

Predicting Success on the Maryland School Assessment in Reading with Measures of Academic Progress in Reading

Huafang Zhao, Ph.D.

Background

Every spring, Maryland students in Grades 3–8 take the Maryland School Assessment (MSA) in reading and mathematics to meet the *No Child Left Behind Act* requirements. Student performance is categorized as basic, proficient, or advanced.

In 2004–2005, the Montgomery County Public Schools (MCPS) began to administer the Measures of Academic Progress—Reading (MAP-R), a computer adaptive achievement test developed by Northwest Evaluation Association (NWEA), to students in Grades 3–5. MAP-R results are reported as scale scores. MCPS students take MAP-R three times a year—fall, winter, and spring. Testing was expanded to include students in Grades 6–8 during the 2005–2006 winter administration.

Methodology

The purpose of this brief is to identify students at risk of scoring basic on the MSA in reading, based on their winter MAP-R scale scores. School staff can then plan instructional support or interventions. The MCPS students who took both the MAP-R in January 2006 and the MSA in reading in March 2006 were included in the study. The study addresses the following two research questions:

1. What are the estimated winter MAP-R cut scores that predict MSA reading success in Grades 3–8?
2. How accurate are the predictions?

A standard setting procedure was used in these analyses. Standard setting procedures can be categorized into test-centered or examinee-centered models (Jaeger, 1993). Angoff’s procedure is an example of a test-centered method (Angoff, 1971), in which judges examine each test item and estimate the probability that a minimally proficient person could correctly answer each test item. The borderline-group procedure (Zieky and Livingston, 1977) is an empirical way of setting standards based on the competency of the examinees. This procedure focuses on examinees’ performance rather than on

test items. Conceptually, borderline students are defined as minimally proficient because they score at or just above the cut score.

The MSA reading cut scores discussed in this study refer to the cut scores between basic and proficient. The Standard Error of Measurement (SEM) for MSA reading cut scores ranges from 10 to 12 scale score points (MSDE, 2005). For example, about 1,571 Grade 3 students had borderline scores, between 377 and 399. Their median MAP-R scale score (190) was used as the estimated MAP-R cut score for predicting MSA reading success.

Students whose MSA scale scores were one SEM below or above MSA cut scores were selected as the borderline group for each grade. The demographic information for the borderline students is presented in Table A1 of the Appendix. Once borderline students were identified, their median MAP-R scores were calculated as the cut scores. Tables were then generated to examine how accurately these cut scores predicted MSA reading proficiency for all MCPS students.

Results

MAP-R Cut Scores for MSA Proficient Students

Table 1 provides MAP-R cut scores produced by NWEA (Bowe and Cronin, 2005) and the borderline-group method. Students below the MAP-R cut scores are at high risk of performing basic on MSA.

Table 1
Estimated MAP-R Cut Scores
to Predict Success on MSA in Reading

	Fall MAP-R Cut Scores by NWEA	Winter MAP-R Cut Scores by Borderline Method
Grade 3	184	190
Grade 4	189	196
Grade 5	200	203
Grade 6	N/A	209
Grade 7	N/A	214
Grade 8	N/A	219

Correlation between the reading MSA and the MAP-R is presented in Table A2 of the Appendix.

Prediction Accuracy of MAP-R Cut Scores

Prediction accuracy is the percentage of students correctly identified (Table 2). A total of 86.1% out of 9,187 Grade 3 students were correctly identified as either proficient (73.6%) or not proficient (12.5%). For Grades 3–5, prediction accuracy ranges from 86.1% to 88.4%, and for Grades 6–8, between 84.7% and 89.5%. Thus, predictions are approximately 85% to 90% accurate.

Table 2
Classification Agreement Between
MAP-R and Reading MSA Proficiency

MAP-R		Spring 2006 MSA Reading	
		Prof	Not Prof
Grade 3 <i>N</i> =9,187	Prof	73.6	5.4
	Not Prof	8.5	12.5
Grade 4 <i>N</i> =9,771	Prof	79.4	3.5
	Not Prof	8.1	9.0
Grade 5 <i>N</i> =9,987	Prof	76.8	5.2
	Not Prof	6.5	11.4
Grade 6 <i>N</i> =9,607	Prof	74.2	4.1
	Not Prof	6.4	15.3
Grade 7 <i>N</i> =9,961	Prof	70.1	5.4
	Not Prof	7.6	16.9
Grade 8 <i>N</i> =10,232	Prof	65.7	6.3
	Not Prof	9.0	19.0

A review for race/ethnicity findings indicate that prediction accuracy for African American and Hispanic students ranges from 75.3% to 84.4% (Appendix Table A3), slightly lower than the 87.2% to 94.2% for White and Asian American students. One of the possible reasons is that there are higher percentages of African American and Hispanic students (Table A1) who scored close to the MSA reading cut scores. These students are more likely to be misidentified.

