Nebraska Through-Year Research Study Report: Linking Study Between NSCAS and MAP Growth based on Common Item Linking

July 23,2021

NWEA Operational Content and Psychometrics

Table of Contents

List of Tables

List of Figures

Executive Summary

To ensure a successful transition to a through-year assessment that capitalizes on the benefits of MAP Growth while also meeting the state requirements for identifying proficiency, a link must be provided between the Nebraska Student-Centered Assessment System (NSCAS) and MAP Growth scales. Whereas equipercentile linking was used to produce the Rasch Unit (RIT) scores for the Spring 2021 Phase 1 Pilot administration, NWEA has been investigating various linking approaches for the Winter Pilot and beyond.

The common item linking study presented in this document is the most recent study conducted as part of the investigation into the NSCAS and MAP Growth link. The goal is to investigate the degree to which NSCAS items could be brought onto the RIT scale and achieve comparable results. To conduct the common item linking study, a set of MAP Growth items was selected and embedded at the end of the Spring 2021 Phase 1 Pilot test forms for ELA and mathematics. Student responses from the 2021 administrations of MAP Growth and the NSCAS Phase 1 Pilot were then used to conduct the study.

IRT linked RIT scores, as well as two sets of equipercentile linked RIT scores, were computed and compared to scores from the NSCAS Spring 2021 Phase 1 Pilot and to the RIT scores that are based on the merged data (i.e., the NSCAS Spring 2021 Phase 1 Pilot data merged with the Spring 2021 MAP Growth results from Nebraska students). Based on the score comparison results, NWEA recommends moving forward with the IRT linked RIT score with the Mean/Sigma (MS) transformation (using items from the Reading Vocabulary and Reading Comprehension items only for ELA) for the 2022 Winter Pilot and beyond.

1. Introduction

Nebraska and NWEA are collaborating on a through-year assessment system design that capitalizes on the benefits of MAP Growth while also meeting the state requirements for identifying proficiency. In Spring 2021, students received both a Nebraska Student-Centered Assessment System (NSCAS) scale score and a MAP Growth Rasch Unit (RIT) score. A link must be provided between the NSCAS and MAP Growth scales to allow for a through-year RIT score to be generated. Whereas equipercentile linking was used to produce scores for the Spring 2021 Phase 1 Pilot administration, NWEA has been investigating various linking approaches to implement a more efficient method for the 2022 Winter Pilot and beyond. This document presents the results of a common item linking study using the item response theory (IRT) method and Spring 2021 testing data to improve the linking between NSCAS and MAP Growth.

1.1. Linking Analyses

Several analyses have been conducted to investigate the link between NSCAS and MAP Growth. In February 2020, NWEA conducted a common person linking study to link the scales between NSCAS and MAP Growth to investigate the degree to which MAP Growth items could be brought onto the same scale and achieve comparable results using the Spring 2019 MAP Growth and NSCAS archival data (NWEA, 2020a). Based on the anomalous findings of this linking study, NWEA proposed investigating further into common person linking and implementing a common item linking study for the following administration year.

In October 2020, NWEA conducted a follow-up common person linking study that sought an improved approach to linking the two vertically scaled assessments (NWEA, 2020b). Eight conditions were evaluated based on three variables: (1) the time elapsed between the two tests (all spring vs. 30-day data), (2) calibration across grades vs. by grade, and (3) calibration (fixed vs. estimated). The Condition 4 results (all spring data, by-grade calibration, and estimated calibration method) performed better than the other linking procedures. Both the effect sizes and percentage differences were much larger with the across-grades calibration compared to the by-grade calibration for both datasets and for both calibration methods. Based on this pattern found from this follow-up study, transformation constants were obtained for each grade and subsequently used in the common item linking study.

In November 2020, NWEA examined whether the linking parameters from Condition 4 could be applied to the Spring 2021 administration by applying those linking constants to the simulation results of the 2021 test design through the transformation of the ability estimates. These results were presented at the November 18, 2020, TAC meeting (NWEA, 2020c, pp. 13–15). The results from the equipercentile linking study between MAP Growth and NSCAS (NWEA, 2020d) were also applied to the simulation results as part of this investigation. The use of equipercentile linking produced scores from the simulation results that were more closely aligned to the MAP Growth score distributions overall. Thus, for reporting on the RIT scale in 2021, NWEA recommended using the equipercentile linking results to provide a better comparison with MAP Growth scores across the academic year. NDE approved it for 2021.

The common item linking study presented in this document is the most recent study conducted as part of the investigation into the NSCAS and MAP Growth link. Whereas the common person linking study used data that included a set of students taking both tests, the common item linking study used a set of common items from both tests. The goal is to investigate the degree to which NSCAS items could be brought onto the RIT scale and achieve comparable results to provide a link that would allow NWEA to report out on the RIT scale rather than use the equipercentile linking results in the future.

1.2. Spring 2021 Phase 1 Pilot Test Design

The Spring 2021 NSCAS ELA and Mathematics administration consisted of 35 items: 23 operational adaptive items and 12 non-operational items (seven field test and five MAP Growth). The test design required that items be selected based on student ability from the full item pool for each grade and content area (ELA and mathematics). Thus, each student received a unique test event based on their ability as they moved through the test. This test design allowed NWEA to conduct the needed field testing in Spring 2021 to support the through-year solution and to complete the common item linking study. Since this test was shorter than the previous NSCAS design in terms of operational items, the results are not comparable to previous summative assessments, although it was designed to provide feedback to schools and districts on student ability levels in Spring 2021.

2. Method

To conduct the common item linking study, a set of MAP Growth items were selected and embedded at the end of the NSCAS Spring 2021 Phase 1 Pilot test forms for ELA and mathematics. Student responses from the 2021 administrations of both the Pilot and MAP Growth assessments were then used to conduct the study.

2.1. Embedded MAP Growth Items

NSCAS and MAP Growth use different item players, which means ELA reading passages are formatted differently. Mathematics items have different calculator rules regarding when calculators can be used and what calculator types can be used. Item display settings such as color, text font, and layout are also different. Therefore, a subset of items on the MAP Growth tests that are the least different in formatting from NSCAS were selected for the common item linking study by the NWEA Content and Psychometric Solutions teams. These MAP Growth linking items were then placed at the end of the Spring 2021 Phase 1 Pilot test forms.[1] Following NDE's approval, NWEA selected the most NSCAS-like items in the MAP Growth item pool and placed them at the end of the forms as follows:

1. Create a MAP Growth item pool for each grade that aligns to Nebraska's College and Career Ready Standards.

2. Select the NSCAS-like items from this item pool to form the Nebraska MAP Growth item pool. MAP Growth Reading items must have passages, and the MAP Growth Mathematics items must meet current NSCAS calculator rules.
    - Select MAP Growth items based on the percentage of items for each reporting category of the Nebraska MAP Growth pool.
    - Select approximately 150 items per grade.

3. Place these MAP Growth items at the end of the forms (i.e., Items 31–35), as opposed to embedding them in the typical field test slots within the operational test.

Table 2.1 presents the number of embedded MAP Growth items selected for the item pool for each grade. These items did not contribute to operational scores. For NSCAS ELA, the target number of items for MAP Growth Reading and Language Usage was 110 and 40, respectively. Where MAP Growth Reading items could not meet the target number of 110, more Language Usage items were included. The minimum n-count for each MAP Growth item was 750. Items were originally selected for the 2020 testing that was cancelled due to COVID. Some items were replaced for the NSCAS Spring 2021 Phase 1 Pilot forms due to item retirement or a change in alignment.

[1] These recommendations were provided by NWEA in "Memo to NDE_2020 Embedding MAP Growth Items 2019-12-19" and approved by NDE.

Table 2.1. Number of Embedded MAP Growth Items in the Spring 2021 Phase 1 Pilot

Grade	#Embedded MAP Growth Items			
	ELA*			
	Reading	Language Usage	Total	Mathematics
3	89	61	150	150
4	113	40	153	150
5	112	40	152	150
6	110	40	150	150
7	88	61	149	150
8	106	40	146	150
Total	618	282	900	900

*ELA = MAP Growth Reading+Language Usage. 61 Language Usage items were selected for Grades 3 and 7 where the number of Reading items were less than the targeted number of 110.

To demonstrate how the MAP Growth items were administered during the Spring 2021 Phase 1 Pilot, NWEA ran the 2021 simulations with these MAP Growth linking items (NWEA, 2021). The following constraints were imposed for the MAP Growth items:

- The total number of MAP Growth linking items for each student is 5.
- Each student gets MAP Growth linking items at the end of the test.
- MAP Growth linking items are not included for calculating student scores.
- The maximum number of passages is 1.
- The minimum number of items per passage is 3.
- The maximum number of items per reporting category is 2 or 3.
- The targeted minimum number of students for each MAP Growth item is 750.
- Students are pseudo-randomly assigned to each MAP Growth item.

2.2. Data

The following NSCAS and MAP Growth assessments were linked. Language Usage is not included because the data were not included in the equipercentile linking or the score comparisons (see Section 3.3 of this report for more details).

- ELA_RD = NSCAS ELA, MAP Growth Reading
- MA_MA = NSCAS Mathematics, MAP Growth Mathematics

Table 2.2 describes the data sets used in this study. Data from the NSCAS Spring 2021 Phase 1 Pilot assessments in ELA and mathematics were used to calibrate the embedded MAP Growth items in the common item linking study and compare achievement level distributions based on students' NSCAS scores and linked RIT scores. The Spring 2021 NSCAS and the Spring 2021 MAP Growth results from Nebraska students were merged by students to compare the RIT and linked RIT scores. To merge the data, each student's NSCAS testing record was matched to their MAP Growth score using their student ID. Only students who took both the MAP Growth and NSCAS assessments in Spring 2021 were included in the study sample. This merged data were also used to run the 2021 equipercentile linking.

Table 2.2. Data used in this Study

Data	Description	Uses
MAP Growth ("Before Merge")	Spring 2021 MAP Growth data file from Nebraska students	• Merge with the NSCAS Spring 2021 Phase 1 Pilot data to generate the RIT data set
NSCAS ("Before Merge")	Spring 2021 Phase 1 Pilot data that include both the operational NSCAS and embedded MAP Growth items	• Calibrate the MAP Growth items in the IRT common item linking study • Calculate the IRT linked RIT scores • Compare NSCAS vs. linked RIT achievement level distributions (Appendix E)
RIT ("After Merge")	Spring 2021 MAP Growth data from Nebraska students merged with the Spring 2021 Phase 1 Pilot data	• Run the 2021 equipercentile linking to calculate the equipercentile linked RIT scores • Compare RIT vs. linked RIT effect size (Appendix D)

Table 2.3 presents the number of students in the Spring 2021 NSCAS ELA and Mathematics student population and the Nebraska MAP Growth Reading and Mathematics student population ("Before Merge"). It then presents the number of total students in the merged data ("After Merge"). About 13,000 or more students were merged per grade, with 65–85% NSCAS students and 93–94% MAP Growth students being merged. Demographics of the merged students are representative of the Nebraska population, as indicated by the percentage of students in terms of sex and race that are all within a 5% difference (i.e., a maximum percent difference of -3.34% is observed for Grade 7 Mathematics), as reported in Appendix A.

Table 2.3. Study Sample Before and After Merging

	#Students in Study Sample			%Merge	
	Before Merge				
Grade	NSCAS	MAP Growth	After Merge	NSCAS	MAP Growth
ELA_RD					
3	21,621	19,697	18,442	85.30	93.63
4	21,551	16,597	15,462	71.75	93.16
5	22,046	16,888	15,761	71.49	93.33
6	22,157	17,313	16,242	73.30	93.81
7	21,960	16,021	14,873	67.73	92.83
8	20,572	14,511	13,503	65.64	93.05
MA_MA					
3	21,482	16,854	15,609	72.66	92.61
4	21,605	16,584	15,548	71.96	93.75
5	22,130	16,908	15,897	71.83	94.02
6	22,167	16,758	15,687	70.77	93.61
7	22,017	15,396	14,345	65.15	93.17
8	20,611	14,239	13,316	64.61	93.52

Table 2.4 presents descriptive statistics of the NSCAS student scores before merging compared to the descriptive statistics of the merged data to see the representativeness of these students in terms of scores. NSCAS scores are lower for the samples after merging compared to the NSCAS population across content areas and grades. The maximum was observed for ELA_RD Grade 7, with a difference of -5 and an effect size of -0.07. Based on the criterion of effect size < 0.1 used in year-to-year evaluations of pre-equating and post-equating evaluations, the samples are sufficiently similar in terms of student scores.

