College Readiness LINKING STUDY

A Study of the Alignment of the RIT Scales of NWEA's MAP® Assessments with the College Readiness Benchmarks of EXPLORE®, PLAN®, and ACT®

December 2011

(updated January 17, 2012)

A STUDY OF THE ALIGNMENT OF THE RIT SCALES OF NWEA'S MAP® ASSESSMENTS WITH THE COLLEGE READINESS BENCHMARKS OF EXPLORE®, PLAN®, AND ACT®

DECEMBER 2011

Recently, NWEA completed a study to examine the predictive relationship between the RIT scales of NWEA's MAP® assessments in reading, language usage, and mathematics to the college readiness benchmarks of the EXPLORE, PLAN, and ACT achievement tests in reading, English, and mathematics. The EXPLORE, PLAN, and ACT also offer tests in science achievement, but these tests were not included in the current study. The objective of this study was to identify cut scores on the MAP reading, language usage, and general mathematics tests that correspond to the published college readiness benchmarks on the EXPLORE, PLAN, and ACT assessments (ACT, 2010). A secondary objective was to create a series of probability tables that estimate the likelihood of meeting the designated college readiness benchmark, given an observed MAP score.

To conduct the study, we linked together individual EXPLORE, PLAN, and ACT scale scores and NWEA MAP assessment RIT scores for a sample of students who had completed both exams in the same (or a comparable) subject. EXPLORE, PLAN, and ACT scores were provided by NWEA partnering school districts and individually linked to those students' MAP assessment RIT scores from the same (or the prior) testing season. In all, the sample contained over 108,000 matched pairs of scores from 26,000 students from 140 schools in three states. All valid matched data (i.e., data with valid scores and linking IDs) from the resulting sample were included in the analyses; no attempt was made to rebalance the sample in order to simulate a state- or nationally-representative population.

Visual examinations of scatter plots of the data revealed curvilinear relationships between the MAP scale scores and the EXPLORE, PLAN and ACT scale scores. Consequently, a series of curvilinear (quadratic) regression models were fitted to the data, using MAP RIT scores as the single predictor of performance on each of the college readiness tests. MAP assessments in reading and language usage were both fit to predictive models of performance on college readiness tests of English and Reading. MAP mathematics was used to predict mathematics college readiness. In all, fifteen predictive models were fitted. See the Methodology Appendix for a more detailed description of the methods used.

Table Sets 1 and 2 show the estimated cut scores, or minimum equivalent RIT scores corresponding to the college readiness benchmarks on EXPLORE, PLAN, or ACT when taken in the same (spring) or prior (fall) testing seasons. Also shown are the NWEA normative percentile ranks[1] associated with these MAP

[1] Percentile ranks are based on NWEA's 2011 norming study.

Updated on January 17, 2012

cut scores. These tables can be used to identify students who might benefit from additional assistance, or who may be at risk of failing to meet these benchmarks. The percentile ranks also provide an indicator of the difficulty of these benchmarks, relative to a nationally representative norming sample. In general, the MAP cut scores associated with the college readiness benchmarks in English range from the 40th to 60th percentiles on the MAP reading and language usage tests, while the college readiness benchmarks in reading and mathematics are higher, ranging primarily from the 70th to 80th percentiles on the MAP assessments reading and mathematics tests, respectively.

The tables in Table Set 3 show the estimated probability of a student meeting the designated college readiness benchmark, based on that student's RIT score taken in the same testing season. These tables provide empirical information about the likelihood of meeting or exceeding the designated college readiness benchmark, given an observed MAP RIT score.

The tables in Table Set 4 show the correlation coefficients and the goodness-of-fit statistics for the regression models used to predict MAP and the corresponding college readiness tests. These statistics show the degree to which MAP scores accurately predicted the PLAN, EXPLORE, and ACT scale scores of the study sample. The reported models show a moderately high correlation between MAP RIT scores and the scores on the college readiness benchmark tests, with correlations ranging from .66 to .87, but primarily in the range of .75-.80. In general, values at or near 1.0 suggest a perfect predictive relationship, whereas values near 0.0 indicate no predictive relationship. Goodness-of-fit statistics indicate that substantial variation within the observed college readiness benchmark scores can be predicted by MAP RIT scores, with values ranging from 44-76% of observed variance.

