category 1					Difference (Baseline - Sample) for Score 1			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	78	20.34	15.74	0.30	69.60	-1.31	0.90	-3.40	0.50
Sample 30%	78	20.00	15.62	0.20	69.10	-0.97	0.71	-2.70	0.50
Sample 40%	78	19.91	15.55	0.20	68.80	-0.88	0.64	-2.80	0.40
Sample 50%	78	19.76	15.47	0.30	68.50	-0.73	0.52	-2.40	0.20
Baseline	78	19.03	15.19	0.30	67.20				

	Score Category 2					Difference (Baseline - Sample) for Score 2			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	37	8.37	8.71	0.10	36.30	-0.69	0.66	-3.20	0.30
Sample 30%	37	8.22	8.67	0.10	36.20	-0.54	0.61	-2.70	0.40
Sample 40%	37	8.18	8.61	0.10	36.40	-0.50	0.54	-2.50	0.20
Sample 50%	37	8.15	8.59	0.10	36.30	-0.47	0.49	-2.20	0.00
Baseline	37	7.68	8.13	0.10	34.80				

	Score Category 3					Difference (Baseline - Sample) for Score 3			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	37	3.27	3.29	0.10	12.40	-0.35	0.38	-1.30	0.00
Sample 30%	37	3.19	3.22	0.10	11.80	-0.27	0.32	-1.20	0.10
Sample 40%	37	3.17	3.22	0.00	11.90	-0.25	0.30	-1.10	0.10
Sample 50%	37	3.15	3.21	0.00	12.10	-0.23	0.29	-1.00	0.10
Baseline	37	2.92	2.94	0.10	11.10				

	Score Category 4					Difference (Baseline - Sample) for Score 4			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	24	2.39	2.70	0.00	8.50	-0.23	0.33	-1.10	0.10
Sample 30%	24	2.33	2.64	0.00	8.40	-0.17	0.27	-0.80	0.10
Sample 40%	24	2.30	2.63	0.00	8.20	-0.15	0.26	-0.70	0.10
Sample 50%	24	2.30	2.59	0.00	8.10	-0.14	0.22	-0.60	0.10
Baseline	24	2.16	2.39	0.00	7.60				

	Score Category 5					Difference (Baseline - Sample) for Score 5			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	12	2.68	2.98	0.10	8.60	-0.33	0.42	-1.10	0.10
Sample 30%	12	2.59	2.86	0.10	8.20	-0.25	0.31	-0.80	0.00
Sample 40%	12	2.60	2.84	0.10	8.30	-0.26	0.28	-0.80	0.00
Sample 50%	12	2.58	2.84	0.10	8.30	-0.24	0.28	-0.80	0.00
Baseline	12	2.34	2.56	0.10	7.50				

	Score Category 6					Difference (Baseline - Sample) for Score 6			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	12	1.52	2.43	0.00	6.80	-0.18	0.34	-1.00	0.00
Sample 30%	12	1.45	2.35	0.00	6.50	-0.11	0.26	-0.70	0.10
Sample 40%	12	1.43	2.28	0.00	6.30	-0.09	0.18	-0.50	0.10
Sample 50%	12	1.46	2.29	0.00	6.20	-0.12	0.20	-0.60	0.10

Baseline	12	1.34	2.10	0.00	5.80	

Table F.7 EOY Item Percent Correct for the Equating Data Sets for Integrated Math II

EOY Item Percent Correct								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	49	0.204	0.181	0.003	0.583	0.99696	-0.594	-0.120
Sample 30%	49	0.219	0.185	0.006	0.603	0.99896	-0.187	-0.038
Sample 40%	49	0.223	0.185	0.007	0.610	0.99909	-0.103	-0.021
Sample 50%	49	0.230	0.187	0.009	0.618	0.99946	0.092	0.019
Baseline	49	0.226	0.187	0.011	0.621			

Table F.8 EOY Item Mean for the Equating Data Sets for Integrated Math II

EOY Item Mean								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	49	0.306	0.305	0.003	1.389	0.99744	-0.508	-0.103
Sample 30%	49	0.328	0.315	0.006	1.443	0.99908	-0.160	-0.032
Sample 40%	49	0.333	0.317	0.007	1.466	0.99920	-0.086	-0.017
Sample 50%	49	0.344	0.324	0.009	1.499	0.99958	0.078	0.016
Baseline	49	0.339	0.321	0.011	1.468			

Table F.9 EOY Response Category Distributions for the Equating Data Sets for Integrated Math II

	Score Category 0					Difference (Baseline - Sample) for Score 0			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	49	73.27	23.02	17.40	99.70	-2.45	1.83	-6.80	0.80
Sample 30%	49	71.64	23.25	16.00	99.40	-0.82	1.21	-4.10	2.50
Sample 40%	49	71.30	23.28	15.50	99.30	-0.48	1.07	-3.40	1.70
Sample 50%	49	70.51	23.47	14.60	99.10	0.31	0.73	-1.60	2.70
Baseline	49	70.82	23.56	15.20	98.90				

	Score Category 1					Difference (Baseline - Sample) for Score 1			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	49	23.21	18.49	0.30	58.30	1.74	1.79	-1.70	6.00
Sample 30%	49	24.33	18.50	0.60	60.30	0.62	1.08	-2.80	3.10
Sample 40%	49	24.55	18.42	0.70	61.00	0.40	0.92	-2.00	3.00
Sample 50%	49	25.11	18.47	0.90	61.80	-0.16	0.72	-2.70	1.60
Baseline	49	24.95	18.60	1.10	62.10				

	Score Category 2					Difference (Baseline - Sample) for Score 2			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	18	8.65	10.29	0.10	33.90	1.76	1.44	0.00	5.10
Sample 30%	18	9.83	10.77	0.20	37.30	0.58	0.71	-0.50	1.80
Sample 40%	18	10.13	10.95	0.20	37.80	0.28	0.69	-0.90	1.70
Sample 50%	18	10.71	11.46	0.30	40.70	-0.30	0.59	-1.70	0.80
Baseline	18	10.41	11.33	0.20	39.00				

Score Category 3						Difference (Baseline - Sample) for Score 3			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	3	5.53	5.61	1.30	11.90	0.83	0.65	0.20	1.50
Sample 30%	3	6.40	6.04	2.10	13.30	-0.03	0.15	-0.20	0.10
Sample 40%	3	6.63	6.19	2.20	13.70	-0.27	0.15	-0.40	-0.10
Sample 50%	3	6.93	6.70	2.20	14.60	-0.57	0.57	-1.20	-0.10
Baseline	3	6.37	6.14	2.10	13.40				