Table 2.4. Study Sample NSCAS Score Comparison Before and After Merging

Grade	NSCAS Before Merge			NSCAS After Merge			Difference (After Merge-Before Merge)	Effect Size
	N	Mean	SD	N	Mean	SD		
ELA_RD								
3	21,621	2469	85.20	18,442	2468	85.64	-1	-0.01
4	21,551	2503	81.78	15,462	2500	82.70	-3	-0.04
5	22,046	2516	79.83	15,761	2515	80.10	-1	-0.01
6	22,157	2528	77.46	16,242	2525	77.29	-3	-0.04
7	21,960	2539	74.28	14,873	2534	75.14	-5	-0.07
8	20,572	2556	72.30	13,503	2553	72.55	-3	-0.04
MA_MA								
3	21,482	1186	76.64	15,609	1184	77.04	-2	-0.03
4	21,605	1213	73.74	15,548	1213	73.98	0	<0.01
5	22,130	1229	71.50	15,897	1227	70.78	-2	-0.03
6	22,167	1239	72.56	15,687	1236	71.75	-3	-0.04
7	22,017	1246	67.97	14,345	1244	68.07	-2	-0.03
8	20,611	1260	71.28	13,316	1257	70.81	-3	-0.04

3. Results

3.1. NSCAS Anchor Items

The first step of the common item linking procedure was to determine the NSCAS anchor items by reviewing and comparing plots of the ICCs and the distribution of student responses for each item. To illustrate what these plots look like, Appendix B presents example plots of ICC and student responses for selected items. One dichotomous and one polytomous item examples are included for either case of anchors or non-anchors to highlight how these plots were used for selecting anchors. The number of student responses and the item-total correlation were also considered. Table 3.1 presents the total number of NSCAS operational items and the number of anchor items used in calibrating the MAP Growth items. Out of approximately 500 operational items, 63–95 items were selected as anchors for each grade and content area.

Table 3.1. Number of NSCAS Anchor Items used for MAP Growth Calibration

	#NSCAS Items	
Grade	Operational	Anchor
ELA		
3	590	63
4	579	78
5	508	65
6	518	67
7	478	95
8	553	90
MA		
3	540	69
4	418	62
5	432	69
6	537	86
7	457	70
8	435	77

3.2. Transformation Constants

Once the embedded MAP Growth items were calibrated while fixing the NSCAS anchor items, their item parameters were verified to ensure that they align with the distribution of student responses. Items were removed if they had a low item-total correlation (<0.2) or positive distractor correlation (>0.05). The remaining items were then used to obtain the transformation constants using STUIRT. Table 3.2 presents these results, including the number of embedded MAP Growth items removed from the analysis and the number of items used in STUIRT to obtain the transformation constants.

The table also presents the correlation between the two sets of MAP Growth items (i.e., the MAP Growth original bank value vs. the newly calibrated value for this study). The correlation values are higher than 0.90, except for ELA_RD Grade 8 (0.89) and MA_MA Grade 4 (0.88). These values are sufficiently high to consider the linked scores to be a concordance (Dorans, 1999).

Table 3.2. Number of Embedded MAP Growth Items used for Transformation

Grade	#Embdded MAP Growth Items			Correlation between Two Sets of Item Parameter Estimates
	Total	Removed	Included in STUIRT	
ELA_RD				
3	89	1	88	0.93
4	113	2	111	0.93
5	112	7	105	0.93
6	110	5	105	0.91
7	88	7	81	0.93
8	106	6	100	0.89
MA_MA				
3	150	4	146	0.93
4	150	6	144	0.88
5	150	6	144	0.92
6	150	11	139	0.93
7	150	10	140	0.94
8	150	29	121	0.90

Table 3.3 presents the transformation constants calculated. The two sets of item parameters were plotted. Where differences are found, the MM and MS lines seem more different, while the HB and SL lines are in the middle (i.e., between the MM and MS lines). Further, the MS transformation seems to reflect some outlier items at both tails, if any. MM is included for comparison as it has previously been used in Nebraska, including in the common person linking study.

Table 3.3. Transformation Constants

Grade	Method	ELA_RD		MA_MA	
		Slope	Intercept	Slope	Intercept
3	Mean/Mean (MM)	1.0000	0.1219	1.0000	0.6908
	Mean/Sigma (MS)	1.0052	0.1249	0.9967	0.6859
	Haebara (HB)	1.0027	0.1227	1.0045	0.6890
	Stocking-Lord (SL)	1.0015	0.1211	1.0004	0.6822
4	Mean/Mean (MM)	1.0000	0.3497	1.0000	1.2708
	Mean/Sigma (MS)	0.9293	0.3257	1.2460	1.3367
	Haebara (HB)	0.9742	0.3387	1.0876	1.3246
	Stocking-Lord (SL)	0.9844	0.3394	1.0559	1.3193
5	Mean/Mean (MM)	1.0000	0.8093	1.0000	1.6404
	Mean/Sigma (MS)	0.9544	0.7960	1.3230	1.7079
	Haebara (HB)	0.9893	0.8028	1.1820	1.7342
	Stocking-Lord (SL)	0.9939	0.8017	1.1323	1.7511
6	Mean/Mean (MM)	1.0000	1.0308	1.0000	1.9669
	Mean/Sigma (MS)	0.9519	1.0363	1.2554	1.8600
	Haebara (HB)	0.9811	1.0315	1.1594	1.9339
	Stocking-Lord (SL)	0.9876	1.0264	1.1098	1.9711

Grade	Method	ELA_RD		MA_MA	
		Slope	Intercept	Slope	Intercept
7	Mean/Mean (MM)	1.0000	1.0551	1.0000	2.1887
	Mean/Sigma (MS)	1.0455	1.0373	1.3398	2.0295
	Haebara (HB)	1.0241	1.0485	1.2032	2.1331
	Stocking-Lord (SL)	1.0141	1.0500	1.1465	2.1803
8	Mean/Mean (MM)	1.0000	1.2371	1.0000	2.7904
	Mean/Sigma (MS)	1.0948	1.1786	1.3579	2.3666
	Haebara (HB)	1.0264	1.2196	1.2487	2.5928
	Stocking-Lord (SL)	1.0125	1.2263	1.2509	2.8041

3.3. Score Comparisons

The following six linked RIT scores were used for comparing to the NSCAS and RIT data:

1. IRT linked RIT (MM) calculated with two reporting category items (for ELA)
2. IRT linked RIT (MS) calculated with two reporting category items (for ELA)
3. IRT linked RIT (MM) calculated with all reporting category items
4. IRT linked RIT (MS) calculated with all reporting category items
5. Equipercentile linked RIT (2021 data)
6. Equipercentile linked RIT (2019 data)

Based on the observations from Appendix C, NWEA decided to focus on the MM and MS results only for comparing and analyzing the results. Language Usage was also not included. Nebraska students typically take MAP Growth Reading and Mathematics, while only about a quarter of the student population takes the Language Usage assessment. Also, the correlations between the two sets of MAP Growth item parameter estimates (shown in Table 3.2) is lower for Language Usage (approximately 0.70) compared to Reading and Mathematics (0.88 or higher).

The NSCAS ELA Grades 3–8 assessments include three reporting categories: Reading Vocabulary, Reading Comprehension, and Writing Skills. However, MAP Growth Reading only includes the first two reporting categories, while MAP Growth Language Usage includes the writing items. The IRT linked RIT calculation initially used all NSCAS ELA items (i.e., all three reporting categories). However, to better match the construct of the NSCAS ELA and MAP Growth Reading assessments, NWEA also computed the IRT linked RIT for ELA using only the two reporting categories of Reading Vocabulary and Reading Comprehension. Results for the IRT linked RIT are therefore presented in two different ways for ELA: (1) with the two reporting categories only and (2) with all reporting categories.

Furthermore, based on the 2021 NSCAS data, there was a larger than expected number of students with low linked RIT scores who received the LOSS + 2 minimum score.[2] Further investigation showed that while most of these students responded to all 35 items, they had very low raw scores and had shorter test duration than the general population of students taking the test. Based on these results, NWEA believes that there is a possible student engagement issue for these scores and decided to remove them from all subsequent analyses.

[2] These results were provided by NWEA in “Memo to NDE_Linked RIT 2021-06-04.”

3.3.1. RIT vs. Linked RIT

Appendix C presents the score comparison between RIT and linked RIT using the merged data. Table 3.4 presents the effect sizes between the RIT and linked RIT scores (i.e., the last column in the Appendix C tables). Effect sizes of 0.1 or higher in absolute value were flagged. Overall, the following results are observed:

- Effect sizes are very small for the equipercentile linked RIT using the 2021 data. This is expected because two current score distributions were used to link them (i.e., Spring 2021 Phase 1 Pilot and MAP Growth), while linear transformations were applied for the IRT linked RIT scores and the score distribution from 2019 was used for the equipercentile linked RIT using the 2019 data.
- Effect sizes are similar across the linked RIT scores for ELA_RD (using both two reporting categories only and all reporting categories), with the exception of the equipercentile linked RIT using the 2021 data.
- Excluding the equipercentile linked RIT using the 2021 data, the effect sizes are flagged that are higher or equal to 0.1 in absolute value. Considering that this effect size criterion is very conservative and has been used for the NSCAS operational year-to-year consistency check, it may need to be relaxed.
- The effect sizes between the 2021 and 2019 equipercentile linked RIT scores are more similar for MA_MA than for ELA_RD.
- The effect sizes are smaller for the equipercentile linked RIT scores compared to all the IRT linked RIT scores.
- Effect sizes are similar between the MM and MS transformation methods.
- Effect sizes are slightly smaller for “Two Reporting Categories” than for “All Reporting Categories” in ELA_RD for several grades. For example, the MS effect size is -0.18 for Grade 4 when only two reporting categories were used, while it is -0.22 when all three reporting categories were used.

Table 3.4. Effect Sizes between RIT and Linked RIT Scores

		RIT vs. IRT Linked RIT				RIT vs. Equipercentile Linked RIT	
		Two Reporting Categories		All Reporting Categories			
Grade	N	MM	MS	MM	MS	2021	2019
ELA_RD							
3	18,442	-0.16	-0.16	-0.16	-0.16	<0.01	-0.18
4	15,462	-0.17	-0.18	-0.20	-0.22	<0.01	-0.11
5	15,761	-0.16	-0.17	-0.13	-0.14	<0.01	-0.12
6	16,242	-0.09	-0.09	-0.09	-0.10	<0.01	-0.10
7	14,873	-0.19	-0.18	-0.18	-0.17	<0.01	-0.04
8	13,503	-0.15	-0.15	-0.13	-0.11	-0.01	-0.03
MA_MA							
3	15,609	–	–	0.11	0.10	<0.01	0.00
4	15,548	–	–	0.26	0.31	<0.01	-0.06
5	15,897	–	–	0.13	0.25	<0.01	-0.10
6	15,687	–	–	0.20	0.22	<0.01	-0.06
7	14,345	–	–	0.13	0.18	<0.01	-0.01
8	13,316	–	–	0.34	0.27	<0.01	-0.05

3.3.2. NSCAS vs. Linked RIT

Appendix D presents the comparison between the NSCAS Spring 2021 Phase 1 Pilot data and students' linked RIT scores in terms of achievement level distribution. Descriptive statistics of scores could not be computed because the NSCAS and linked RIT scores are on different scales. For linked RIT scores, cut scores were also placed on the RIT scale to determine the achievement level distributions. For the IRT linked RIT, transformation constants were applied to the NSCAS theta cuts and scaled to the RIT scale using MAP Growth scaling values (i.e., slope = 10 and intercept = 200). For the equipercentile linked RIT, RIT scores corresponding to the NSCAS cuts were obtained from the equipercentile conversion. Table 3.5 presents the NSCAS and linked RIT cut scores. Overall, the following results are observed:

- No linked RIT scores are flagged with more than 5% difference in the Developing level.
- The percentage difference is not close to zero for equipercentile linked RIT using the 2021 data, unlike the effect sizes. This is because the conversion from NSCAS scores to RIT are not 1:1. For example, the NSCAS ELA Grade 4 On Track cut is 2500, which was converted to a RIT score of 206. However, NSCAS scores of 2496–2499 are also converted to 206, so their achievement level changed from Developing to On Track.
- The percentage differences are similar between the two equipercentile linked RIT scores for most cases.
- The percentage differences are similar between the four IRT linked RIT scores in general, although they are slightly lower for MS than MM.

