Adjustments made to the Results of the NWEA RIT Scale Minnesota Comprehensive Assessment Alignment Study

July, 2004

John Cronin, Ph.D. – Northwest Evaluation Association

Recently NWEA completed a project to connect the scale of the MCA and BST with NWEA's RIT scale. Six Minnesota systems participated in the study, using test information from a group of over 13,000 students enrolled in third, fifth, and eighth grades who took these Minnesota Assessments and NWEA tests in the spring of 2003. Information from these tests was used in a comprehensive study to identify the capacity of the RIT scale to predict success on the Minnesota Assessments and to identify performance level scores on the RIT scale that would indicate a good chance of success on this test.

After the announcement of the study results, the Minnesota Department of Education informed schools that the MCA scales and proficiency cut points needed to be changed. Based on their announced changes, we have made adjustments in our estimated RIT cut scores for the MCA tests.

We made these adjustments from tables provided by the Mounds View School District. These tables linked the new scales scores to the original scale scores for the MCA. We entered the new cut scores into our prior regression estimates to estimate the changes in the projected cut points. While this method may not produce quite as precise a result as we might have achieved by rerunning the study, we believed the methodology we applied would result in no meaningful loss in the accuracy of prediction. Tables 1 and 2 show the original estimate, the adjustments, and the new estimates.

Contact John Cronin (503) 624-1951 at Northwest Evaluation Association if you wish further information or have questions about this study.

Table 1 – Best estimate of RIT scores
linked to the performance levels on the Minnesota Comprehensive Assessment and Basic Skills Test - Reading

(the change from the prior estimated cut score is noted in parentheses – estimated cut scores for non-tested years are in blue)

Reading to Minnesota Reading					
Grade	I	IIa	IIb	III	IV
3 (MCA)	<185 (+3)	185 (+3)	196 (+3)	201 (+1)	215 (+1)
4	<191	191	202	206	221
5 (MCA)	<196 (+3)	196 (+3)	207 (+3)	211 (+1)	226 (+1)
6			210		
7			214		
8 (BST)			218		

Table 2– Best estimate of RIT scores
linked to the performance levels on the Minnesota Comprehensive Assessment and Basic Skills Test – Mathematics

(estimated cut scores for years not evaluated are in blue – 4th grade score projects 5th grade IIb level on MCA – 6th and 7th grade score projects "pass" on 8th grade BST)

Mathematics to Minnesota Mathematics					
Grade	I	IIa	IIb	III	IV
3 (MCA)	<186 (+3)	186 (+3)	200 (+2)	203 (0)	217 (0)
4	<195	195	206	211	227
5 (MCA)	<202 (+6)	202 (+6)	211 (+1)	218 (0)	237 (0)
6			217		
7			224		
8 (BST)			231		

Aligning the NWEA RIT Scale with the Minnesota Comprehensive Assessment and Minnesota Basic Skills Test

December 2003

John Cronin, Ph.D. and Branin Bowe – Northwest Evaluation Association

Each year, Minnesota students in grades three and five participate in statewide assessment using the *Minnesota Comprehensive Assessment (MCA).* Students in grade eight take the Minnesota Basic Skills Test (BST). 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.

The Minnesota Department of Families and Learning has developed scales that are used to assign students to one of five performance levels on the MCA. These are, from the lowest cut score to the highest, I, IIa, IIb, III, and IV. For purposes of NCLB, level IIb is considered the level that represents satisfactory performance. The Basic Skills Test is a two level, pass/fail assessment.

Many students who attend school in Minnesota 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.

The versions of NWEA tests in use in Minnesota have been specifically aligned to match the content of local and Minnesota state curriculum standards. Because of this, we believe there is a good match in content between the NWEA tests and the curriculum standards being used in Minnesota.

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. 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 have now been completed in Arizona, California, Colorado, Idaho, Illinois, Illinois, Iowa, Minnesota, Montana, Oregon, South Carolina, Texas Wyoming, and Washington. 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 Minnesota assessments?
- What RIT scores correspond to various performance levels on the Minnesota scales?
- How well can *passing* performance on the Minnesota assessments (level IIb on the MCA and a score of pass on the BST) be predicted from RIT scores when NWEA assessments are administered in the same time frame?

Method

Participating School Systems

An e-mail solicitation was sent in October, 2003 to all Minnesota school systems who had two or more seasons of experience with NWEA testing prior to spring 2003 to secure participants for the study. Based on the response from this solicitation, spring 2003 MCA and NWEA student assessment records were collected from seven school districts. Spring 2003 BST and NWEA student assessment records were collected from two of these seven school districts.

Data Preparation

Student level test records from spring 2003 Minnesota testing and spring 2003 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 the reading analysis for this study.

Table 1
Reading Tests Included by Grade

	MCA		BST	
2002-3 grade in school	**3**	**5**	**8**	**Total**
Reading	5129	5466	2513	13108
Mathematics	5206	5506	2491	13203

We had enough student records at each grade to adequately cover the breadth of the scale and perform a robust analysis. We had fewer districts participate in the grade eight study than normal for these kinds of analysis. If the alignment of curriculum in these two districts were to differ substantially from the level of curriculum alignment present in other districts, there might be some minor changes in the predicted cut scores. 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 Minnesota assessment scores at each grade to determine how closely the scores on the NWEA test correlated with same subject scores on the corresponding Minnesota assessment. Simple bivariate correlation coefficients were computed among these scores.