Score Category 4						Difference (Baseline - Sample) for Score 4			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	3	0.20	0.10	0.10	0.30	0.17	0.06	0.10	0.20
Sample 30%	3	0.33	0.21	0.10	0.50	0.03	0.06	0.00	0.10
Sample 40%	3	0.40	0.26	0.10	0.60	-0.03	0.12	-0.10	0.10
Sample 50%	3	0.43	0.29	0.10	0.60	-0.07	0.15	-0.20	0.10
Baseline	3	0.37	0.15	0.20	0.50				

Table F.10 PBA Item Percent Correct for the Equating Data Sets for Integrated Math II

PBA Item Percent Correct								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	28	0.132	0.141	0.000	0.496	0.99736	-0.605	-0.162
Sample 30%	28	0.152	0.148	0.001	0.516	0.99736	-0.082	-0.022
Sample 40%	28	0.157	0.149	0.002	0.515	0.99659	0.047	0.013
Sample 50%	28	0.165	0.152	0.003	0.532	0.99696	0.251	0.067
Baseline	28	0.155	0.148	0.003	0.521			

Table F.11 PBA Item Mean for the Equating Data Sets for Integrated Math II

PBA Item Mean								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	28	0.241	0.296	0.000	1.488	0.99428	-0.613	-0.164
Sample 30%	28	0.279	0.308	0.001	1.548	0.99804	-0.131	-0.035
Sample 40%	28	0.289	0.306	0.002	1.544	0.99846	-0.022	-0.006
Sample 50%	28	0.308	0.318	0.003	1.595	0.99857	0.204	0.054
Baseline	28	0.290	0.312	0.003	1.563			

Table F.12 PBA Response Category Distributions for the Equating Data Sets for Integrated Math II

Score Category 0						Difference (Baseline - Sample) for Score 0			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	28	81.46	19.09	11.90	100.00	-3.01	1.56	-7.20	-0.30
Sample 30%	28	78.90	19.35	10.80	99.90	-0.45	1.35	-2.60	2.40
Sample 40%	28	78.31	19.18	11.10	99.80	0.14	1.43	-2.20	2.90
Sample 50%	28	77.25	19.36	10.40	99.70	1.20	1.28	-1.20	4.30
Baseline	28	78.45	19.24	11.50	99.70				

Score Category 1						Difference (Baseline - Sample) for Score 1			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	28	14.78	13.89	0.00	44.20	2.05	2.07	-3.70	7.10
Sample 30%	28	16.73	14.58	0.10	49.00	0.10	1.27	-2.40	2.30
Sample 40%	28	17.18	14.65	0.20	50.20	-0.35	1.32	-2.90	2.20
Sample 50%	28	17.76	14.89	0.30	52.10	-0.93	1.28	-4.30	1.20
Baseline	28	16.83	14.61	0.30	49.90				

Score Category 2						Difference (Baseline - Sample) for Score 2			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	13	5.38	11.74	0.00	43.50	0.86	0.54	0.10	1.80
Sample 30%	13	5.84	11.81	0.10	44.20	0.41	0.43	-0.20	1.20
Sample 40%	13	5.95	11.57	0.20	43.60	0.29	0.44	-0.30	1.10
Sample 50%	13	6.45	12.02	0.20	45.50	-0.21	0.43	-1.20	0.50
Baseline	13	6.25	11.76	0.20	44.30				

Score Category 3						Difference (Baseline - Sample) for Score 3			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	13	2.05	3.15	0.00	9.40	0.68	0.82	0.00	3.20
Sample 30%	13	2.52	3.55	0.00	10.70	0.21	0.37	-0.30	1.10
Sample 40%	13	2.59	3.60	0.00	11.00	0.14	0.29	-0.30	0.80
Sample 50%	13	2.92	3.88	0.00	12.20	-0.18	0.25	-0.70	0.10
Baseline	13	2.73	3.72	0.00	11.80				

Score Category 4						Difference (Baseline - Sample) for Score 4			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	6	0.83	0.50	0.00	1.30	0.38	0.31	0.00	0.80
Sample 30%	6	1.22	0.74	0.00	2.00	0.00	0.22	-0.40	0.20
Sample 40%	6	1.30	0.72	0.00	2.10	-0.08	0.31	-0.50	0.40
Sample 50%	6	1.45	0.77	0.00	2.20	-0.23	0.23	-0.60	0.00
Baseline	6	1.22	0.70	0.00	2.10				

Score Category 5						Difference (Baseline - Sample) for Score 5			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	4	0.70	0.42	0.20	1.20	0.60	0.29	0.20	0.90
Sample 30%	4	1.15	0.66	0.30	1.90	0.15	0.26	-0.10	0.50
Sample 40%	4	1.13	0.63	0.50	2.00	0.18	0.15	0.00	0.30
Sample 50%	4	1.38	0.71	0.80	2.40	-0.08	0.15	-0.30	0.00
Baseline	4	1.30	0.57	0.80	2.10				

Score Category 6						Difference (Baseline - Sample) for Score 6			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	4	0.28	0.22	0.10	0.60	0.43	0.40	0.00	0.90
Sample 30%	4	0.53	0.29	0.20	0.90	0.18	0.34	-0.10	0.60
Sample 40%	4	0.73	0.42	0.10	1.00	-0.03	0.17	-0.20	0.20
Sample 50%	4	0.80	0.47	0.10	1.10	-0.10	0.12	-0.20	0.00

Baseline	4	0.70	0.42	0.10	1.10	

Table F.13 EOY Item Percent Correct for the Equating Data Sets for ELA/L Grade 6

EOY Item Percent Correct								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	73	0.427	0.142	0.159	0.758	0.99977	0.300	0.050
Sample 30%	73	0.426	0.142	0.158	0.755	0.99982	0.236	0.039
Sample 40%	73	0.424	0.141	0.157	0.751	0.99989	0.144	0.024
Sample 50%	73	0.423	0.141	0.157	0.750	0.99990	0.139	0.023
Baseline	73	0.420	0.140	0.155	0.745			

Table F.14 EOY Item Mean for the Equating Data Sets for ELA/L Grade 6

EOY Item Mean								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	73	0.854	0.284	0.318	1.516	0.99976	0.298	0.049
Sample 30%	73	0.851	0.284	0.317	1.511	0.99983	0.230	0.038
Sample 40%	73	0.847	0.282	0.314	1.502	0.99988	0.143	0.024
Sample 50%	73	0.847	0.282	0.313	1.499	0.99991	0.136	0.022
Baseline	73	0.841	0.279	0.310	1.490			

Table F.15 EOY Response Category Distributions for the Equating Data Sets for ELA/L Grade 6