Table 3.5. NSCAS and Linked RIT Cut Scores

			IRT Linked RIT				Equipercentile Linked RIT			
	NSCAS		MM		MS		2021		2019	
Grade	OT	CCR	OT	CCR	OT	CCR	OT	CCR	OT	CCR
ELA_RD										
3	2477	2557	198	209	198	209	200	215	201	215
4	2500	2582	203	215	203	214	206	221	206	221
5	2531	2599	212	222	212	221	215	228	215	227
6	2543	2603	216	225	216	224	219	230	219	230
7	2556	2630	218	228	218	229	222	236	222	237
8	2561	2632	221	231	221	232	224	238	223	237
MA_MA										
3	1190	1286	205	223	205	222	204	221	205	220
4	1222	1317	217	234	218	240	214	233	215	231
5	1236	1331	223	240	226	249	222	242	223	243
6	1244	1342	228	246	229	251	226	246	226	246
7	1247	1346	230	248	232	256	232	255	231	252
8	1264	1365	240	258	239	264	237	260	236	260

4. Discussion

Effect sizes are very small for the equipercentile linked RIT using the 2021 data, which is the result of the two current score distributions of the Spring 2021 Phase 1 Pilot and MAP Growth being used. This implies that to have these small effect sizes, a large number of students needs to take both tests for each term. However, with the transition to a through-year assessment, it is desired that students receive a linked RIT score without taking two tests. Considering the differences observed between the two equipercentile linked RIT scores (i.e., 2019 vs. 2021), the IRT linked RIT scores produced fairly similar results to the equipercentile linked RIT using the 2019 data. Therefore, based on the comparisons between the RIT and NSCAS scores, NWEA recommends that IRT linked RIT with the MS transformation be used for the Nebraska through-year assessments, using items from the two reading reporting categories only for ELA (i.e., Reading Vocabulary and Reading Comprehension).

4.1. Recommended IRT Linked RIT (MS) Analyses

Table 4.1 presents the descriptive statistics of the recommended IRT linked RIT (MS) based on only two reporting categories for ELA_RD and all reporting categories for MA_MA, as well as the Fall 2020 RIT and the Spring 2021 RIT.

Figure 4.1 and Figure 4.2 plot the means across grades, with the recommended IRT linked RIT (MS) based on only two reporting categories for ELA_RD and all reporting categories for MA_MA, as well as the RIT (Spring 2021) and the equipercentile linked RIT scores based on the 2019 data that were part of the reported scores for the Spring 2021 Phase 1 Pilot. The figures show that the means are similar for all three RITs, particularly more for ELA_RD than for MA_MA.

Table 4.1. Descriptive Statistics of RIT and Linked RIT Scores

	RIT (Fall 2020) *					RIT (Spring 2021)*					IRT Linked RIT (MS) (Recommendation)			
Grade	**N**	**Mean**	**SD**	**Min.**	**Max.**	**N**	**Mean**	**SD**	**Min.**	**Max.**	**Mean**	**SD**	**Min.**	**Max.**
ELA_RD														
3	16,719	189.71	15.82	140	239	18,442	198.98	15.76	135	245	196.61	13.18	137	238
4	13,995	199.24	15.08	145	249	15,462	206.08	15.28	140	260	203.59	12.06	154	255
5	14,209	206.62	14.70	147	250	15,761	211.71	14.93	145	262	209.44	11.74	161	255
6	14,333	212.07	14.31	152	254	16,242	215.00	15.08	156	261	213.74	11.40	165	269
7	13,183	215.66	14.52	155	261	14,873	217.58	15.45	154	267	215.07	11.73	167	256
8	11,935	219.37	14.76	154	267	13,503	221.27	15.51	151	274	219.19	11.87	174	264
MA_MA														
3	14,106	188.78	12.76	121	250	15,609	202.49	14.22	138	266	203.95	13.98	171	256
4	14,122	199.78	13.54	134	256	15,548	211.21	15.64	139	269	216.23	16.78	171	281
5	14,379	209.23	14.39	135	310	15,897	219.38	17.21	144	289	223.59	17.05	174	292
6	13,951	215.48	14.12	141	276	15,687	223.27	16.78	146	288	226.88	16.40	180	294
7	12,725	222.44	15.38	146	283	14,345	227.94	17.85	138	307	231.05	16.61	185	303
8	11,722	228.39	16.50	146	297	13,316	232.81	19.15	136	316	237.72	17.52	187	310

*The Fall 2020 RIT results used merged data from Fall 2020 MAP Growth, Spring 2021 MAP Growth, and Spring 2021 NSCAS. The Spring 2021 RIT results used merged data from Spring 2021 MAP Growth and NSCAS MAP Growth. The merged Spring 2021 data were also used for the recommended IRT linked RIT (MS).

Figure 4.1. Mean Scores—ELA_RD

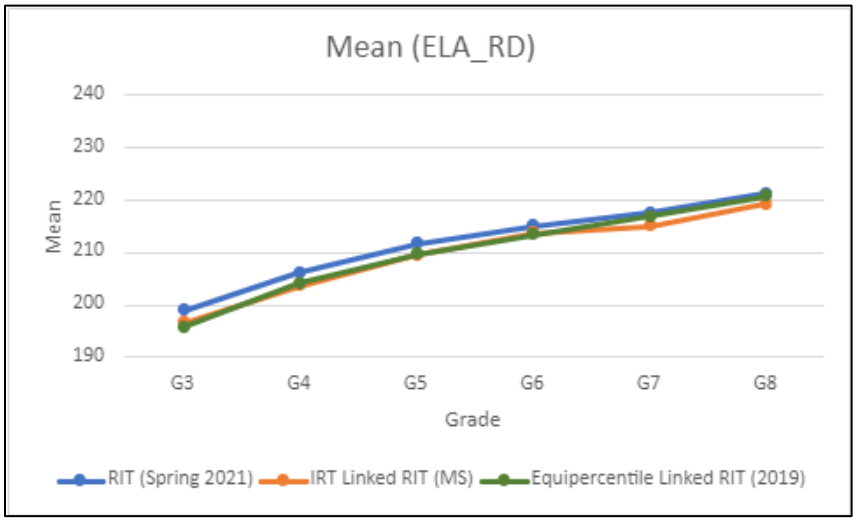


Figure 4.2. Mean Scores—MA_MA

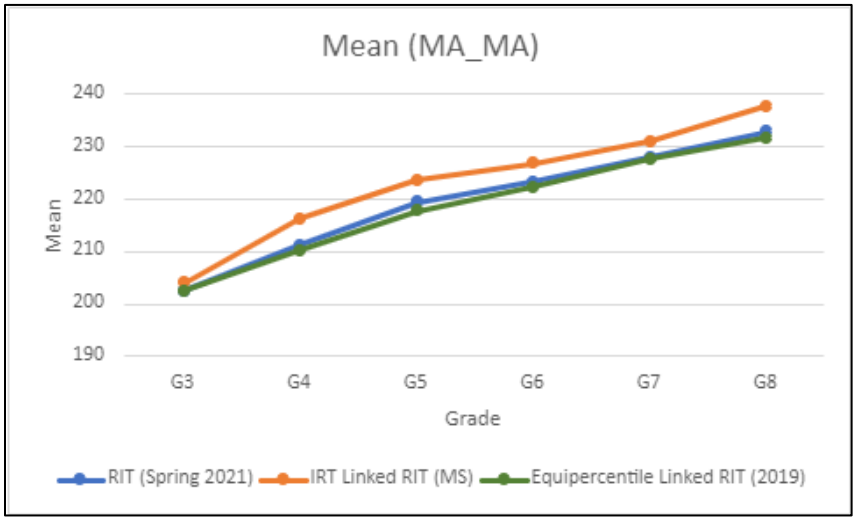


Table 4.2 presents the achievement level distributions, including the distributions for NSCAS for comparison. The percentage of students at each achievement level are very similar between IRT linked RIT (MS) and equipercentile linked RIT using 2019 data that were part of the reported scores for the Spring 2021 Phase 1 Pilot.

Table 4.2. Recommended vs. Spring 2021 Phase 1 Pilot Linked RIT Achievement Level Distributions

Grade	N (Before Merge)	NSCAS			IRT Linked RIT (MS) (Recommendation)			Equipercentile Linked RIT (2019 Data)		
		%Dev	%OT	%CCR	%Dev	%OT	%CCR	%Dev	%OT	%CCR
ELA_RD										
3	21,621	49.5	36.1	14.4	48.4	34.0	17.6	48.6	36.1	15.3
4	21,551	45.9	36.8	17.3	42.6	35.7	21.7	44.9	37.1	18.0
5	22,046	53.8	31.5	14.8	52.9	29.9	17.2	51.8	33.2	15.0
6	22,157	54.0	30.2	15.8	51.9	28.2	19.9	54.0	29.3	16.7
7	21,960	55.1	35.9	9.0	52.6	35.8	11.6	54.6	35.5	9.8
8	20,572	49.1	37.9	13.0	49.0	37.1	13.9	47.3	38.6	14.0
MA_MA										
3	21,482	52.2	38.3	9.5	51.0	38.5	10.5	49.5	40.3	10.2
4	21,605	54.2	37.7	8.2	52.5	39.1	8.4	51.9	39.5	8.5
5	22,130	54.3	38.2	7.6	54.0	38.5	7.6	52.7	39.2	8.0
6	22,167	52.7	39.2	8.1	52.5	39.0	8.5	51.4	39.9	8.7
7	22,017	53.7	38.4	7.9	53.3	38.6	8.1	53.0	39.1	7.9
8	20,611	54.5	37.8	7.7	52.4	39.3	8.3	54.5	37.8	7.7

To further compare these methods, Appendix E presents plots of three score distributions (i.e., RIT, NSCAS, and linked RIT) on the theta metric using the merged data for the recommended IRT linked RIT (MS) (2019 data).

4.2. Further Considerations

Although NWEA is recommending the IRT linked RIT with the MS transformation, there are areas of further consideration. First, Table 4.1 shows that the tails of the distribution are being pulled in with the linked RIT as compared to the RIT. One possible reason for this is that NSCAS uses only on-grade items, while MAP Growth uses both on- and off-grade items. Including off-grade items in the through-year assessment may move student scores at both tails closer to that of the MAP Growth distribution. Also, the NSCAS LOSS may need to be adjusted to be lower, and the NSCAS HOSS may need to be higher when the new scale is set in 2022. The updates to the LOSS and HOSS are more needed considering approximately 100 students were piled at the calculated LOSS in 2021.

The administration dates may need to be considered as well. Using 30 days between one test's end and the other test's start date, approximately70% of students took both MAP Growth Reading and NSCAS ELA and 80% of students took MAP Growth Mathematics and NSCAS Mathematics in Spring 2019 and Spring 2021. If data with this much time between administrations are used, it may impact linking and scoring results. Although the IRT linked RIT are different than the RIT, especially at both tails, scores overlap between the RIT and linked RIT in general, considering that the mean SEM for RIT is approximately 3 for each grade and content area.

For ELA, one more set of scores was computed using only Reading Vocabulary and Reading Comprehension items to better match the construct of the MAP Growth Reading and NSCAS ELA assessments, since the NSCAS ELA assessments also include a third reporting category of Writing Skills that is included in MAP Growth Language Usage and not Reading. However, the construct differences between NSCAS ELA and MAP Growth Reading still exist. MAP Growth Reading items are more stand-alone items, while all NSCAS reading items are associated with passages. Furthermore, in general, NSCAS has more items per passage. All MAP Growth passages have at least one item associated, and only 50% of students see passages with three items while the minimum number of items per passage is set to four for NSCAS.

5. References

Dorans, N. J. (1999). *Correspondences between ACT and SAT I scores*. College Board Report nos. 99–1, 99–2. College Entrance Examination Board.

Kim, S., & Kolen, M. J. (2004). *STUIRT: A computer program for scale transformation under unidimensional item response theory models*. University of Iowa.

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

NWEA. (2020a, February). *Nebraska through-year research study report: Linking study between NSCAS and MAP Growth based on common person linking*. Report presented to NDE.

NWEA. (2020b, October). *Nebraska through-year research study report: Linking study between NSCAS and MAP Growth based on common person linking-Follow-up Analyses*. Report presented to NDE.

NWEA. (2020c, November). *NWEA document 1a: Spring 2021 alternative test design—follow-up analyses*. Document presented to the Nebraska TAC on November 18, 2020.

NWEA. (2020d, September). *Linking study report: Predicting performance on NSCAS General Summative assessments based on NWEA MAP Growth scores*. https://www.nwea.org/resources/nebraska-linking-study/

NWEA. (2021, January). *Constraint-based engine simulation report for the spring 2021 NSCAS ELA and Mathematics assessments.* Report presented to NDE.

