Aligning the NWEA RIT Scale with the South Carolina High School Assessment Program

John Cronin, Ph.D

August 2004

Northwest Evaluation Association
5885 SW Meadows Road, Suite 200
Lake Oswego, OR 97035-3526

www.nwea.org
Tel 503-624-1951
Fax 503-639-7873

Aligning the NWEA RIT Scale with the South Carolina High School Assessment Program (HSAP)

John Cronin, Ph.D.

August, 2004

Each year, South Carolina students participate in testing as part of the South Carolina assessment program. Students in grades 3 through 8 take the Palmetto Achievement Challenge Tests (PACT) in English/Language Arts and Mathematics. Students in grade 10 take the High School Assessment Program (HSAP) in English/Language Arts and mathematics. These tests serve as an important measure of student achievement for the state's accountability system. Results from these assessments are used to make state-level decisions concerning education, to meet *Adequate Yearly Progress* (AYP) reporting requirements of the *No Child Left Behind Act* (NCLB), and to inform schools and school districts of their performance. In addition, students must achieve Level 2 performance on the HSAP in order to graduate from high school.

The South Carolina Department of Education has developed scales that are used to assign students to one of four performance levels on the HSAP. Level 2 is considered the level that represents passing performance.

Many students who attend school in South Carolina also take tests developed in cooperation with the Northwest Evaluation Association (NWEA). These tests report student performance on a single, cross-grade scale, which NWEA calls the RIT scale. This scale was developed using Rasch scaling methodologies. RIT-based tests are used to inform a variety of educational decisions at the district, school, and classroom level. They are also used to monitor academic growth of students and cohorts. Districts choose whether to include these assessments in their local assessment programs. They are not state mandated.

In order to use the two testing systems to support each other, an alignment of the scores from the state and RIT-based tests is as important as the curriculum alignment. NWEA has now conducted three studies to establish the alignment of cut scores between the PACT and NWEA tests. The current study is intended to establish aligned cut scores on the RIT scale for HSAP assessments.

The current study is one of an ongoing series of studies that are being conducted to identify the relationships between NWEA tests and state-mandated assessments. Studies in seventeen states have now been completed. For purposes of this study we focused on examining the relationships between HSAP and NWEA assessments in reading and mathematics only.

The primary questions addressed in this study are:

- To what extent do the same subject scores for the NWEA test correlate to the content-similar subjects on the HSAP tests?
- What RIT scores correspond to various performance levels on the HSAP tests?
- How well can *passing* performance on the South Carolina assessments be predicted from RIT scores when NWEA assessments are administered in the same time frame?

Method

Participating School Systems

Students from the Horry County, Richland 2, and Charleston County school systems participated in this study.

Data Preparation

For purposes of studying NWEA test alignment with the HSAP, 10th grade student level test records from spring 2004 HSAP testing and spring 2004 NWEA assessments were matched using district assigned student ID numbers. Matched records were then screened to remove invalid scores. Table I shows the number of student records included in this study.

Table 1 – Reading and Mathematics Tests Included by Grade

Subject	Students
Reading	3749
Language Usage	3552
Mathematics	3538

We had enough student records at each grade to adequately cover the breadth of the scale and perform a robust analysis near the passing score for this assessment. Because the study involved a small number of districts, we recommend that schools validate our estimates by cross-checking their own students' performance against our cut scores.

Analyses

Pearson correlations

The initial analyses focused on the relationships among the NWEA and South Carolina assessment scores at each grade to determine how closely the scores on the NWEA test correlated with same subject scores on the HSAP. Simple bivariate correlation coefficients were computed among these scores.

Linking HSAP scores to the RIT scales

Three methods of estimating cut scores for HSAP levels were used. The most straightforward was simple linear regression (HSAPpred =a(RIT) + c). Since we sometimes observe departures from a linear relationship on the lower and upper ends of state test scales, a second order regression model was also used (HSAPpred=a(RIT2) + b(RIT) + c). For each of these methods, the RIT score was determined by substituting the appropriate HSAP score for $HSAP_{pred}$ and solving the equation for RIT.

A fixed-parameter Rasch model was also used to estimate RIT cut scores. In this method, the HSAP performance level was treated as a test item. The assumption is that the performance level 'item' should contain all the information about the difficulty of the test. Student abilities (RIT scores) were the 'fixed parameter' used to anchor the difficulty estimate of the 'status' item to the RIT scale. The resulting 'difficulty estimate' was taken as the RIT cut score for this method. This is referred to as the Rasch Status on Standard (or simply Rasch SOS) method.