	Response Category 0					Difference (Baseline - Sample) Response 0			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	73	43.76	15.67	1.10	76.90	0.72	0.46	-0.50	2.00
Sample 30%	73	43.93	15.66	1.10	76.90	0.55	0.39	-0.40	1.50
Sample 40%	73	44.13	15.64	1.10	77.10	0.35	0.30	-0.30	1.00
Sample 50%	73	44.13	15.63	1.10	77.20	0.35	0.26	-0.30	1.20
Baseline	73	44.48	15.53	1.30	77.50				

	Response Category 1					Difference (Baseline - Sample) Response 1			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	73	27.03	15.28	6.00	64.40	-0.04	0.47	-1.60	2.10
Sample 30%	73	27.00	15.25	6.00	64.40	-0.02	0.38	-1.20	1.60
Sample 40%	73	27.02	15.24	5.90	64.60	-0.03	0.26	-0.70	0.90
Sample 50%	73	27.04	15.27	5.80	64.60	-0.06	0.22	-0.70	0.50
Baseline	73	26.98	15.18	5.80	64.80				

	Response Category 2					Difference (Baseline - Sample) Response 2			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	73	29.21	16.60	1.80	70.20	-0.68	0.46	-2.20	0.20
Sample 30%	73	29.07	16.53	1.80	69.80	-0.54	0.38	-1.70	0.10
Sample 40%	73	28.85	16.45	1.70	69.30	-0.32	0.29	-1.10	0.20
Sample 50%	73	28.83	16.43	1.70	69.10	-0.30	0.25	-1.10	0.10
Baseline	73	28.53	16.27	1.60	68.30				

Table F.16 PBA Item Percent Correct for the Equating Data Sets for ELA/L Grade 6

PBA Item Percent Correct								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	69	0.470	0.152	0.191	0.783	0.99940	0.356	0.061
Sample 30%	69	0.468	0.152	0.190	0.780	0.99944	0.284	0.048
Sample 40%	69	0.466	0.151	0.191	0.778	0.99969	0.186	0.032
Sample 50%	69	0.466	0.151	0.192	0.776	0.99972	0.181	0.031
Baseline	69	0.461	0.150	0.193	0.772			

Table F.17 PBA Item Mean for the Equating Data Sets for ELA/L Grade 6

PBA Item Mean								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	69	1.633	1.737	0.383	6.863	0.99989	0.182	0.031
Sample 30%	69	1.625	1.725	0.381	6.812	0.99990	0.154	0.026
Sample 40%	69	1.609	1.697	0.383	6.731	0.99991	0.101	0.017
Sample 50%	69	1.611	1.701	0.385	6.738	0.99994	0.105	0.018
Baseline	69	1.580	1.643	0.386	6.582			

Table F.18 PBA Response Category Distributions for the Equating Data Sets for ELA/L Grade 6

	Response Category 0					Difference (Baseline - Sample) Response 0			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	69	35.64	17.57	6.50	76.70	0.94	0.58	-0.30	2.80
Sample 30%	69	35.82	17.58	6.60	76.80	0.77	0.58	-0.40	2.60
Sample 40%	69	36.03	17.53	6.80	76.60	0.56	0.49	-0.20	2.40
Sample 50%	69	36.04	17.53	6.90	76.60	0.55	0.47	-0.20	2.30
Baseline	69	36.59	17.29	7.40	76.40				

	Response Category 1					Difference (Baseline - Sample) Response 1			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	69	23.78	17.38	3.70	76.50	-0.04	0.40	-1.00	0.50
Sample 30%	69	23.80	17.36	3.80	76.70	-0.06	0.34	-0.90	0.50
Sample 40%	69	23.83	17.32	3.80	76.60	-0.09	0.24	-0.60	0.30
Sample 50%	69	23.83	17.31	3.80	76.70	-0.09	0.21	-0.50	0.30
Baseline	69	23.74	17.21	4.00	76.70				

	Response Category 2					Difference (Baseline - Sample) Response 2			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	69	31.22	19.96	0.30	73.50	-0.67	0.53	-1.90	0.10
Sample 30%	69	31.06	19.88	0.30	73.20	-0.51	0.45	-1.60	0.10
Sample 40%	69	30.86	19.77	0.30	72.90	-0.31	0.33	-1.40	0.20
Sample 50%	69	30.84	19.73	0.30	72.70	-0.29	0.28	-1.30	0.10
Baseline	69	30.55	19.54	0.30	72.20				

	Response Category 3					Difference (Baseline - Sample) Response 3			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	1.69	1.89	0.10	5.00	0.14	0.21	-0.10	0.40
Sample 30%	9	1.73	1.96	0.10	5.20	0.10	0.13	0.00	0.30
Sample 40%	9	1.77	2.00	0.10	5.20	0.07	0.09	0.00	0.20
Sample 50%	9	1.77	2.00	0.10	5.20	0.07	0.09	0.00	0.20
Baseline	9	1.83	2.08	0.10	5.40				

	Response Category 4					Difference (Baseline - Sample) Response 4			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	7.87	7.11	2.30	20.10	0.21	0.23	-0.10	0.70
Sample 30%	9	7.93	7.09	2.40	20.20	0.14	0.25	-0.20	0.60
Sample 40%	9	8.00	7.16	2.40	20.40	0.08	0.20	-0.20	0.40
Sample 50%	9	8.00	7.16	2.50	20.40	0.08	0.17	-0.10	0.40
Baseline	9	8.08	7.29	2.50	20.80				

	Response Category 5					Difference (Baseline - Sample) Response 5			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	18.47	11.83	2.70	33.40	0.63	0.68	-0.20	1.90
Sample 30%	9	18.64	11.95	2.70	33.50	0.46	0.52	-0.10	1.40
Sample 40%	9	18.90	12.08	2.70	33.80	0.20	0.34	-0.40	0.70
Sample 50%	9	18.93	12.07	2.80	33.70	0.17	0.29	-0.30	0.60
Baseline	9	19.10	12.14	2.80	33.40				

	Response Category 6					Difference (Baseline - Sample) Response 6			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	3.13	2.20	0.50	6.00	-0.26	0.31	-0.80	0.10
Sample 30%	9	3.10	2.16	0.60	5.90	-0.22	0.24	-0.60	0.00
Sample 40%	9	3.06	2.12	0.60	5.80	-0.18	0.20	-0.50	0.00
Sample 50%	9	3.02	2.10	0.60	5.80	-0.14	0.17	-0.50	0.00
Baseline	9	2.88	1.97	0.60	5.60				