The tables in Table Set 5 show the accuracy of the estimated MAP cut scores in predicting whether students met or exceeded the corresponding college readiness benchmark for the study sample. In general, the estimated MAP cut scores accurately predicted whether or not students would meet the EXPLORE, PLAN, and ACT benchmarks with 75-90% accuracy. Among incorrect predictions, false negatives (students who were incorrectly predicted NOT to meet the college readiness benchmark) outnumbered false positives (students predicted to meet readiness benchmarks but who failed to do so).

The estimated MAP cut scores in this report provide a basis for making useful predictions about students' likely college readiness status, as measured by EXPLORE, PLAN, and ACT, when MAP is taken within the same (or nearly the same) testing season. However, MAP is not designed to measure identical content as the ACT assessments. MAP assessments are aligned to each state's curriculum standards rather than the curriculum standards of the ACT. Thus while the tests measure much content that would be similar, they do not share a common design. Knowledge of a student's MAP score permits fairly accurate predictions about a student's probable college readiness status, as measured by EXPLORE, PLAN, or ACT.

TABLE SET 1 – MINIMUM ESTIMATED SAME-SEASON (SPRING) RIT CUT SCORES CORRESPONDING TO COLLEGE READINESS BENCHMARKS[2]

MAP Reading RIT Score as Predictor – Same Season								
	Cut Scores and Normative Percentile Ranks on MAP Corresponding to College Readiness Benchmarks							
Grade	**Reading College Readiness Test**	Benchmark	MAP Cut Score	MAP Normative Percentile Rank	**English College Readiness Test**	Benchmark	MAP Cut Score	MAP Normative Percentile Rank
8	EXPLORE Reading	15	230	70	EXPLORE English	13	220	44
10	PLAN Reading	17	234	73	PLAN English	15	227	58
11	ACT Reading	21	237	77	ACT English	18	232	68

MAP Language Usage RIT Score as Predictor – Same Season								
	Cut Scores and Normative Percentile Ranks on MAP Corresponding to College Readiness Benchmarks							
Grade	**Reading College Readiness Test**	Benchmark	MAP Cut Score	MAP Normative Percentile Rank	**English College Readiness Test**	Benchmark	MAP Cut Score	MAP Normative Percentile Rank
8	EXPLORE Reading	15	229	72	EXPLORE English	13	219	43
10	PLAN Reading	17	232	73	PLAN English	15	225	56
11	ACT Reading	21	234	76	ACT English	18	228	62

MAP Mathematics RIT Score as Predictor – Same Season				
	Cut Scores and Normative Percentile Ranks on MAP Corresponding to College Readiness Benchmarks			
Grade	**Mathematics College Readiness Test**	Benchmark	MAP Cut Score	MAP Normative Percentile Rank
8	EXPLORE Math	17	245	72
10	PLAN Math	19	251	77
11	ACT Math	22	258	83

[2] The MAP cut scores shown in these tables are the **minimum** estimated scores. Meeting the minimum MAP cut score corresponds to a 50% probability of achieving that benchmark. Use the probabilities in Table Set 3 to determine the appropriate 'target' scores for a desired level of certainty.

TABLE SET 2 – MINIMUM ESTIMATED PRIOR-SEASON (FALL) RIT CUT SCORES CORRESPONDING TO COLLEGE READINESS BENCHMARKS[3]

MAP Reading RIT Score as Predictor – Prior Season								
	Cut Scores and Normative Percentile Ranks on MAP Corresponding to College Readiness Benchmarks							
Grade	**Reading College Readiness Test**	Benchmark	MAP Cut Score	MAP Normative Percentile Rank	**English College Readiness Test**	Benchmark	MAP Cut Score	MAP Normative Percentile Rank
8	EXPLORE Reading	15	227	70	EXPLORE English	13	217	44
10	PLAN Reading	17	232	73	PLAN English	15	226	58
11	ACT Reading	21	236	77	ACT English	18	231	68

MAP Language Usage RIT Score as Predictor – Prior Season								
	Cut Scores and Normative Percentile Ranks on MAP Corresponding to College Readiness Benchmarks							
Grade	**Reading College Readiness Test**	Benchmark	MAP Cut Score	MAP Normative Percentile Rank	**English College Readiness Test**	Benchmark	MAP Cut Score	MAP Normative Percentile Rank
8	EXPLORE Reading	15	226	72	EXPLORE English	13	216	43
10	PLAN Reading	17	230	73	PLAN English	15	224	56
11	ACT Reading	21	233	76	ACT English	18	227	62