Predicting HSAP performance levels from RIT scores

RIT scores were first used to predict whether students were likely to achieve performance at or above the passing score (Level 2) on the HSAP. The predictions of HSAP performance were compared to observed performance in 2 X 2 contingency tables. A *prediction index* score was generated to measure the ratio of Type I error to accurate prediction of proficiency status. This score is expressed as

1-(Number of Type I errors/Number of correct predictions)

Higher prediction index numbers generally show more accurate prediction with lower levels of Type I error. Type I error occurs when NWEA assessments predict that a student will achieve above a passing level of performance when the student actually achieves a failing score. This index was generated for the linear, second order, and Rasch SOS methodologies. In general, the highest prediction index score was used to select the RIT cut score to be adapted as the recommended RIT score we would associate with achieving the passing standard on the corresponding HSAP assessment for the particular grade level and subject area. We do make exceptions to this rule when the estimated score produces high accuracy rates but inordinately large numbers of Type II errors. This condition indicates a greatly overestimated cut score, so we select a method that produces a more balanced Type I to Type II error ratio in these instances.

In addition, we evaluated the accuracy of predictions of HSAP levels based on observed RIT scores. The predictions of HSAP level performance were compared to observed performance in 4 X 4 contingency tables. Once again a prediction index score was generated to provide an estimate of accuracy.

Content Validity

Formal comparisons of the content of NWEA and the HSAP were not conducted for purposes of this study. The standards used to construct the NWEA Assessments were the same as those used for the South Carolina assessments. Both NWEA assessments and the South Carolina assessments include multiple-choice items. The HSAP also includes short answer and extended response questions. Results from our previous fifteen studies indicate that the addition of items in alternate formats generally does not, by itself, materially affect the ability of the NWEA test to generate reasonably accurate predictions of performance levels.

Results

Descriptive Statistics

Table 2 reviews descriptive statistics for the HSAP and NWEA assessments. The median RIT scores for this sample in reading and language usage are near the median for the NWEA norm population. The median RIT score in mathematics, however, is 11 points below the median for the NWEA norm population. The difference in mathematics is large and its potential impact on the accuracy of our estimates merits discussion.

Normal distributions around a nationally-normed mean are desirable but not necessarily essential when conducting alignment studies. It is more important that the sample provide reasonable numbers of students who perform at all levels on the test scales so that the statistical methods applied have an adequately large sample to derive good estimates of performance levels. In this case we had reasonably large representations of students who performed at all performance levels.

It is fair to say, however, that school districts with large numbers of low performing students may align their curriculum differently to the state standards. There may also be other, hard to know factors, related

to this phenomenon that may influence alignment. That's why we recommend that school systems test the application of the study results in their own setting to validate the predicted cut score's accuracy.

It should also be noted that the participating districts all used NWEA's general mathematics test for purposes of this study. The NWEA norms for grade 10 reflect the performance not only of students who have taken the general mathematics test, but also students who have taken NWEA's end of course tests in Algebra I, Geometry, and Algebra II. This may be one reason why the median scores in mathematics are lower relative to its respective norm than the reading and language usage scores.

Table 2 – Means, Standard Deviations, and Medians for the HSAP and NWEA assessments

Grade	NWEA Reading	NWEA Lanaguage Usage	HSAP English/Language Arts	NWEA Mathematics	HSAP Mathematics
N	3749	3552	3749	3538	3538
Mean	224.67	222.78	226.59	237.87	223.72
Median	227	224	228	239	221
Std. Deviation	15.258	12.862	23.638	17.453	26.645

Pearson correlations

Table 3 shows the results of this analysis for each grade. Concurrent validity was tested by examining same subject Pearson correlations between the NWEA and HSAP. Same subject correlations were high, ranging from .78 to .85, numbers that suggest the tests were generally measuring the same constructs. Discriminant validity was tested by examining same subject Pearson correlations next to correlations for the alternate subject (math against reading). In all cases the same subject correlations were higher than correlations against the alternate subject.

