Achievement and Growth Norms for Course-Specific MAP® Growth™ Algebra 1, Geometry, and Algebra 2 Tests

July 2019

Yeow Meng Thum, Ph.D.
Wei He, Ph.D.

Suggested citation: Thum, Y. M., & He, W. (2019). *Achievement and growth norms for course-specific MAP® Growth™ Algebra 1, Geometry, and Algebra 2 tests*. NWEA Research Report. Portland, OR: NWEA.

Table of Contents

List of Tables

List of Figures

Acknowledgements

This paper benefitted from reviews by colleagues, most notably Dr. Nate Jensen, Dr. Andy Hegedus, and Dr. Patrick Meyer; contributions from Bill Noland; and the expert copy-editing efforts of Kelly Larson and Debbie Bauman. The study also benefited from the quality check effort by Dr. Emily Bo and Dr. Sylvia Li.

1. Introduction

1.1. Purpose of the Study

This report describes the norming procedures used to produce the user norms—including fall, winter, and spring achievement and fall-to-spring growth norms—for the course-specific MAP® Growth™ Mathematics tests in Algebra 1, Geometry, and Algebra 2. Specifically, the norming procedures include the selection of a norming sample and the use of a model-based approach (i.e., a multivariate true score model that factors out known imprecision of scores) to generate the norms. This report also provides snapshots and an explanation of the resulting achievement and growth norms.

1.2. Overview of the Course-Specific MAP Growth Mathematics Tests

In August 2017, NWEA® released two suites of course-specific MAP Growth Mathematics tests aligned to the NWEA standards and the Common Core State Standards (CCSS) for Mathematics. The NWEA-aligned tests replaced the older NWEA Mathematics End-of-Course (EOC) tests that had been used previously. These NWEA-aligned tests use the same instructional areas and subareas as their EOC predecessors. The blueprints of the NWEA-aligned tests are designed to capture current trends and accepted best practices in Mathematics curricula. The CCSS-aligned tests were developed by following the recommended course content from Appendix A of the CCSS for Mathematics (NGA Center for Best Practices & CCSSO, 2010). The CCSS-aligned tests have some overlap in content with the NWEA suite, but these tests are organized with different instructional areas and subareas. Both the NWEA and CCSS suite of tests include items that assess course pre-requisites to better assess specific course readiness.

Following the 2017 release, there was an increased awareness of course-specific MAP Growth tests by partners. This awareness created an increased demand for the development of course-specific tests aligned directly to different state standards. In August 2018, NWEA released course-specific tests for Florida, Missouri, Texas, and Virginia. The design of these tests was based on the blueprints established by the respective state boards of education for each EOC examination within that state. Like the CCSS- and NWEA-aligned tests, these tests also include items that assess course prerequisites to better assess specific course readiness.

Within each course there is a certain degree of overlap in content assessed when compared across the three versions of course-specific tests (NWEA, CCSS, and state specific). The Algebra 1 content assessed in the NWEA suite is similar to the content assessed in the CCSS- and state-specific versions of Algebra 1. This also holds true for Algebra 2 and Geometry content across suites. Table 1.1 summarizes the course-specific MAP Growth Mathematic tests included in this study.

Table 1.1. Course-Specific MAP Growth Mathematic Tests Included in this Study

Algebra 1	Geometry	Algebra 2
Growth: Algebra 1 CCSS 2010	Growth: Geometry CCSS 2010	Growth: Algebra 2 CCSS 2010
Growth: Algebra 1 NWEA 2017	Growth: Geometry NWEA 2017	Growth: Algebra 2 NWEA 2017
Growth: Algebra 1 FL 2014	Growth: Geometry FL 2014	Growth: Algebra 2 MO 2016
Growth: Algebra 1 MO 2016		
Growth: Algebra 1 TX 2012		
Growth: Algebra 1 VA 2016		

Like other MAP Growth assessments, the course-specific MAP Growth tests are item-level computerized adaptive tests (CATs) in which items that yield the best information about an examinee's interim ability are sequentially selected for administration. The Rasch model, an item response theory (IRT) model commonly used in large-scale assessments, is used for scaling items and scoring the tests. A randomesque item exposure control procedure described in Kingsbury and Zara (1989) is used to select one out of several items that provides the best information about an examinee. To ensure that the content of a test matches the intended test blueprint, the tests employ a content-balancing method that selects items from the least represented instructional area according to its target administration value specified in the test blueprint (Kingsbury & Zara, 1991). The maximum likelihood estimation (MLE) method is used to estimate abilities for these variable-length tests ranging from 41 to 43 items.

These course-specific tests share the same scale as the regular MAP Growth Mathematics tests. In particular, their scores are also expressed as Rasch Unit (RIT). However, a score of 220 on a course-specific test should not be used interchangeably with a score of 220 on MAP Growth Mathematics because they test different subject domains.

Different from the prior NWEA EOC tests taken only at the end of a course, these course-specific tests can be administered multiple times throughout the school year, typically in the fall, winter, and spring. This allows for student growth to be evaluated in a content area over the duration of a course. The adaptive nature of these assessments yields much greater measurement precision than a traditional linear test of similar length, making these course-specific tests well suited for measuring growth.

2. Methodology

Norms describe the performance of students relative to a target population. In status norms, a student's performance on the test is associated with a percentile ranking that shows how well the student performed in a content area compared to students in the norming group. The relative evaluation of a student's growth from one period to another (e.g., from fall to spring) is provided by suitably constructed growth norms. This section describes the methods used in this study to select the norming sample and generate the achievement and growth norms.

2.1. Norming Sample Selection

Unlike the nationally representative norms described in the 2015 MAP Growth norms study (Thum & Hauser, 2015), this norming study was designed and conducted to support inferences about the student's performance in MAP Growth Algebra 1, Geometry, and Algebra 2, respectively, with reference to students who took these tests from Fall 2017 to Spring 2019.

Most U.S. public high school students must earn at least three credits of Mathematics to meet graduation requirements. The typical pathway includes Algebra 1, Geometry, and Algebra 2, offered in that order to students in Grades 9, 10, and 11 consecutively. The length of each course is typically a year. However, some middle school students, typically advanced students, often take these tests, and some high school students, typically low-performing students, take these courses in the upper grades of high school. Table 2.1 reports the number of test events in each subject across grades, terms, and school years. It reflects the course-taking sequence that most students took Algebra 1, Geometry, and Algebra 2 in Grades 9, 10, and 11, respectively, but also suggests students who took these tests were enrolled in grades between 6 and 12.

To make sure this norming study represented all students who took a subject-specific mathematics course, students in Grades 6–12 who took a course-specific test in either the 2017 or 2018 school year were included. That is, Grades 6–12 students who took Algebra 1, Geometry, and Algebra 2 in either 2017 or 2018, but not both school years, were used as the norming samples in each subject. This approach compares the results of a student to fellow students who has taken the same course, thus best preserving a consistent vertical scale interpretation of scores and the relative percentile comparisons among all students taking a test. If a student has a higher score than another student, they will also receive the higher percentile rank regardless of the grade in which the student is enrolled. For example, on the score scale, a RIT score of 210 always indicates higher relative performance than a RIT score of 200.

This norming sample selection approach resulted in 747,936 course-specific MAP Growth test events administered to 342,821 students from 50 states between Fall 2017 and Spring 2019 (i.e., the first two years after the course-specific Mathematics tests were released). Among these test events, 452,942 were from 230,725 students who took Algebra 1, 190,292 were from 96,966 students who took Geometry, and 104,702 were from 54,270 students who took Algebra 2, as shown in Table 2.1.

Table 2.1. Number of Test Events from Fall 2017 to Spring 2019

Course-Specific Test	Grade	Number of Test Events									
		2017 School Year			2018 School Year			2017+2018 School Year			
		Fall 2017	Winter 2018	Spring 2018	Fall 2018	Winter 2019	Spring 2019	Fall	Winter	Spring	Total
Algebra 1	6	98	59	137	221	259	316	319	318	453	**1,090**
	7	2,997	2,549	3,595	5,863	5,773	6,349	8,860	8,322	9,944	**27,126**
	8	14,048	13,914	18,783	26,943	24,295	25,989	40,991	38,209	44,772	**123,972**
	9	34,894	25,691	29,637	60,110	49,078	51,128	95,004	74,769	80,765	**250,538**
	10	7,410	4,908	5,032	7,088	6,069	5,804	14,498	10,977	10,836	**36,311**
	11	2,084	1,414	1,288	2,304	1,858	1,576	4,388	3,272	2,864	**10,524**
	12	664	414	333	826	711	433	1,490	1,125	766	**3,381**
	Total	**62,195**	**48,949**	**5,8805**	**103,355**	**88,043**	**91,595**	**165,550**	**136,992**	**150,400**	**452,942**
Geometry	6	8	8	5	18	12	8	26	20	13	**59**
	7	63	27	59	115	158	133	178	185	192	**555**
	8	2,097	1,595	2,207	4,013	3,815	4,187	6,110	5,410	6,394	**17,914**
	9	5,339	4,585	5,356	9,264	6,922	8,741	14,603	11,507	14,097	**40,207**
	10	15,096	12,463	13,735	27,742	20,826	24,213	42,838	33,289	37,948	**114,075**
	11	2,382	1,889	1,748	3,553	2,736	2,884	5,935	4,625	4,632	**15,192**
	12	350	228	263	661	420	368	1,011	648	631	**2,290**
	Total	**25,335**	**20,795**	**23,373**	**45,366**	**34,889**	**40,534**	**70,701**	**55,684**	**63,907**	**190,292**
Algebra 2	6	2	6	4	9	7	2	11	13	6	**30**
	7	13	21	7	18	17	16	31	38	23	**92**
	8	156	176	159	182	247	195	338	423	354	**1,115**
	9	794	569	815	1,955	1,035	1,712	2,749	1,604	2,527	**6,880**
	10	4,701	3,830	4,492	9,820	7,224	8,579	14,521	11,054	13,071	**38,646**
	11	6,378	6,048	5,039	12,819	9,751	10,436	19,197	15,799	15,475	**50,471**
	12	1,083	909	620	2,043	1,550	1,263	3,126	2,459	1,883	**7,468**
	Total	**13,127**	**11,559**	**11,136**	**26,846**	**19,831**	**22,203**	**39,973**	**31,390**	**33,339**	**104,702**
										Grand Total	**747,936**

2.2. Building Achievement and Growth Scales

The norming procedure was a model-based approach employing a multivariate true score model that factors out known imprecision of scores from the fall, winter, and spring test scores of examinees in the selected norming population. This procedure provided norms for student achievement status for each term and growth norms for students' gains between fall and spring.