Linking MCA/BST scores to the RIT scales. Three methods of estimating cut scores for MCA/BST levels (abbreviated MINN) were used. The most straightforward was simple linear regression ($MINN_{pred} = a(RIT) + c$). 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 ($MINN_{pred} = a(RIT^2) + b(RIT) + c$). For each of these methods, the RIT score was determined by substituting the appropriate MINN score for $MINN_{pred}$ and solving the equation for RIT.

A fixed-parameter Rasch model was also used to estimate RIT cut scores. In this method, MINN performance level was treated as a MINN 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 Minnesota performance levels from RIT scores. RIT scores were first used to predict whether students were likely to achieve performance at or above the IIb performance level (1420 scale score) on the MCA or a passing score on the BST (600). We make the estimates from this level in order to maintain consistency with prior studies of state test alignment, which make comparisons based on the NCLB reported performance level. This allows us to make accurate comparisons of our alignment with different state tests.

The predictions of Minnesota 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 in fact achieves a failing score. This index was generated for the linear, second order, and Rasch methodologies. In general, the highest prediction index score was used to select the RIT cut score to be adapted as the *official* RIT score we would associate with achieving the pass standard on corresponding Minnesota test for the particular grade level and subject area. We do make exceptions 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 MCA levels based on observed RIT scores. The predictions of MCA level performance were compared to observed performance in 5 X 5 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 Minnesota tests were not conducted for purposes of this study. The standards used to construct the NWEA Assessments were the same as those used for the Minnesota assessments. Both NWEA assessments and the Minnesota assessments include multiple-choice items. The Minnesota assessments also include short answer and extended response questions.

Results

Descriptive Statistics

Table 2 reviews descriptive statistics for the Minnesota and NWEA assessments. The median RIT scores for this sample are well above those for the NWEA norm population. In reading, median scores range from 5 to 7 points above the spring median for the NWEA norm group, depending on grade level. In mathematics, the differences range from 7 to 10 points above the NWEA norm group. This distribution can impact the predictions in a cut score alignment study. Our primary concern is that the predicted MCA cut scores for students scoring in the I and IIa levels are generated from smaller numbers of students than we would have

generated from a normal distribution. This does not seriously affect predictions at the IIa level, which was estimated from about 600 students per grade. It may have some affect at level I, a level achieved by only 188 students at grade 3 and 288 students at grade 5.

Table 2
Means, Standard Deviations, and Medians for Minnesota and NWEA assessments

Grade	3	5	8
Minnesota Reading (MCA Grades 3 & 5 BST Grade 8)			
N	5129	5456	2513
Mean	1531.00	1590.67	659.31
Median	1550	1600	653
SD	184.50	202.33	57.36
NWEA Reading			
N	5129	5456	2513
Mean	202.29	216.51	231.00
Median	204	217	232
SD	14.40	13.84	13.20
Minnesota Math (MCA Grades 3 & 5 BST Grade 8)			
N	5142	5456	2492
Mean	1564.09	1555.72	638.48
Median	1550	1550	642
SD	237.56	193.12	54.56
NWEA Math			
N	5142	5456	2492
Mean	206.97	223.41	243.68
Median	208	224	245
SD	13.70	16.23	17.97

Pearson correlations

Tables 3 to 5 show the results of this analysis for each grade. Same subject Pearson correlations between the NWEA and MCA assessments were quite high, ranging from .77 (grade 3 mathematics) to .83 (grade 5 reading and mathematics). Same subject correlations between the NWEA and BST assessment were also quite high.

Table 3
Pearson Correlations for MCA and NWEA assessments by Subject - Reading

	MCA Grade 3	MCA Grade 5
NWEA Grade 3	.82	
NWEA Grade 5		.83

Table 4
Pearson Correlations for MCA and NWEA assessments by Subject - Mathematics