Thum, Y. M., & Kuhfeld, M. (2020). *NWEA 2020 MAP Growth achievement status and growth norms for students and schools*. NWEA Research Report. https://teach.mapnwea.org/impl/normsResearchStudy.pdf

Appendix A: Student Sample Demographics

Table A.1. Student Sample Demographic Comparison Before and After Merging—Grade 3

	Before Merge				After Merge				%Difference (After - Before)	
	ELA		MA		ELA_RD		MA_MA			
Demographic Subgroup	N	%	N	%	N	%	N	%	ELA_RD	MA
Total	21,621	100.00	21,482	100.00	18,442	100.00	15,609	100.00	–	–
Sex										
Female	10,558	48.83	10,494	48.85	8,999	48.80	7,667	49.12	-0.03	0.27
Male	11,063	51.17	10,988	51.15	9,443	51.20	7,942	50.88	0.03	-0.27
Race										
American Indian/Alaska Native	284	1.31	271	1.26	205	1.11	183	1.17	-0.20	-0.09
Asian	693	3.21	688	3.20	630	3.42	507	3.25	0.21	0.05
Black	1,284	5.94	1,259	5.86	1,188	6.44	997	6.39	0.50	0.53
Hispanic	4,156	19.23	4,115	19.16	3,731	20.23	3,331	21.34	1.00	2.18
Native Hawaiian/Other Pacific Islander	36	0.17	36	0.17	33	0.18	33	0.21	0.01	0.04
White	14,169	65.55	14,127	65.77	11,790	63.93	9,944	63.71	-1.62	-2.06
Two or More Races	994	4.60	982	4.57	864	4.69	613	3.93	0.09	-0.64

Table A.2. Student Sample Demographic Comparison Before and After Merging—Grade 4

	Before Merge				After Merge				%Difference (After - Before)	
	ELA		MA		ELA_RD		MA_MA			
Demographic Subgroup	N	%	N	%	N	%	N	%	ELA_RD	MA
Total	21,551	100.00	21,605	100.00	15,462	100.00	15,548	100.00	–	–
Sex										
Female	10,516	48.8	10,530	48.74	7,575	48.99	7,607	48.93	0.19	0.19
Male	11,035	51.2	11,075	51.26	7,887	51.01	7,941	51.07	-0.19	-0.19
Race										
American Indian/Alaska Native	249	1.16	251	1.16	153	0.99	154	0.99	-0.17	-0.17
Asian	651	3.02	651	3.01	485	3.14	488	3.14	0.12	0.13
Black	1,220	5.66	1,233	5.71	941	6.09	960	6.18	0.43	0.47
Hispanic	4,228	19.62	4,256	19.7	3,378	21.85	3,402	21.88	2.23	2.18
Native Hawaiian/Other Pacific Islander	35	0.16	35	0.16	28	0.18	27	0.17	0.02	0.01
White	14,232	66.04	14,234	65.89	9,923	64.18	9,957	64.05	-1.86	-1.84
Two or More Races	934	4.33	943	4.37	552	3.57	558	3.59	-0.76	-0.78

Table A.3. Student Sample Demographic Comparison Before and After Merging—Grade 5

	Before Merge				After Merge				%Difference (After - Before)	
	ELA		MA		ELA_RD		MA_MA			
Demographic Subgroup	N	%	N	%	N	%	N	%	ELA_RD	MA
Total	22,046	100.00	22,130	100.00	15,761	100.00	15,897	100.00	–	–
Sex										
Female	10,713	48.59	10,725	48.46	7,684	48.75	7,734	48.65	0.16	0.19
Male	11,333	51.41	11,405	51.54	8,077	51.25	8,163	51.35	-0.16	-0.19
Race										
American Indian/Alaska Native	273	1.24	277	1.25	153	0.97	160	1.01	-0.27	-0.24
Asian	631	2.86	633	2.86	480	3.05	482	3.03	0.19	0.17
Black	1,328	6.02	1,336	6.04	1,030	6.54	1,034	6.50	0.52	0.46
Hispanic	4,342	19.70	4,373	19.76	3,427	21.74	3,443	21.66	2.04	1.90
Native Hawaiian/Other Pacific Islander	34	0.15	34	0.15	29	0.18	29	0.18	0.03	0.03
White	14,485	65.71	14,519	65.61	10,032	63.65	10,133	63.75	-2.06	-1.86
Two or More Races	951	4.31	956	4.32	609	3.86	615	3.87	-0.45	-0.45

Table A.4. Student Sample Demographic Comparison Before and After Merging—Grade 6

	Before Merge				After Merge				%Difference (After - Before)	
	ELA		MA		ELA_RD		MA_MA			
Demographic Subgroup	N	%	N	%	N	%	N	%	ELA_RD	MA
Total	22,157	100.00	22,167	100.00	16,242	100.00	15,687	100.00	–	–
Sex										
Female	10,796	48.73	10,796	48.70	7,854	48.36	7,606	48.49	-0.37	-0.21
Male	11,361	51.27	11,371	51.30	8,388	51.64	8,081	51.51	0.37	0.21
Race										
American Indian/Alaska Native	284	1.28	273	1.23	175	1.08	165	1.05	-0.20	-0.18
Asian	582	2.63	583	2.63	468	2.88	434	2.77	0.25	0.14
Black	1,286	5.80	1,279	5.77	1,012	6.23	941	6.00	0.43	0.23
Hispanic	4,458	20.12	4,477	20.20	3,705	22.81	3,609	23.01	2.69	2.81
Native Hawaiian/Other Pacific Islander	33	0.15	32	0.14	27	0.17	26	0.17	0.02	0.03
White	14,613	65.95	14,621	65.96	10,273	63.25	9,986	63.66	-2.70	-2.30
Two or More Races	901	4.07	902	4.07	582	3.58	526	3.35	-0.49	-0.72

Table A.5. Student Sample Demographic Comparison Before and After Merging—Grade 7

	Before Merge				After Merge				%Difference (After - Before)	
	ELA		MA		ELA_RD		MA_MA			
Demographic Subgroup	N	%	N	%	N	%	N	%	ELA_RD	MA
Total	21,960	100.00	22,017	100.00	14,873	100.00	14,345	100.00	–	–
Sex										
Female	10,615	48.34	10,636	48.31	7,152	48.09	6,917	48.22	-0.25	-0.09
Male	11,345	51.66	11,381	51.69	7,721	51.91	7,428	51.78	0.25	0.09
Race										
American Indian/Alaska Native	267	1.22	269	1.22	150	1.01	146	1.02	-0.21	-0.20
Asian	586	2.67	591	2.68	435	2.92	411	2.87	0.25	0.19
Black	1,260	5.74	1,270	5.77	957	6.43	890	6.20	0.69	0.43
Hispanic	4,124	18.78	4,153	18.86	3,245	21.82	3,123	21.77	3.04	2.91
Native Hawaiian/Other Pacific Islander	35	0.16	35	0.16	25	0.17	25	0.17	0.01	0.01
White	14,783	67.32	14,794	67.20	9,516	63.98	9,282	64.71	-3.34	-2.49
Two or More Races	904	4.12	903	4.10	545	3.66	467	3.26	-0.46	-0.84

Table A.6. Student Sample Demographic Comparison Before and After Merging—Grade 8

	Before Merge				After Merge				%Difference (After - Before)	
	ELA		MA		ELA_RD		MA_MA			
Demographic Subgroup	N	%	N	%	N	%	N	%	ELA_RD	MA
Total	20,572	100.00	20,611	100.00	13,503	100.00	13,316	100.00	–	–
Sex										
Female	9,844	47.85	9,858	47.83	6,412	47.49	6,335	47.57	-0.36	-0.26
Male	10,728	52.15	10,753	52.17	7,091	52.51	6,981	52.43	0.36	0.26
Race										
American Indian/Alaska Native	279	1.36	275	1.33	156	1.16	149	1.12	-0.20	-0.21
Asian	494	2.40	496	2.41	343	2.54	342	2.57	0.14	0.16
Black	1,173	5.70	1,191	5.78	846	6.27	838	6.29	0.57	0.51
Hispanic	3,887	18.90	3,918	19.01	2,887	21.38	2,875	21.59	2.48	2.58
Native Hawaiian/Other Pacific Islander	37	0.18	39	0.19	27	0.20	28	0.21	0.02	0.02
White	13,923	67.69	13,921	67.55	8,774	64.98	8,655	65.00	-2.71	-2.55
Two or More Races	776	3.77	769	3.73	469	3.47	428	3.21	-0.30	-0.52

Appendix B: Example Plots of ICC and Student Responses

Figure B.1. Example Plot of ICC and Student Responses—NSCAS Math Grade 3 Dichotomous Item Not Selected as an Anchor

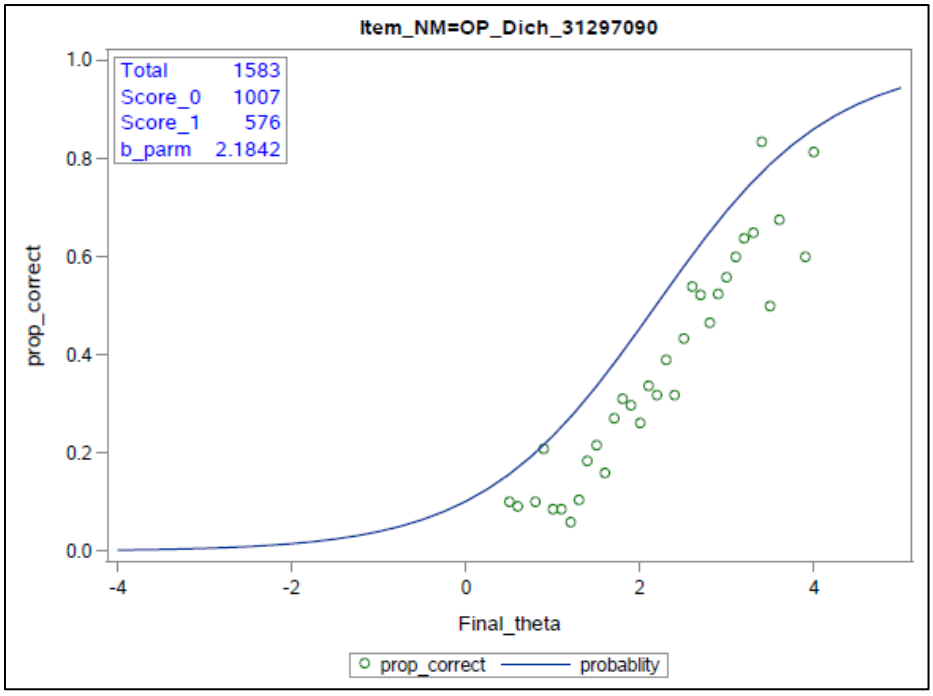


Figure B.2. Example Plot of ICC and Student Responses—NSCAS Math Grade 3 Dichotomous Item Selected as an Anchor

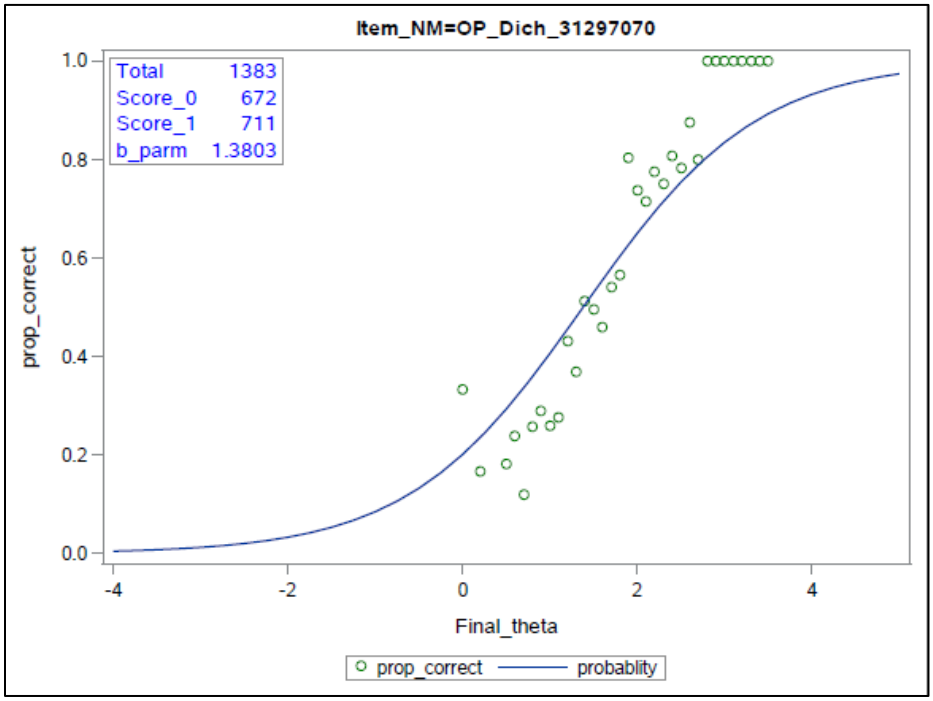


Figure B.3. Example Plot of ICC and Student Responses—NSCAS Math Grade 3 Polytomous Item Not Selected as an Anchor