This norming approach recognizes that a model of learning growth supplies the basis for making simultaneous inferences about achievement and growth (Thum & Hauser, 2015). In this setting, a multivariate true score model is considered for fall, winter, and spring test scores of examinees in the user population for each test. The true score model is defined in Equation 1:

$$\begin{bmatrix} y_{1i} & y_{2i} & y_{3i} \end{bmatrix} = \begin{bmatrix} \mu_{1i} & \mu_{2i} & \mu_{3i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1i} & \varepsilon_{2i} & \varepsilon_{3i} \end{bmatrix} , \quad (1)$$

where:

- $\mathbf{y}_i = (y_{1i}, y_{2i}, y_{3i})$ are the observed scores for examinee i.
- $\boldsymbol{\mu}_i = (\mu_{1j}, \mu_{2j}, \mu_{3j})$ are the true scores for examinee i.
- $\boldsymbol{\varepsilon}_i = (\varepsilon_{1i}, \varepsilon_{2i}, \varepsilon_{3i})$ are the error scores for examinee i.

The analysis considers the imprecision of observed scores by introducing the observed standard errors of measurement (SEMs) of each score (s_{1j}, s_{2j}, s_{3j}) into the model, such that:

$$\mathrm{Var}(\varepsilon_{1j}) = s_{1j}^2 ,\ \mathrm{Var}(\varepsilon_{2j}) = s_{2j}^2 ,\ \text{and } \mathrm{Var}(\varepsilon_{3j}) = s_{3j}^2 . \quad (2)$$

True scores of examinees are assumed to be distributed as a multivariate normal distribution in the user population:

$$\boldsymbol{\mu}_i \sim \mathrm{MVN}[\boldsymbol{\gamma}_i, \mathbf{T}] \quad (3)$$

Restricted maximum likelihood estimates, $\hat{\boldsymbol{\gamma}}$, $\mathrm{Var}(\hat{\boldsymbol{\gamma}})$, and $\hat{\mathbf{T}}$, are easily obtained by standard statistics packages such as HLM7 or SAS Proc Mixed. These estimates define the joint distribution of predicted fall, winter, and spring scores defined in Equation 4 in the user norming population:

$$\hat{\boldsymbol{\mu}}_i \sim \mathrm{MVN}\left[\hat{\boldsymbol{\gamma}}, \mathrm{Var}(\hat{\boldsymbol{\gamma}}) + \hat{\mathbf{T}}\right] \quad (4)$$

This joint distribution provides the basis for constructing achievement and growth norms. Achievement norms for fall, winter, and spring scores $(\hat{\mu}_{1j}, \hat{\mu}_{2j}, \hat{\mu}_{3j})$ are derived from the predicted marginal distributions, as are the marginal fall-to-spring growth norms $(\hat{\mu}_{3j} - \hat{\mu}_{1j})$. Fall -to-spring conditional gains for examinees, with a specific fall score $\hat{\mu}_{1j}$, are obtained as the predicted distribution of $(\hat{\mu}_{3j} - \hat{\mu}_{1j} \mid \hat{\mu}_{1j})$.

3. Results

3.1. Summary Statistics

Table 3.1 presents the mean and standard deviation (SD) of RIT test scores for students in Grades 6–12 with at least 100 test events, along with the overall mean and SD of RIT scores for the norming samples in each subject. With few exceptions, average test scores decreased as grades increased for each course-specific test. Lower-grade students (i.e., Grades 6–8 students) tend to perform better than the students of the grade at which a course is usually targeted, and the upper-grade high school students tend to perform worse than the students of the grade at which a course is usually targeted. Grade 7 students achieved the highest average test score in Algebra 1 and Geometry, and Grade 8 students achieved the highest average test score in Algebra 2. In general, lower-grade students tend to grow more fall to spring for almost all grades compared with high school students. By and large, higher self-selection on ability or readiness in the earlier grade levels is quite evident from the cross-grade data.

Table 3.1. Summary Descriptive Statistics of Sample Test Scores

Grade		Algebra 1			Geometry			Algebra 2		
		Fall	Winter	Spring	Fall	Winter	Spring	Fall	Winter	Spring
6	Mean	230.50	245.06	253.67						
	SD	25.92	23.42	24.82						
	N	319	318	453						
7	Mean	244.48	250.56	257.89	261.54	261.26	268.09			
	SD	11.37	11.66	13.54	19.88	18.51	18.00			
	N	8,860	8,322	9,944	178	185	192			
8	Mean	239.67	244.90	250.58	250.80	256.40	265.78	260.89	256.74	272.67
	SD	12.01	13.16	15.07	12.50	11.93	12.89	21.87	22.01	19.59
	N	40,991	38,209	44,772	6,110	5,410	6,394	338	423	354
9	Mean	227.76	230.33	234.71	243.48	248.16	254.43	251.28	257.11	260.79
	SD	15.52	16.22	17.00	13.89	14.89	16.07	16.98	16.72	18.79
	N	95,004	74,769	80,765	14,603	11,507	14,097	2,749	1,604	2,527
10	Mean	225.14	227.86	230.35	230.96	233.59	238.76	245.89	249.48	253.78
	SD	17.72	18.28	19.01	13.61	14.82	15.73	14.99	16.44	17.13
	N	14,498	10,977	10,836	42,838	33,289	37,948	14,521	11,054	13,071
11	Mean	223.76	225.84	228.03	224.63	227.71	230.53	236.67	238.89	241.82
	SD	17.61	18.46	18.57	13.24	14.23	15.05	14.67	15.66	15.93
	N	4,388	3,272	2,864	5,935	4,625	4,632	19,197	15,799	15,475
12	Mean	224.33	226.60	228.80	222.68	226.04	228.64	234.00	235.20	236.66
	SD	17.86	17.60	18.07	14.20	15.14	15.29	15.65	16.17	17.09
	N	1,490	1,125	766	1,011	648	631	3,126	2,459	1,883
Overall	**Mean**	**231.24**	**235.32**	**240.55**	**234.69**	**238.34**	**244.31**	**241.01**	**243.50**	**248.00**
	SD	**16.12**	**17.31**	**18.73**	**15.62**	**17.03**	**18.59**	**16.17**	**17.39**	**18.39**
	N	**165,550**	**136,992**	**150,400**	**70,701**	**55,684**	**63,907**	**39,973**	**31,390**	**33,339**

3.2. Normality Assumption

Inferences based on the multivariate true score models relied on the reasonableness of the joint normality assumption of score components for their validity. Normality was examined from different perspectives such as quantile-quantile (Q-Q) plots, cumulative distribution function (CDF) curves for RIT scores, and residuals from model estimation. Figure 3.1, Figure 3.2, and Figure 3.3 present a series of graphs including histograms, Q-Q plots, and CDF curves based on RIT scores (left panel of the figure) and residuals from model estimation (right panel of the figure) for Algebra 1. The Q-Q plots indicate that most of the data fall close to the 45-degree reference line except at the very low and high ends, suggesting that normality was a reasonably good approximation. The two CDF curves also reasonably overlap with each other. These observations hold true for both RIT score and residuals for the true score model. In general, these graphs support the assumption of marginal normality for the Algebra 1 test. Normality assumptions of the model also seemed reasonable for Algebra 1 upon examining the scatterplots in Figure 3.4 for each pair of RIT scores and residuals from model estimation.

The same graphs in Appendix A for Geometry and Algebra 2 resemble those for Algebra 1, suggesting that normality is also a reasonably good approximation for those tests.

