

MONTGOMERY COUNTY PUBLIC SCHOOLS, ROCKVILLE, MARYLAND

**Linking Early Literacy Assessments to State and National Assessments
Reading Outcomes in Grade 3**

Helen Wang, Ph.D.
Shihching Liu, M.A.

Executive Summary

To examine the impact of the upcoming Comprehensive PreK-3 Literacy Policy to prioritize development in core components of foundational literacy skills and reading comprehension for students from all backgrounds, designed by Maryland State Board of Education (Maryland State Board of Education, 2024), the Department of Shared Accountability conducted a study to support identification of students who may not be on track to meet grade-level literacy proficiency on Grade 3 state and national standardized assessment.

This study examined how the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) and its Spanish version, Lectura, administered to primary students in MCPS can serve as predictors of later reading outcomes. This research focused on the predictive relationship between DIBELS scores of Grade 1 (2022–2023) and Grade 2 (2023–2024) and subsequent student performance on two Grade 3 assessments, which were the Maryland Comprehensive Assessment Program English Language Arts and Literacy (MCAP ELA/L) and the spring Measures of Academic Progress Reading (MAP-R) in the 2024–2025 school year.

Method and Analysis

A single-group linking method, a common approach for studies that connect assessments taken by the same group of students, was adopted to explore predictive validity between Grades 1 and 2 spring DIBELS scores and Grade 3 reading outcomes. Linking samples were created from students who completed both spring DIBELS (in Grade 1 or 2) and Grade 3 MCAP ELA/L and spring MAP-R, respectively. Research questions were designed to investigate the relationship of spring DIBELS scores in Grades 1 and 2 with each of the outcome measures in Grade 3: MCAP ELA/L and spring MAP-R. Various statistical procedures, including Logistic regression, were applied to identify DIBELS composite score thresholds that predicted a high probability (75% and 80%) of achieving reading proficiency on the Grade 3 assessments.

Results

The analysis revealed moderately strong predictive relationships between spring DIBELS performance and later reading outcomes. The spring DIBELS composite score thresholds in Grades 1 and 2 were then identified that corresponded to a high likelihood (75% and 80%) of students meeting reading proficiency standards on the Grade 3 MCAP ELA/L and spring MAP-R assessments. The threshold-analysis results for two outcome measures were highly consistent.

Conclusion

The spring DIBELS composite score thresholds established from this study offer valuable guidance for MCPS administrators and teachers. Specifically, predicted reading outcomes can help inform instructional planning and early intervention strategies in primary grades to support students' reading development and improve their readiness for Grade 3 reading assessments. Particularly, the results would support identification of students who may not be on track toward meeting reading proficiency expectation in Grade 3 in alignment with the state literacy policy.

Background

In the context that the Maryland students have experienced a decline in reading performance since 2013, the Maryland State Board of Education (Maryland State Board of Education, 2024) has designated the Comprehensive PreK-3 Literacy Policy to prioritize development in core components of foundational literacy skills and reading comprehension for students from all backgrounds. It has been the district's priority to promote primary students' progress toward reading proficiency in Grade 3 and later grades, as measured by state and national standardized assessments, in order to align with the state Common Core State Standards for college and career readiness (CCR).

MCPS uses the DIBELS and Lectura assessments as research-based screening tools to measure empirically validated foundational skills of primary students related to their reading ability (Amplify, 2020). The administration of DIBELS and Lectura in the district is aligned with the Maryland State Board of Education's advocacy for the statewide assessment tools that are used to understand student mastery of foundational literacy skills. Using the universal screening tools help accurately and reliably identify students who show early signs of reading difficulties and who may be at risk for poor learning outcomes in foundational reading skills (Maryland State Board of Education, 2025).