MAP Mathematics RIT Score as Predictor – Prior Season				
	Cut Scores and Normative Percentile Ranks on MAP Corresponding to College Readiness Benchmarks			
Grade	**Mathematics College Readiness Test**	Benchmark	MAP Cut Score	MAP Normative Percentile Rank
8	EXPLORE Math	17	240	72
10	PLAN Math	19	248	77
11	ACT Math	22	255	83

[3] The MAP cut scores shown in these tables are the **minimum** estimated scores. Meeting the minimum MAP cut score corresponds to a 50% probability of achieving that benchmark. Use the probabilities in Table Set 3 to determine the appropriate 'target' scores for a desired level of certainty.

TABLE SET 3 –PROBABILITY OF MEETING OR EXCEEDING COLLEGE READINESS BENCHMARK IN SAME SEASON (SPRING), BY STUDENT GRADE AND RIT SCORE RANGE

	MAP Reading RIT Score as Predictor					
	Reading Benchmark			English Benchmark		
MAP Reading	EXPLORE	PLAN	ACT	EXPLORE	PLAN	ACT
RIT Range	8th Grade	10th Grade	11th Grade	8th Grade	10th Grade	11th Grade
145	0%	0%	0%	0%	0%	0%
150	0%	0%	0%	0%	0%	0%
155	0%	0%	0%	0%	0%	0%
160	0%	0%	0%	0%	0%	0%
165	0%	0%	0%	0%	0%	0%
170	0%	0%	0%	0%	0%	0%
175	0%	0%	0%	1%	0%	0%
180	0%	0%	0%	1%	0%	0%
185	0%	1%	0%	1%	2%	0%
190	0%	1%	0%	4%	2%	0%
195	2%	1%	0%	5%	7%	0%
200	2%	2%	0%	10%	10%	0%
205	2%	2%	1%	14%	12%	1%
210	6%	6%	3%	25%	18%	6%
215	14%	10%	6%	39%	30%	11%
220	27%	17%	10%	61%	44%	22%
225	45%	30%	16%	77%	60%	39%
230	64%	44%	32%	91%	76%	56%
235	82%	66%	54%	99%	90%	79%
240	96%	82%	80%	100%	96%	93%
245	100%	92%	93%	100%	100%	99%
250	100%	100%	96%	100%	100%	100%
255	100%	100%	100%	100%	100%	100%
260	100%	100%	100%	100%	100%	100%
265	100%	100%	100%	100%	100%	100%
270	100%	100%	100%	100%	100%	100%
275	100%	100%	100%	100%	100%	100%
280	100%	100%	100%	100%	100%	100%
285	100%	100%	100%	100%	100%	100%
290	100%	100%	100%	100%	100%	100%
295	100%	100%	100%	100%	100%	100%
300	100%	100%	100%	100%	100%	100%

*Note: This table shows the proportion of students in the study sample who, based on a MAP reading score taken during the same (season), met the associated college readiness benchmark. Example: an eighth grade student scoring 220 on a MAP reading test taken during the same season would have a 27% chance of meeting the EXPLORE college readiness benchmark in reading, and about a 61% chance of meeting the EXPLORE college readiness benchmark in English.

MAP Language Usage RIT Score as Predictor						
	Reading Benchmark			English Benchmark		
MAP Language Usage	EXPLORE	PLAN	ACT	EXPLORE	PLAN	ACT
RIT Range	8th Grade	10th Grade	11th Grade	8th Grade	10th Grade	11th Grade
130	0%	0%	0%	0%	0%	0%
135	0%	0%	0%	0%	0%	0%
140	0%	0%	0%	0%	0%	0%
145	0%	0%	0%	0%	0%	0%
150	0%	0%	0%	0%	0%	0%
155	0%	0%	0%	0%	0%	0%
160	0%	0%	0%	0%	0%	0%
165	0%	0%	0%	0%	0%	0%
170	0%	0%	0%	0%	0%	0%
175	0%	0%	0%	0%	0%	0%
180	0%	0%	0%	0%	0%	0%
185	0%	0%	0%	1%	0%	0%
190	0%	0%	0%	1%	0%	0%
195	1%	3%	0%	2%	5%	0%
200	2%	3%	0%	8%	11%	0%
205	3%	4%	0%	11%	14%	4%
210	8%	4%	4%	24%	16%	4%
215	16%	11%	6%	39%	32%	18%
220	27%	22%	12%	64%	50%	34%
225	47%	36%	24%	81%	72%	60%
230	67%	56%	38%	97%	85%	82%
235	84%	76%	77%	100%	95%	92%
240	96%	91%	84%	100%	100%	100%
245	100%	100%	100%	100%	100%	100%
250	100%	100%	100%	100%	100%	100%
255	100%	100%	100%	100%	100%	100%
260	100%	100%	100%	100%	100%	100%
265	100%	100%	100%	100%	100%	100%
270	100%	100%	100%	100%	100%	100%
275	100%	100%	100%	100%	100%	100%
280	100%	100%	100%	100%	100%	100%
285	100%	100%	100%	100%	100%	100%
290	100%	100%	100%	100%	100%	100%
295	100%	100%	100%	100%	100%	100%
300	100%	100%	100%	100%	100%	100%