Figure 3.1. Histograms, Q-Q Plots, and CDFs for Algebra 1 Fall Scores

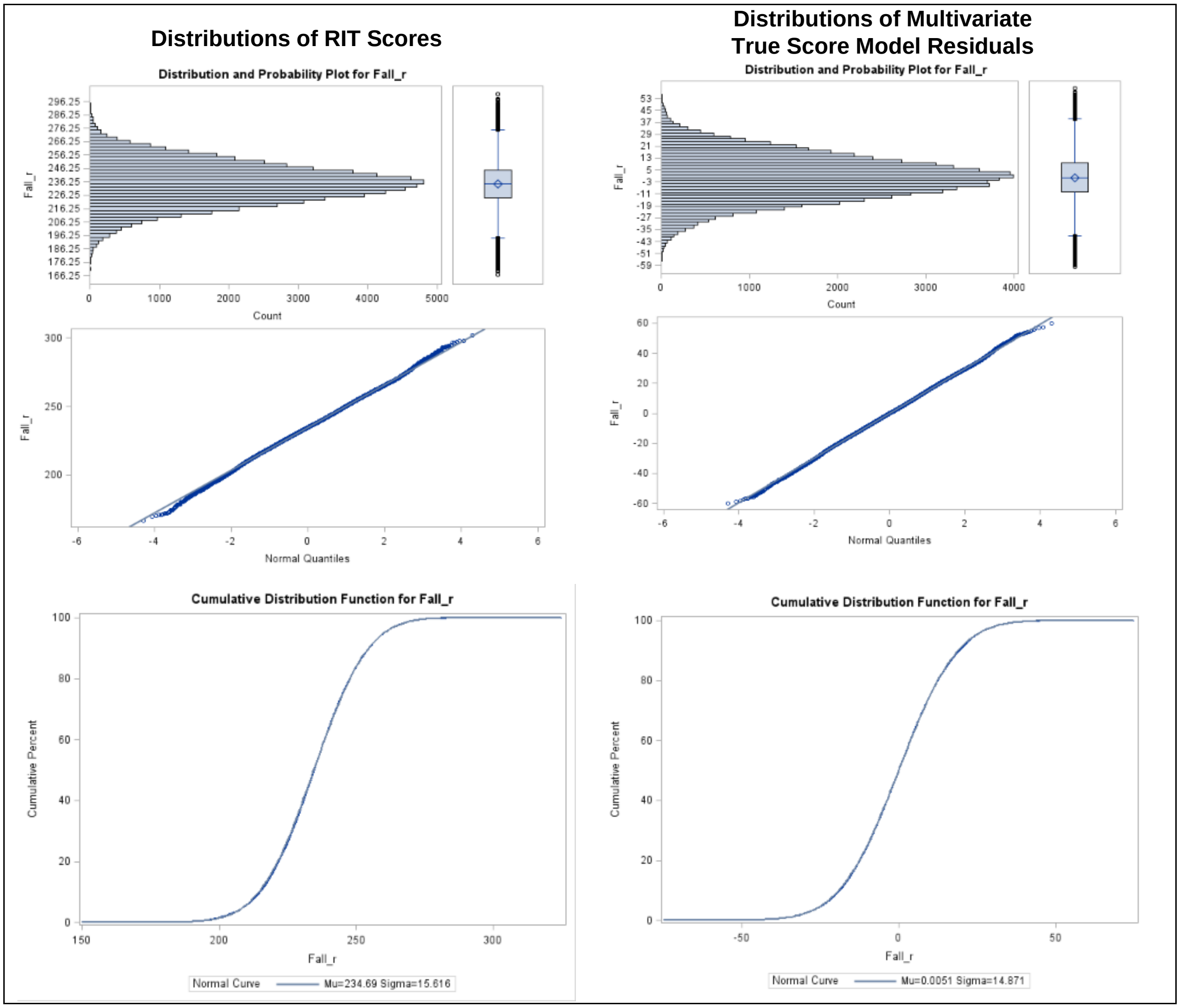


Figure 3.2. Histograms, Q-Q Plots, and CDFs for Algebra 1 Winter Scores

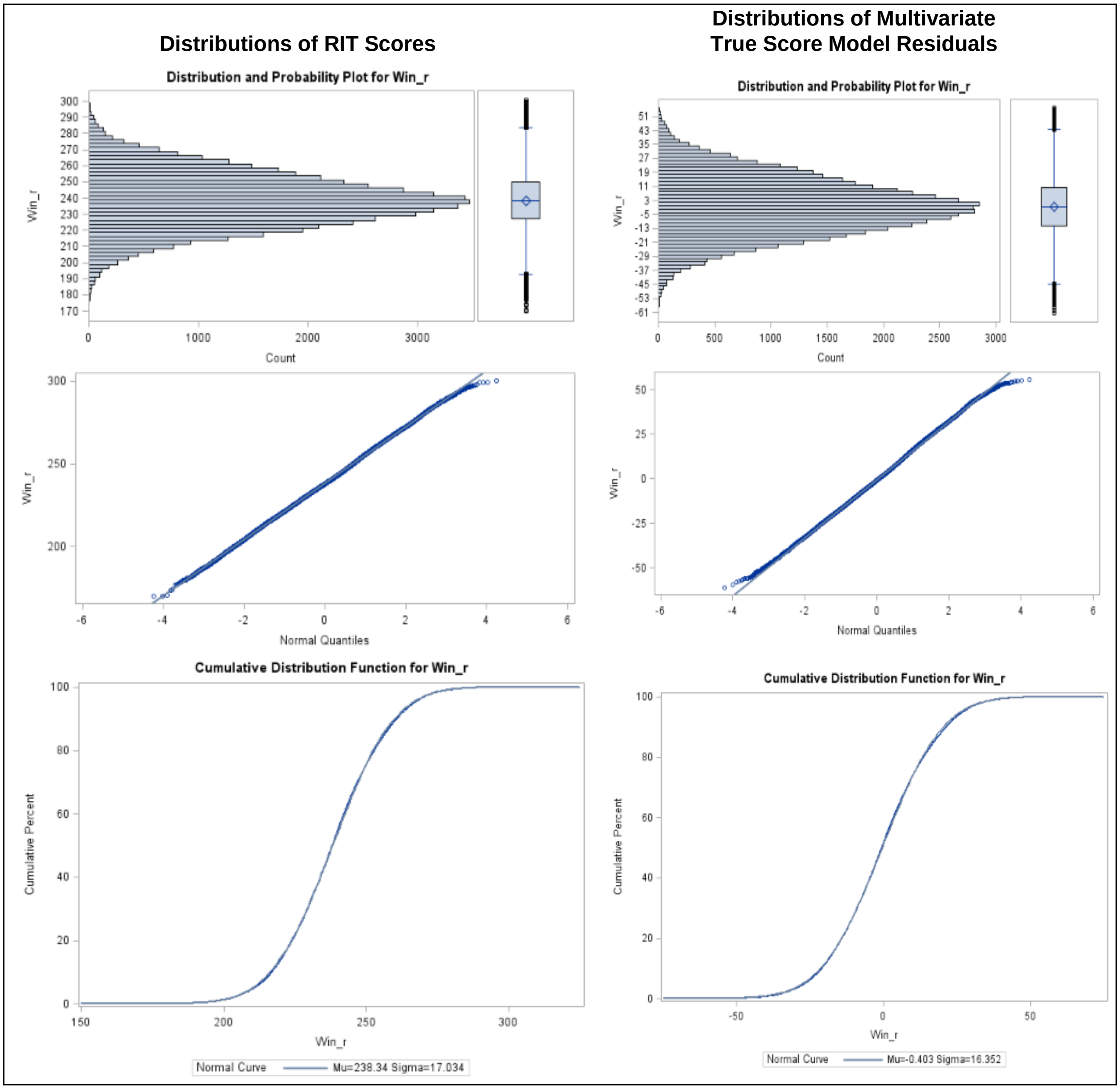


Figure 3.3. Histograms, Q-Q Plots, and CDFs for Algebra 1 Spring Scores

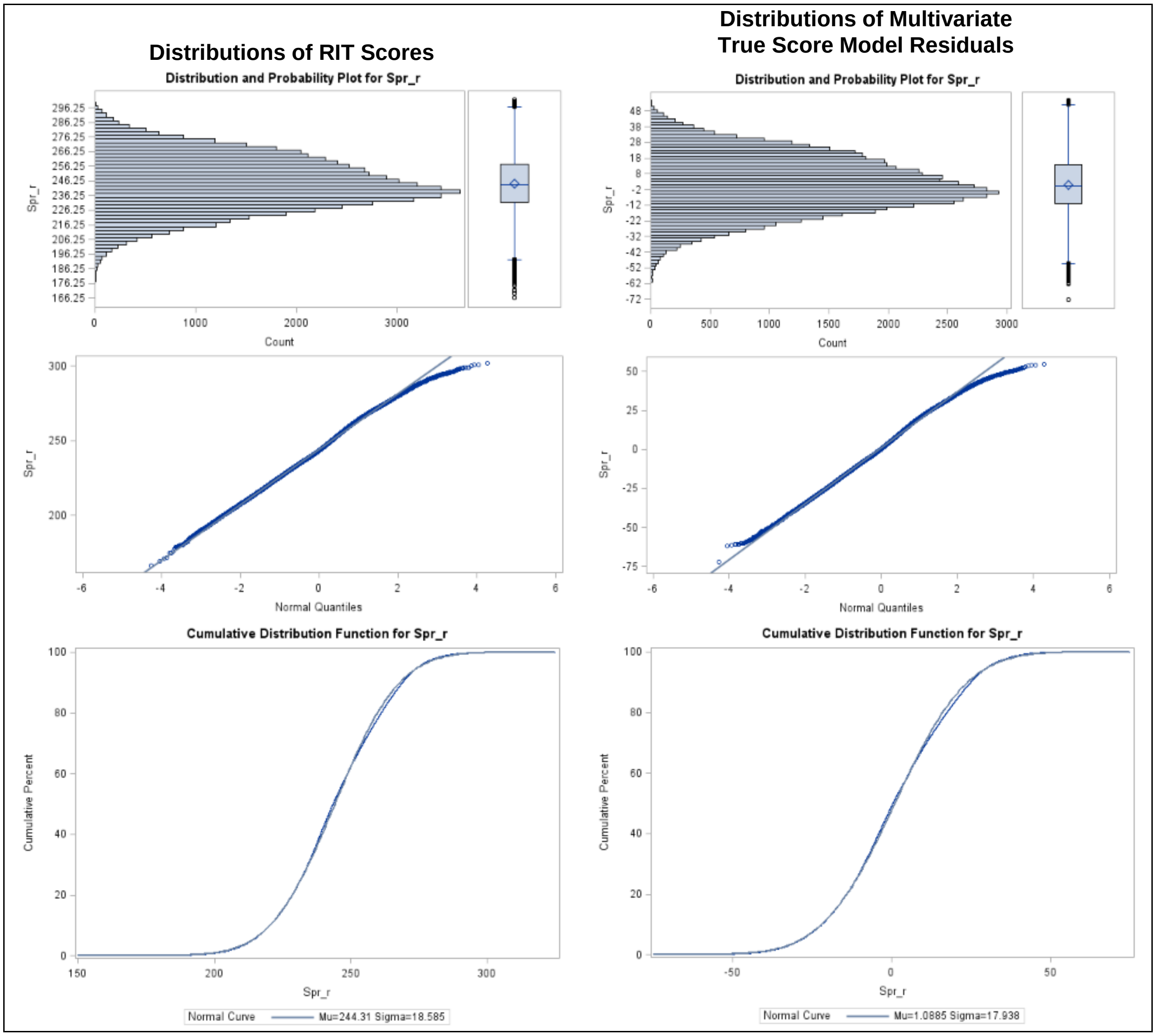
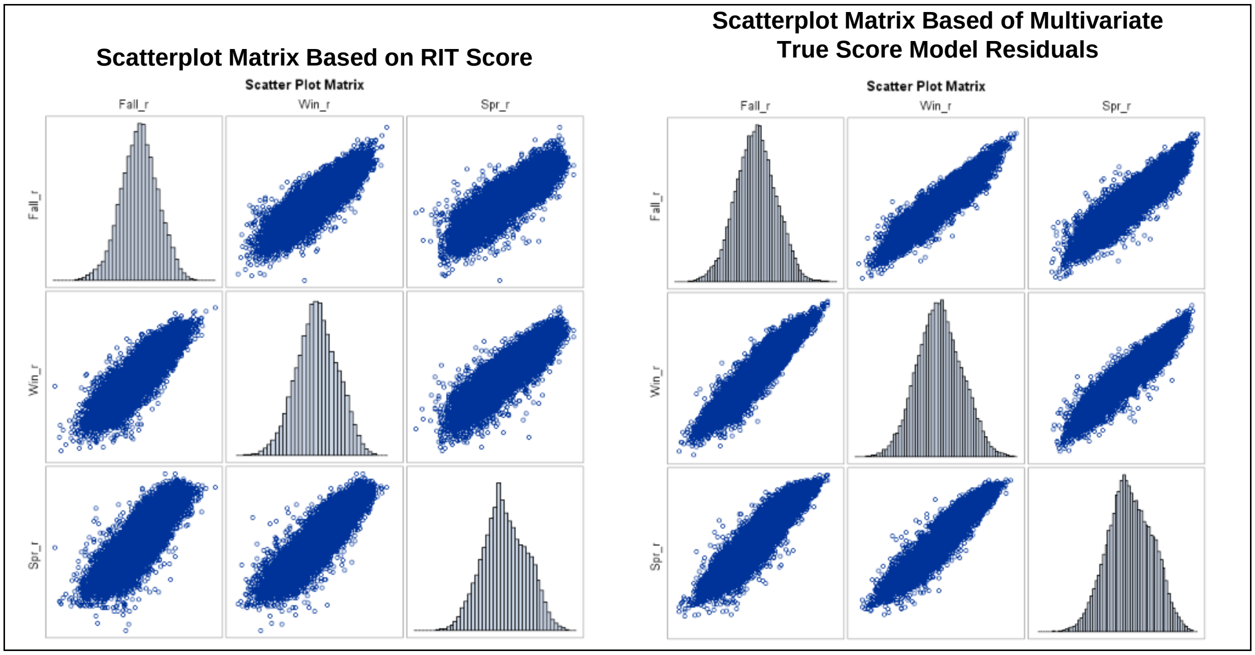


Figure 3.4. Scatterplot Matrix among Fall, Winter, and Spring Scores for Grade 9 Test Takers for Algebra 1



3.3. Pearson Correlation Coefficients

Table 3.2 presents the relationship of scores between administrations in the form of Pearson correlation coefficients (*r*) using observed RIT scores and estimates from the true score models (i.e., correlations between scores in fall vs. winter, fall vs. spring, and winter vs. spring). The bolded coefficients were computed based on the estimates from the true score models, whereas the non-bolded coefficients were computed based on the observed RIT scores. Specifically, correlations between true scores in the user population were given by the correlations between random effects estimated by the true score models. These coefficients are more appropriate than the observed bivariate correlation coefficients to be used to evaluate the magnitude of score relationship due to the missingness in the observed data and the imprecision of observed scores. As shown in the table, the Pearson correlation coefficients computed based on the estimates from the true score models are above 0.90 for almost all tests, suggesting that scores from each administration were strongly correlated. The correlation coefficients based on the estimates from the true score models are corrected for attenuation (e.g., Bock & Petersen, 1975) and are therefore higher than those from the observed scores.

Table 3.2. Pearson Correlation Coefficients (r) among Fall, Winter, and Spring Scores

Course-Specific Test	*r*		
	Fall, Winter	Fall, Spring	Winter, Spring
Algebra 1	**0.91**	**0.88**	**0.92**
	0.86	0.82	0.87
Geometry	**0.92**	**0.91**	**0.93**
	0.88	0.86	0.89
Algebra 2	**0.91**	**0.89**	**0.91**
	0.85	0.83	0.86

*Bolded coefficients are correlations corrected for attenuation.

3.4. Status and Growth Norms