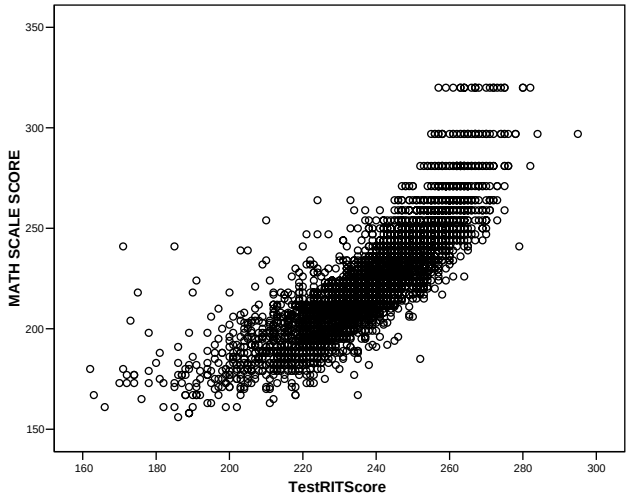
Table 3 – Pearson Correlations for HSAP and NWEA assessments by Subject

	NWEA Reading	NWEA Language Usage	HSAP English/Language Arts	NWEA Mathematics	HSAP Mathematics
NWEA Reading	1		.781		.658
NWEA Language Usage		1	.786		.673
HSAP English/Language Arts	.781	.786	1		.737
NWEA Mathematics			.716	1	.847
HSAP Mathematics			.737	.847	1

Same subject correlations are shaded

Analysis of scatterplots suggested that relationships might be somewhat curvolinear, and that some of the scale relationships might break down slightly near the lower end of the scales, possibly indicating a floor effect on the HSAP. Figure 1 provides an example from the mathematics sample that illustrates both the scale relationships and the evidence of some breakdown in correlation near the bottom of the HSAP Scale. For example, note that students achieving scores near 190 on the HSAP scale, achieve scale scores that range from about 170 to nearly 240 on the NWEA test. One possible explanation for this is that the NWEA test, because it is adaptive as opposed to single form, has the capacity to more accurately measure performance at the low end of the performance spectrum.

Figure 1 – Scatterplot depicting Grade 8 NWEA mathematics RIT agains the HSAP mathematics scale score



Linking HSAP performance level cut scores to the RIT scale

The primary purpose of this study was to estimate the RIT scale scores that most closely correspond to the cut scores for different performance levels on the HSAP. This information allows schools to identify students who may need additional support to reach state standards. It can also help schools identify students who are performing well enough that they are ready to tackle work beyond what the state standards require.

Table 4 shows several estimations of the Spring 2004 RIT score that correspond to the cut scores for the various performance levels on the HSAP scales. As a rule the three methodologies came to similar estimates of cut scores for each of the performance levels, although the Rasch SOS methodology did produce somewhat higher estimates of the RIT score required to pass Level 2.

Table 4 – Estimated points on the RIT scale equating to the minimum scores (rounded) for performance levels on the HSAP

	Linear Regression				Second-order Regression				Rasch Status-on-Standard			
	1	2	3	4	1	2	3	4	1	2	3	4
Reading	<204	204	222	236	<205	205	225	237	<209	209	223	234
Language Usage	<206	206	221	233	<205	205	221	232	<210	210	222	231
Mathematics	<220	220	236	251	<220	220	237	250	<223	223	237	250

Predicting HSAP pass-fail status from RIT scores

Once the cut scores were estimated from the three methods, we evaluated each possible cut score to determine how accurately it predicted students' actual performance on the corresponding HSAP assessment. The most accurate method of prediction was generally used to derive the best estimate of RIT cut scores that equate to the different HSAP performance levels. A *prediction index* statistic (described on page 3) scored the accuracy of prediction.

For this study, we first assessed the accuracy of the RIT scale in correctly predicting whether students are likely to reach the *passing* level on the corresponding HSAP test. Next we assessed the accuracy with which the RIT predicted level assignment on this test. Use of the prediction index statistic helped assure that the method chosen produced a high ratio of accurate passing predictions relative to Type I errors. Type I errors occur when the RIT scale predicts a passing score for a student who actually fails the assessment. These types of errors raise particular concern because they fail to identify students who might need additional support and resources in order to achieve their targets. A high prediction index number indicates that the test maximizes accuracy of prediction while minimizing Type I errors.

In these kinds of studies we want to emphasize that prediction is not used to foretell an inevitable future for the student, rather it is used to help schools plan for instruction and offer appropriate interventions to children who need additional support to be successful.

Table 5 shows the results of the analysis. All methods considered were highly accurate (better than 88%) in predicting pass-fail against the Level 2 cut score. Although all methods produced prediction index scores above .900, the Rasch SOS method generated fewer Type I errors and higher prediction index scores for all subjects.