*Note:
This table shows the proportion of students in the study sample who, based on a MAP language usage score taken during the same (season), met the associated college readiness benchmark. Example: an eighth grade student scoring 220 on a MAP language usage test taken during the same season would have a 27% chance of meeting the EXPLORE college readiness benchmark in reading, and about a 64% chance of meeting the EXPLORE college readiness benchmark in English.

Updated on January 17, 2012

MAP Mathematics RIT Score as Predictor			
	Mathematics Benchmark		
MAP Mathematics	EXPLORE	PLAN	ACT
RIT Range	8th Grade	10th Grade	11th Grade
130	0%	0%	0%
135	0%	0%	0%
140	0%	0%	0%
145	0%	0%	0%
150	0%	0%	0%
155	0%	0%	0%
160	0%	0%	0%
165	0%	0%	0%
170	0%	0%	0%
175	0%	0%	0%
180	0%	0%	0%
185	0%	0%	0%
190	0%	0%	0%
195	0%	0%	0%
200	0%	0%	0%
205	0%	0%	0%
210	0%	0%	0%
215	1%	0%	0%
220	3%	1%	0%
225	7%	1%	0%
230	18%	3%	0%
235	33%	6%	1%
240	53%	18%	2%
245	70%	35%	17%
250	85%	55%	37%
255	93%	75%	62%
260	99%	92%	84%
265	100%	95%	99%
270	100%	100%	100%
275	100%	100%	100%
280	100%	100%	100%
285	100%	100%	100%
290	100%	100%	100%
295	100%	100%	100%
300	100%	100%	100%

*Note:
This table shows the proportion of students in the study sample who, based on a MAP mathematics score taken during the same (season), met the associated college readiness benchmark. Example: an eighth grade student scoring 240 on a MAP mathematics test taken during the same season would have a 53% chance of meeting the EXPLORE college readiness benchmark in mathematics.

Updated on January 17, 2012

TABLE SET 4 – CORRELATIONS BETWEEN MAP AND COLLEGE READINESS TEST SCORES AND REGRESSION MODEL GOODNESS OF FIT STATISTICS[4]

	MAP Reading Test as Predictor					
Grade	College Readiness Test	Correlations	Goodness of Fit	College Readiness Test	Correlations	Goodness of Fit
8	EXPLORE Reading	0.743	55.2%	EXPLORE English	.785	61.6%
10	PLAN Reading	0.686	47.0%	PLAN English	.731	53.4%
11	ACT Reading	0.779	60.7%	ACT English	.800	64.1%

	MAP Language Usage Test as Predictor					
Grade	College Readiness Test	Correlations	Goodness of Fit	College Readiness Test	Correlations	Goodness of Fit
8	EXPLORE Reading	.714	51.0%	EXPLORE English	.804	64.6%
10	PLAN Reading	.662	43.8%	PLAN English	.745	55.5%
11	ACT Reading	.764	58.4%	ACT English	.837	70.0%

	MAP Mathematics Test as Predictor		
Grade	College Readiness Test	Correlations	Goodness of Fit
8	EXPLORE Mathematics	.825	68.0%
10	PLAN Mathematics	.802	64.3%
11	ACT Mathematics	.870	75.7%

[4] These correlations are comparable to Pearson's *r* values, except that they denote the extent to which the two scales are related by a quadratic function. Correlations range from 0 to 1, where 0 indicates no correlation between college readiness test scores and MAP scores, while 1 indicates a completely correlational relationship between scores on the two tests. Goodness of fit statistics indicate the percentage of observed variance accounted for by the quadratic regression model, with 100% indicating that college readiness test scores can be predicted with full reliability, whereas 0% indicates no predictability between the two assessments.