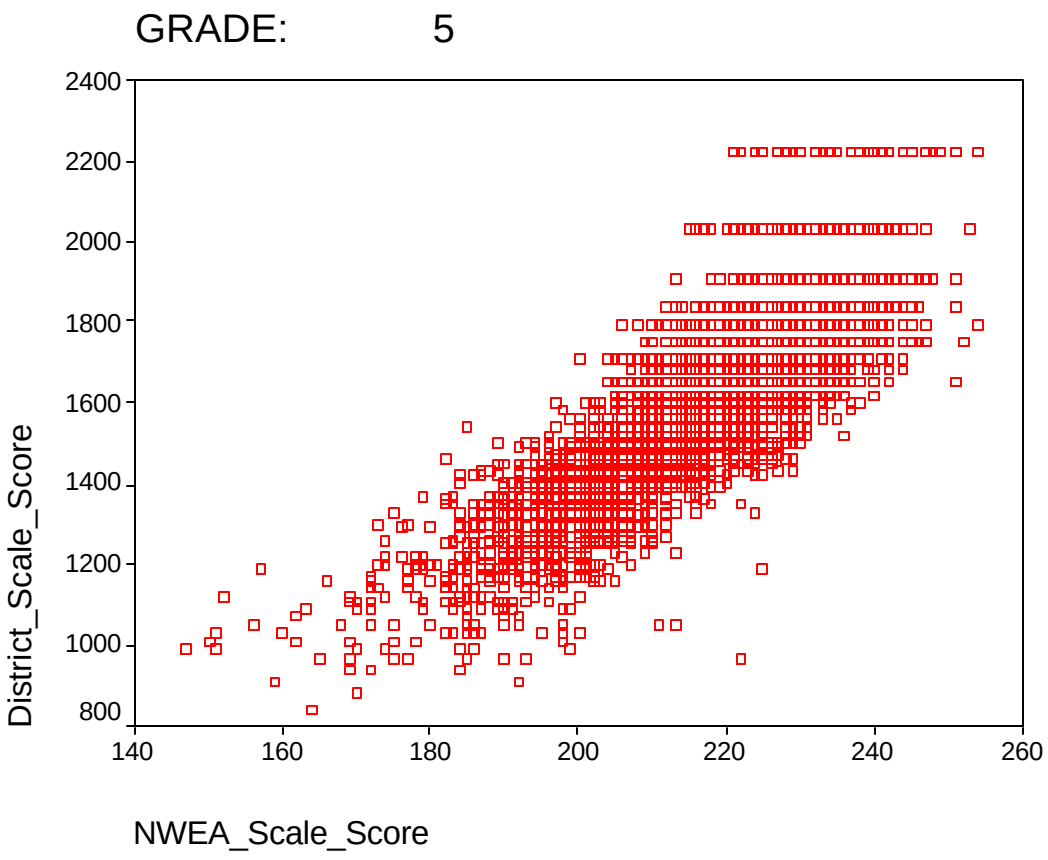
	MCA Grade 3	MCA Grade 5
NWEA Grade 3	.77	
NWEA Grade 5		.83

Table 5 - Pearson Correlations for BST and NWEA assessments

	BST Reading	BST Math
NWEA Reading	.77	
NWEA Math		.85

Analysis of scatterplots suggested that the scale relationships were curvilinear and revealed some tendency for correlations to break down near the lower end of the two scales, possibly indicating a floor effect, on both the MCA and BST. Figure 1 provides an example that illustrates both the scale relationships and the breakdown in correlation near the bottom of the scales reasonably well. Readers might also note that "rows" near the top of the MCA scale, which are a product of successively larger differences in scale scores as the student's raw score increases near the top of the scale.

Figure 1 – Scatterplot depicting Grade 5 MCA math plotted against Grade 5 NWEA math



Linking MCA 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 MCA and the *pass* cut score for the BST. 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 6 shows several estimations of the Spring 2003 RIT score that correspond to the cut scores for the various performance levels on the MCA scale. Table 7 shows several estimations of the Spring 2003 RIT score that correspond to the minimum passing score for the BST.

Table 6
Estimated points on the RIT scale equating to the minimum scores for the performance levels on the MCA (Rounded) -Reading

	Linear regression					Second-order regression					Rasch status-on-standard				
Reading	I	IIa	**IIb**	III	IV	I	IIa	**IIb**	III	IV	I	IIa	**IIb**	III	IV
Grade 3	<178	178	**192**	200	218	<178	178	**192**	200	218	<183	183	**193**	200	214
Grade 5	<189	189	**202**	209	226	<189	189	**202**	209	226	<193	193	**204**	210	225
Mathematics	I	IIa	**IIb**	III	IV	I	IIa	**IIb**	III	IV	I	IIa	**IIb**	III	IV
Grade 3	<180	180	**196**	203	221	<180	180	**196**	202	220	<183	183	**198**	204	217
Grade 5	<188	188	**210**	218	240	<190	190	**210**	218	239	<196	196	**210**	218	237

Table 7
Estimated points on the RIT scale equating to the minimum scores for the performance levels on the BST (Rounded) -Mathematics

	Linear regression	Second-order regression	Rasch status-on-standard
Reading	214	214	218
Mathematics	229	228	231

Predicting Minnesota assessment 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 Minnesota assessments. The most accurate method of prediction was generally used to derive the best estimate of RIT cut scores that equate to the different MCA performance levels and the BST minimum passing score. 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 either level IIb or above on the MCA or a score of pass on the BST. Next we assessed the accuracy with which the RIT predicted level assignment on MCA. 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. The *prediction index* statistic is used for this purpose. 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 isn't 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. For purposes of the *No Child Left Behind Act* , schools are judged on their ability to move children to the IIb level and beyond. RIT scores can provide teachers with advance notice about students who may not reach these goals on the Minnesota assessment that corresponds to their grade level.

Table 8 shows the results for reading. When using the methodology chosen to produce the most reasonable estimate, the rate of accurate prediction ranged from about 89% and 93%. Type I error ranged between about 4% and 7% and prediction index scores were above .930 at every grade. The results suggest that the NWEA reading assessment predicts *pass-fail* status on the Minnesota reading assessments well.

Table 9 shows the results for mathematics. The accuracy of proficiency status predictions ranged from about 88% to about 92% when using the best methodology. Type 1 errors ranged from about 5% to 7%. Prediction index scores were all above .920 when using the best method. The results again suggest a high rate of prediction for all grades.