Table 5 – Accuracy of the RIT scale in predicting HSAP Pass/Fail Status

Reading	Cut Score	Accuracy	Type I Error	Prediction Index
Linear	204	90.23%	7.33%	0.919
Second Order	205	90.01%	7.75%	0.914
Rasch SOS	209*	90.23%	5.43%	0.940
Language Usage	**Cut Score**	**Accuracy**	**Type I Error**	**Prediction Index**
Linear	206	90.17%	7.49%	0.917
Second Order	205	89.95%	8.08%	0.910
Rasch SOS	210*	89.92%	5.57%	0.938
Mathematics	**Cut Score**	**Accuracy**	**Type I Error**	**Prediction Index**
Linear	220	88.24%	8.17%	0.907
Second Order	220	88.24%	8.17%	0.907
Rasch SOS	223*	88.64%	5.96%	0.933

* Indicates methodology chosen for recommended estimate

Table 6 summarizes the accuracy of prediction for this study relative to other state alignment studies. Prediction index scores for South Carolina are somewhat higher than average in reading and language usage and slightly lower than average for mathematics. Nevertheless, these rates of correct prediction are easily high enough to provide useful information to educators who are planning instruction to ensure all students perform at a level that meets the standards.

Table 6 – Prediction Indices (Based on Proficiency Status) for Previous NWEA State Alignment Studies

State	Reading	State	Language	State	Math
Texas	.974	Texas	.968	Texas	.970
Washington	.971	South Carolina Exit	.938	Wyoming	.961
Minnesota	.944	California	.913	Colorado '01	.957
South Carolina Exit	.940	Indiana '01	.907	Washington	.949
Wyoming	.931	Colorado '03	.903	Illinois	.946
Colorado '03	.931	Indiana '03	.894	Colorado '03	.943
Illinois	.928	Arizona	.874	South Carolina '03	.943
California	.925			Minnesota	.936
Arizona	.912			Washington	.936
Colorado '01	.910			South Carolina Exit	.933
Nevada	.902			Arizona	.919
South Carolina '03	.902			California	.910
Indiana '01	.902			Indiana '01	.899
Indiana '03	.900			Nevada	.866
Washington	.886			Indiana '03	.860

* Texas results were generated by a study of over 1,000 per grade from a single school district.

Predicting HSAP Performance Levels from RIT Scores

The HSAP reports four levels of performance. Three cut scores are set to define these four levels. Analyzing the capacity of RIT scores to predict students' HSAP performance levels can help educators triangulate information about student performance on their state test, assuring that instructional plans and interventions are adequately reinforced by data. Predictions of performance level are not as accurate as the predictions of proficiency status. This is true in part because tests vary in their ability to measure students at the highest and lowest performance levels.

When predicting performance levels, a case is identified as accurate when the performance level assigned by the HSAP and RIT score are the same. A Type I error occurs when the RIT score assigns a performance level that is higher than the student actually achieved on the state test. For example, if the

RIT score projects Level 3 performance for the student and the HSAP result is Level 2, we declare the case a Type I error because the RIT score overestimated performance.

Table 7 – Accuracy of the RIT scale in predicting HSAP performance level

Reading	Accuracy	Type I Error	Prediction Index	Level 4 found	Level 1 Found
Linear	59.71%	35.86%	0.641	64.4%	44.7%
Second Order	59.29%	35.13%	0.649*	64.4%	41.6%
Rasch SOS	60.81%	22.09%	0.637	74.8%*	59.1%*
Language Usage	**Accuracy**	**Type I Error**	**Prediction Index**	**Level 4 found**	**Level 1 Found**
Linear	59.85%	36.36%	0.636*	61.9%	46.9%
Second Order	60.14%	38.86%	0.611	67.6%	42.7%
Rasch SOS	60.16%	22.58%	0.625	77.9%*	60.5%*
Mathematics	**Accuracy**	**Type I Error**	**Prediction Index**	**Level 4 found**	**Level 1 Found**
Linear	64.33%	33.83%	0.662	74.7%	56.0%
Second Order	64.90%	33.19%	0.668	79.1%*	56.0%
Rasch SOS	65.49%	19.33%	0.705*	79.1%*	67.9%*

* Indicates methodology chosen for recommended estimate

The results reported in table 7 show that second order regression produced the best overall estimate of performance level for reading, while linear regression method produced the best estimate for NWEA's language usage assessment, and the Rasch SOS method produced the best estimate in mathematics. The Rasch SOS was generally more successful than the other methods in finding the most students performing at the lowest and highest performance levels.