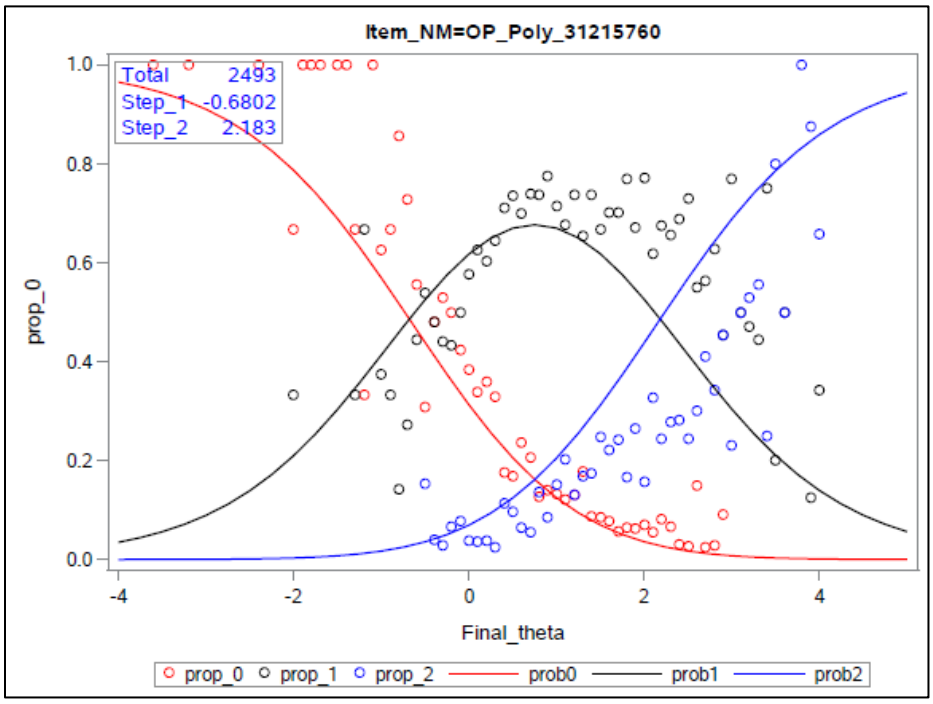
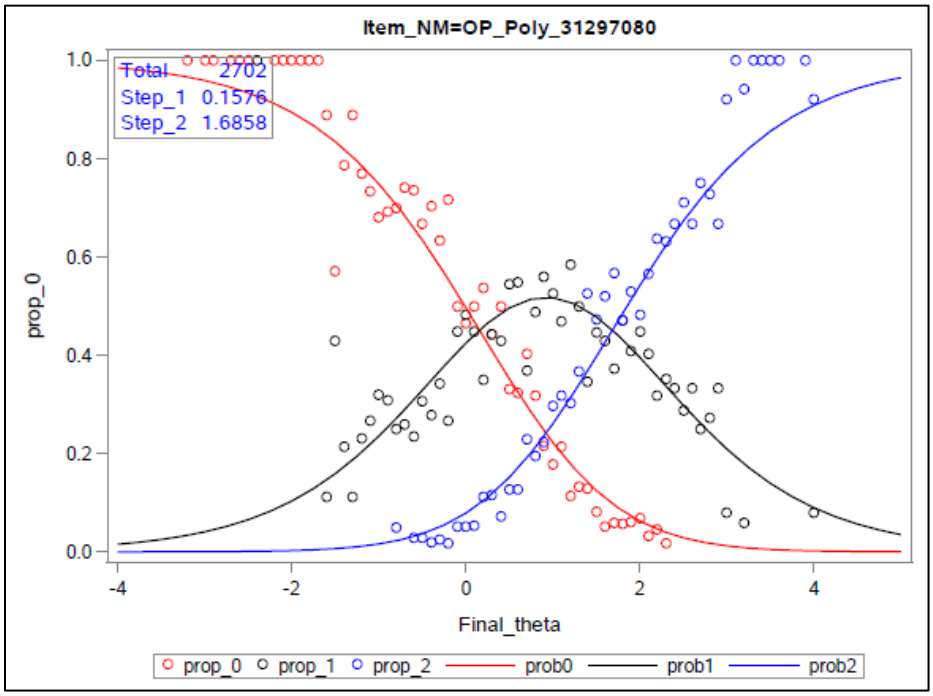


Figure B.4. Example Plot of ICC and Student Responses—NSCAS Math Grade 3 Polytomous Item Selected as an Anchor



Appendix C: RIT vs. Linked RIT Scores

Table C.1. RIT vs. IRT Linked RIT—Two Reporting Categories Only

Grade	N (After Merge)	RIT				IRT Linked RIT				Mean Difference (Linked RIT - RIT)	Effect Size*
		Mean	SD	Min.	Max.	Mean	SD	Min.	Max.		
ELA_RD (MM)											
3	18,442	198.98	15.76	135	245	196.60	13.13	137	237	-2.37	-0.16
4	15,462	206.08	15.28	140	260	203.72	12.57	153	257	-2.36	-0.17
5	15,761	211.71	14.93	145	262	209.60	12.09	160	256	-2.12	-0.16
6	16,242	215.00	15.08	156	261	213.80	11.74	164	270	-1.20	-0.09
7	14,873	217.58	15.45	154	267	215.07	11.42	168	256	-2.52	-0.19
8	13,503	221.27	15.51	151	274	219.26	11.21	177	261	-2.02	-0.15
ELA_RD (MS)											
3	18,442	198.98	15.76	135	245	196.61	13.18	137	238	-2.37	-0.16
4	15,462	206.08	15.28	140	260	203.59	12.06	154	255	-2.49	-0.18
5	15,761	211.71	14.93	145	262	209.44	11.74	161	255	-2.27	-0.17
6	16,242	215.00	15.08	156	261	213.74	11.40	165	269	-1.26	-0.09
7	14,873	217.58	15.45	154	267	215.07	11.73	167	256	-2.51	-0.18
8	13,503	221.27	15.51	151	274	219.19	11.87	174	264	-2.09	-0.15

*Results are highlighted if abs (effect size) ≥ 0.1.

Table C.2. RIT vs. IRT Linked RIT (MM)—All Reporting Categories

Grade	N (After Merge)	RIT				IRT Linked RIT (MM)				Mean Difference (Linked RIT - RIT)	Effect Size*
		Mean	SD	Min.	Max.	Mean	SD	Min.	Max.		
ELA_RD											
3	18,442	198.98	15.76	135	245	196.80	11.82	163	236	-2.18	-0.16
4	15,462	206.08	15.28	140	260	203.44	11.41	169	250	-2.64	-0.20
5	15,761	211.71	14.93	145	262	210.05	11.06	178	248	-1.66	-0.13
6	16,242	215.00	15.08	156	261	213.81	10.67	182	260	-1.20	-0.09
7	14,873	217.58	15.45	154	267	215.27	10.37	183	251	-2.31	-0.18
8	13,503	221.27	15.51	151	274	219.63	10.01	187	252	-1.64	-0.13
MA_MA											
3	15,609	202.49	14.22	138	266	204.03	14.03	171	256	1.54	0.11
4	15,548	211.21	15.64	139	269	215.05	13.47	179	267	3.84	0.26
5	15,897	219.38	17.21	144	289	221.37	12.89	184	273	1.99	0.13
6	15,687	223.27	16.78	146	288	226.24	13.06	189	280	2.97	0.20
7	14,345	227.94	17.85	138	307	229.87	12.40	195	284	1.92	0.13
8	13,316	232.81	19.15	136	316	238.29	12.90	201	292	5.48	0.34

*Results are highlighted if abs (effect size) ≥ 0.1.

Table C.3. RIT vs. IRT Linked RIT (MS)—All Reporting Categories

Grade	N (After Merge)	RIT				IRT Linked RIT (MS)				Mean Difference (Linked RIT - RIT)	Effect Size*
		Mean	SD	Min.	Max.	Mean	SD	Min.	Max.		
ELA_RD											
3	18,442	198.98	15.76	135	245	196.77	11.88	163	236	-2.20	-0.16
4	15,462	206.08	15.28	140	260	203.25	10.62	172	247	-2.83	-0.22
5	15,761	211.71	14.93	145	262	209.87	10.55	179	246	-1.85	-0.14
6	16,242	215.00	15.08	156	261	213.66	10.15	183	258	-1.34	-0.10
7	14,873	217.58	15.45	154	267	215.31	10.85	182	253	-2.27	-0.17
8	13,503	221.27	15.51	151	274	219.77	10.96	184	255	-1.51	-0.11
MA_MA											
3	15,609	202.49	14.22	138	266	203.95	13.98	171	256	1.46	0.10
4	15,548	211.21	15.64	139	269	216.23	16.78	171	281	5.02	0.31
5	15,897	219.38	17.21	144	289	223.59	17.05	174	292	4.21	0.25
6	15,687	223.27	16.78	146	288	226.88	16.40	180	294	3.60	0.22
7	14,345	227.94	17.85	138	307	231.05	16.61	185	303	3.10	0.18
8	13,316	232.81	19.15	136	316	237.72	17.52	187	310	4.91	0.27

*Results are highlighted if abs (effect size) ≥ 0.1.

Table C.4. RIT vs. Linked RIT—Equipercentile Linked RIT (2021 Data)

Grade	N (After Merge)	RIT				Equipercentile Linked RIT (2021)				Mean Difference (Linked RIT - RIT)	Effect Size
		Mean	SD	Min.	Max.	Mean	SD	Min.	Max.		
ELA_RD											
3	18,442	198.98	15.76	135	245	199.04	15.62	135	245	0.06	<0.01
4	15,462	206.08	15.28	140	260	206.07	15.24	140	260	-0.01	<0.01
5	15,761	211.71	14.93	145	262	211.71	14.93	145	262	0.00	<0.01
6	16,242	215.00	15.08	156	261	214.99	14.97	156	261	-0.02	<0.01
7	14,873	217.58	15.45	154	267	217.53	15.33	154	267	-0.05	<0.01
8	13,503	221.27	15.51	151	274	221.19	15.50	151	274	-0.09	-0.01
MA_MA											
3	15,609	202.49	14.22	138	266	202.52	14.33	139	252	0.03	<0.01
4	15,548	211.21	15.64	139	269	211.26	15.60	139	266	0.04	<0.01
5	15,897	219.38	17.21	144	289	219.36	17.37	144	284	-0.02	<0.01
6	15,687	223.27	16.78	146	288	223.26	16.83	146	286	-0.02	<0.01
7	14,345	227.94	17.85	138	307	227.93	18.09	138	291	-0.02	<0.01
8	13,316	232.81	19.15	136	316	232.80	19.19	138	305	-0.01	<0.01

Table C.5. RIT vs. Linked RIT using the Merged Data—Equipercentile Linked RIT (2019 Data)

Grade	N (After Merge)	RIT				Equipercentile Linked RIT (2019)				Mean Difference (Linked RIT - RIT)	Effect Size*
		Mean	SD	Min.	Max.	Mean	SD	Min.	Max.		
ELA_RD											
3	18,442	198.98	15.76	135	245	195.74	20.40	103	246	-3.23	-0.18
4	15,462	206.08	15.28	140	260	204.16	18.48	143	334	-1.92	-0.11
5	15,761	211.71	14.93	145	262	209.67	19.28	142	256	-2.04	-0.12
6	16,242	215.00	15.08	156	261	213.38	18.26	148	343	-1.63	-0.10
7	14,873	217.58	15.45	154	267	216.89	16.28	144	291	-0.69	-0.04
8	13,503	221.27	15.51	151	274	220.75	17.97	150	266	-0.52	-0.03
MA_MA											
3	15,609	202.49	14.22	138	266	202.44	15.09	136	256	-0.05	0.00
4	15,548	211.21	15.64	139	269	210.21	18.00	138	268	-1.00	-0.06
5	15,897	219.38	17.21	144	289	217.66	18.92	140	279	-1.73	-0.10
6	15,687	223.27	16.78	146	288	222.18	18.89	136	285	-1.09	-0.06
7	14,345	227.94	17.85	138	307	227.72	20.77	143	290	-0.23	-0.01
8	13,316	232.81	19.15	136	316	231.71	20.99	139	299	-1.09	-0.05

*Results are highlighted if abs (effect size) ≥ 0.1.