	Score Category 7					Difference (Baseline - Sample) for Score 7			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	3.44	3.91	0.40	9.60	-0.06	0.11	-0.20	0.20
Sample 30%	9	3.47	3.97	0.40	9.80	-0.08	0.08	-0.20	0.10
Sample 40%	9	3.48	4.01	0.30	9.80	-0.09	0.06	-0.20	0.00
Sample 50%	9	3.46	4.03	0.30	9.80	-0.07	0.07	-0.20	0.00
Baseline	9	3.39	3.97	0.30	9.60				

	Score Category 8					Difference (Baseline - Sample) for Score 8			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	6.29	9.04	0.10	19.90	-0.26	0.35	-0.90	0.00
Sample 30%	9	6.26	9.00	0.10	19.90	-0.22	0.29	-0.70	0.00
Sample 40%	9	6.21	8.91	0.10	19.70	-0.18	0.20	-0.50	0.00
Sample 50%	9	6.18	8.91	0.10	19.70	-0.14	0.20	-0.50	0.00

Baseline	9	6.03	8.72	0.10	19.40				

	Score Category 9					Difference (Baseline - Sample) for Score 9			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	2.10	0.55	1.40	3.10	-0.01	0.12	-0.20	0.20
Sample 30%	9	2.12	0.57	1.40	3.20	-0.03	0.11	-0.20	0.20
Sample 40%	9	2.12	0.56	1.40	3.20	-0.03	0.10	-0.10	0.20
Sample 50%	9	2.16	0.59	1.40	3.20	-0.07	0.07	-0.20	0.00
Baseline	9	2.09	0.59	1.40	3.20				

	Score Category 10					Difference (Baseline - Sample) for Score 10			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	12.67	9.09	0.70	21.80	-0.30	0.67	-1.60	0.70
Sample 30%	9	12.64	9.09	0.70	21.80	-0.28	0.68	-1.60	0.60
Sample 40%	9	12.48	8.95	0.80	21.40	-0.11	0.55	-1.20	0.60
Sample 50%	9	12.58	8.98	0.80	21.10	-0.21	0.43	-0.90	0.30
Baseline	9	12.37	8.83	0.80	20.40				

	Score Category 11					Difference (Baseline - Sample) for Score 11			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	3.87	3.58	0.70	9.30	-0.39	0.35	-1.00	0.00
Sample 30%	9	3.83	3.58	0.70	9.30	-0.36	0.34	-1.00	0.00
Sample 40%	9	3.70	3.49	0.70	9.00	-0.22	0.25	-0.70	0.00
Sample 50%	9	3.69	3.49	0.70	9.00	-0.21	0.23	-0.70	0.00
Baseline	9	3.48	3.29	0.70	8.30				

	Score Category 12					Difference (Baseline - Sample) for Score 12			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	2.52	3.56	0.10	8.10	-0.27	0.39	-0.90	0.00
Sample 30%	9	2.47	3.49	0.10	7.90	-0.21	0.32	-0.70	0.00
Sample 40%	9	2.40	3.38	0.10	7.70	-0.14	0.22	-0.50	0.00
Sample 50%	9	2.38	3.35	0.10	7.60	-0.12	0.19	-0.40	0.00
Baseline	9	2.26	3.17	0.10	7.20				

	Score Category 13					Difference (Baseline - Sample) for Score 13			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	0.10	0.09	0.00	0.20	-0.01	0.03	-0.10	0.00
Sample 30%	9	0.10	0.09	0.00	0.20	-0.01	0.03	-0.10	0.00
Sample 40%	9	0.10	0.09	0.00	0.20	-0.01	0.03	-0.10	0.00
Sample 50%	9	0.10	0.09	0.00	0.20	-0.01	0.03	-0.10	0.00
Baseline	9	0.09	0.08	0.00	0.20				

	Score Category 14					Difference (Baseline - Sample) for Score 14			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	1.31	0.71	0.60	2.80	-0.10	0.07	-0.20	0.00
Sample 30%	9	1.30	0.71	0.50	2.80	-0.09	0.06	-0.20	0.00
Sample 40%	9	1.23	0.68	0.50	2.70	-0.02	0.04	-0.10	0.00

Sample 50%	9	1.29	0.69	0.50	2.80	-0.08	0.07	-0.20	0.00
Baseline	9	1.21	0.68	0.50	2.70				

	Score Category 15					Difference (Baseline - Sample) for Score 15			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	5.99	1.03	4.80	7.60	-0.77	0.21	-1.00	-0.40
Sample 30%	9	5.86	1.04	4.70	7.50	-0.63	0.20	-0.90	-0.30
Sample 40%	9	5.61	0.99	4.50	7.20	-0.39	0.17	-0.60	-0.10
Sample 50%	9	5.59	0.96	4.50	7.10	-0.37	0.14	-0.60	-0.20
Baseline	9	5.22	0.95	3.90	6.80				

	Score Category 16					Difference (Baseline - Sample) for Score 16			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	6	0.20	0.13	0.10	0.40	-0.02	0.04	-0.10	0.00
Sample 30%	6	0.20	0.13	0.10	0.40	-0.02	0.04	-0.10	0.00
Sample 40%	6	0.20	0.13	0.10	0.40	-0.02	0.04	-0.10	0.00
Sample 50%	6	0.20	0.13	0.10	0.40	-0.02	0.04	-0.10	0.00
Baseline	6	0.18	0.10	0.10	0.30				

	Score Category 17*					Difference (Baseline - Sample) for Score 17			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sample 30%	6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sample 40%	6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sample 50%	6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Baseline	6	0.00	0.00	0.00	0.00				

	Score Category 18					Difference (Baseline - Sample) for Score 18			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	6	0.13	0.08	0.00	0.20	0.00	0.06	-0.10	0.10
Sample 30%	6	0.13	0.08	0.00	0.20	0.00	0.06	-0.10	0.10
Sample 40%	6	0.12	0.08	0.00	0.20	0.02	0.04	0.00	0.10
Sample 50%	6	0.12	0.08	0.00	0.20	0.02	0.04	0.00	0.10
Baseline	6	0.13	0.05	0.10	0.20				

	Score Category 19					Difference (Baseline - Sample) for Score 19			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	6	2.92	0.81	1.90	3.90	-0.45	0.31	-0.90	-0.10
Sample 30%	6	2.77	0.77	1.80	3.70	-0.30	0.27	-0.70	0.00
Sample 40%	6	2.67	0.71	1.80	3.60	-0.20	0.19	-0.50	0.00
Sample 50%	6	2.70	0.68	1.90	3.70	-0.23	0.12	-0.40	-0.10
Baseline	6	2.47	0.59	1.80	3.40				

*The score point of 17 has zero students due to 17 being difficult to attain when combining the traits. It requires students to get almost zero points on one trait and full credit on other traits.

