

Predicting Grade 5 Students' Performance on the Maryland School Assessment in Reading

Vasuki Rethinam, Ph.D. and Carol J. Schatz, Ph.D.

Executive Summary

This research brief describes a model that correctly predicts reading performance (basic vs. proficient or above) on the Maryland School Assessment (MSA) for approximately 93% of Grade 5 students. This brief can be used to help staff better understand the student reports shown on Data Warehouse for prediction models.

Three recommendations are:

1. The predicted scores should be used as one of the tools to gauge students' instructional needs because predictions provide a best estimate of student's performance score rather than the student's predetermined score.
2. Predictions should be used in combination with other student information to inform decisions about instruction, intervention, and acceleration.
3. The Office of Shared Accountability (OSA) should continue to monitor model performance and update the model as needed because the changes in student cohorts, assessments, and instructional practices could affect the accuracy of the prediction.

Background

Goal 1 of the Montgomery County Public Schools (MCPS) strategic plan is to ensure success for every student (MCPS, 2006). Although MCPS is making progress toward attainment of this goal, one area that continues to be the focus of improvement efforts is students' reading performance in elementary school.

MCPS administrators and instructors increasingly are invested in identifying tools and methods that can help determine the likelihood of a student attaining proficiency on the MSA. Having this information at the beginning of the school year can facilitate instructional planning for all students and allow interventions to begin earlier for students in need of assistance.

This brief is an extension of a study predicting MSA reading scores in middle schools (Schatz, Gheen, & Rethinam, 2008). Using student information available before the beginning of the school year, several indicators are integrated to predict performance with more accuracy than from a single indicator. This study addresses the following two questions:

1. Can a model be developed to predict Grade 5 students' MSA reading scores and performance levels using indicators available in MCPS databases?
2. How accurate are the performance-level predictions for all students and for different student groups?

Methodology

The sample for this study consisted of 8,821 Grade 5 students in MCPS who took the 2008 MSA in reading and for whom predictor data were available.

The prediction model was developed in two steps. The first step was to predict Grade 5 students' MSA reading scale scores. The second step was to identify two "cuts" for the predicted scores to categorize predicted MSA performance levels (basic, proficient, and advanced).

The variables used to develop the prediction model were selected because of their hypothesized association with reading achievement (Abedi, 2002; Cunningham & Stanovich, 2003; Dunn, Kadane, & Garrow, 2003; Hoffman, 1991; McCoy & Reynolds, 1999) and their availability in MCPS databases. The following variables were included to predict Grade 5 students' MSA reading scores:

- Grade 4 MSA reading score
- Grade 4 spring MAP-R score
- Grade 3 MSA reading score
- Grade 4 MSA mathematics score
- Grade 4 reading level on report card.

Grade 4 attendance rate also was examined as a predictor but was not included in the final model because it did not improve prediction accuracy.

To include as many students as possible in the sample, Grade 3 MSA reading scores were estimated for 629 Grade 5 students who did not have one (Cohen, Cohen, West, & Aiken, 2003; Graham & Hofer, 2000). Generalized estimating equations (GEE)—regression-based technique—were used to estimate the final model parameters.

Cut scores for assigning students to predicted performance levels were chosen based on the resulting model accuracy and prediction errors for all demographic groups. The cut scores selected for predicted scores were set slightly higher than the cut scores for the Grade 5 MSA reading established by the Maryland State Department of Education (MSDE). This was done to underestimate rather than overestimate students' predicted performance levels.

Results

Prediction accuracy is calculated as the percentage of students whose predicted Grade 5 MSA reading performance level matches their actual performance level. Prediction errors occur when students perform below or above their predicted level.

Basic vs. proficient or above. Table 1 shows the accuracy and error rates by demographic group for predicting two MSA performance levels—basic vs. proficient or above.