TABLE 5 – PERCENTAGE OF STUDENTS WHOSE PASS STATUS WAS ACCURATELY PREDICTED BY THEIR MAP PERFORMANCE USING REPORTED CUT SCORES[5]

MAP Reading Test as Predictor										
	Percentage of Sample whose College Readiness Status was Accurately Predicted by MAP Score									
Grade	College Readiness Test	Sample Size	Percentage Correctly Predicted	Percentage of False Positives	Percentage of False Negatives	College Readiness Test	Sample Size	Percentage Correctly Predicted	Percentage of False Positives	Percentage of False Negatives
8	EXPLORE Reading	12704	81%	5%	14%	EXPLORE English	12776	79%	10%	10%
10	PLAN Reading	9593	79%	7%	14%	PLAN English	9625	75%	9%	16%
11	ACT Reading	2817	84%	5%	11%	ACT English	2825	80%	7%	13%

MAP Language Usage as Predictor										
	Percentage of Sample whose College Readiness Status was Accurately Predicted by MAP Score									
Grade	College Readiness Test	Sample Size	Percentage Correctly Predicted	Percentage of False Positives	Percentage of False Negatives	College Readiness Test	Sample Size	Percentage Correctly Predicted	Percentage of False Positives	Percentage of False Negatives
8	EXPLORE Reading	10876	81%	5%	14%	EXPLORE English	10938	80%	11%	9%
10	PLAN Reading	4804	78%	5%	17%	PLAN English	4865	75%	8%	17%
11	ACT Reading	780	83%	4%	12%	ACT English	786	80%	6%	14%

MAP Mathematics as Predictor					
	Percentage of Sample whose College Readiness Status was Accurately Predicted by MAP Score				
Grade	College Readiness Test	Sample Size	Percentage Correctly Predicted	Percentage of False Positives	Percentage of False Negatives
8	EXPLORE Mathematics	12753	82%	4%	14%
10	PLAN Mathematics	9516	86%	4%	9%
11	ACT Mathematics	2948	91%	1%	7%

[5] Correct predictions refer to the percentage of students in the study sample whose MAP scores accurately indicated their college readiness status on the college readiness test. False positives indicate the percentage of students predicted to be college ready, but who failed to meet the college readiness benchmark. False negatives indicate the percentage incorrectly predicted to fail to meet the college readiness benchmark, but who did.

REFERENCES

ACT. 2010. *Issues in College Readiness: What are ACT's College Readiness Benchmarks* (IC 050805090). Retrieved from ACT website: http://www.act.org/research/policymakers/pdf/benchmarks.pdf

Northwest Evaluation Association. 2011. *RIT Scale Norms*. Portland, OR: Northwest Evaluation Association.

APPENDIX 1: METHODOLOGY

This linking study examines the concurrent relationship between EXPLORE/PLAN/ACT and MAP assessments with the goal of publishing benchmarks on the RIT scale that are predictive of the ACT's college readiness benchmarks (ACT, 2010).

Appendix Table 1 describes the three ACT assessments with their respective college readiness targets by content area.

Appendix Table 1 – ACT College Readiness Cut Points

	EXPLORE		PLAN	ACT
	Grade 8	Grade 9	Grade 10	Grade 11
English	13	14	15	18
Math	17	18	19	22
Reading	15	16	17	21

Study Sample

NWEA solicited all known partner districts that administer both the EXPLORE/PLAN/ACT and MAP assessments to participate in this study. While not every eligible partner participated, the final study sample was large enough (unique total student n=29,417) to proceed with the analysis. Appendix Table 2 contains the distinct sample count.

Appendix Table 2 – Sample Counts

Test	Grade	Unique State Count	Unique District Count	Unique School Count	Unique Student Count
ACT	11	3	9	36	3680
Explore	8	3	5	51	11822
Plan	10	3	7	54	13915

Each district's EXPLORE/PLAN/ACT test records were matched to their corresponding MAP data via a robust matching algorithm housed in NWEA's Growth Research Database. The bulk of the study used matched students who took both assessments in same term. In the cases where students took the NWEA MAP test in different terms, we employed the following prioritized matching process.