Table F.19 EOY Item Percent Correct for the Equating Data Sets for Math 5

EOY Item Percent Correct								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	144	0.431	0.234	0.065	0.928	0.99977	0.174	0.020
Sample 30%	144	0.428	0.234	0.064	0.926	0.99981	0.081	0.010
Sample 40%	144	0.431	0.234	0.068	0.929	0.99988	0.174	0.021
Sample 50%	144	0.430	0.234	0.068	0.928	0.99988	0.162	0.019
Baseline	144	0.426	0.233	0.066	0.922			

Table F.20 EOY Item Mean for the Equating Data Sets for Math 5

EOY Item Mean								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	144	0.503	0.289	0.065	1.542	0.99979	0.173	0.020
Sample 30%	144	0.500	0.288	0.064	1.540	0.99984	0.084	0.010
Sample 40%	144	0.504	0.289	0.068	1.552	0.99986	0.178	0.021
Sample 50%	144	0.503	0.289	0.068	1.549	0.99985	0.167	0.020
Baseline	144	0.498	0.286	0.066	1.529			

Table F.21 EOY Response Category Distributions for the Equating Data Sets for Math 5

	Response Category 0					Difference (Baseline - Sample) Response 0			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	144	54.43	23.11	7.20	93.50	0.50	0.54	-1.80	1.90
Sample 30%	144	54.69	23.07	7.40	93.60	0.24	0.48	-1.90	1.50
Sample 40%	144	54.43	23.09	7.10	93.20	0.50	0.41	-1.00	1.60
Sample 50%	144	54.47	23.03	7.20	93.20	0.46	0.40	-1.00	1.50
Baseline	144	54.93	22.95	7.80	93.40				

	Response Category 1					Difference (Baseline - Sample) Response 1			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	144	40.79	23.67	3.40	92.80	-0.40	0.54	-1.90	1.80
Sample 30%	144	40.58	23.59	3.50	92.60	-0.19	0.47	-1.50	1.90
Sample 40%	144	40.78	23.64	3.50	92.90	-0.39	0.41	-1.30	1.00
Sample 50%	144	40.75	23.58	3.60	92.80	-0.36	0.39	-1.30	1.00
Baseline	144	40.39	23.45	3.50	92.20				

	Response Category 2					Difference (Baseline - Sample) Response 2			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	31	22.17	19.52	2.60	72.90	-0.42	0.41	-1.20	0.30
Sample 30%	31	21.97	19.50	2.50	72.90	-0.23	0.33	-0.90	0.40
Sample 40%	31	22.25	19.61	2.60	73.40	-0.50	0.40	-1.40	0.10
Sample 50%	31	22.22	19.56	2.60	73.30	-0.47	0.37	-1.30	0.10
Baseline	31	21.75	19.32	2.60	72.20				

Table F.22 PBA Item Percent Correct for the Equating Data Sets for Math 5

PBA Item Percent Correct								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	62	0.397	0.224	0.033	0.880	0.99955	-0.028	-0.005
Sample 30%	62	0.396	0.224	0.033	0.878	0.99967	-0.070	-0.013
Sample 40%	62	0.400	0.224	0.035	0.879	0.99980	0.030	0.005
Sample 50%	62	0.400	0.224	0.034	0.879	0.99987	0.040	0.007
Baseline	62	0.398	0.225	0.033	0.879			

Table F.23 PBA Item Mean for the Equating Data Sets for Math 5

PBA Item Mean								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	62	0.749	0.520	0.091	2.571	0.99980	-0.001	0.000
Sample 30%	62	0.744	0.517	0.089	2.561	0.99985	-0.046	-0.008
Sample 40%	62	0.753	0.521	0.092	2.596	0.99991	0.046	0.008
Sample 50%	62	0.754	0.522	0.092	2.601	0.99994	0.053	0.010
Baseline	62	0.749	0.518	0.093	2.578			

Table F.24 PBA Response Category Distributions for the Equating Data Sets for Math 5

	Response Category 0					Difference (Baseline - Sample) Response 0			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	62	47.92	22.18	7.70	93.90	-0.03	0.85	-2.10	2.80
Sample 30%	62	48.13	22.18	7.70	94.00	-0.24	0.72	-1.90	2.10
Sample 40%	62	47.69	22.09	7.50	93.80	0.19	0.54	-1.50	1.60
Sample 50%	62	47.66	22.10	7.70	93.80	0.22	0.44	-1.20	1.30
Baseline	62	47.89	22.07	8.10	94.00				

	Response Category 1					Difference (Baseline - Sample) Response 1			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	62	37.89	22.66	2.30	88.00	0.08	0.82	-3.10	2.10
Sample 30%	62	37.80	22.62	2.30	87.80	0.17	0.74	-2.60	1.90
Sample 40%	62	38.00	22.69	2.30	87.90	-0.04	0.48	-1.60	1.50
Sample 50%	62	38.02	22.72	2.30	87.90	-0.06	0.39	-1.40	1.20
Baseline	62	37.96	22.71	2.30	87.90				

	Response Category 2					Difference (Baseline - Sample) Response 2			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	36	15.27	8.87	1.60	34.20	-0.09	0.24	-0.60	0.50
Sample 30%	36	15.17	8.82	1.60	34.20	0.01	0.24	-0.40	0.60
Sample 40%	36	15.31	8.86	1.70	34.50	-0.13	0.19	-0.60	0.30
Sample 50%	36	15.34	8.90	1.70	34.50	-0.17	0.18	-0.50	0.20
Baseline	36	15.18	8.84	1.60	34.30				

	Response Category 3					Difference (Baseline - Sample) Response 3			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	25	8.00	7.16	1.20	32.90	-0.01	0.24	-0.50	0.40

Sample 30%	25	7.93	7.15	1.10	32.70	0.06	0.20	-0.30	0.50
Sample 40%	25	8.10	7.13	1.20	32.70	-0.10	0.17	-0.50	0.20
Sample 50%	25	8.08	7.14	1.20	32.70	-0.09	0.15	-0.40	0.30
Baseline	25	7.99	7.08	1.20	32.40				

	Response Category 4					Difference (Baseline - Sample) Response 4			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	11	7.48	7.78	1.30	22.20	0.01	0.27	-0.60	0.40
Sample 30%	11	7.43	7.74	1.30	22.10	0.06	0.12	-0.10	0.30
Sample 40%	11	7.61	7.85	1.30	22.70	-0.12	0.19	-0.60	0.10
Sample 50%	11	7.63	7.84	1.30	22.60	-0.14	0.20	-0.50	0.10
Baseline	11	7.49	7.77	1.30	22.10				