Discussion

The winter MAP-R cut scores produced by the borderline-group method predicted MCPS student performance on the reading MSA reasonably well with 85% to 90% accuracy. The high classification agreement makes MAP-R cut scores useful for school staff. The proposed MAP-R cut scores can help predict success on the MSA in reading. Students below the estimated MAP-R cut scores may need instructional intervention.

The prediction deliberately overidentified basic student performance to allow school staff the opportunity to intervene. As a result, only 6.3% or fewer students were identified as proficient who later scored basic on the reading MSA from Grades 3–8.

Additional analyses on the 2004–2005 cohort show consistent results for elementary schools. In middle schools, only one year of data were available so it is impossible to validate the results. It is necessary to examine additional middle school cohorts in the future to determine if the proposed MAP-R cut scores yield similar prediction results for middle schools from year to year.

References

Angoff, W. H. (1971). Scales, norms and equivalent scores. In R. L. Thorndike (Ed.) *Educational Measurement*. Washington DC: American Council on Education.

Bowe, B. & Cronin, J. (2005). *Aligning the NWEA RIT Scale with the Maryland School Assessment (MSA)*. Portland, OR: Northwest Education Association.

Jaeger, R. (1993). *Certification of student competency*. Educational measurement (3rd edition). Phoenix, AZ: Oryx Press.

Maryland State Department of Education (2005). *Maryland School Assessment Technical Reports*. Baltimore, MD.

Zieky, M. J., & Livingston, S. A. (1977). *Manual for setting standards on the basic skills assessment test*. Princeton, NJ: Educational Testing Service.

Appendix

Table A1
Number and Percentage of Students in Borderline Group by Grade and Ethnicity

	Total <i>N</i>	Female %	Male %	Asian %	African American %	Hispanic %	White %	FARMS*	Sp. Edu.**	ELL***
Grade 3	1,571	44.6	55.4	10.3	32.7	32.5	24.2	42.3	15.7	13.8
Grade 4	1,208	44.5	55.5	9.7	35.3	34.0	20.5	47.4	23.1	14.0
Grade 5	1,745	46.0	54.0	9.6	35.6	32.4	22.1	45.3	22.9	13.1
Grade 6	1,217	49.3	50.7	9.9	35.3	32.3	22.1	41.9	19.3	6.9
Grade 7	1,778	47.4	52.6	10.6	34.4	28.2	26.7	34.1	16.4	4.7
Grade 8	2,060	46.7	53.3	12.4	31.8	26.6	29.0	32.1	14.6	5.1

* FARMS refers to students who receive Free and Reduced-price Meals System services.

** Sp. Edu. refers to students who receive special education services.

*** ELL refers to students who are English language learners.

Table A2
Correlation Coefficients
Between Winter 2005 MAP-R and
Spring 2006 Reading MSA Scale Scores

	<i>N</i>	<i>r</i>
Grade 3	9,187	0.76
Grade 4	9,771	0.76
Grade 5	9,987	0.78
Grade 6	9,607	0.80
Grade 7	9,961	0.78
Grade 8	10,232	0.77

Table A3
Prediction Percentages Between MAP-R and Reading MSA by Race/Ethnicity and Grade

	Proficient on Both	Basic on Both	Proficient on MAP-R Basic on MSA	Proficient on MSA Basic on MAP-R	% Accurately Predicted
Grade 3					
African American	57.5	23.2	8.8	10.5	80.7
Asian	84.5	6.3	4.2	5.0	90.8
Hispanic	54.1	21.2	7.3	17.4	75.3
White	87.5	5.0	3.0	4.5	92.5
Grade 4					
African American	64.7	16.5	6.4	12.4	81.2
Asian	89.4	3.9	2.1	4.6	93.3
Hispanic	64.4	15.8	4.8	15.0	80.2
White	90.6	3.6	2.0	3.9	94.2
Grade 5					
African American	60.5	19.9	9.2	10.4	80.4
Asian	87.2	6.3	3.4	3.2	93.5
Hispanic	58.8	21.9	7.3	11.9	80.7
White	90.3	3.9	2.7	3.2	94.2
Grade 6					
African American	56.8	27.6	5.8	9.9	84.4
Asian	84.9	6.8	3.2	5.1	91.7
Hispanic	52.5	30.6	6.5	10.3	83.1
White	89.0	5.1	2.5	3.4	94.1
Grade 7					
African American	50.5	31	7.5	11.1	81.5
Asian	81.9	8.0	4.3	5.9	89.9
Hispanic	46.9	33.5	8.0	11.6	80.4
White	86.5	5.2	3.6	4.7	91.7
Grade 8					
African American	45.7	32.4	8.8	13.1	78.1
Asian	76.5	10.7	5.2	7.6	87.2
Hispanic	39.9	38.3	8.6	13.2	78.2
White	82.9	7.2	4.3	5.6	90.1