For about two decades, MCPS has administered the Measures of Academic Progress (MAP) assessments in Grades 3-8, which serves the purpose of informing instruction, projecting proficiency, and identifying students for intervention. Moreover, MAP is integral to the work throughout the school district, supporting everything from individual and classroom goal setting to districtwide evaluation and strategic planning. MAP assessments were also used to support identification of students who were likely to be proficient in reading on the state's assessments. In an earlier MCPS study that linked spring MAP and Maryland Comprehensive Assessment Program (MCAP), strong or very strong validity evidence in concurrent relationships between the assessments was found in both reading and mathematics across Grades 3 to 8 (Wang, 2023). Based on the linking study, concordance tables were established to find scores on spring MAP that corresponded to scores on MCAP associated with the CCR benchmark. The score identified for Grade 3 spring MAP-Reading (MAP-R) and its corresponding score for the CCR benchmark (750 or higher aligned to performance level 3 or higher) on Maryland Comprehensive Assessment Program English Language Arts and Literacy (MCAP ELA/L) were used as the reading proficiency standards in the current study.

Purpose and Scope

The purpose of the study was to predict reading proficiency in Grade 3 and identify primary students who, based on composite score thresholds on spring DIBELS, may not be on track to meeting reading proficiency on MCAP ELA/L and spring MAP-R at the end of Grade 3. Data obtained from the study can be put in place to ensure all students' reading needs met through receiving additional reading instruction and intervention at school.

The study was initially planned to link assessments of interest for Grade 3 students in 2024–2025, examining the predictive relationship of spring DIBELS and spring Lectura assessments they took in primary grades with MCAP ELA/L and spring MAP-R assessments they took later in Grade 3. However, the exploratory analysis revealed that predictive relationships between spring Lectura and MCAP ELA/L and spring MAP-R were not strong enough for making connection of the assessments. Therefore, students who only took spring Lectura in Grades 1 and 2 were not included in the analysis of literacy proficiency prediction. In addition, there was no DIBELS data for Kindergarten to be used for this study because DIBELS assessments were not administered in 2021–2022.

Research Questions

1. To what extent were spring DIBELS assessments from Grades 1 and 2 correlated with MCAP ELA/L overall scale scores and spring MAP-R Rasch Unit (MAP-R RIT) scores in Grade 3?
2. What were the spring DIBELS composite score thresholds in Grades 1 and 2 that were associated with 75-percent and 80-percent probabilities to attain CCR benchmark score on Grade 3 MCAP ELA/L assessment and to earn Grade 3 spring MAP-R proficiency score established by the MCPS linking study?
3. How accurately did the established spring DIBELS composite score thresholds in Grades 1 and 2 predict attainment of reading proficiency on MCAP ELA/L and spring MAP-R in Grade 3? How did prediction accuracy vary by student race/ethnicity and service receipt?

Methodology

The study adopted a single-group linking method that related two assessments of interest. The design, commonly used for linking studies, requires a sample of students who took both assessments of interest (i.e., spring DIBELS in Grade 2 and MCAP ELA/L in Grade 3). The advantage of this design is its capability to control for differential proficiency of test takers (Dorans et al., 2010).

Data and Measures

The study incorporated the following assessment data for Grade 3 students in 2024–2025 and their assessment data from their primary grades:

- Spring 2023–2024 DIBELS composite scores as predictor data in Grade 2
- Spring 2022–2023 DIBELS composite scores as predictor data in Grade 1
- 2024–2025 MCAP ELA/L overall scale scores as predicted outcome data in Grade 3 at the benchmark score of 750 on a grade-level scale (650 to 850), that corresponds to the state CCR benchmark for performance level 3 or higher.
- Spring 2024–2025 MAP-R RIT scores (100 to 300) as predicted outcome data in Grade 3 at the benchmark score of 197 for reading proficiency, which was identified through the MCPS linking study for MAP and MCAP (Wang, 2023).

DIBELS as Predictor

MCPS administers the state-approved DIBELS assessment three times per year at the beginning (fall), middle (winter), and end (spring) of the school year. The test developer, Amplify (2020), indicated that the recently updated 8th version of DIBELS was among the most effective reading screening tools available. The screening tool is widely used in various educational settings and by educators focused on improving literacy for all students. With consistent grade-level subtests, DIBELS provides empirically validated measures of reading skill acquisition related to general reading outcomes. The composite score, including scaled scores on individual subtests of DIBELS, provides estimates of students' overall performance relative to each subtest, and can be used to predict reading proficiency in current and future grade levels (Amplify, 2020).