	Response Category 5					Difference (Baseline - Sample) Response 5			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	4	6.73	4.55	1.50	12.60	0.07	0.17	-0.10	0.30
Sample 30%	4	6.65	4.47	1.50	12.40	0.15	0.19	-0.10	0.30
Sample 40%	4	6.83	4.59	1.50	12.70	-0.03	0.10	-0.10	0.10
Sample 50%	4	6.88	4.58	1.50	12.70	-0.08	0.10	-0.20	0.00
Baseline	4	6.80	4.63	1.40	12.70				

	Response Category 6					Difference (Baseline - Sample) Response 6			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	4	5.38	6.23	0.60	14.40	-0.02	0.05	-0.10	0.00
Sample 30%	4	5.25	6.16	0.60	14.20	0.10	0.08	0.00	0.20
Sample 40%	4	5.45	6.31	0.60	14.60	-0.10	0.14	-0.30	0.00
Sample 50%	4	5.45	6.31	0.60	14.60	-0.10	0.14	-0.30	0.00
Baseline	4	5.35	6.18	0.60	14.30				

Table F.25 EOY Item Percent Correct for the Equating Data Sets for ELA/L Grade 3

EOY Item Percent Correct								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	36	0.422	0.153	0.129	0.702	0.99983	-0.098	-0.023
Sample 30%	36	0.423	0.154	0.130	0.704	0.99987	-0.050	-0.012
Sample 40%	36	0.424	0.154	0.130	0.704	0.99993	-0.018	-0.004
Sample 50%	36	0.424	0.154	0.130	0.704	0.99991	-0.038	-0.009
Baseline	36	0.425	0.154	0.130	0.706			

Table F.26 EOY Item Mean for the Equating Data Sets for ELA/L Grade 3

EOY Item Mean								
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	36	0.843	0.307	0.259	1.405	0.99981	-0.099	-0.023
Sample 30%	36	0.847	0.308	0.259	1.408	0.99988	-0.051	-0.012
Sample 40%	36	0.849	0.308	0.259	1.408	0.99993	-0.019	-0.005
Sample 50%	36	0.848	0.307	0.260	1.408	0.99992	-0.038	-0.009
Baseline	36	0.850	0.308	0.259	1.411			

Table F.27 EOY Response Category Distributions for the Equating Data Sets for ELA/L Grade 3

	Score Category 0					Difference (Baseline - Sample) for Score 0			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	36	45.51	17.43	14.90	79.60	-0.34	0.33	-1.40	0.30
Sample 30%	36	45.34	17.47	14.70	79.70	-0.17	0.27	-1.10	0.40
Sample 40%	36	45.23	17.46	14.70	79.50	-0.06	0.20	-0.60	0.30
Sample 50%	36	45.28	17.43	14.80	79.30	-0.11	0.20	-0.70	0.20
Baseline	36	45.17	17.45	14.70	79.40				

	Score Category 1					Difference (Baseline - Sample) for Score 1			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	36	24.67	14.48	5.60	59.00	-0.05	0.28	-0.80	0.50
Sample 30%	36	24.66	14.47	5.60	58.90	-0.04	0.22	-0.50	0.40
Sample 40%	36	24.66	14.45	5.60	58.70	-0.04	0.17	-0.40	0.30
Sample 50%	36	24.68	14.42	5.70	58.60	-0.06	0.14	-0.40	0.30
Baseline	36	24.62	14.47	5.60	58.60				

	Score Category 2					Difference (Baseline - Sample) for Score 2			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	36	29.81	16.45	2.20	63.90	0.39	0.35	-0.10	1.20
Sample 30%	36	30.00	16.54	2.10	64.00	0.21	0.26	-0.30	0.90
Sample 40%	36	30.11	16.54	2.20	63.90	0.10	0.21	-0.30	0.60
Sample 50%	36	30.04	16.50	2.10	63.80	0.17	0.23	-0.30	0.80
Baseline	36	30.21	16.59	2.10	64.00				

Table F.28 PBA Item Percent Correct for the Equating Data Sets for ELA/L Grade 3

	PBA Item Percent Correct							
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	60	0.428	0.168	0.155	0.821	0.99978	-0.119	-0.022
Sample 30%	60	0.429	0.168	0.157	0.822	0.99978	-0.056	-0.010
Sample 40%	60	0.431	0.169	0.159	0.823	0.99991	-0.017	-0.003
Sample 50%	60	0.430	0.168	0.159	0.822	0.99994	-0.049	-0.009
Baseline	60	0.431	0.169	0.161	0.822			

Table F.29 PBA Item Mean for the Equating Data Sets for ELA/L Grade 3

	PBA Item Mean							
	N	Mean	STD	Min	Max	Correlation	T	Cohen's D
Sample 25%	60	1.295	0.978	0.310	4.363	0.99981	-0.030	-0.006
Sample 30%	60	1.302	0.986	0.313	4.401	0.99983	0.009	0.002
Sample 40%	60	1.302	0.978	0.317	4.349	0.99989	0.010	0.002
Sample 50%	60	1.296	0.968	0.318	4.304	0.99992	-0.025	-0.004
Baseline	60	1.300	0.969	0.322	4.311			

Table F.30 PBA Response Category Distributions for the Equating Data Sets for ELA/L Grade 3

	Score Category 0					Difference (Baseline - Sample) for Score 0			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	60	40.13	16.87	8.00	82.40	-0.36	0.40	-1.20	0.80
Sample 30%	60	39.91	16.87	8.00	82.20	-0.14	0.38	-0.80	1.00
Sample 40%	60	39.78	16.81	7.90	81.90	-0.01	0.29	-0.50	0.70
Sample 50%	60	39.86	16.76	7.80	81.90	-0.09	0.23	-0.70	0.50
Baseline	60	39.77	16.68	7.90	81.70				

	Score Category 1					Difference (Baseline - Sample) for Score 1			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	60	21.69	11.92	4.30	53.70	-0.05	0.28	-0.90	0.50
Sample 30%	60	21.70	11.95	4.30	53.80	-0.06	0.26	-0.70	0.40
Sample 40%	60	21.71	11.94	4.40	53.80	-0.07	0.23	-0.70	0.40
Sample 50%	60	21.77	11.93	4.40	53.90	-0.14	0.17	-0.60	0.20
Baseline	60	21.64	11.92	4.50	53.70				

	Score Category 2					Difference (Baseline - Sample) for Score 2			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	60	30.68	19.30	1.40	73.50	0.37	0.39	-0.70	1.10
Sample 30%	60	30.83	19.37	1.40	73.60	0.22	0.33	-0.50	1.00
Sample 40%	60	30.97	19.42	1.40	73.70	0.08	0.26	-0.60	0.70
Sample 50%	60	30.87	19.37	1.40	73.50	0.18	0.22	-0.40	0.60
Baseline	60	31.05	19.46	1.20	73.70				