The Rasch method of calibration generally produced more accurate predictions of proficiency status than either regression method in all subjects, although all three methods worked reasonably well.

Table 8
Accuracy of the RIT scale in predicting MCA/BST reading pass-fail status

Grade 3 - MCA	**Cut Score**	**Accuracy**	**Type I Error**	**Prediction Index**
Linear	192	88.79%	6.67%	.925
Second Order	192	88.79%	6.67%	.925
Rasch SOS	193	88.81%	5.95%	.933*
Grade 5 – MCA	**Cut Score**	**Accuracy**	**Type I Error**	**Prediction Index**
Linear	202	91.90%	5.69%	.938
Second Order	202	91.90%	5.69%	.938
Rasch SOS	204	91.88%	4.21%	.954*
Grade 8 – BST	**Cut Score**	**Accuracy**	**Type I Error**	**Prediction Index**
Linear	214	93.00%	5.41%	.942
Second Order	214	93.00%	5.41%	.942
Rasch SOS	218	92.80%	3.82%	.959*

* Indicates methodology chosen for recommended estimate

Table 9
Accuracy of the RIT scale in predicting MCA/BST pass-fail status - mathematics

Grade 3	Cut Score	Accuracy	Type I Error	Prediction Index
Linear	196	87.53%	8.82%	.899
Second Order	196	87.53%	8.82%	.899
Rasch SOS	198	87.65%	6.93%	.921*
Grade 5	**Cut Score**	**Accuracy**	**Type I Error**	**Prediction Index**
Linear	210	90.99%	4.81%	.947*
Second Order	210	90.99%	4.81%	.947*
Rasch SOS	210	90.99%	4.81%	.947*
Grade 8	**Cut Score**	**Accuracy**	**Type I Error**	**Prediction Index**
Linear	229	91.45%	6.34%	.931
Second Order	228	91.25%	7.07%	.923
Rasch SOS	231	91.69%	5.14%	.944*

* Indicates methodology chosen for recommended estimate

Table 10 summarizes the accuracy of prediction for this study relative to other state alignment studies. To date, NWEA has used the prediction index to estimate the accuracy with which the RIT score predicts pass-fail performance for state tests in Arizona, California, Colorado, Indiana, Illinois, Minnesota, South Carolina, Wyoming, and Washington. The prediction indices for Minnesota are high relative to those we have generated for other states in reading. Some portion of this might be explained by the high performance of the students (when students pass the vast majority of the time it becomes easier to predict a pass; it is also easy to predict failure when the vast majority of students fail a test). The index number is slightly below average in mathematics, primarily because of the slightly lower prediction index average for third grade. The indices show overall that NWEA assessments correctly predict pass-fail status for about 17 students for each type I error in reading and 15 students for each Type I error in mathematics. These are more than adequate for the assessments to be useful for identifying students who are at risk against state standards.

Table 10
Prediction Indices (Based on Proficiency Status)
for Previous NWEA State Alignment Studies

State	Reading	State	Lanaguage	State	Math
Texas	.974	Texas	.968	Texas	.970
Washington	.971	Indiana '01	.907	Wyoming	.961
Minnesota	**.944**	Colorado '03	.903	Colorado '01	.957
Wyoming	.931	Indiana '03	.894	Washington	.949
Colorado '03	.931	Arizona	.874	Illinois	.946
Illinois	.928			California	.944
California*	.921			Colorado '03	.943
Arizona	.912			South Carolina	.943
Colorado '01	.910			**Minnesota**	**.936**
South Carolina	.902			Washington	.936
Indiana '01	.902			Arizona	.919
Indiana '03	.900			Indiana '01	.899
Washington	.886			Indiana '03	.860

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

Predicting MCA Performance Levels from RIT Scores

The MCA reports five levels of performance. Four cut scores are set to define these five levels. Analyzing the capacity of RIT scores to predict students' MCA 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 student performance varies some from test to test and because tests vary in their ability to measure students at the high and low performance level. This is especially true when predicting the lowest performance level, which was achieved by only about 7% of the students in the study. The accuracy of performance level prediction declines as additional levels are assigned, and Minnesota's use of five levels makes correct prediction of the performance level considerably more difficult than would be the case in most other states.

When predicting performance levels, a case is identified as accurate when the performance level assigned by the MCA 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 MCA. For example, if the RIT score projects a level IV performance for the student and the MCA test result is level III, we declare the case a Type I error because the RIT score overestimated performance.

Table 11
Accuracy of the RIT scale in predicting MCA performance level - reading

Grade 3	**Accuracy**	**Type I Error**	**Prediction Index**	**% Level V found**	**% Level I found**
Linear	56.64%	21.45%	.621*	43.5%	51.4%
Second Order	56.64%	21.45%	.621*	43.5%	51.4%
Rasch SOS	56.91%	23.18%	.593	62.7%*	68.8%*
Grade 5	**Accuracy**	**Type I Error**	**Prediction Index**	**% Level V found**	**% Level I found**
Linear	61.06%	22.48%	.632	67.5%	46.5%
Second Order	61.06%	22.48%	.632	67.5%	46.5%
Rasch SOS	60.69%	21.03%	.653*	71.9%*	61.9%*