Figure 3.5, Figure 3.6, and Figure 3.7 present snapshots of the status and fall-to-spring growth norms and their associated percentiles for each test, as well as the expected fall-to-spring gain and SD of predicted growth score. For ease of presentation, not every possible percentile is provided in these figures. The numbers in the yellow box under “Spring Percentile and Score” indicate spring status norms and their corresponding percentiles. The rest of the numbers in the mixed color box indicate the growth percentiles associated with fall-to-spring growth scores. The meanings of the figures’ acronyms are as follows:

- Fall-Spring Cond. Growth Norms = Fall-to-spring conditional growth norms
- Mean = Expected fall-to-spring growth given a fall score
- SD = Standard deviation of fall-to-spring growth given a fall score

Figure 3.5. Snapshot of Status and Growth Norms for Algebra 1

	Term			Fall-Spring Cond. Growth		Spring Percentile and Score																		
						5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95
Percentile	Fall	Winter	Spring	Mean	SD	207	214	219	223	226	229	231	234	236	239	241	244	246	249	252	255	259	263	270
5	205	206	207	6.4	9.0	34	64	81	90	95	98	99	99	99	99	99	99	99	99	99	99	99	99	99
10	210	213	214	6.8	9.0	14	37	58	73	83	90	94	97	98	99	99	99	99	99	99	99	99	99	99
15	214	217	219	7.0	9.0	6	22	40	56	70	80	87	92	95	97	99	99	99	99	99	99	99	99	99
20	217	220	223	7.1	9.0	3	13	27	42	56	68	78	85	91	94	97	98	99	99	99	99	99	99	99
25	220	223	226	7.3	9.0	1	7	17	30	43	56	67	77	84	90	94	96	98	99	99	99	99	99	99
30	223	226	229	7.4	9.0	1	4	11	21	33	45	57	67	76	84	89	94	96	98	99	99	99	99	99
35	225	228	231	7.5	9.0	1	2	7	14	24	35	46	57	68	77	84	90	94	97	98	99	99	99	99
40	227	231	234	7.6	9.0	1	1	4	10	17	26	37	48	58	68	77	84	90	94	97	99	99	99	99
45	229	233	236	7.7	9.0	1	1	3	6	12	19	28	38	49	59	69	78	85	91	95	98	99	99	99
50	231	235	239	7.9	9.0	1	1	1	4	8	13	21	30	40	50	61	71	79	87	92	96	99	99	99
55	233	237	241	8.0	9.0	1	1	1	2	5	9	15	22	31	41	51	62	72	81	88	94	98	99	99
60	235	239	244	8.1	9.0	1	1	1	1	3	6	10	16	23	32	42	53	63	74	83	91	96	99	99
65	237	242	246	8.2	9.0	1	1	1	1	2	3	6	10	16	24	33	43	54	65	76	86	93	98	99
70	239	244	249	8.3	9.0	1	1	1	1	1	2	4	6	11	16	24	33	44	55	67	79	89	96	99
75	242	247	252	8.4	9.0	1	1	1	1	1	1	2	4	6	10	16	23	33	44	57	70	83	93	99
80	244	250	255	8.6	9.0	1	1	1	1	1	1	1	2	3	6	10	15	22	32	44	58	73	87	97
85	248	253	259	8.7	9.0	1	1	1	1	1	1	1	1	1	3	5	8	13	20	31	44	60	78	94
90	251	257	263	9.0	9.0	1	1	1	1	1	1	1	1	1	1	2	3	6	10	17	27	42	63	86
95	257	264	270	9.3	9.0	1	1	1	1	1	1	1	1	1	1	1	1	1	3	5	10	19	36	66