Priority 1 – MAP given in **SAME** term as EXPLORE/PLAN/ACT
Priority 2 - MAP test given one term **BEFORE** EXPLORE/PLAN/ACT
Priority 3 - MAP test given two terms **BEFORE** EXPLORE/PLAN/ACT
Priority 4 - MAP test given one term **AFTER** EXPLORE/PLAN/ACT
Priority 5 - Map test given two terms **AFTER** EXPLORE/PLAN/ACT

In order to ensure comparable RIT scores, we took the MAP percentile associated with the RIT score and substituted the RIT score associated with that percentile in the term the EXPLORE/PLAN/ACT was administered. For example, a winter test score under Priority 2 with a percentile of 75 would be substituted for the corresponding spring RIT score associated with the 75th percentile. Appendix Table 3 contains the distribution of unique students by the EXPLORE/PLAN/ACT test, MAP Grade and prioritized matching scheme.

Appendix Table 3 – Prioritized Matching Count (Unique Students) by ACT Test and MAP Grade

		Priority Matching				
		P1 - Exact	P2 - 1 Term Prior	P3 - 2 Terms Prior	P4 - 1 Term After	Total
EXPLORE/PLAN/ACT Test and MAP Grade	ACT 11	2912	405	357	6	3680
	PLAN 10	4456	70	6587	709	11,822
	EXPLORE 8	13,443	22	356	94	13,915
	Total	20,811	497	7300	809	29,417

Analysis

The goal of the analysis is to find the statistical model that best describes the scale relationship between EXPLORE/PLAN/ACT and the NWEA MAP assessments. We tested multiple models including Ordinary Least Squares Regression (Linear and Quadratic) and Hierarchical Linear Model (HLM). The best model was determined by correlation (*r)* and overall model fit.

Akaike Information Criterion (AIC) is the measure of model fit we used. Generally speaking, the AIC examines the tradeoffs between model accuracy and complexity whereby the model with the lowest AIC value is said to be the most parsimonious

Separate linear (1a) and quadratic (1b) regression routines were run for each relationship model.

$$ACT = a + X\beta + e \quad \text{(1a)}$$

where X= RIT

$$ACT = a + X\beta + X^2\beta + e \quad \text{(1b)}$$

where X= *RIT*

Based on the correlations and AIC fit statistics, the quadratic regression best described the shape of the scale relationship (see Appendix Table 4).

Appendix Table 4 – Initial Model Correlations

Test	Model	Linear Regression		Quadratic Regression	
		r	*AIC*	*r*	*AIC*
ACT	**Language Usage to English**	0.71	2400	0.77	2237
ACT	**Language Usage to Reading**	0.64	2450	0.70	2330
ACT	**Math to Math**	0.73	6549	0.81	5686
ACT	**Reading to English**	0.69	8457	0.75	7948
ACT	**Reading to Reading**	0.65	8513	0.72	8013
PLAN	**Language Usage to English**	0.65	12264	0.68	11813
PLAN	**Language Usage to Reading**	0.55	13674	0.59	13306
PLAN	**Math to Math**	0.71	21874	0.74	20855
PLAN	**Reading to English**	0.63	24102	0.67	23299
PLAN	**Reading to Reading**	0.57	26343	0.62	25418
EXPLORE	**Language Usage to English**	0.70	25133	0.75	23440
EXPLORE	**Language Usage to Reading**	0.59	24728	0.66	23204
EXPLORE	**Math to Math**	0.76	24878	0.77	24615
EXPLORE	**Reading to English**	0.69	29685	0.73	28068
EXPLORE	**Reading to Reading**	0.62	28324	0.69	26121

The next step in the analytic process was to determine whether any between-school variation existed in our scale relationships by running a simple unconstrained HLM model (2):

$$ACT_{ij} = \gamma_{00} + \mu_{0j} + r_{ij} \qquad (2)$$

ACT_{ij} is the ACT/PLAN/EXPLORE score for student *i* in school *j;*

γ_{00} is the grand mean (students within schools)

μ_{0j} variance in intercept **between** schools

r_{ij} **within** school variance

Appendix Table 4 contains the Intraclass Correlation Coefficient (ICC) for each test and model. The ICC measures the proportion of variance in the dependent variable ACT_{ij} that is accounted for by our grouping structure.