* indicates most accurate method

Table 12
Accuracy of the RIT scale in predicting MCA performance level - mathematics

Grade 3	**Accuracy**	**Type I Error**	**Prediction Index**	**% Level V found**	**% Level I found**
Linear	54.29%	24.54%	.548*	45.3%	46.3%
Second Order	54.62%	26.27%	.519	48.7%	46.3%
Rasch SOS	53.63%	25.96%	.516	63.3%*	56.3%*
Grade 5	**Accuracy**	**Type I Error**	**Prediction Index**	**% Level V found**	**% Level I found**
Linear	59.99%	20.13%	.664*	51.7%	24.6%
Second Order	60.27%	20.42%	.661	55.0%	30.3%
Rasch SOS	60.71%	21.27%	.650	64.9%*	58.5%*

* indicates most accurate method for this purpose

The results reported in tables 11 and 12 suggest that the Rasch method produced the most accurate predictions for students at the highest and lowest performance level, while the linear

regression method more often accurately predicted students in the middle three performance levels.

NWEA has reported estimated performance level assignments for prior studies conducted in Arizona, Colorado, Illinois, Indiana, Minnesota, South Carolina, Texas and Washington. Table 13 compares the accuracy with which these tests predict performance level. The results show that the MCA performance index scores are lower than those for other studies conducted to date. This is primarily a product of the fact that Minnesota uses five performance levels while the other states use 3 or 4. This means that Minnesota predictions must match within a 25 cell matrix (5 X 5 performance levels), while predictions in all other states matched within either a 9 cell matrix (3 X 3 performance levels) or a 16 cell matrix (4 X 4 performance levels).

Given the difficulty of this type of prediction, we were satisfied that performance level predictions were reasonably accurate. In addition, the model accurately identified between 56% and 72% of the students who performed in the highest and lowest performance categories, an identification rate that was equal or superior to what we have found in other studies.

Table 13
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 '03	.860	Illinois	.888
Colorado	.840	Colorado	.808
Illinois	.804	Washington	.805
South Carolina	.757	Indiana '03	.804
Arizona	.756	South Carolina	.764
Washington	.698	Arizona	.756
Minnesota	**.627**	**Minnesota**	**.611**

Best estimates of MCA performance level cut scores

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

- For the level IIb RIT score, we selected the methodology that produced the highest performance index score in predicting "pass/fail" alone.
- For the level III RIT score, which differentiates the below standard and warning level we used the methodology that produced the highest performance index score for predicting performance levels.
- For the level I, IIa, and IV RIT score, we selected the cut score that correctly predicted the largest proportion of students who actually achieved level I or V performance on the MCA.

Tables 14 and 15 summarize the recommended cut scores for each performance level on the MCA and for pass/fail status on the BST. Based on NWEA student growth norms, the table also includes estimated scores that would indicate "on-track" performance for students in grade 4 on the MCA and for grades 6 and 7 relative to the BST. These scores are in blue text.

Table 14
Projected Reading RIT Scores Equivalent to Performance Levels on Minnesota MCA/BST
(estimated scores for years not tested are in blue)

Grade	Level I			Level IIa	Level IIb			Level III			Level IV		
	Score Range	% of pop. Identified	Method	Cut Score	Cut Score	Performance Index	Method	Cut Score	Performance Index	Method	Cut Score	% of pop. Identified	Method
3	<183	68.8%	Rasch	183	193	.933	Rasch	200	.621	Linear	214	62.7%	Rasch
4	<188			188	199			205			220		
5	<193	61.9%	Rasch	193	204	.954	Rasch	210	.653	Rasch	225	71.9%	Rasch
6					209								
7					214								
8					218	.958	Rasch						

Table 15
Projected Mathematics RIT Scores Equivalent to Performance Levels on Minnesota MCA/BST
(estimated scores for years not tested are in blue)

Grade	Level I			Level IIa	Level IIb			Level III			Level IV		
	Score Range	% of pop. Identified	Method	Cut Score	Cut Score	Performance Index	Method	Cut Score	Performance Index	Method	Cut Score	% of pop. Identified	Method
3	<183	56.3%	Rasch	183	198	.921	Rasch	203	.548	Linear	217	63.3%	Rasch
4	<190			190	204			211			228		
5	<196	58.5%	Rasch	196	210	.947	Rasch	218	.664	Linear	237	64.9%	Rasch
6					217								
7					224								
8					231	.944	Rasch						

Using RIT scores to estimate student probability of achieving passing performance on the MCA/BST

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 tests 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 MCA or BST. 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 Minnesota standards so they can focus their time with these students on providing new challenges that better suit their current needs.

Tables 16 and 17 show the proportion of students at each RIT level who earned scores at or above the level IIb on MCA reading and mathematics assessments or a score of *pass* on the BST during the same spring. Using Table 16 as an example, we find that 32% of the 3rd grade students who achieved a reading RIT score between 185 and 189 went on to achieve a passing score on the MCA reading assessment. A 3rd grade teacher with ten students performing in this range would know that only about three in ten of these kinds of students will meet standard on the MCA unless they work harder, receive more focused instruction, or have access to additional resources.

On the other hand, almost all (about 98%) of third grade students performing at a 205 to 209 level achieved proficiency on MCA/BST reading. Teachers should feel free to focus their efforts with these students on new and more difficult challenges than the basic third grade standards might provide.