Figure 3.6. Snapshot of Status and Growth Norms for Geometry

	Term			Fall-Spring Cond. Growth		Spring Percentile and Score																		
						5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95
Percentile	Fall	Winter	Spring	Mean	SD	212	219	224	227	230	233	236	238	241	243	246	248	251	253	256	259	263	268	274
5	209	211	212	5.8	8.0	36	69	86	94	97	99	99	99	99	99	99	99	99	99	99	99	99	99	99
10	215	217	219	6.4	8.0	13	39	62	77	87	93	97	98	99	99	99	99	99	99	99	99	99	99	99
15	219	221	224	6.8	8.0	5	21	41	59	73	84	90	95	97	99	99	99	99	99	99	99	99	99	99
20	222	225	227	7.1	8.0	2	11	26	43	58	71	81	89	93	96	98	99	99	99	99	99	99	99	99
25	224	227	230	7.4	8.0	1	6	16	29	44	58	70	80	87	93	96	98	99	99	99	99	99	99	99
30	227	230	233	7.7	8.0	1	3	9	19	32	45	58	70	80	87	92	96	98	99	99	99	99	99	99
35	229	232	236	7.9	8.0	1	1	5	12	22	34	47	59	70	80	87	92	96	98	99	99	99	99	99
40	231	235	238	8.1	8.0	1	1	3	7	15	24	36	48	60	71	80	87	93	96	98	99	99	99	99
45	233	237	241	8.3	8.0	1	1	2	4	9	17	26	37	49	61	71	81	88	93	97	99	99	99	99
50	235	239	243	8.5	8.0	1	1	1	2	6	11	18	28	38	50	62	73	82	89	94	98	99	99	99
55	237	241	246	8.7	8.0	1	1	1	1	3	7	12	19	29	39	51	63	74	83	91	96	99	99	99
60	239	243	248	8.9	8.0	1	1	1	1	2	4	7	13	20	29	40	52	64	76	85	93	97	99	99
65	241	245	251	9.1	8.0	1	1	1	1	1	2	4	8	13	21	30	41	53	66	78	88	95	99	99
70	243	248	253	9.4	8.0	1	1	1	1	1	1	2	4	8	13	21	30	42	55	68	81	91	97	99
75	245	250	256	9.6	8.0	1	1	1	1	1	1	1	2	4	8	13	20	30	42	56	71	84	94	99
80	248	253	259	9.9	8.0	1	1	1	1	1	1	1	1	2	4	7	12	19	29	42	57	74	89	98
85	251	256	263	10.2	8.0	1	1	1	1	1	1	1	1	1	1	3	5	10	16	27	41	59	79	95
90	255	260	268	10.6	8.0	1	1	1	1	1	1	1	1	1	1	1	2	3	7	13	23	39	61	87
95	260	266	274	11.2	8.0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	3	6	14	31	64

Figure 3.7. Snapshot of Status and Growth Norms for Algebra 2

						Spring Percentile and Score																		
	Term			Fall-Spring Cond. Growth		5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95
Percentile	Fall	Winter	Spring	Mean	SD	216	223	227	231	234	237	239	242	244	246	249	251	254	256	259	262	266	270	277
5	214	215	216	5.2	8.4	35	66	83	92	96	98	99	99	99	99	99	99	99	99	99	99	99	99	99
10	220	222	223	5.3	8.4	13	38	59	75	85	92	95	98	99	99	99	99	99	99	99	99	99	99	99
15	224	226	227	5.4	8.4	6	22	40	57	71	81	88	93	96	98	99	99	99	99	99	99	99	99	99
20	227	229	231	5.4	8.4	2	12	26	42	57	69	79	87	92	95	97	99	99	99	99	99	99	99	99
25	230	232	234	5.5	8.4	1	7	17	30	44	57	68	78	85	91	95	97	99	99	99	99	99	99	99
30	232	235	237	5.5	8.4	1	4	10	20	32	45	57	68	78	85	91	95	97	99	99	99	99	99	99
35	235	237	239	5.6	8.4	1	2	6	14	23	35	46	58	69	78	85	91	95	97	99	99	99	99	99
40	237	240	242	5.6	8.4	1	1	4	9	16	25	36	48	59	69	78	86	91	95	98	99	99	99	99
45	239	242	244	5.7	8.4	1	1	2	5	11	18	27	38	49	60	70	79	87	92	96	98	99	99	99
50	241	244	246	5.7	8.4	1	1	1	3	7	12	20	29	39	50	61	71	80	88	93	97	99	99	99
55	243	246	249	5.7	8.4	1	1	1	2	4	8	14	21	30	40	51	62	73	82	89	95	98	99	99
60	245	249	251	5.8	8.4	1	1	1	1	2	5	9	14	22	31	41	52	64	75	84	91	96	99	99
65	247	251	254	5.8	8.4	1	1	1	1	1	3	5	9	15	22	32	42	54	66	77	87	94	98	99
70	249	253	256	5.9	8.4	1	1	1	1	1	1	3	6	9	15	22	32	43	55	68	80	90	96	99
75	252	256	259	5.9	8.4	1	1	1	1	1	1	1	3	5	9	15	22	32	43	56	70	83	93	99
80	254	259	262	6.0	8.4	1	1	1	1	1	1	1	1	3	5	8	14	21	31	43	58	74	88	98
85	258	262	266	6.0	8.4	1	1	1	1	1	1	1	1	1	2	4	7	12	19	29	43	60	79	95
90	262	267	270	6.1	8.4	1	1	1	1	1	1	1	1	1	1	1	2	5	9	15	25	41	62	87
95	268	273	277	6.2	8.4	1	1	1	1	1	1	1	1	1	1	1	1	1	2	4	8	17	34	65

Using the norms for the Algebra 1 test in Figure 3.5 as an example, the 55th achievement percentile scores for fall, winter, and spring are 233, 237, and 241, respectively. The expected fall-to-spring gain for a student who starts in the fall at the 55th percentile score (233) is 8 with an associated SD of growth of 9. This indicates that students who perform at the 55th percentile in the fall test tend to gain 8 RITs of growth, on average, from fall to spring.

Figure 3.5 allows the reader to normatively evaluate the actual gain a student may have made from fall to spring. For example, if a student who scores 233 in the fall (55th percentile) obtains a score of 244 in the spring (60th percentile), this student has improved 11 RITs from fall to spring. Locating the intersection in Figure 3.5, corresponding to the row where the achievement percentile is 55 and the column where the spring score percentile is 60, the 11 fall-to-spring RIT gain puts this student at the 62nd percentile in the specific growth scale.

Recall that the reference group for each test consisted of students who received instruction in that course. This implies that, in the example explained above, if a student obtains a score of 233 on the Fall Algebra 1 test regardless of the grade they are in, this student has performed better than 55 percent of the students who take the Algebra 1 test. Further, if this student obtains a score of 244 on the Algebra 1 test in the spring, improving by 11 RITs from fall to spring, this student has made better progress than 62 percent of the students whose fall scores are 233. The interpretation of the scores also follows suit for the Geometry and Algebra 2 tests.

4. Conclusion and Discussion

This study documents the procedure used to develop the achievement status and growth user norms for course-specific MAP Growth Mathematics tests in Algebra 1, Geometry, and Algebra 2. The cross-grade data used in this norming study reveal a more realistic picture in taking these advanced course-specific mathematics courses in U.S. schools. Specifically, while most students take these courses at a target high school grade, the cross-grade data clearly indicate that the students who take these tests are enrolled in both middle and high schools and students in middle school exhibit the higher self-selection on ability or readiness, and vice versa for upper-grade high school students. To account for this, this norming study used Grades 6–12 students who took these three subject tests in either 2017 or 2018 as the norming samples. This approach is believed to provide an accurate description of a student's achievement relative to the other students who take the same course at the same time.

Since these course-specific tests have been used in the field for only two years and will grow over time, the data used in this study is limited, and therefore so is the generalizability of the study results. Given the available evidence employed to construct these norms, users should exercise caution about the limited generalizability of the inferences supported by the results presented in this report. For example, instructional decisions that rely on inferences about the normative performance of students are likely to be less precise. Similarly, the lower precision in these norms should be factored into secondary or derived uses of student normative scores such as teacher or school accountability.

While NWEA will continue to improve these norms as more data become available, these norms offer a first attempt to schools, teachers, or parents to interpret and understand how students are performing at a point in time and over the course of the year in a specific mathematics subject. Educators may want to combine this normative information with other evidence about student performance in making placement decisions or other major instructional or programmatic decisions.