NWEA has reported estimated performance level assignments for prior studies conducted in 11 states. Table 8 compares the accuracy with which these tests predict performance level. The results show the HSAP performance index scores are below the median in both reading and mathematics.

Table 8 – Prediction index scores by performance level assignment for previous NWEA state alignment Studies

State	Reading	State	Math
Washington	.874	Washington	.928
Texas	.868	Texas	.900
Indiana	.860	Illinois	.888
Colorado	.840	Colorado	.808
Illinois	.804	Washington	.805
Nevada	.776	Indiana	.804
South Carolina '03	.757	South Carolina '03	.764
Arizona	.756	Arizona	.756
Washington	.698	Nevada	.742
South Carolina Exit	.649	South Carolina Exit	.705
Minnesota	.627	Minnesota	.611
California	.600	California	.565

Best estimates of HSAP performance level cut scores

To estimate the RIT scores that best predict the cut scores for the various South Carolina performance levels we did the following:

- For the Level 2 RIT score, we selected the methodology that produced the highest performance index score in predicting "pass/fail" alone.
- For the Level 3 RIT score, we selected the methodology that produced the highest performance index score for predicting the level of performance.
- For the Level 4 RIT score, we selected the cut scores that correctly predicted the largest proportion of students who actually achieved this level of performance on the HSAP.

Table 9 summarizes the recommended cut scores for each performance level on the HSAP.

Table 9 – Projected RIT Scores Equivalent to Performance Levels on HSAP

	1			2	3			4		
	Score Range	% of pop. identified	Method	Cut Score	Cut Score	Perf. Index	Method	Cut Score	% of pop. Identified	Method
Reading	<209	59.1%	Rasch	209	224	.649	Second Order	234	74.8%	Rasch
Language Usage	<210	60.5%	Rasch	210	221	.636	Linear	230	77.9%	Rasch
Mathematics	<223	67.9%	Rasch	223	237	.705	Rasch	250	79.1%	Rasch/Second Order

Using RIT scores to estimate student probability of achieving passing performance on the HSAP

Helping students pass the state test is not the primary reason our members use NWEA assessments. We hope they are used to provide teachers information that will allow them to improve the learning of all students. Nevertheless, state test results are important and failing to do well on them can have deleterious effects on students and their schools. Because of this, we believed educators would benefit from knowing more about the probability that a student's RIT score would lead to a passing score on the HSAP. This would allow educators to more reliably identify students who will need additional resources to reach this level of performance. Equally important, however, it will allow educators to know which students are "safe" against South Carolina standards so they can focus their time with these students on providing new challenges that better suit their current needs.

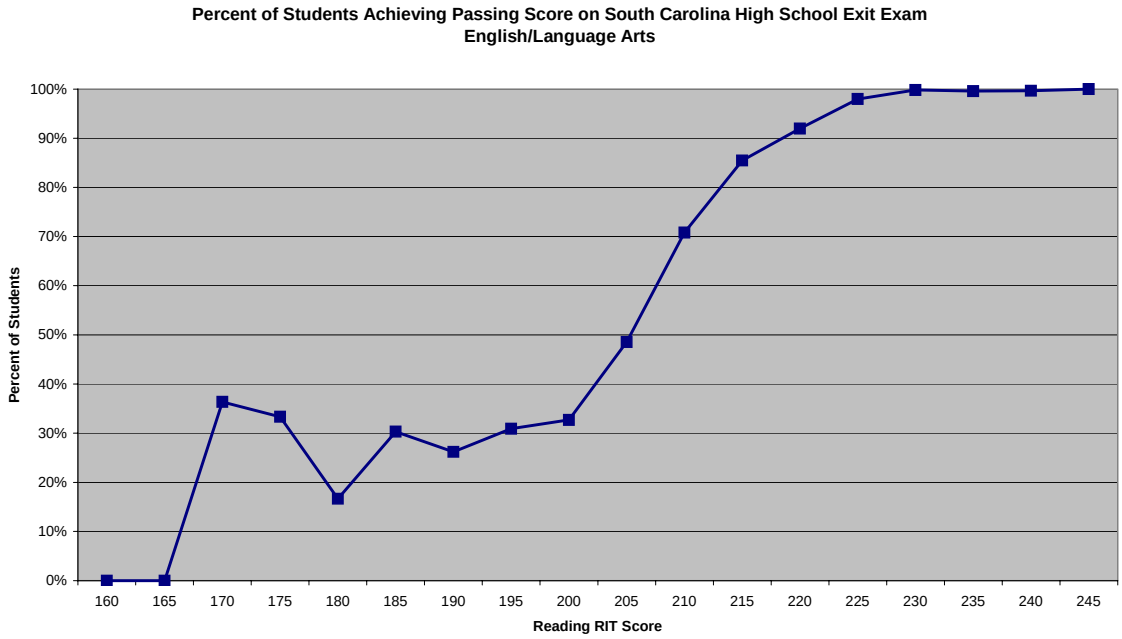
Table 10 shows the proportion of students at each 5 point RIT level who earned scores at or above the Level 2 on the HSAP ELA and mathematics assessments. Using reading as an example, we find that about 31% of the students who achieved a reading RIT score between 195 and 200 went on to achieve a passing score on the HSAP ELA assessment. An English with ten students performing in this range would know that only about three in ten of these students will be proficient on the HSAP unless they work harder, receive more focused instruction, or have access to additional resources.