	Score Category 3					Difference (Baseline - Sample) for Score 3			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	1.90	1.61	0.50	4.70	-0.06	0.11	-0.20	0.20
Sample 30%	9	1.87	1.61	0.40	4.70	-0.02	0.10	-0.10	0.20
Sample 40%	9	1.87	1.58	0.40	4.60	-0.02	0.13	-0.10	0.30
Sample 50%	9	1.87	1.65	0.40	4.90	-0.02	0.07	-0.10	0.10
Baseline	9	1.84	1.66	0.40	4.90				

	Score Category 4					Difference (Baseline - Sample) for Score 4			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	10.88	12.43	1.60	31.40	0.41	0.45	0.00	1.40
Sample 30%	9	10.92	12.46	1.70	31.30	0.37	0.37	0.00	1.10
Sample 40%	9	11.03	12.54	1.80	31.70	0.26	0.36	-0.10	1.00
Sample 50%	9	11.14	12.58	1.80	31.70	0.14	0.25	-0.10	0.60
Baseline	9	11.29	12.70	1.90	31.60				

	Score Category 5					Difference (Baseline - Sample) for Score 5			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	16.54	6.27	5.00	23.20	0.44	0.78	-0.60	1.80
Sample 30%	9	16.67	6.39	4.90	23.40	0.32	0.67	-0.50	1.40
Sample 40%	9	16.77	6.46	4.90	23.50	0.22	0.58	-0.60	1.20

Sample 50%	9	16.74	6.46	4.80	23.20	0.24	0.40	-0.30	1.00
Baseline	9	16.99	6.63	4.70	22.90				

	Score Category 6					Difference (Baseline - Sample) for Score 6			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	5.40	4.24	0.30	10.90	-0.23	0.25	-0.70	0.00
Sample 30%	9	5.43	4.24	0.30	10.80	-0.27	0.23	-0.60	0.00
Sample 40%	9	5.32	4.16	0.30	10.60	-0.16	0.15	-0.40	0.00
Sample 50%	9	5.26	4.11	0.30	10.50	-0.09	0.14	-0.30	0.00
Baseline	9	5.17	4.03	0.30	10.20				

	Score Category 7					Difference (Baseline - Sample) for Score 7			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	2.20	1.71	0.60	6.50	-0.07	0.09	-0.20	0.10
Sample 30%	9	2.18	1.65	0.60	6.30	-0.04	0.13	-0.20	0.20
Sample 40%	9	2.14	1.63	0.50	6.20	-0.01	0.14	-0.20	0.20
Sample 50%	9	2.12	1.64	0.50	6.20	0.01	0.13	-0.10	0.20
Baseline	9	2.13	1.70	0.50	6.40				

	Score Category 8					Difference (Baseline - Sample) for Score 8			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	4.17	6.34	0.10	16.60	0.09	0.21	-0.10	0.60
Sample 30%	9	4.24	6.42	0.10	16.90	0.01	0.23	-0.40	0.50
Sample 40%	9	4.24	6.41	0.10	16.80	0.01	0.18	-0.30	0.40
Sample 50%	9	4.20	6.33	0.10	16.50	0.06	0.15	-0.10	0.40
Baseline	9	4.26	6.41	0.10	16.50				

	Score Category 9					Difference (Baseline - Sample) for Score 9			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	1.32	0.57	0.80	2.70	-0.01	0.13	-0.20	0.20
Sample 30%	9	1.33	0.57	0.80	2.70	-0.02	0.12	-0.20	0.20
Sample 40%	9	1.32	0.59	0.80	2.80	-0.01	0.08	-0.10	0.10
Sample 50%	9	1.29	0.60	0.80	2.80	0.02	0.04	0.00	0.10
Baseline	9	1.31	0.63	0.80	2.90				

	Score Category 10					Difference (Baseline - Sample) for Score 10			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	4.00	3.34	0.10	9.70	-0.02	0.18	-0.40	0.20
Sample 30%	9	4.03	3.35	0.10	9.70	-0.06	0.19	-0.40	0.20
Sample 40%	9	4.00	3.34	0.10	9.70	-0.02	0.17	-0.40	0.20
Sample 50%	9	3.92	3.29	0.10	9.60	0.06	0.14	-0.20	0.30
Baseline	9	3.98	3.35	0.10	9.90				

	Score Category 11					Difference (Baseline - Sample) for Score 11			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	1.18	0.55	0.50	2.10	-0.09	0.08	-0.20	0.00
Sample 30%	9	1.18	0.55	0.50	2.10	-0.09	0.08	-0.20	0.00

Sample 40%	9	1.17	0.56	0.50	2.10	-0.08	0.07	-0.20	0.00
Sample 50%	9	1.12	0.55	0.50	2.10	-0.03	0.05	-0.10	0.00
Baseline	9	1.09	0.51	0.50	2.00				

	Score Category 12					Difference (Baseline - Sample) for Score 12			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	9	1.43	1.93	0.00	4.60	-0.06	0.10	-0.20	0.10
Sample 30%	9	1.52	2.04	0.00	4.90	-0.14	0.20	-0.50	0.00
Sample 40%	9	1.46	1.93	0.00	4.60	-0.08	0.11	-0.30	0.00
Sample 50%	9	1.41	1.87	0.00	4.50	-0.03	0.09	-0.20	0.10
Baseline	9	1.38	1.85	0.00	4.40				

	Score Category 13					Difference (Baseline - Sample) for Score 13			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	6	0.05	0.08	0.00	0.20	-0.02	0.04	-0.10	0.00
Sample 30%	6	0.03	0.05	0.00	0.10	0.00	0.00	0.00	0.00
Sample 40%	6	0.03	0.05	0.00	0.10	0.00	0.00	0.00	0.00
Sample 50%	6	0.03	0.05	0.00	0.10	0.00	0.00	0.00	0.00
Baseline	6	0.03	0.05	0.00	0.10				

	Score Category 14					Difference (Baseline - Sample) for Score 14			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	6	0.32	0.34	0.10	1.00	-0.02	0.04	-0.10	0.00
Sample 30%	6	0.32	0.34	0.10	1.00	-0.02	0.04	-0.10	0.00
Sample 40%	6	0.30	0.35	0.10	1.00	0.00	0.00	0.00	0.00
Sample 50%	6	0.30	0.35	0.10	1.00	0.00	0.00	0.00	0.00
Baseline	6	0.30	0.35	0.10	1.00				

	Score Category 15					Difference (Baseline - Sample) for Score 15			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	6	1.03	0.81	0.40	2.60	-0.12	0.10	-0.30	0.00
Sample 30%	6	1.07	0.85	0.40	2.70	-0.15	0.14	-0.40	0.00
Sample 40%	6	1.05	0.81	0.40	2.60	-0.13	0.10	-0.30	0.00
Sample 50%	6	0.97	0.79	0.30	2.50	-0.05	0.08	-0.20	0.00
Baseline	6	0.92	0.72	0.30	2.30				