Appendix D: NSCAS vs. Linked RIT Achievement Level Distributions

Table D.1. NSCAS vs. IRT Linked RIT Achievement Level Distributions—Two Reporting Categories Only

Grade	N (Before Merge)	NSCAS			IRT Linked RIT			Difference (Linked RIT – NSCAS)		
		%Dev	%OT	%CCR	%Dev	%OT	%CCR	%Dev	%OT	%CCR
ELA_RD (MM)										
3	21,621	49.5	36.1	14.4	48.4	34.1	17.5	-1.1	-2.0	3.1
4	21,551	45.9	36.8	17.3	42.2	37.1	20.7	-3.7	0.3	3.4
5	22,046	53.8	31.5	14.8	52.1	32.1	15.8	-1.7	0.6	1.0
6	22,157	54.0	30.2	15.8	51.4	30.5	18.1	-2.6	0.3	2.3
7	21,960	55.1	35.9	9.0	53.0	34.1	13.0	-2.1	-1.8	4.0
8	20,572	49.1	37.9	13.0	49.2	35.9	14.9	0.1	-2.0	1.9
ELA_RD (MS)										
3	21,621	49.5	36.1	14.4	48.4	34.0	17.6	-1.1	-2.1	3.2
4	21,551	45.9	36.8	17.3	42.6	35.7	21.7	-3.3	-1.1	4.4
5	22,046	53.8	31.5	14.8	52.9	29.9	17.2	-0.9	-1.6	2.4
6	22,157	54.0	30.2	15.8	51.9	28.2	19.9	-2.1	-2.0	4.1
7	21,960	55.1	35.9	9.0	52.6	35.8	11.6	-2.5	-0.1	2.6
8	20,572	49.1	37.9	13.0	49.0	37.1	13.9	-0.1	-0.8	0.9

Table D.2. NSCAS vs. IRT Linked RIT (MM) Achievement Level Distributions—All Reporting Categories

Grade	N (Before Merge)	NSCAS			IRT Linked RIT (MM)			Difference (Linked RIT – NSCAS)		
		%Dev	%OT	%CCR	%Dev	%OT	%CCR	%Dev	%OT	%CCR
ELA_RD										
3	21,621	49.5	36.1	14.4	47.7	36.7	15.6	-1.8	0.6	1.2
4	21,551	45.9	36.8	17.3	42.9	39.4	17.8	-3.0	2.6	0.5
5	22,046	53.8	31.5	14.8	50.7	34.4	15.0	-3.1	2.9	0.2
6	22,157	54.0	30.2	15.8	50.7	33.5	15.8	-3.3	3.3	0.0
7	21,960	55.1	35.9	9.0	51.9	37.2	10.9	-3.2	1.3	1.9
8	20,572	49.1	37.9	13.0	48.3	38.8	13.0	-0.8	0.9	0.0
MA_MA										
3	21,482	52.2	38.3	9.5	50.8	39.7	9.5	-1.4	1.4	0.0
4	21,605	54.2	37.7	8.2	53.6	37.7	8.6	-0.6	0.0	0.4
5	22,130	54.3	38.2	7.6	53.1	38.7	8.2	-1.2	0.5	0.6
6	22,167	52.7	39.2	8.1	52.5	39.5	8.0	-0.2	0.3	-0.1
7	22,017	53.7	38.4	7.9	50.7	40.6	8.7	-3.0	2.2	0.8
8	20,611	54.5	37.8	7.7	54.5	37.4	8.1	0.0	-0.4	0.4

Table D.3. NSCAS vs. IRT Linked RIT (MS) Achievement Level Distributions—All Reporting Categories

Grade	N (Before Merge)	NSCAS			IRT Linked RIT (MS)			Difference (Linked RIT – NSCAS)		
		%Dev	%OT	%CCR	%Dev	%OT	%CCR	%Dev	%OT	%CCR
ELA_RD										
3	21,621	49.5	36.1	14.4	47.7	36.7	15.6	-1.8	0.6	1.2
4	21,551	45.9	36.8	17.3	43.3	39.0	17.8	-2.6	2.2	0.5
5	22,046	53.8	31.5	14.8	51.8	32.5	15.6	-2.0	1.0	0.8
6	22,157	54.0	30.2	15.8	51.7	31.6	16.7	-2.3	1.4	0.9
7	21,960	55.1	35.9	9.0	51.4	38.6	10.0	-3.7	2.7	1.0
8	20,572	49.1	37.9	13.0	47.3	39.5	13.2	-1.8	1.6	0.2
MA_MA										
3	21,482	52.2	38.3	9.5	51.0	38.5	10.5	-1.2	0.2	1.0
4	21,605	54.2	37.7	8.2	52.5	39.1	8.4	-1.7	1.4	0.2
5	22,130	54.3	38.2	7.6	54.0	38.5	7.6	-0.3	0.3	0.0
6	22,167	52.7	39.2	8.1	52.5	39.0	8.5	-0.2	-0.2	0.4
7	22,017	53.7	38.4	7.9	53.3	38.6	8.1	-0.4	0.2	0.2
8	20,611	54.5	37.8	7.7	52.4	39.3	8.3	-2.1	1.5	0.6

Table D.4. NSCAS vs. Equipercentile Linked RIT Achievement Level Distributions (2021 Data)

Grade	N (Before Merge)	NSCAS			Equipercentile Linked RIT (2021 Data)			Difference (Linked RIT – NSCAS)		
		%Dev	%OT	%CCR	%Dev	%OT	%CCR	%Dev	%OT	%CCR
ELA_RD										
3	21,621	49.5	36.1	14.4	49.5	35.2	15.3	0.0	-0.9	0.9
4	21,551	45.9	36.8	17.3	43.9	38.6	17.5	-2.0	1.8	0.2
5	22,046	53.8	31.5	14.8	52.7	32.3	15.0	-1.1	0.8	0.2
6	22,157	54.0	30.2	15.8	54.0	30.2	15.8	0.0	0.0	0.0
7	21,960	55.1	35.9	9.0	53.8	37.2	9.0	-1.3	1.3	0.0
8	20,572	49.1	37.9	13.0	47.3	38.1	14.5	-1.8	0.2	1.5
MA_MA										
3	21,482	52.2	38.3	9.5	50.8	39.5	9.8	-1.4	1.2	0.3
4	21,605	54.2	37.7	8.2	54.2	36.7	9.2	0.0	-1.0	1.0
5	22,130	54.3	38.2	7.6	53.6	38.4	8.0	-0.7	0.2	0.4
6	22,167	52.7	39.2	8.1	50.8	40.6	8.7	-1.9	1.4	0.6
7	22,017	53.7	38.4	7.9	53.0	38.5	8.6	-0.7	0.1	0.7
8	20,611	54.5	37.8	7.7	52.8	39.3	7.9	-1.7	1.5	0.2

Table D.5. NSCAS vs. Equipercentile Linked RIT Achievement Level Distributions (2019 Data)

Grade	N (Before Merge)	NSCAS			Equipercentile Linked RIT (2019 Data)			Difference (Linked RIT – NSCAS)		
		%Dev	%OT	%CCR	%Dev	%OT	%CCR	%Dev	%OT	%CCR
ELA_RD										
3	21,621	49.5	36.1	14.4	48.6	36.1	15.3	-0.9	0.0	0.9
4	21,551	45.9	36.8	17.3	44.9	37.1	18.0	-1.0	0.3	0.7
5	22,046	53.8	31.5	14.8	51.8	33.2	15.0	-2.0	1.7	0.2
6	22,157	54.0	30.2	15.8	54.0	29.3	16.7	0.0	-0.9	0.9
7	21,960	55.1	35.9	9.0	54.6	35.5	9.8	-0.5	-0.4	0.8
8	20,572	49.1	37.9	13.0	47.3	38.6	14.0	-1.8	0.7	1.0
MA_MA										
3	21,482	52.2	38.3	9.5	49.5	40.3	10.2	-2.7	2.0	0.7
4	21,605	54.2	37.7	8.2	51.9	39.5	8.5	-2.3	1.8	0.3
5	22,130	54.3	38.2	7.6	52.7	39.2	8.0	-1.6	1.0	0.4
6	22,167	52.7	39.2	8.1	51.4	39.9	8.7	-1.3	0.7	0.6
7	22,017	53.7	38.4	7.9	53.0	39.1	7.9	-0.7	0.7	0.0
8	20,611	54.5	37.8	7.7	54.5	37.8	7.7	0.0	0.0	0.0

Appendix E: Score Distribution Plots: IRT Linked RIT (MS)

The ELA_RD plots in this appendix use items from the two reading reporting categories only, whereas the MA_MA plots use all reporting categories.