5. References

Bock, R. D., & Petersen, A. C. (1975). A multivariate correction for attenuation. *Biometrika, 62*(3), 673–678.

National Governors Association (NGA) Center for Best Practices & Council of Chief State School Officers (CCSSO). (2010). *Common core state standards for mathematics, Appendix A: Designing high school mathematics courses based on the common core state standards*. Washington, D.C.: Authors. Retrieved from http://www.corestandards.org/assets/CCSSI_Mathematics_Appendix_A.pdf.

Kingsbury, G. G., & Zara, A. (1989). Procedures for selecting items for computerized adaptive tests. *Applied Measurement in Education, 2*(4), 359–375.

Kingsbury, G. G., & Zara, A. (1991). A comparison of procedures for content-sensitive item selection in computerized adaptive tests. *Applied Measurement in Education, 4*(3), 241–261.

Thum, Y. M., & Hauser, C. H. (2015). *NWEA 2015 MAP norms for student and school achievement status and growth*. NWEA Research Report. Portland, OR: NWEA.

Appendix A: Normality Assumption Graphs for Geometry and Algebra 2

Figure A.1. Histograms, Q-Q Plots, and CDFs for Geometry Fall Score

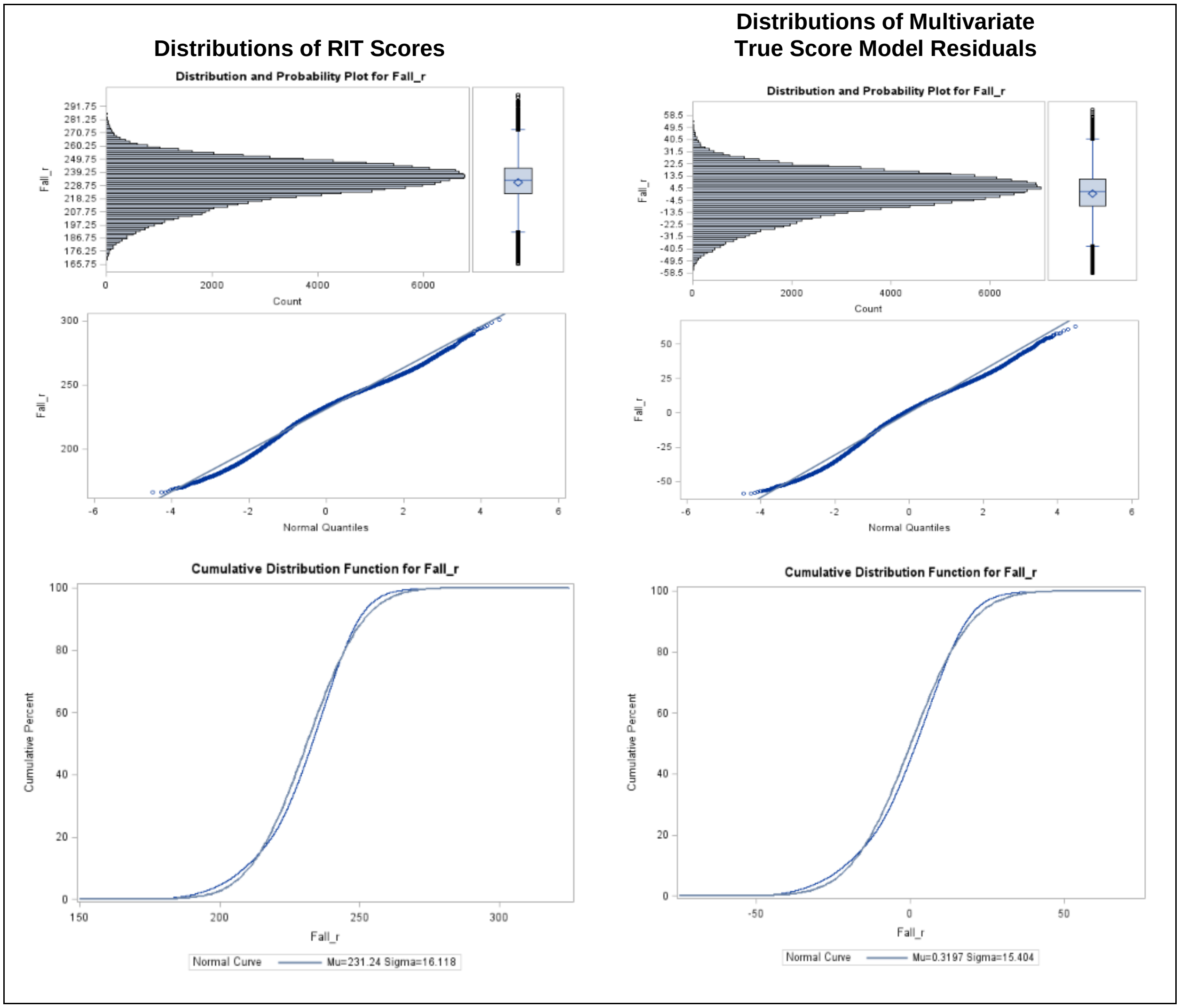


Figure A.2. Histograms, Q-Q Plots, and CDFs for Geometry Winter Scores

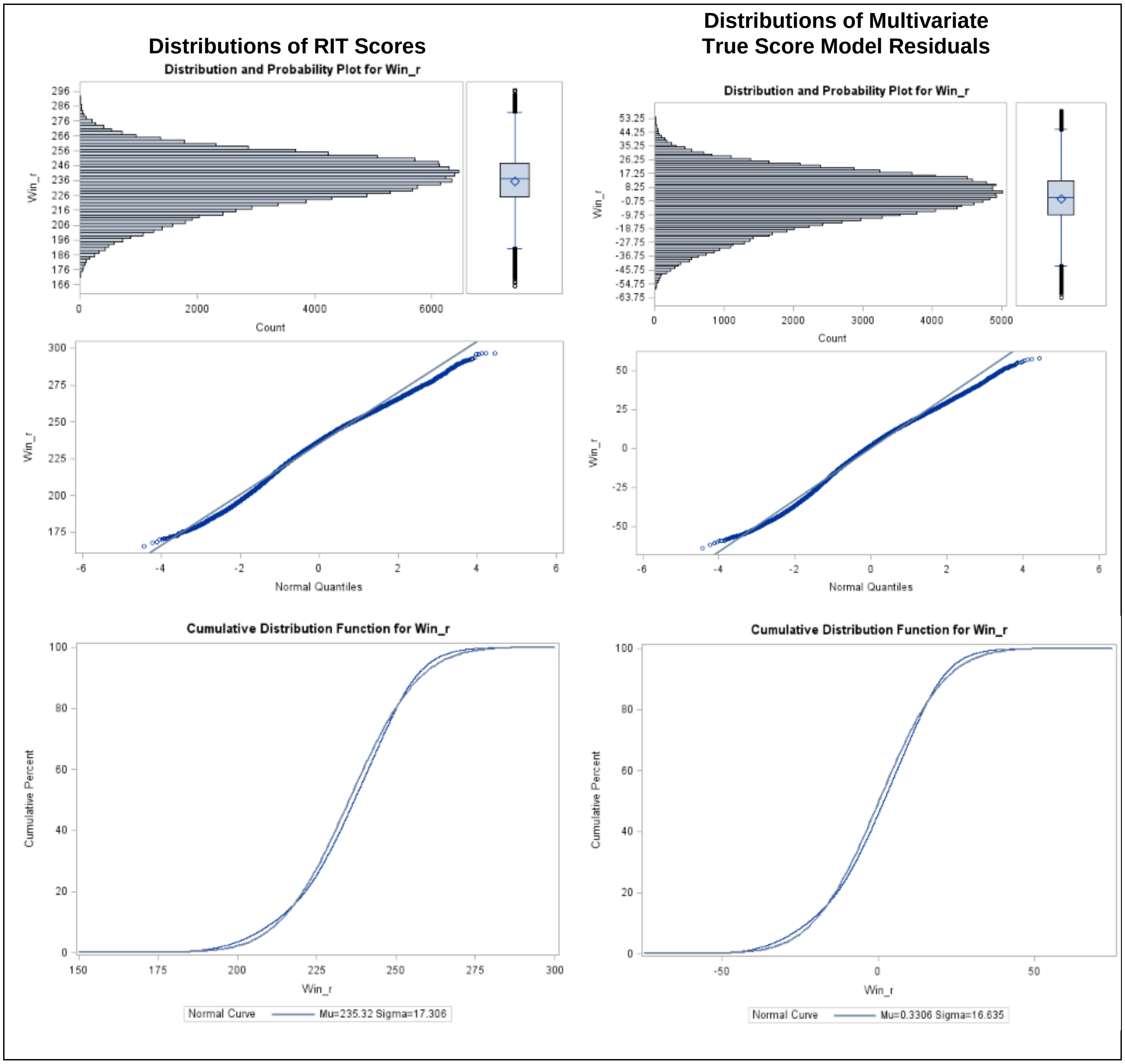


Figure A.3. Histograms, Q-Q Plots, and CDFs for Geometry Spring Scores

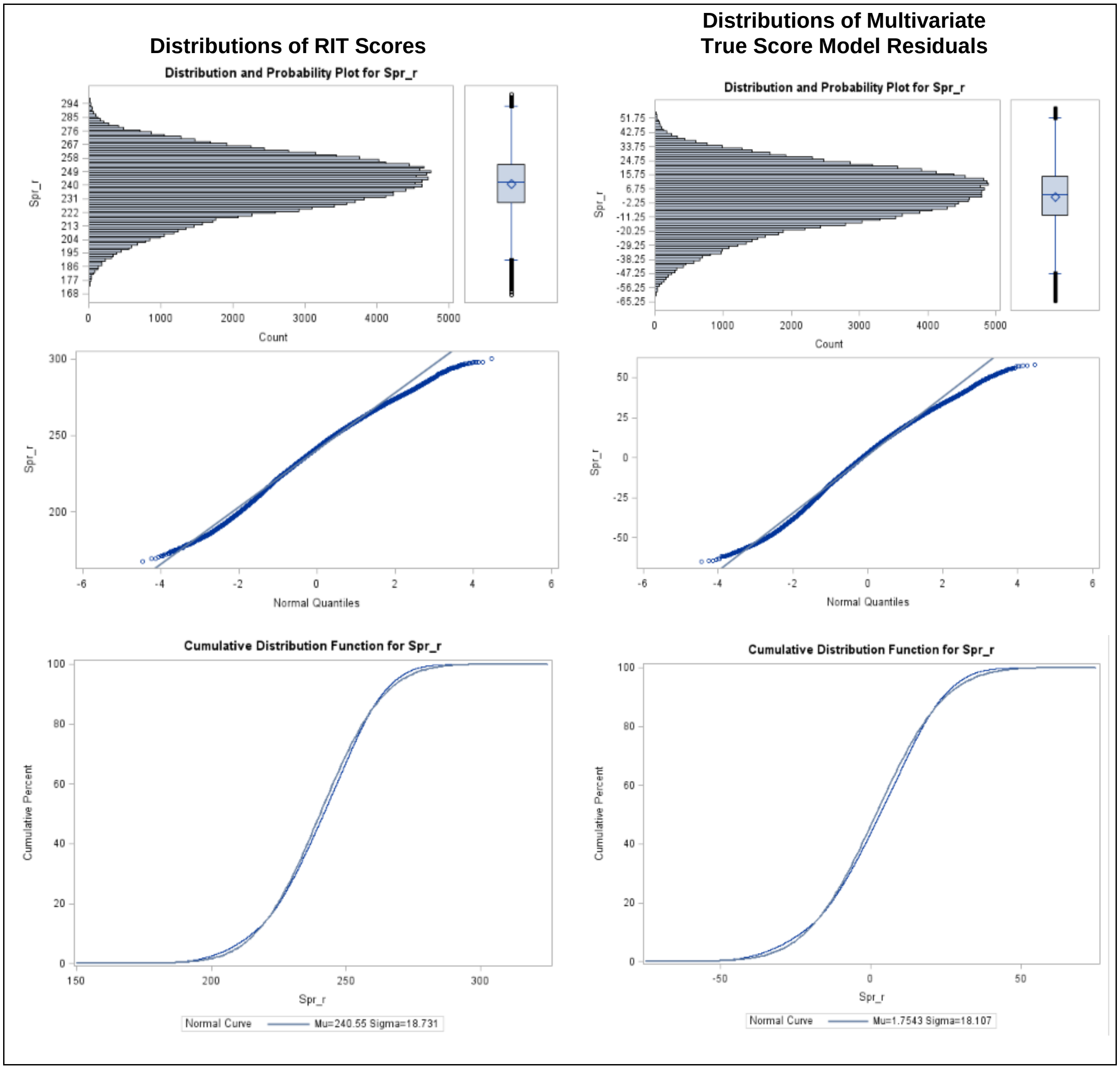


Figure A.4. Scatterplot Matrix among Fall, Winter, and Spring Scores for Geometry Test Takers