Figures 1 and 2 are graphic depictions of the data in the tables.

Table 16
Proportion of students passing the MCA and BST based on same spring RIT score - Reading

RIT	Grade 3	Grade 5	Grade 8
160	0.0%		
165	4.5%		
170	7.9%		
175	9.1%	0.0%	
180	19.5%	5.0%	
185	32.0%	8.4%	0.0%
190	52.4%	9.9%	11.1%
195	79.1%	27.9%	10.7%
200	91.1%	45.3%	15.8%
205	97.9%	77.9%	14.1%
210	99.8%	92.8%	28.0%
215	99.8%	98.1%	60.5%
220	100.0%	99.4%	76.1%
225		99.9%	93.5%
230		99.8%	98.4%
235		100.0%	99.8%
240			100.0%

Figure 1 - Proportion of students passing the MCA and BST based on same spring RIT score - Reading

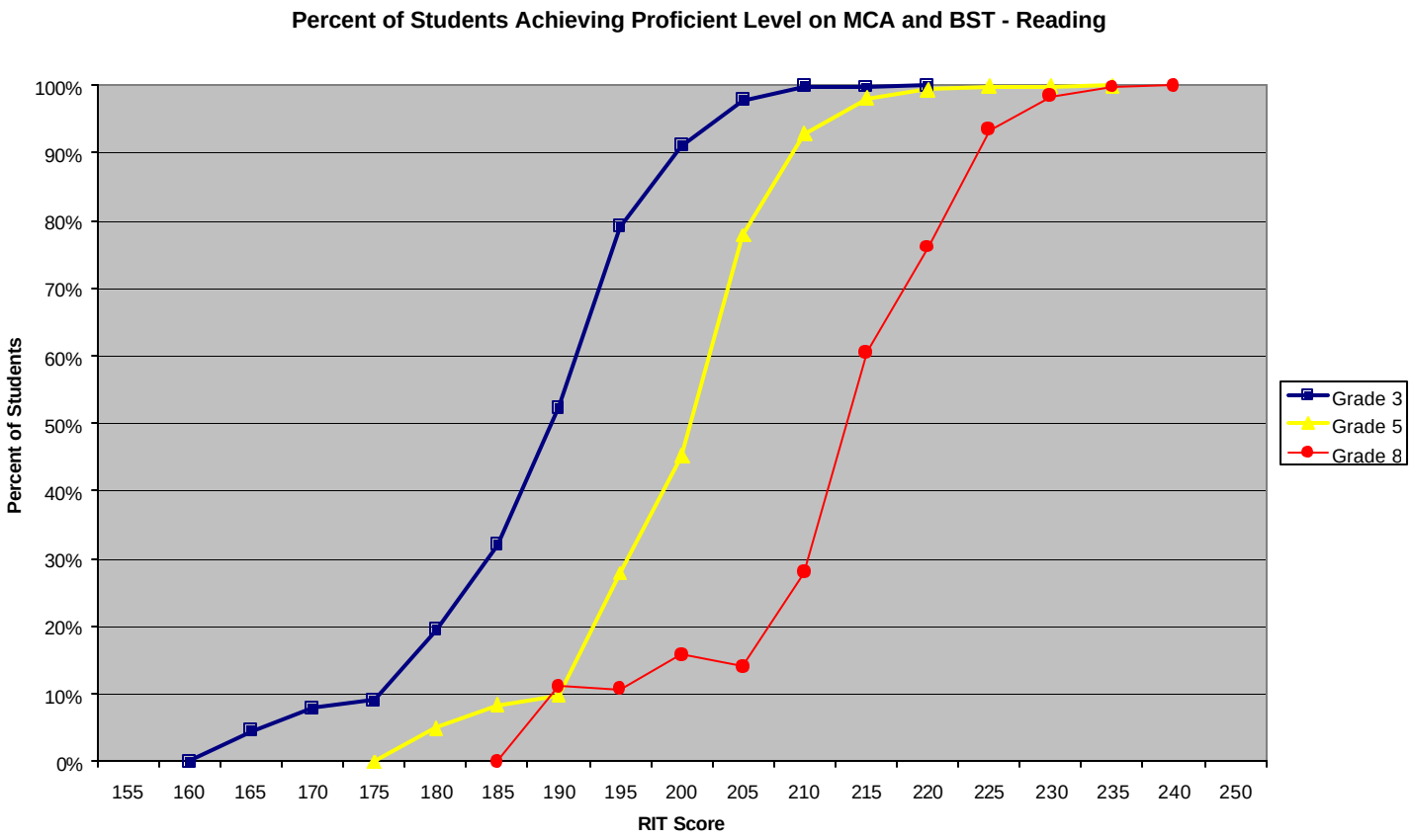
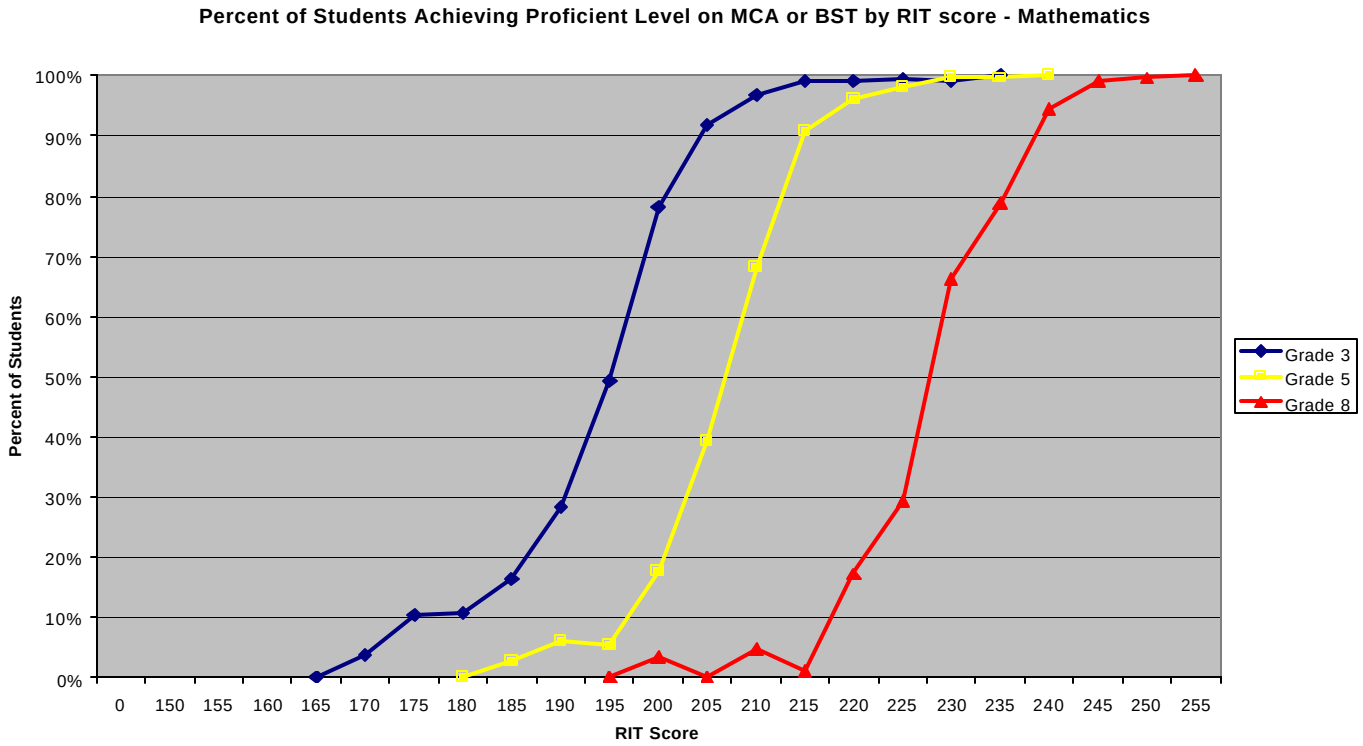