Figure E.1. Score Distribution Plot, IRT Linked RIT—ELA_RD, Grade 3

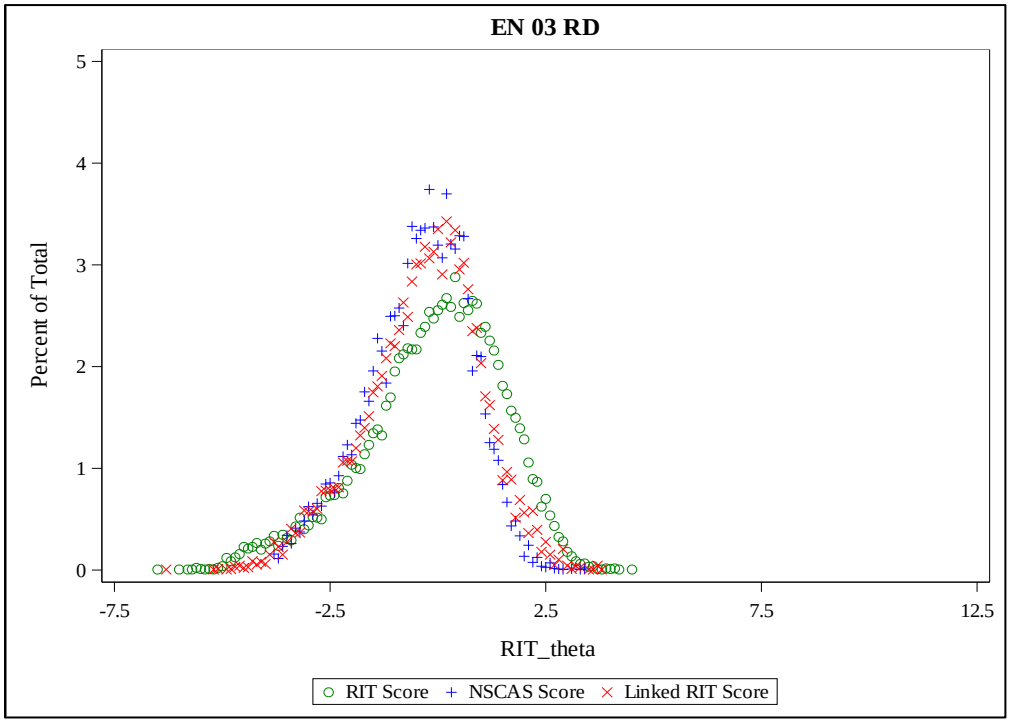


Figure E.2. Score Distribution Plot, IRT Linked RIT—ELA_RD, Grade 4

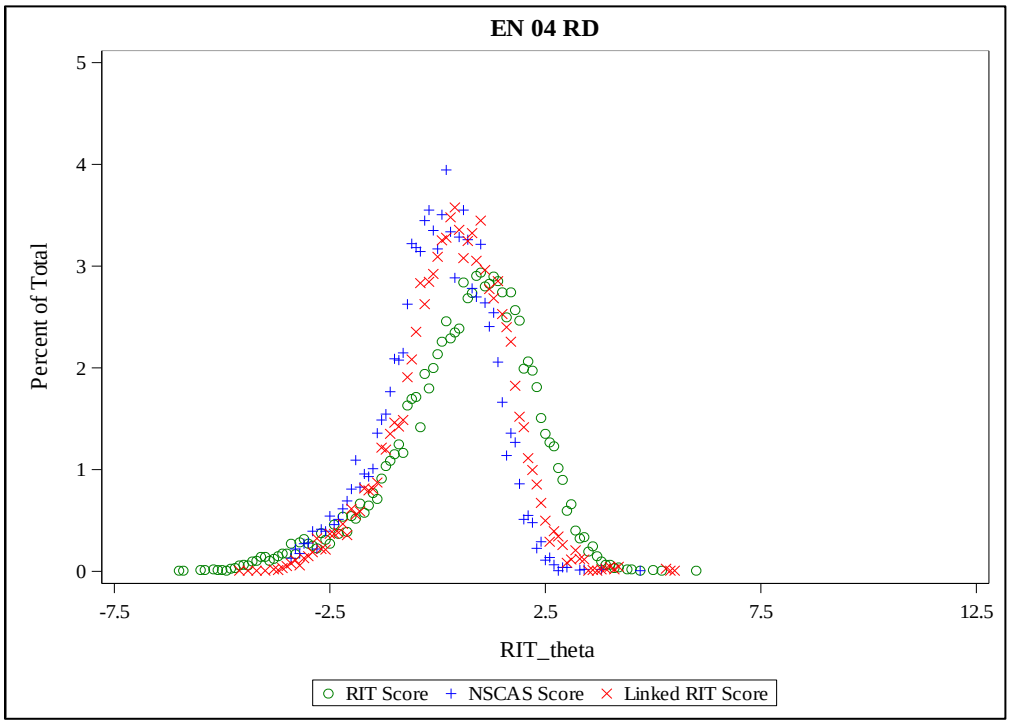


Figure E.3. Score Distribution Plot, IRT Linked RIT—ELA_RD, Grade 5

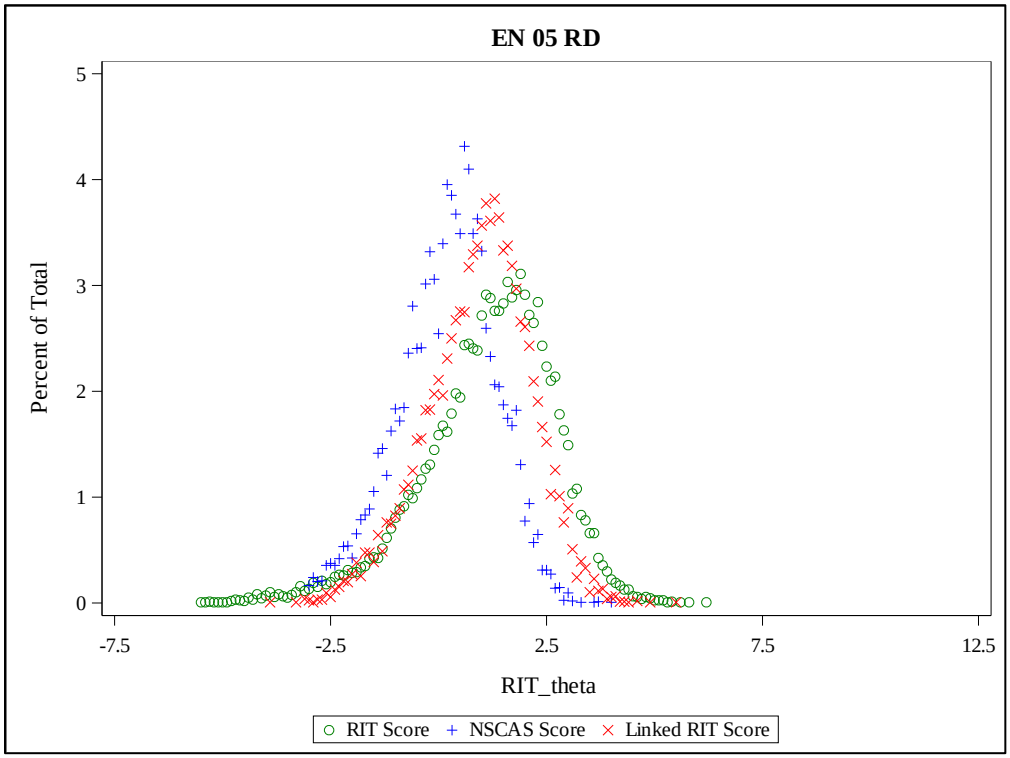


Figure E.4. Score Distribution Plot, IRT Linked RIT—ELA_RD, Grade 6

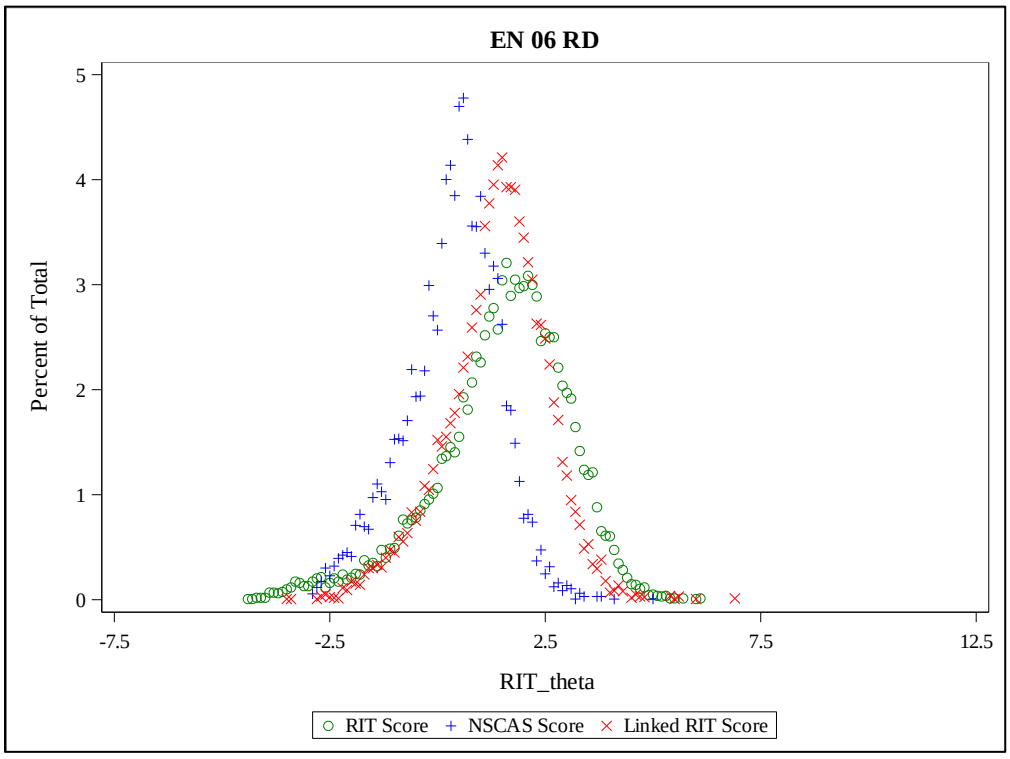


Figure E.5. Score Distribution Plot, IRT Linked RIT—ELA_RD, Grade 7

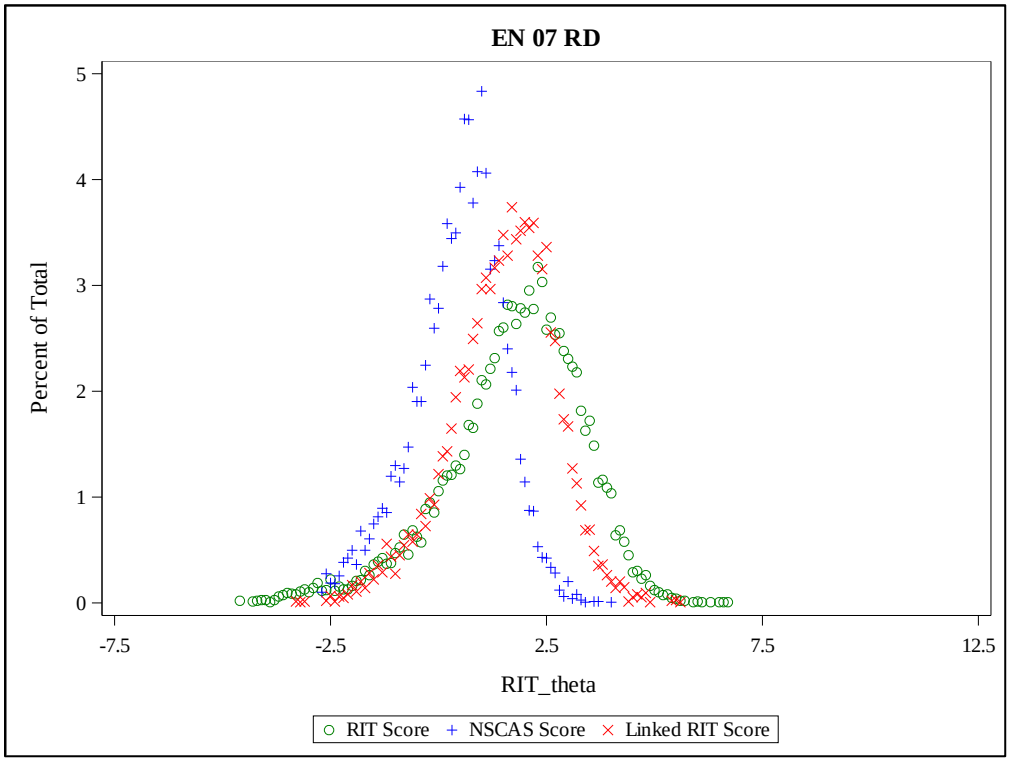


Figure E.6. Score Distribution Plot, IRT Linked RIT—ELA_RD, Grade 8