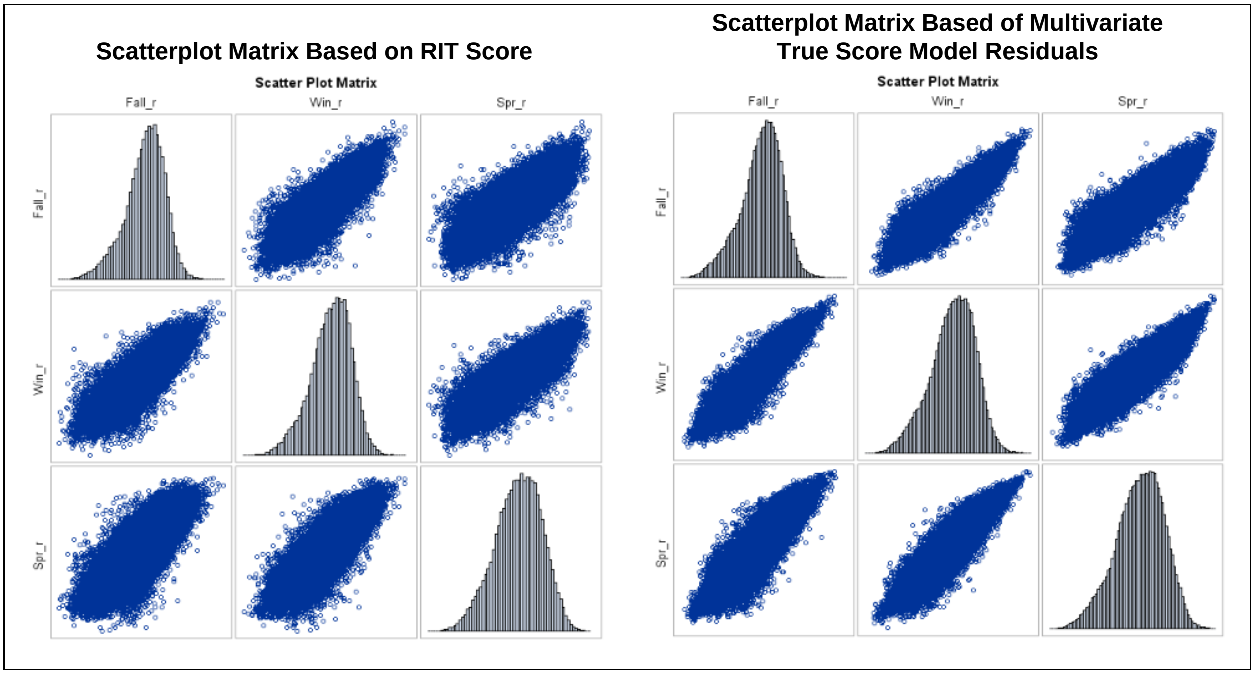


Figure A.5. Histograms, Q-Q Plots, and CDFs for Algebra 2 Fall Score

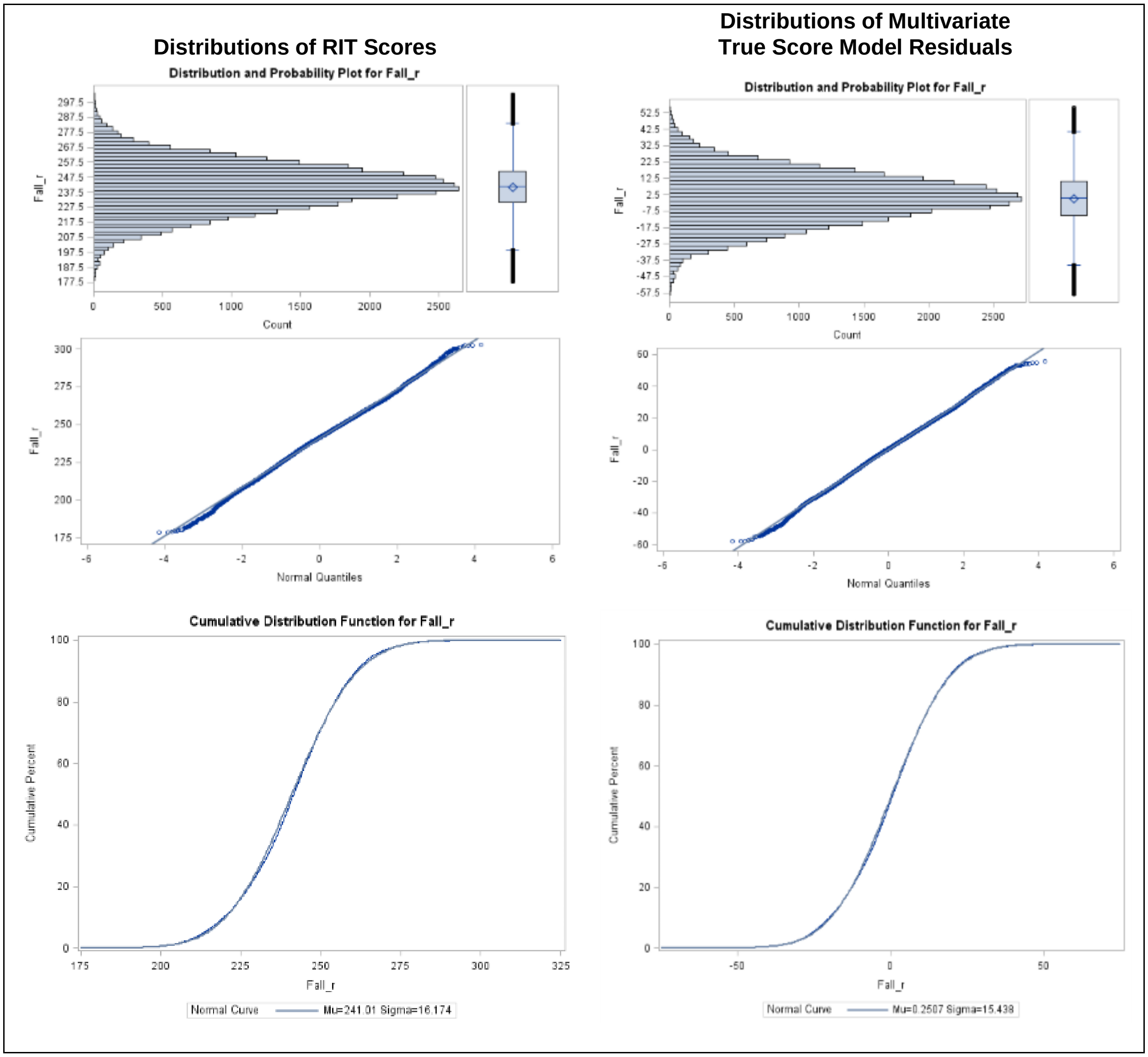


Figure A.6. Histograms, Q-Q Plots, and CDFs for Algebra 2 Winter Scores