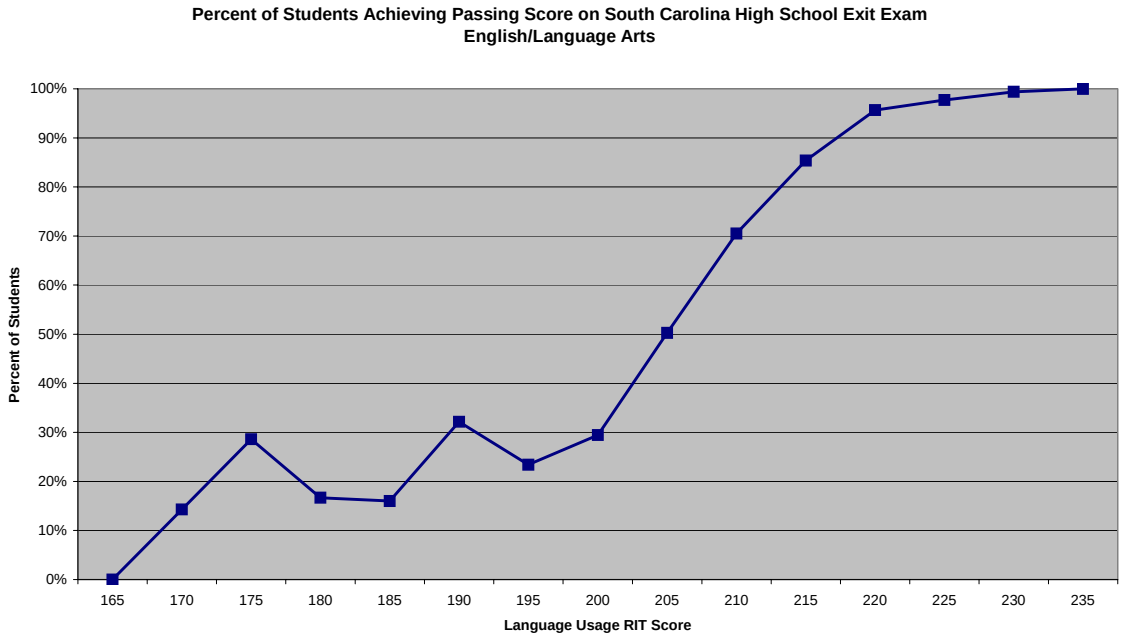
On the other hand, about 92% of students performing in the 220 to 225 range achieved proficiency on the South Carolina ELA assessment. Teachers should feel free to focus their efforts with these students on content and skills that go beyond the minimum expectations for performance.