Appendix Table 5– Intraclass Correlation Coefficients (ICC) by Test and Model

Test	Model	t_{00}	σ^2	ICC
ACT	Language Usage to English	9.650	30.166	0.242
ACT	Language Usage to Reading	6.490	29.035	0.183
ACT	Math to Math	2.458	16.348	0.131
ACT	Reading to English	6.669	29.763	0.183
ACT	Reading to Reading	4.610	28.564	0.139
Plan	Language Usage to English	5.478	17.651	0.237
Plan	Language Usage to Reading	6.175	19.496	0.241
Plan	Math to Math	5.388	16.303	0.248
Plan	Reading to English	4.734	16.944	0.218
Plan	Reading to Reading	4.992	18.878	0.209
Explore	Language Usage to English	0.979	16.793	0.055
Explore	Language Usage to Reading	0.649	12.878	0.048
Explore	Math to Math	0.825	14.619	0.053
Explore	Reading to English	0.980	16.434	0.056
Explore	Reading to Reading	0.787	12.815	0.058

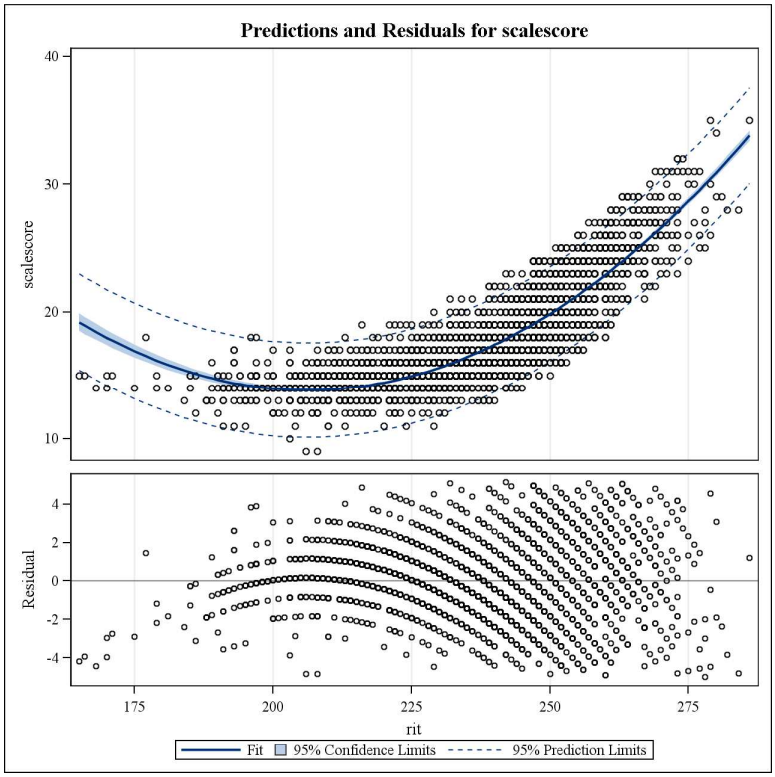
The ICCs offer a somewhat conflicting picture on the appropriateness of using a multilevel model in the case of this study. For instance, the EXPLORE assessments have the least amount of between-group variance (less than six percent) and the nearly the most number grouping levels (between 29 and 51 individual schools depending on the model). While no well-established ICC thresholds exist per se, it would appear the EXPLORE would not be a good candidate when compared to the ACT and PLAN ICCs.

One explanation for the observed differences could be related to the specific analytic sample used. While we could have employed two separate methods (Quadratic for EXPLORE, HLM for ACT and PLAN),

we felt the quadratic model offered transparency and consistency while maintaining good model fit characteristics.