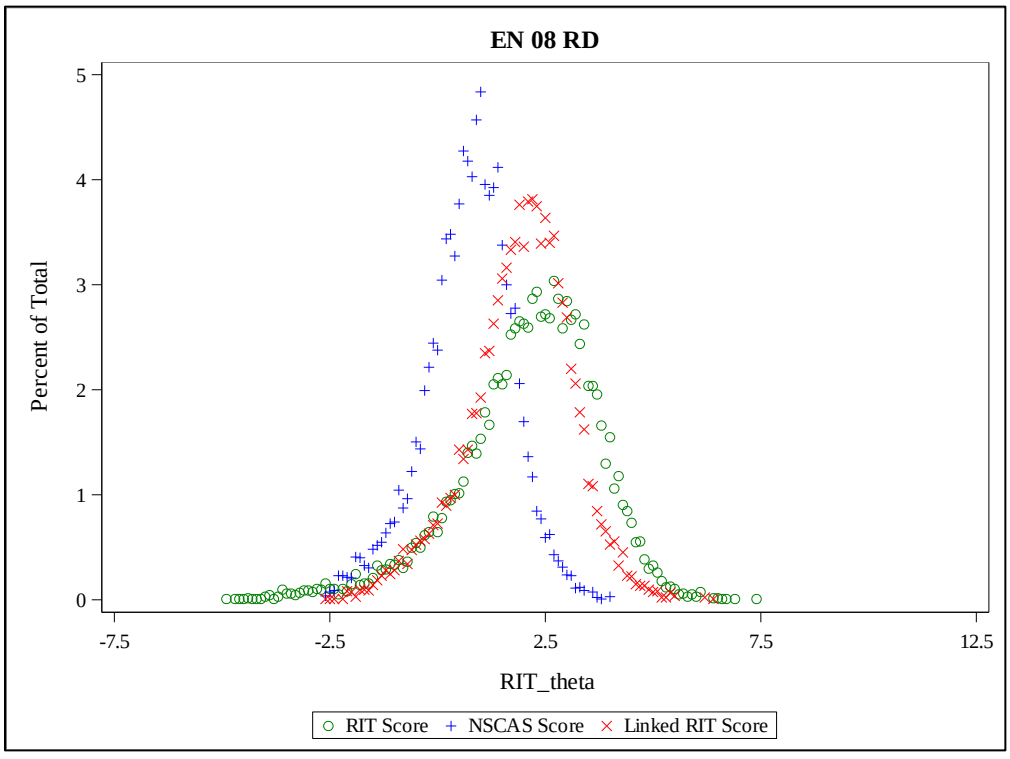


Figure E.7. Score Distribution Plot, IRT Linked RIT—MA_MA, Grade 3

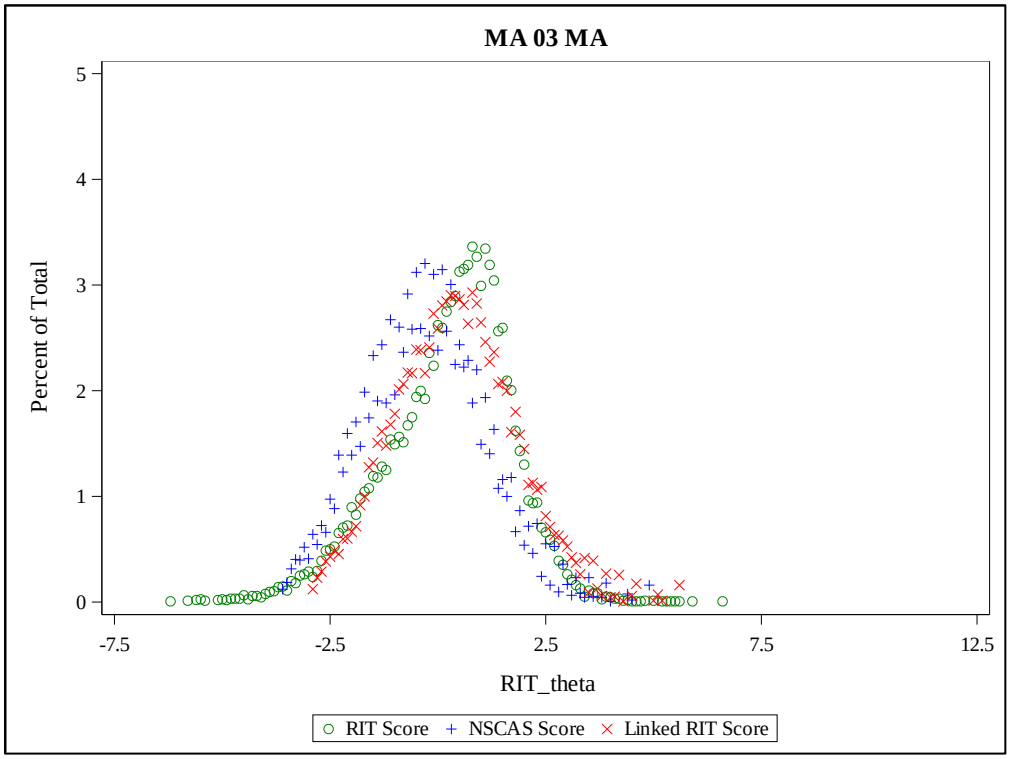


Figure E.8. Score Distribution Plot, IRT Linked RIT—MA_MA, Grade 4

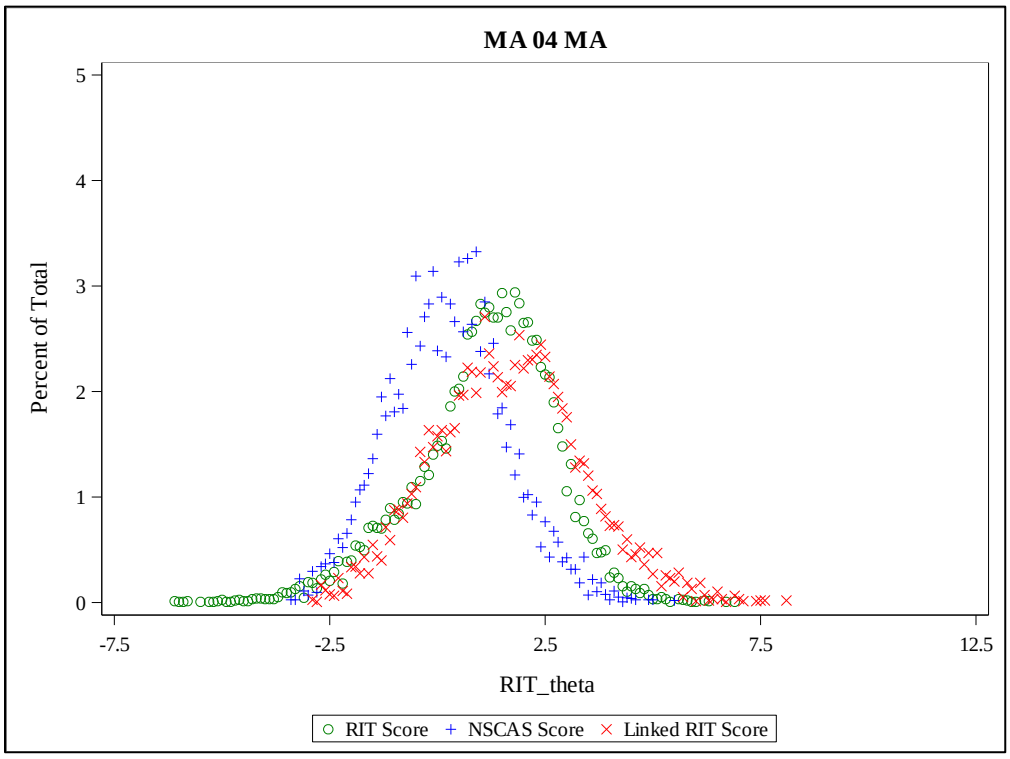


Figure E.9. Score Distribution Plot, IRT Linked RIT—MA_MA, Grade 5

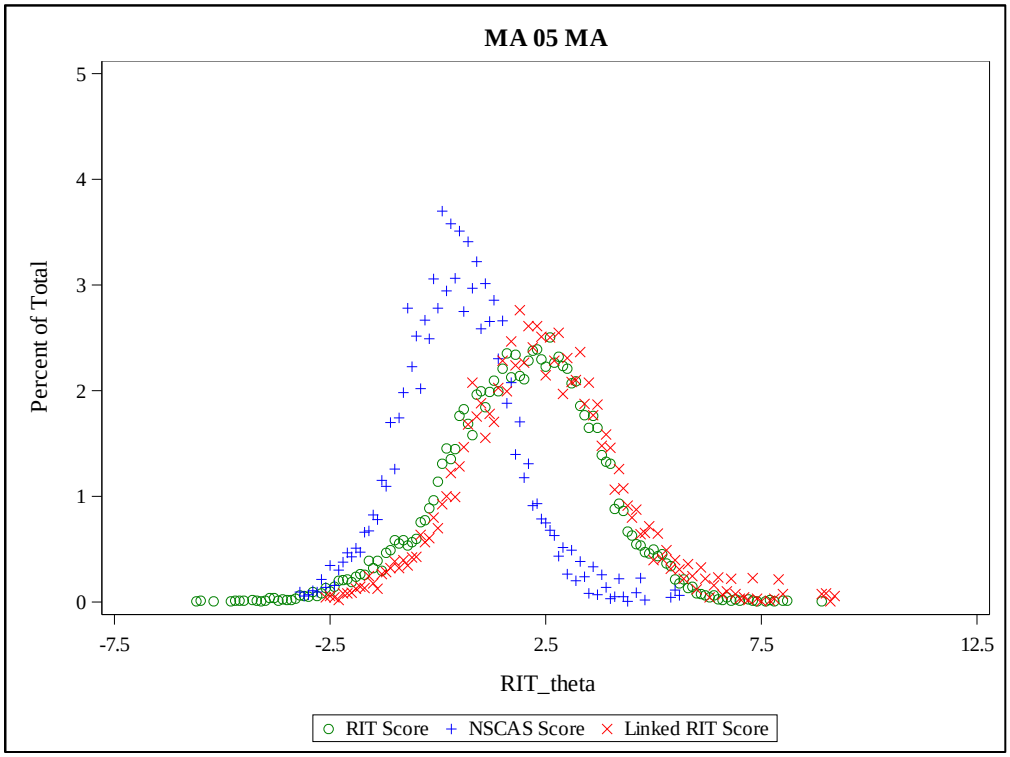


Figure E.10. Score Distribution Plot, IRT Linked RIT—MA_MA, Grade 6

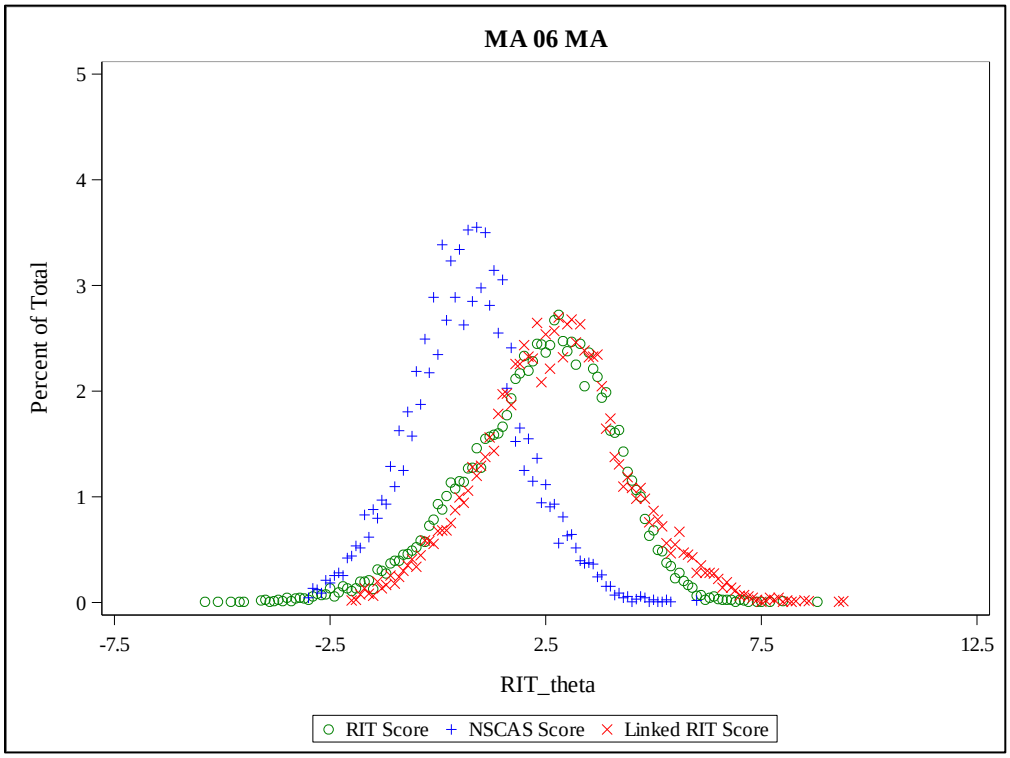


Figure E.11. Score Distribution Plot, IRT Linked RIT—MA_MA, Grade 7

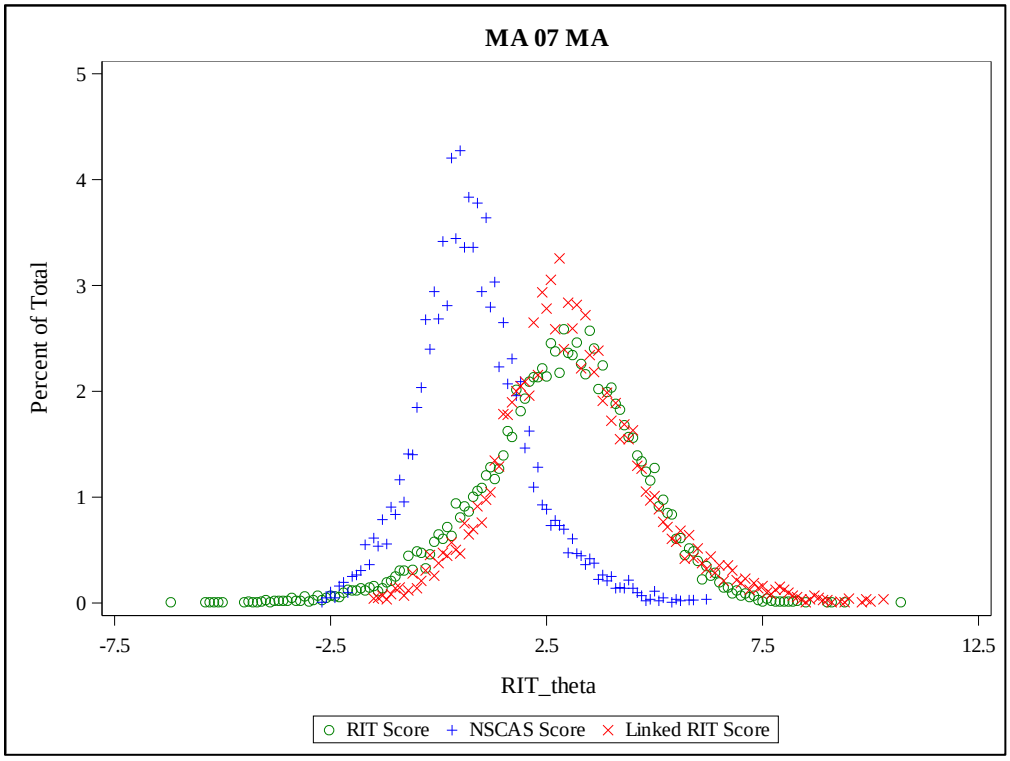


Figure E.12. Score Distribution Plot, IRT Linked RIT—MA_MA, Grade 8