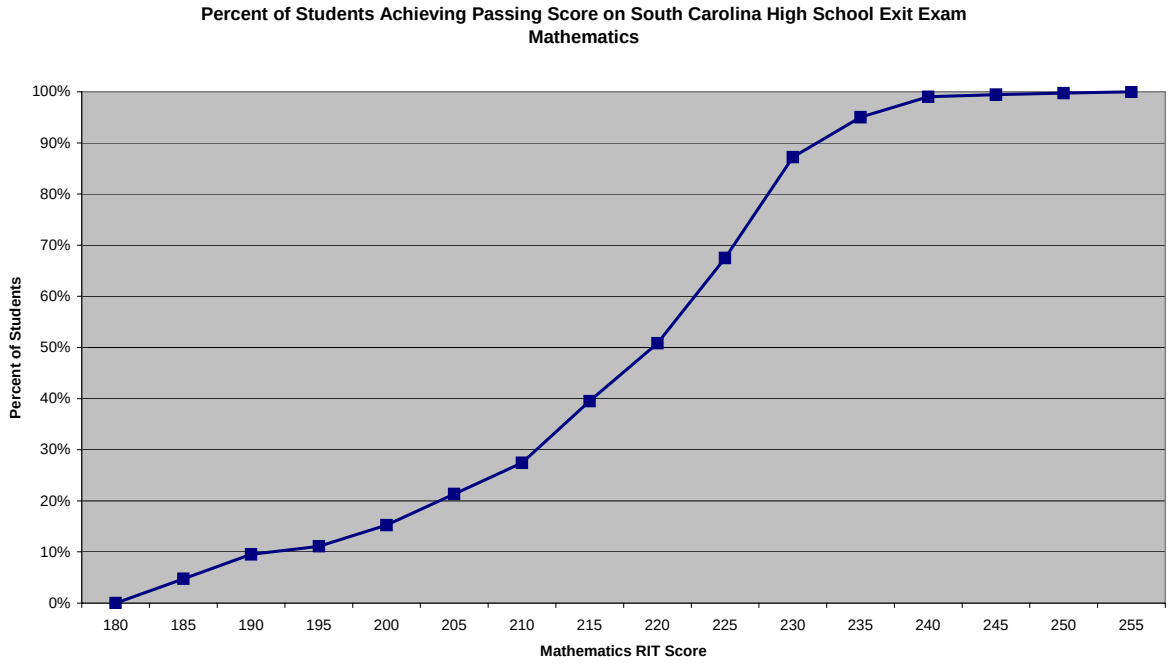
Figures 2, 3, and 4 are graphic depictions of the data in the tables.

Table 10 – Proportion of students passing the HSAP based on same spring RIT score

RIT	Reading	Language Usage	Mathematics
165	0.00%	0.00%	
170	36.36%	14.29%	
175	33.33%	28.57%	
180	16.67%	16.67%	0.00%
185	30.30%	16.00%	4.76%
190	26.19%	32.14%	9.52%
195	30.91%	23.38%	11.11%
200	32.71%	29.46%	15.25%
205	48.54%	50.28%	21.33%
210	70.82%	70.51%	27.43%
215	85.46%	85.40%	39.49%
220	91.95%	95.67%	50.81%
225	97.98%	97.72%	67.49%
230	99.80%	99.42%	87.23%
235	99.60%	100.00%	95.01%
240	99.68%		99.03%
245	100.00%		99.44%
245			99.73%
250			100.00%
255			0.00%


Percent of Students Achieving Passing Score on South Carolina High School Exit Exam
English/Language Arts
Percent of Students
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245
Reading RIT Score


Percent of Students Achieving Passing Score on South Carolina High School Exit Exam
English/Language Arts
Percent of Students
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
165 170 175 180 185 190 195 200 205 210 215 220 225 230 235
Language Usage RIT Score


Percent of Students Achieving Passing Score on South Carolina High School Exit Exam
Mathematics
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Percent of Students
180
185
190
195
200
205
210
215
220
225
230
235
240
245
250
255
Mathematics RIT Score

Comparing South Carolina HSAP standards with the estimated standards reported in other state test alignment studies

Northwest Evaluation Association tests have been aligned with the cut scores for the state high school standards and/or proficiency tests in eight states. To get an estimate of the difficulty of the HSAP in relation to other state tests, we evaluated the standard defined as the NCLB passing score and compared it to the cut score representing the same standard in these other states.

The results are summarized in Table 11. South Carolina's cut scores in reading are lower than five of the eight states studied. The cut scores in mathematics are the lowest of any state studied. We would recommend caution about drawing any judgments about the quality of South Carolina's standards from that information. States establish standards for different purposes. States also attach different stakes to their standards. Some states, Oregon might be an example, set their high school standards prior to the adoption of NCLB. In Oregon's case, these standards were set at a level they believe appropriate for students pursuing some form of post-secondary education. In addition, Oregon does not require that students pass these standards as a condition for graduation. This confluence of factors explain why the Oregon standard was set relatively high.

Other states, California would be an example, established high school performance standards after the passage of NCLB. They were not necessarily intended to reflect performance needed to pursue post-secondary education. They were intended to be a prerequisite for graduation, however, although the state has postponed the requirement for now. Given that the standards were implemented with the intention that all students would be required to achieve this level of performance, it is not a surprise that the California standard is not as rigorous as Oregon's.