Core components of DIBELS administered to primary students include foundational reading skills: phonological awareness, phonics and vocabulary, accurate and fluent reading, and reading comprehension. The assessment, consisting of many sub-measures, can be used to regularly monitor individual students' progress toward success in reading and identify students with early warning signs of reading difficulties who need additional instructional support and intervention. In addition, DIBELS measures are aligned with many of the Common Core State Standards, especially those for foundational reading skills. Therefore, DIBELS data also can help develop grade-level and districtwide plans associated with enhancement of curriculum, validation of instructional objectives, and recommendation of targeted intervention programs to students at different grade levels, which would ultimately lead to adequate progress toward meeting the system-level literacy benchmark goals at current and later grades.

MCAP ELA/L as Outcome

MCAP ELA/L, administered to Grades 3 to 8 at a period toward the end of the school year, is the statewide assessment that measures student progress towards proficiency upon the state CCR standards with regard to the skills and knowledge necessary to be college and career ready in literacy. The assessment focuses on reading of literary and informational passages, while students demonstrate their reading comprehension and literacy skills through responding to text-based questions and writing prompts. Based on overall scale scores of the assessment, students are placed

in a proficiency continuum classified as beginning learner (level 1), developing learner (level 2), proficient (level 3), and distinguished learner (level 4). Students who perform at levels 3 and 4 are proficient in the assessed content and meet or exceed the benchmark score of 750 for performance level 3.

DIBELS scores in Grades 1 and 2 were used to determine the extent to which they predict proficiency on Grade 3 reading outcome as measured by the MCAP ELA/L. The Grade 3 MCAP ELA/L assessment measures students' foundational reading skills and ability in understanding and analyzing informational and literary texts.

MAP-R as Outcome

MAP-R, developed and nationally normed by the Northwest Evaluation Association (NWEA), is a computer-adaptive assessment administered to students in Grades 3 to 8 in MCPS. The assessment, used as a student achievement measure in the district, is designed to be aligned with the objectives of the state's CCR standards and the Common Core State Standards. It measures various reading areas, including word recognition, reading comprehension, inferential or interpretive comprehension, evaluative comprehension, literary responses or analysis, and general reading. Vertically equated MAP-R RIT scores can be used to measure student progress and achievement in reading proficiency and help adjust instructional practices accordingly.

In this study, Grade 3 students' spring MAP-R data were used as a reading outcome measure that was predicted by the students' early reading performance on spring DIBELS in Grades 1 and 2. Grade 3 MAP-R measures how well students can read, comprehend, analyze informational texts, understanding the author's message, identifying the main idea, and analyzing the text's structure and features, based on the foundational vocabulary use and reading comprehension skills acquired in their primary grades.

Samples

The selected study design used the linking samples, which comprised students who had completed both assessments of interest. Specifically, the linking samples were generated to include Grade 3 students who took both MCAP ELA/L in 2024–2025 and spring DIBELS in their primary grades in 2023–2024 (Grade 2) and 2022–2023 (Grade 1), respectively. The samples were also created for Grade 3 student who took spring MAP-R in Grade 3 and spring DIBELS in the primary grades. Table 1 presents information about four linking samples used for the study.

Table 1
Linking Sample Sizes and Mean Scores of Assessments

Linking Samples		N Students	Mean Scores		
			Spring DIBELS Composite Score	MCAP ELA/L Overall Scale Score	Spring MAP-R RIT Score
Linking sample 1	Students taking Grade 2 spring DIBELS and Grade 3 MCAP ELA/L	10,621	454.9	755.0	--
Linking sample 2	Students taking Grade 1 spring DIBELS and Grade 3 MCAP ELA/L	9,871	474.2	755.5	--
Linking sample 3	Students taking Grade 2 spring DIBELS and Grade 3 spring MAP-R	10,581	454.8	--	197.3
Linking sample 4	Students taking Grade 1 spring DIBELS and spring Grade 3 MAP-R	9,825	474.3	--	197.9

Analyses

The linkages of spring DIBELS composite scores in Grades 1 and 2 with MCAP ELA/L overall scale scores and spring MAP-R RIT scores in Grade 3, respectively, were realized through various statistical procedures, such as Pearson correlation and Logistic regression procedures.