Table 1
Number of Students and Percent Predicted
Basic vs. Proficient or Above for 2008
Grade 5 MSA Reading, by Demographic Group

Basic vs. Proficient or Above	Total N	Correct Pred. %	Worse	Better
			than Pred. %	than Pred. %
All Students	8,821	92.6	3.1	4.3
Male	4,525	91.4	4.0	4.6
Female	4,296	93.9	2.2	3.8
African Am.	2,002	87.0	5.6	7.3
Asian Am.	1,384	96.5	1.8	1.7
Hispanic	1,811	86.1	5.2	8.7
White	3,601	97.6	1.2	1.3
FARMS ¹	2,401	84.9	5.7	9.4
Special Education	1,092	78.8	5.5	15.8
ESOL ²	833	79.8	6.6	13.6

Note. American Indian student results not reported separately.

¹ FARMS = Free and Reduced-price Meals System

² ESOL = English for Speakers of Other Languages

The model correctly predicts those two performance levels for 92.6% of the students. Accuracy rates are lowest among groups of students whose predicted scores are near the cut scores. For nearly all student groups, when predictions are wrong, students are more likely to perform better than predicted.

Prediction accuracy was highest for White (97.6%) and Asian American (96.5%) students. The errors of prediction for Asian American and White students are nearly balanced between over-prediction and under-prediction. Prediction accuracy for Hispanic, African American, and groups of students receiving FARMS, ESOL, or special education services was in the range of 79%–87%. However, students in these groups are more likely to perform better than predicted when predictions are wrong.

Advanced vs. proficient or below. The model correctly predicts two MSA performance levels—advanced vs. proficient or below, for 83.9% of all students (Table 2). For nearly all student groups, when predictions are wrong, students are more likely to perform better than predicted.

Table 2
Number of Students and Percent Predicted
Advanced vs. Proficient or Below for 2008
Grade 5 MSA Reading, by Demographic Group

Advanced vs. Proficient or Below	Total N	Correct Pred. %	Worse	Better
			than Pred. %	than Pred. %
All Students	8,821	83.9	6.2	9.9
Male	4,525	83.3	6.7	10.0
Female	4,296	84.6	5.6	9.8
African Am.	2,002	81.8	6.3	11.8
Asian Am.	1,384	85.7	7.2	7.2
Hispanic	1,811	79.8	5.5	14.6
White	3,601	86.4	6.1	7.5
FARMS	2,401	81.0	5.7	13.3
Special Education	1,092	82.4	2.9	14.7
ESOL	833	79.6	4.6	15.8

Note. American Indian student results not reported separately.

Basic vs. proficient vs. advanced. Accuracy rates will drop when predicting three levels of performance rather than two. The error for predicting three levels is at least the sum of the errors at each cut score. Predictions for all three performance levels (basic, proficient, or advanced) are correct for 76.9% of the students, with 13.8% of students attaining a performance level higher than predicted and 9.3% of students earning a performance level below the predicted (see Appendix Table A1).

Discussion

The model achieved high accuracy rates (92.6% for distinguishing between basic vs. proficient and above, and 83.9% for distinguishing between advanced vs. proficient and below), and it achieved reasonably good performance for all demographic groups. This makes it a useful tool for prediction of all three performance levels.

Although the model accuracy rates are high for most student groups, for individual students the predicted score range is estimated to be about plus or minus 20 points of the student's predicted score. This is because other factors, such as students' experience in the current school year, motivation, and engagement may have an influence on students' performance. Student performance observed during the current school year should be weighted more heavily than the prediction from indicators used in this study.

These analyses have been extended to produce MSA reading performance prediction models for Grades 3 and 4. The performance of these models with data from 2007–2008 are provided in the appendix (Table A2 and Table A3).

Recommendation

Staff should interpret the predicted score as a best estimate of the student's performance score rather than the student's predetermined score. Therefore, predictions should be used as just one tool to gauge students' instructional needs. Elementary schools' staff should obtain all the indicators used in the study to predict a student's performance. Students missing one of those scores (except for Grade 3 MSA reading) may not have a predicted score.

When making decisions about interventions, staff should be aware that a student's predicted MSA score is best interpreted within a range or confidence interval of 20 points. If students have a predicted score of basic that is within 20 points of the proficient cut score (as reported in the Data Warehouse) for that grade, they have potential for attaining proficiency. If they have a predicted score of proficient that is within 20 points of the proficient cut score for that grade, they are at risk of performing at the basic level.