In general, standards should be judged on how well they align with the purposes the community has set for establishing them, not purely on how high or low the "bar" is set. One thing the tables make clear is that graduation standards vary widely from state to state and that there is not yet a shared definition of graduation level performance.

Table 11 – Cut scores representing passing level of performance on 8 state high school assessments

Reading			Mathematics		
State	Cut Score	%ile	State	Cut Score	%ile
OR	236	77	WA	257	73
WA	227	53	MT	247	40
ID	224	44	IA	247	40
MT	224	44	OR	245	33
IA	223	42	ID	242	25
SC	209	15	CO	233	14
CO	209	15	CA	232	13
CA	208	14	SC	223	7

Calibration of HSAP standards with standards used for the PACT

Because of the stakes associated with the HSAP standards, schools have an interest in knowing, well before high school begins, which students might need additional support and assistance to achieve the level of learning required by these assessments. Ideally, the performance standards used to represent the Basic level of performance for the PACT would predict, with a reasonable degree of success, passing performance on the HSAP. This would require that the PACT and HSAP standards calibrate in someway.

It is not clear to us whether the PACT and HSAP were designed with this purpose in mind. Therefore, we did not enter into the study with the assumption that prior PACT Basic performance in a grade would predict passing performance on the HSAP in grade 10.

NWEA has conducted three prior studies to estimate the alignment of PACT cut scores with the RIT scale, the most recent being completed simultaneously with this study of the HSAP (Hauser, 2001; Cronin, 2003; Cronin 2004). Based on the results of the most recent study, we believe that students achieving Basic performance on the PACT should easily pass the HSAP assessment in grade 10. In fact, Level 2 performance on the HSAP is generally below the level of performance that would correspond to Basic proficiency on the PACT for grade 8.

Table 12 – Estimated RIT scores aligning with the Basic level of performance on PACT and level 2 performance on HSAP (associated NWEA percentile in parentheses)

Grade	Reading	Language Usage	Mathematics
3	182 (16)	186 (19)	193 (29)
4	194 (22)	197 (24)	202 (31)
5	202 (26)	204 (25)	212 (38)
6	210 (32)	210 (31)	215 (34)
7	210 (24)	211 (26)	223 (39)
8	213 (22)	213 (24)	228 (36)
10 (HSAP)	209 (14)	210 (15)	223 (7)

Table 12 shows that the estimated reading and language usage RIT scores required to project to achieve passing performance on the HSAP English/Language Arts assessment is about the same as the RIT score required to achieve Basic performance on the grade 6 PACT. The estimated mathematics RIT score required to project to achieve passing performance on the HSAP mathematics assessment is about the same as the score required to project to achieve basic performance on the grade 7 PACT. In general, therefore, students who achieve basic performance on PACT should easily achieve the level performance needed to pass the HSAP tests. Based on our prior studies, it seems that the PACT Basic standard is currently more rigorous than Level 2 performance on the HSAP.

Summary and Conclusions

This study investigated the relationship between the scales used for the HSAP assessments and the RIT scales used to report performance on Northwest Evaluation Association tests. The study determined the reading, language usage and mathematics RIT score equivalents for the HSAP performance levels in English/Language Arts and Mathematics. Test records for more than 3,500 students were included in this study.

Three methods generated an estimate of RIT cut scores that could be used to project HSAP performance levels. Rasch SOS methods generally produced the most accurate cut score estimates. Accuracy of predicting HSAP passing performance was above 88% for all subjects when using the best methodology. Type I errors never ranged above 6% when the best methodology was employed.

Readers should exercise some caution about generalizing these results to their own settings. Curricular or instructional differences unique to your districts may influence the accuracy with which the estimated cut scores reflect actual performance in your setting. With this limitation in mind, we would encourage educators to use this data as one tool to inform standards-based decisions.

The information gathered in this study came from measures employing the NWEA RIT Scale. Because all of the research that we have to date indicates that scores generated from computer-based tests and Achievement Level Test (ALT) scores are virtually interchangeable, readers should feel comfortable applying the results of this study in any setting that uses the RIT scale.

We hope that data from this study provides useful information to help South Carolina educators use NWEA assessments to better inform, plan and deliver student instruction. Good information, when matched with the professionalism and commitment of our South Carolina colleagues, will assure that every student has the opportunity to reach their aspirations.

References

Hauser, C. (2002, January). Alignment of the NWEA RIT Scales with the South Carolina *Palmetto Achievement Challenge Tests.*

Cronin, J. (2003, March). Alignment of the NWEA RIT Scales with the South Carolina *Palmetto Achievement Challenge Tests.*