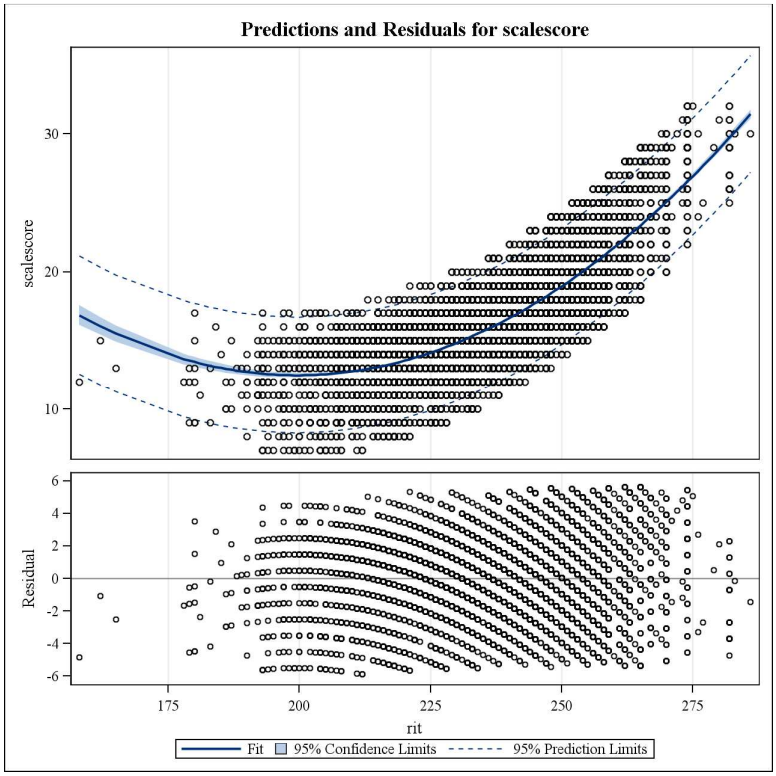
Figures 1 through 3 illustrate the final fitted model for EXPLORE/PLAN/ACT Math to NWEA MAP Math. Each figure contains notes referencing specific sample or estimation characteristics. Please note standardized residuals greater than 2 or less than -2 were removed from the final model to eliminate potential sources of statistical noise. We should also note the MAP assessment measures student performance relative to state content standards rather than discrete college readiness standards. This difference in content alignment could possibility degrade the published regression coefficients between MAP and EXPLORE/PLAN/ACT.

Figure 1 - ACT Math to NWEA Math



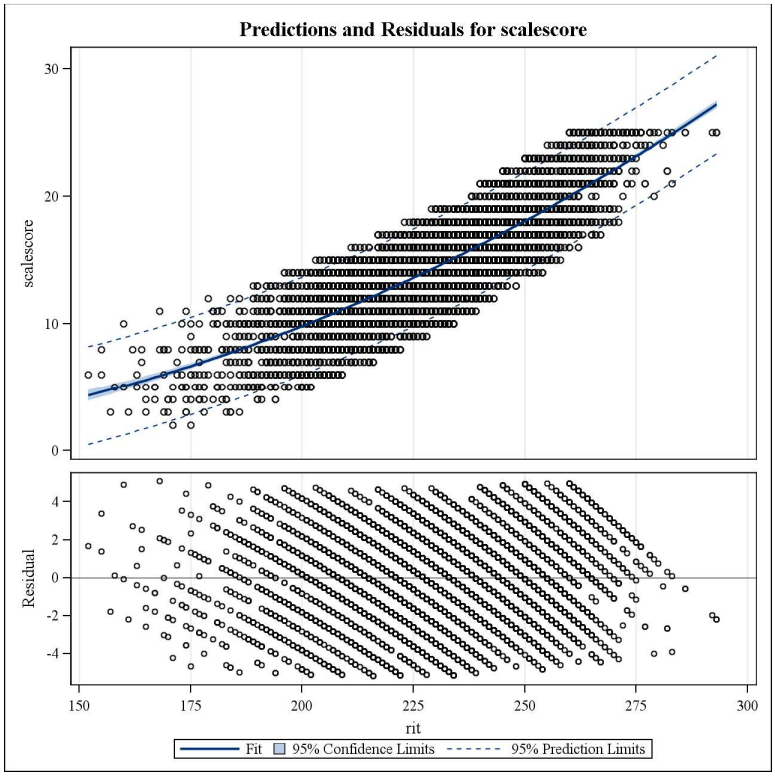
The ACT Math scatter plot and residuals exhibit pronounced curvilinear shapes. The sample also becomes sparse toward the top end of the distribution, making those estimates less reliable.

Figure 2 - PLAN Math to NWEA Math



The PLAN Math scatter plot and residuals display the same curvilinear pattern as the ACT but appears to have more variance along the fit line. Like the ACT, the sample also becomes sparse toward the top end of the distribution, making those estimates less reliable.

Figure 3 - EXPLORE Math to NWEA Math



The EXPLORE Math scatter plot and residuals have a slight curvilinear profile especially when compared to the ACT and PLAN assessments. The EXPLORE assessment has a pronounced ceiling effect in reference to the NWEA MAP assessment, meaning the NWEA assessment has more "stretch" than EXPLORE at the eighth grade level.