Differences between student cohorts and changes in the assessments and/or instructional practices over time could affect the accuracy of the prediction model. OSA should continue to monitor model performance and update the model as needed.

References

- Abedi, J. (2002). Standardized achievement tests and English language learners: Psychometrics issues. *Educational Assessment, 8*, 231–257.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. (3rd ed.). Mahwah, NJ: Erlbaum.
- Cunningham, A. E., & Stanovich, K. E. (2003). Reading matters: How reading engagement influences cognition. In J. Flood, D. Lapp, J. Squire, & J. Jensen (Eds.), *Handbook of research on teaching the English language arts* (2nd ed., pp. 666–675). Mahwah, NJ: Lawrence Erlbaum.
- Dunn, M. C., Kadane, J. B., & Garrow, J. R. (2003). Comparing harm done by mobility and class absence: Missing students and missing data. *Journal of Educational and Behavioral Statistics, 28*, 269–288.
- Graham, J. W., & Hofer, S. M. (2000). Multiple imputation in multivariate research. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches and specific examples* (pp. 201–218). Mahwah, NJ: Lawrence Erlbaum.
- Hoffman, J. V. (1991). Teacher and school effects in learning to read. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 911–950). New York: Longman.
- McCoy, A. R., & Reynolds, A. J. (1999). Grade retention and school performance: An extended investigation. *Journal of School Psychology, 37*, 273–298.
- Montgomery County Public Schools. (2006). *Our Call to Action: Pursuit of Excellence*. Rockville, MD: Author.
- Schatz, C. J., Gheen, M. H., & Rethinam, V. (September 2008). *Research brief: Using lagging indicators to predict Grade 7 performance on the Maryland School Assessment in reading*. Department of Shared Accountability, Montgomery County Public Schools, Maryland.

THIS PAGE WAS LEFT BLANK INTENTIONALLY

Predicting Grade 5 Students' Performance on the Maryland School Assessment in Reading

Appendix

Table A1
Performance of Model¹ for Predicting Performance Levels for Grade 5 Students MSA Reading,
by Demographic Group, 2007–2008

Demographic Group ²	Total <i>N</i>	Predicting Proficient or Above vs. Basic (2 Groups)			Predicting Advanced vs. Proficient or Below (2 Groups)			Predicting Basic vs. Proficient vs. Advanced (3 Groups)		
		Correct Prediction %	Worse than Predicted %	Better than Predicted %	Correct Prediction %	Worse than Predicted %	Better than Predicted %	Correct Prediction %	Worse than Predicted %	Better than Predicted %
All Students	8,821	92.6	3.1	4.3	83.9	6.2	9.9	76.9	9.3	13.8
Male	4,525	91.4	4.0	4.6	83.3	6.7	10.0	75.1	10.7	14.3
Female	4,296	93.9	2.2	3.8	84.6	5.6	9.8	78.9	7.8	13.3
African American	2,002	87.0	5.6	7.3	81.8	6.3	11.8	69.4	12.0	18.6
Asian American	1,384	96.5	1.8	1.7	85.7	7.2	7.2	82.2	9.0	8.8
Hispanic	1,811	86.1	5.2	8.7	79.8	5.5	14.6	66.5	10.7	22.8
White	3,601	97.6	1.2	1.3	86.4	6.1	7.5	84.3	7.2	8.6
FARMS	2,401	84.9	5.7	9.4	81.0	5.7	13.3	66.6	11.4	22.0
Special Education	1,092	78.8	5.5	15.8	82.4	2.9	14.7	62.7	8.3	28.9
ESOL	833	79.8	6.6	13.6	79.6	4.6	15.8	60.3	11.2	28.6

¹ Model predictors include Grade 4 MSA reading, Grade 4 spring MAP-R (Measured Academic Progress-Reading), Grade 3 MSA reading, Grade 4 MSA mathematics, and reading level on last Grade 4 report card.

² American Indian student results not reported separately.