Table F.31 EOY Item-to-Total Correlations for the Equating Data Sets for Algebra 2

	Item-to-Total Correlation					Difference (Baseline - Sample)			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	122	0.436	0.142	0.122	0.748	0.003	0.013	-0.033	0.037
Sample 30%	122	0.434	0.142	0.119	0.744	0.006	0.011	-0.021	0.036
Sample 40%	122	0.434	0.142	0.103	0.739	0.006	0.010	-0.017	0.039
Sample 50%	122	0.436	0.142	0.114	0.741	0.003	0.007	-0.013	0.026
Baseline	122	0.439	0.141	0.124	0.742				

Table F.32 PBA Item-to-Total Correlations for the Equating Data Sets for Algebra 2

		Item-to-Total Correlation				Difference (Baseline - Sample)			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	78	0.495	0.148	0.024	0.774	0.003	0.016	-0.033	0.052
Sample 30%	78	0.491	0.147	0.023	0.772	0.007	0.014	-0.028	0.059
Sample 40%	78	0.491	0.147	0.018	0.774	0.006	0.013	-0.027	0.055
Sample 50%	78	0.493	0.146	0.014	0.775	0.004	0.009	-0.020	0.026
Baseline	78	0.497	0.145	0.021	0.779				

Table F.33 EOY Item-to-Total Correlations for the Equating Data Sets for Integrated Math II

		Item-to-Total Correlation				Difference (Baseline - Sample)			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	49	0.333	0.134	0.012	0.573	0.056	0.049	-0.024	0.145
Sample 30%	49	0.384	0.134	0.071	0.621	0.006	0.027	-0.049	0.072
Sample 40%	49	0.394	0.140	0.059	0.635	-0.004	0.024	-0.065	0.073
Sample 50%	49	0.405	0.133	0.113	0.638	-0.015	0.016	-0.048	0.031
Baseline	49	0.390	0.128	0.101	0.616				

Table F.34 PBA Item-to-Total Correlations for the Equating Data Sets for Integrated Math II

		Item-to-Total Correlation				Difference (Baseline - Sample)			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	28	0.436	0.127	0.099	0.697	0.058	0.046	-0.031	0.129
Sample 30%	28	0.492	0.135	0.109	0.712	0.002	0.036	-0.078	0.094
Sample 40%	28	0.500	0.131	0.157	0.726	-0.006	0.027	-0.052	0.046
Sample 50%	28	0.508	0.128	0.205	0.728	-0.015	0.016	-0.036	0.032
Baseline	28	0.494	0.128	0.203	0.719				

Table F.35 EOY Item-to-Total Correlations for the Equating Data Sets for ELA/L Grade 6

		Item-to-Total Correlation				Difference (Baseline - Sample)			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	73	0.463	0.095	0.190	0.658	0.003	0.006	-0.013	0.014
Sample 30%	73	0.463	0.095	0.190	0.659	0.003	0.005	-0.011	0.012
Sample 40%	73	0.464	0.095	0.190	0.661	0.002	0.004	-0.007	0.011
Sample 50%	73	0.464	0.096	0.188	0.662	0.002	0.003	-0.006	0.009
Baseline	73	0.466	0.096	0.192	0.667				

Table F.36 PBA Item-to-Total Correlations for the Equating Data Sets for ELA/L Grade 6

		Item-to-Total Correlation				Difference (Baseline - Sample)			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	69	0.452	0.173	0.170	0.852	0.006	0.006	-0.007	0.016
Sample 30%	69	0.452	0.173	0.169	0.850	0.005	0.005	-0.005	0.014
Sample 40%	69	0.453	0.172	0.171	0.849	0.005	0.004	-0.003	0.012
Sample 50%	69	0.453	0.172	0.172	0.848	0.005	0.004	-0.003	0.012
Baseline	69	0.458	0.171	0.171	0.849				

Table F.37 EOY Item-to-Total Correlations for the Equating Data Sets for Math 5

		Item-to-Total Correlation				Difference (Baseline - Sample)			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	144	0.462	0.104	0.174	0.724	0.001	0.007	-0.022	0.017
Sample 30%	144	0.462	0.104	0.170	0.724	0.001	0.006	-0.016	0.017
Sample 40%	144	0.463	0.104	0.178	0.726	0.000	0.005	-0.018	0.011
Sample 50%	144	0.463	0.104	0.179	0.727	0.000	0.004	-0.013	0.010
Baseline	144	0.463	0.104	0.176	0.727				

Table F.38 PBA Item-to-Total Correlations for the Equating Data Sets for Math 5

		Item-to-Total Correlation				Difference (Baseline - Sample)			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	62	0.532	0.153	0.209	0.846	0.003	0.005	-0.012	0.014
Sample 30%	62	0.532	0.153	0.209	0.844	0.004	0.005	-0.008	0.014
Sample 40%	62	0.535	0.153	0.215	0.845	0.000	0.004	-0.012	0.013
Sample 50%	62	0.535	0.154	0.212	0.845	0.000	0.004	-0.009	0.016
Baseline	62	0.535	0.153	0.212	0.846				

Table F.39 EOY Item-to-Total Correlations for the Equating Data Sets for ELA/L Grade 3

		Item-to-Total Correlation				Difference (Baseline - Sample)			
EOY	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	36	0.502	0.125	0.201	0.690	0.002	0.005	-0.008	0.011
Sample 30%	36	0.502	0.125	0.198	0.690	0.002	0.004	-0.006	0.009
Sample 40%	36	0.503	0.124	0.200	0.689	0.001	0.002	-0.005	0.006
Sample 50%	36	0.502	0.125	0.202	0.689	0.001	0.003	-0.005	0.006
Baseline	36	0.504	0.125	0.197	0.689				

Table F.40 PBA Item-to-Total Correlations for the Equating Data Sets for ELA/L Grade 3

		Item-to-Total Correlation				Difference (Baseline - Sample)			
PBA	N	Mean	STD	Min	Max	Mean	STD	Min	Max
Sample 25%	60	0.516	0.142	0.142	0.819	0.004	0.004	-0.006	0.016
Sample 30%	60	0.517	0.142	0.146	0.820	0.003	0.004	-0.004	0.014
Sample 40%	60	0.518	0.141	0.149	0.820	0.002	0.003	-0.007	0.010
Sample 50%	60	0.518	0.141	0.150	0.819	0.002	0.003	-0.003	0.007
Baseline	60	0.520	0.141	0.151	0.818				