Table 17
Proportion of students passing the MCA/BST based on same spring RIT score – Mathematics

RIT	Grade 3	Grade 5	Grade 8
160			
165	0.0%		
170	3.9%		
175	10.5%		
180	10.9%	0.0%	
185	16.5%	2.8%	
190	28.5%	6.2%	
195	49.4%	5.4%	0.0%
200	78.2%	17.6%	3.6%
205	92.0%	39.5%	0.0%
210	96.7%	68.3%	4.9%
215	99.1%	91.0%	1.3%
220	99.0%	96.0%	17.3%
225	99.6%	98.1%	29.3%
230	99.2%	99.7%	66.2%
235	100.0%	99.6%	78.8%
240		100.0%	94.5%
245			99.0%
250			99.6%
255			100.0%
260			

Figure 2 - Proportion of students passing the MCA and BST Mathematics based on same spring RIT score – Mathematics



Comparing MCA/BST proficiency 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 proficiency test in fourteen states. To get an estimate of the difficulty of the Minnesota standards in relation to other state tests, we evaluated the standard used as the cut score for *NCLB reporting or passing* and compared it to the cut score representing the same standard in these other states. Although the number of states studied is rapidly increasing, the states studied may not reflect what is typical in regard to these kinds of standards.

The results are summarized in Tables 17 and 18. One finding is that Minnesota's cut scores are relatively similar across grades in both reading and mathematics. When cut scores are calibrated from grade to grade, teachers, students, and parents receive more reliable information about a student's capacity to perform at standard in the future. We would also note that the MCA level IIb and BST cut scores seem to be calibrated to the same level.

A second finding is that Minnesota's cut scores are near the average of that of most of the other states studied, but lower than cut scores estimated from prior studies of the Minnesota assessment. We'd recommend caution about drawing any judgments about the quality of Minnesota's standards from this finding. States establish standards for different purposes. Some states, Washington might be an example, set standards at a level they believe appropriate for students pursuing post-secondary education. Others may set standards at a lower level that reflects the literacy needed to be successful in the workplace. The No Child Left Behind Act requires schools to set targets that would result in all students achieving a ***meets standard or*** proficient level of performance in about 12 years. Some communities in Minnesota are close to achieving this already, but many will have to improve the performance of large proportions of their students to reach this goal. Standards should be judged on how well they align with the purposes the community has set for establishing standards, not purely on how high or low the "bar" is set. One thing the tables make clear is that proficiency standards vary widely from state to state and that proficiency is not yet a concept that has a shared definition, although greater consensus in standard setting seems to be emerging.