Table A2
Performance of Model¹ for Predicting Performance Levels for Grade 4 Students MSA Reading,
by Demographic Group, 2007–2008

Demographic Group ²	Total <i>N</i>	Predicting Proficient or Above vs. Basic (2 Groups)			Predicting Advanced vs. Proficient or Below (2 Groups)			Predicting Basic vs. Proficient vs. Advanced (3 Groups)		
		Correct Prediction %	Worse than Predicted %	Better than Predicted %	Correct Prediction %	Worse than Predicted %	Better than Predicted %	Correct Prediction %	Worse than Predicted %	Better than Predicted %
All Students	8,665	92.9	3.2	3.9	84.6	6.6	8.8	77.5	9.8	12.7
Male	4,388	92.1	3.6	4.3	85.2	7.0	7.8	77.3	10.7	12.1
Female	4,277	93.7	2.8	3.5	84.0	6.1	9.9	77.8	8.9	13.4
African American	1,937	85.8	6.2	8.0	87.9	4.2	7.9	73.6	10.5	15.9
Asian American	1,333	97.8	1.4	0.8	82.4	6.9	10.7	80.3	8.3	11.4
Hispanic	1,689	87.3	5.6	7.1	89.9	3.1	7.0	77.3	8.6	14.1
White	3,683	97.4	1.2	1.3	81.2	9.3	9.5	78.7	10.5	10.8
FARMS	2,287	84.3	6.5	9.2	90.8	2.8	6.3	75.1	9.4	15.6
Special Education	992	78.8	4.9	16.2	92.0	3.0	4.9	71.0	8.0	21.1
ESOL	984	84.0	5.7	10.3	92.9	0.6	6.5	76.9	6.3	16.8

¹ Model predictors include Grade 3 MSA reading, Grade 3 MAP-R (Measured Academic Progress-Reading), Grade 3 MSA mathematics, Grade 2 average NCE Score for TN2 (*TerraNova*, 2nd edition) reading, language, and math, and reading level on last Grade 3 report card.

² American Indian student results not reported separately.

Table A3
Performance of Model¹ for Predicting Performance Levels for Grade 3 Students MSA Reading,
by Demographic Group, 2007–2008

Demographic Group ²	Total <i>N</i>	Predicting Proficient or Above vs. Basic (2 Groups)			Predicting Advanced vs. Proficient or Below (2 Groups)			Predicting Basic vs. Proficient vs. Advanced (3 Groups)		
		Correct Prediction %	Worse than Predicted %	Better than Predicted %	Correct Prediction %	Worse than Predicted %	Better than Predicted %	Correct Prediction %	Worse than Predicted %	Better than Predicted %
All Students	8,676	89.6	4.5	5.9	84.2	6.6	9.2	73.9	11.1	15.0
Male	4,465	88.1	5.5	6.4	85.4	7.0	7.7	73.6	12.5	14.0
Female	4,211	91.1	3.5	5.3	82.9	6.2	10.9	74.2	9.7	16.1
African American	1,912	84.6	7.1	8.3	90.8	4.1	5.1	75.5	11.2	13.3
Asian American	1,343	94.6	3.0	2.5	80.9	8.9	10.1	75.5	11.9	12.6
Hispanic	1,803	80.6	8.4	11.0	91.5	3.2	5.3	72.2	11.5	16.3
White	3,587	95.0	1.8	3.2	78.1	8.8	13.1	73.2	10.6	16.2
FARMS	2,441	80.7	8.4	11.0	93.8	2.2	4.0	74.5	10.5	15.0
Special Education	945	75.4	5.6	18.9	93.2	2.0	4.8	69.2	7.6	23.2
ESOL	1,585	76.8	9.3	13.8	93.4	2.0	4.6	70.5	11.2	18.3

¹ Model predictors include Grade 2 TN2 (*TerraNova*, 2nd edition) reading, Grade 2 TN2 mathematics, Grade 2 TN2 language, Grade 2 M-CLASS benchmark performance which is part of MCPS-Assessment Program in Primary Reading, Grade 2 GT (Gifted and Talented) flag, and reading level on last Grade 2 report card.

² American Indian student results not reported separately.