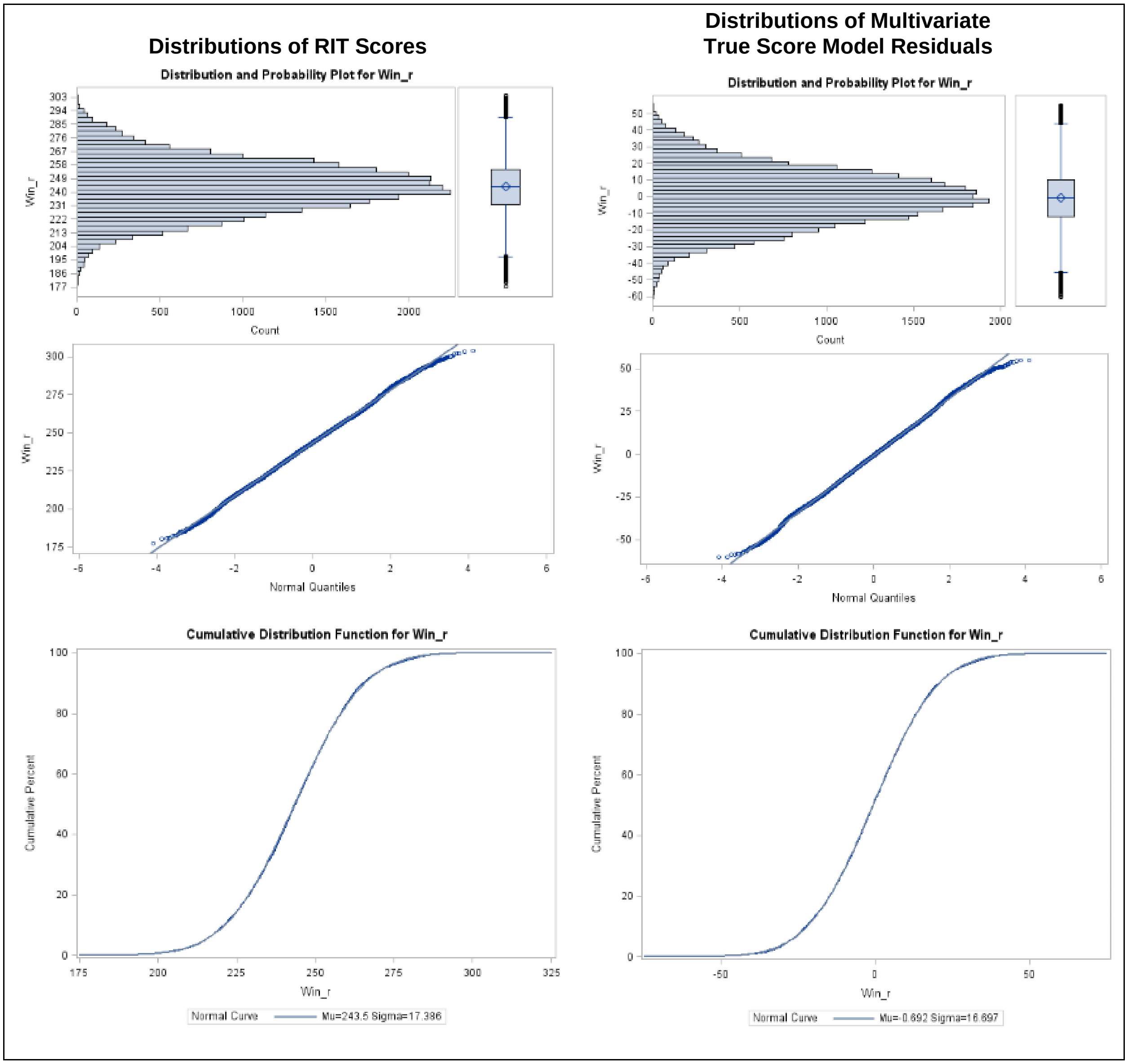


Figure A.7. Histograms, Q-Q Plots, and CDFs for Algebra 2 Spring Scores

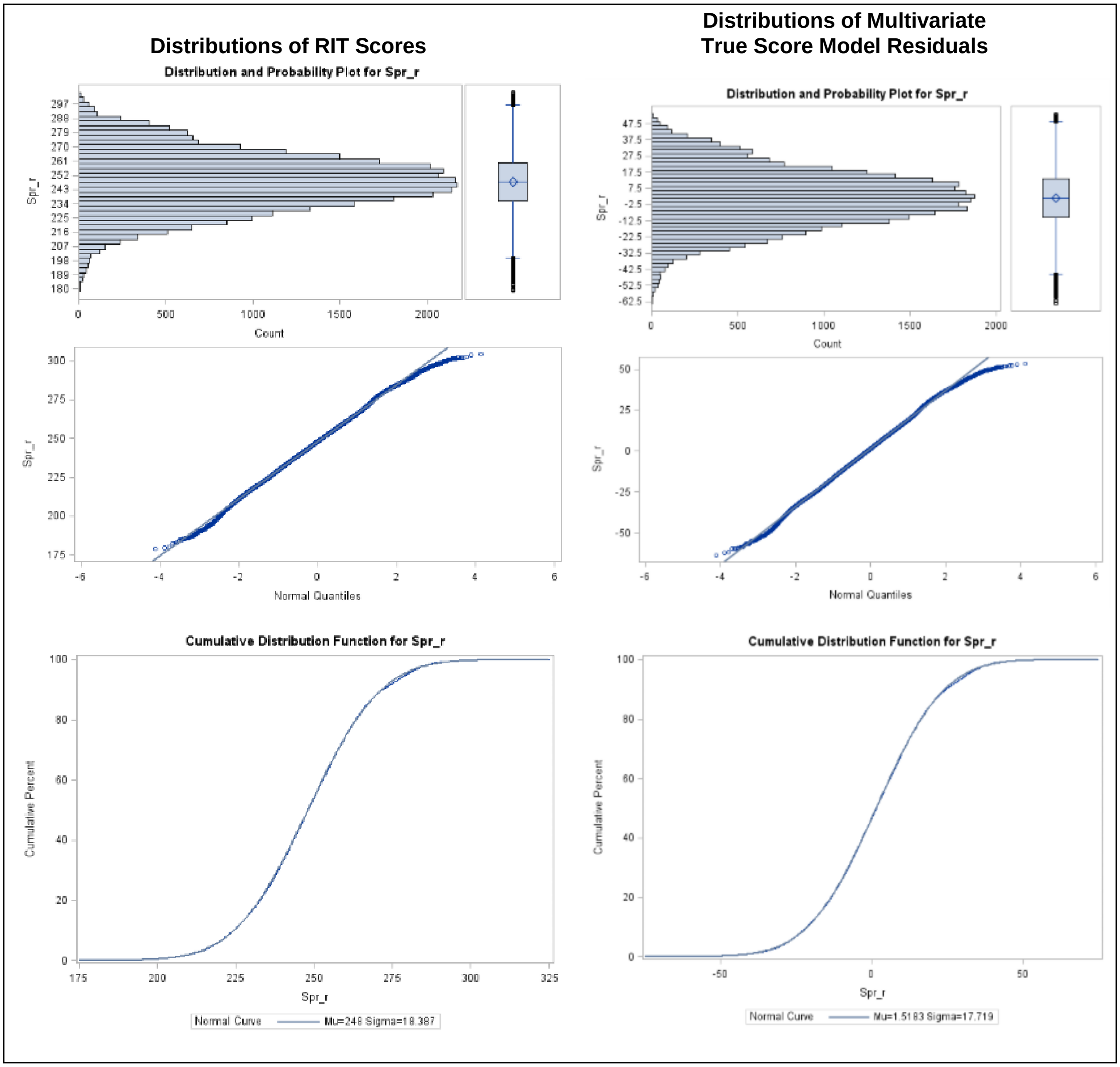


Figure A.8. Scatterplot Matrix among Fall, Winter, and Spring Scores for Algebra 2 Test Takers