Table 18 - Cut scores representing "proficient" or "meets standards" level of performance on 14 state assessments

Reading

Grade 3			Grade 4			Grade 5			Grade 6			Grade 7			Grade 8			Grade 9			Grade 10		
State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile
SC	205	67	WY	214	73	SC	220	73	SC	221	63	SC	227	70	WY	232	74	MT	224	43	OR	236	77
CA	200	51	SC	213	70	CA	214	54	CA	216	46	WA	226	67	SC	230	68	IA	224	43	WA	227	51
MN	193	35	WA	207	53	AZ	210	45	MT	211	35	CA	221	50	OR	227	58	ID	221	37	ID	224	44
OR	193	35	CA	205	46	OR	209	42	ID	211	35	MT	218	43	CA	226	54	CO	204	9	MT	224	44
ID	193	35	ID	200	34	IL	207	37	IN	210	32	IA	216	37	AZ	224	49				IA	223	42
MT	193	35	MT	196	26	MT	206	35	IA	209	30	ID	215	35	IN	219	35				CO	209	15
IL	193	35	IA	196	26	ID	206	35	TX	208	28	TX	210	24	MT	219	35				CA	208	14
IN	192	32	CO	191	18	IA	205	32	CO	197	11	CO	206	18	IA	219	35						
IA	191	31				MN	204	30							ID	218	32						
AZ	190	29				TX	204	30							IL	218	32						
TX	179	13				CO	197	18							MN	218	32						
CO	179	13													CO	206	12						

Mathematics

Grade 3			Grade 4			Grade 5			Grade 6			Grade 7			Grade 8			Grade 9			Grade 10		
State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile
SC	208	75	WY	221	83	SC	227	76	SC	235	78	SC	242	78	WY	257	89	MT	242	47	WA	257	73
CA	204	60	WA	218	76	CA	225	70	CA	230	67	WA	242	78	SC	251	80	IA	241	44	MT	247	40
IN	201	50	SC	217	74	AZ	220	59	IN	221	47	CA	238	70	AZ	248	75	ID	240	42	IA	247	40
OR	199	46	CA	212	59	OR	215	46	ID	219	42	ID	225	44	CA	240	59	CO	235	32	OR	245	33
AZ	199	46	ID	205	39	ID	213	41	IA	218	40	MT	224	42	OR	235	50				ID	242	25
MN	198	42	IA	205	39	MT	212	38	MT	218	40	IA	222	38	ID	233	46				CO	233	14
MT	197	39	MT	205	39	IA	212	38	CO	207	19	TX	221	35	MN	231	42				CA	232	13
IA	197	39				MN	210	33				CO	216	26	IN	231	42						
ID	196	36				IL	210	33							IL	230	40						
IL	193	29				TX	209	31							MT	228	36						
						CO	201	15							IA	228	36						
															CO	225	31						

- Indiana tests students in the fall. Their cut scores were adjusted to reflect equivalent spring performance
- Colorado uses the partially proficient level of performance for NCLB reporting. To maintain consistency we report the level each state uses for NCLB reporting here.
- The Texas estimate is based on the level for proficient performance that will be implemented in 2005.

Using RIT scores and data from this alignment study to set individual growth targets

NWEA encourages educators and parents to collaborate on setting individual growth targets for students based on what we call a "hybrid-growth model". The *meets standard* cut score for each grade reflect benchmarks that students who are "on-target" would meet if they were to achieve the state's benchmark for the *No Child Left Behind Act*. For students who are behind this benchmark, we recommend a growth target that would reflect the norm for their grade and RIT range (see the 2002 NWEA norms study for this information) plus some proportion of the gap between their current performance and the benchmark that the student would try to close during this school year. For those students whose performance is ahead of the benchmark, we suggest a target that reflects their current RIT range norm.

This approach assures that each student has a growth target that is challenging. It also assures that low performing students have targets that will assure they eventually reach proficiency standards. Schools that achieve high rates of success on these kinds of targets will assure that no child is left behind (to borrow a phrase) while also making sure that all children have the opportunity to get ahead, regardless of where they stand against a standard. More information on this approach can be obtained by contacting the research team at NWEA.

Summary and Conclusions

This study investigated the relationship between the scales used for the Minnesota state assessments and the RIT scales used to report performance on Northwest Evaluation Association tests. The study determined RIT score equivalents for MCA/BST passing scores in reading and mathematics. MCA/BST and NWEA test performances for over 7,000 students in grades 3, 5 and 8 were included in this study.

Three methods generated an estimate of RIT cut scores that could be used to project MCA performance levels and BST pass/fail results. Rasch SOS methods generally produced the most accurate cut score estimates. Accuracy of predicting MCA/BST passing performance was above 90% for all grades when using the best methodology. Type I errors were under 6% in reading and mathematics 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 Minnesota educators use NWEA assessments to better inform, plan and deliver student instruction. Good information, when matched with the professionalism and commitment of our Minnesota colleagues, will assure that every student has the opportunity to reach their aspirations.