Analyses for Research Question 1

Predictive relationships were explored with the Pearson correlation analysis:

- Pearson correlation analysis was used to measure the strength and direction of the relationship of spring DIBELS composite scores with spring MCAP ELA/L overall scale scores and MAP-R RIT scores, respectively.
- Correlation coefficients range from -1 to +1 with higher values (those closer to -1 or +1) signifying a stronger correlation. The direction of the relation can be positive (between 0 and 1) or negative (between -1 and 0). A positive correlation means as one score increases, the other score also increases, whereas a negative correlation means as one score increases, the other score decreases.
- A strong or very strong positive correlation indicates predictive validity evidence. The correlation coefficient between assessments of interest would be presented for each of the linking samples.
- It is important to note that correlations do not imply causal relationships between the assessments of interest.

Table 2 provides an interpretation guide for the correlation coefficients, the higher the coefficient the greater the predictive relationship between the assessments of interest.

Table 2

Correlation Interpretation Guide

Size of correlation	Coefficient general interpretation
0.8 to 1.0	Very strong relationship
0.6 to 0.8	Strong relationship
0.4 to 0.6	Moderate relationship
0.2 to 0.4	Weak relationship

Analyses for Research Question 2

Logistic regression analysis was employed to identify predictive thresholds on the spring DIBELS composite scores in Grades 1 and 2. These thresholds correspond to a high probability (75% and 80%) of students later attaining reading proficiency in Grade 3, as measured by two key outcomes: an MCAP ELA/L overall scale score of 750 or higher (using linking samples 1 and 2) and a spring MAP-R RIT score of 197 or higher (using linking samples 3 and 4). The proficiency criterion of 197 on the MAP-R, established through a prior MCPS linking study (Wang, 2023), is aligned with the CCR benchmark score of 750 on the MCAP ELA/L, ensuring consistency in how proficiency is defined for the two assessments.

Analyses for Research Question 3

The prediction accuracy of the thresholds on the spring DIBELS composite scores in Grades 1 and 2, associated with the 75-percent probability of reading success in Grade 3, was estimated to determine whether a student's predicated performance in a primary grade matched his/her actual performance later in Grade 3 on MCAP ELA/L and spring MAP-R. The prediction accuracy of spring DIBELS for MCAP ELA/L and spring MAP-R was assessed by student race/ethnicity and service receipt groups in three categories:

- *Accurate estimation:* Students' predicted performance (whether or not meeting the spring DIBELS composite score thresholds in Grades 1 and 2) agreed with their actual performance (whether or not meeting the reading proficiency criteria on MCAP ELA/L and spring MAP-R in Grade 3).
 - Accurate prediction to be proficient in Grade 3 outcomes
 - Accurate prediction to be not proficient in Grade 3 outcomes
- *Underestimation:* Students who scored below the spring DIBELS composite score thresholds in Grades 1 and 2 actually met or exceeded the proficiency criteria on MCAP ELA/L and spring MAP-R in Grade 3.
- *Overestimation:* Students who scored at or above the spring DIBELS composite score thresholds in Grades 1 and 2 actually did not meet the proficiency criteria on MCAP ELA/L and spring MAP-R in Grade 3.

Results

Research Question 1: To what extent were spring DIBELS assessments from Grades 1 and 2 correlated with MCAP ELA/L overall scale scores and spring MAP-R RIT scores in Grade 3?

To address Research Question 1, a Pearson correlation analysis was conducted to measure the strength of the association between early reading scores and later reading performance. In other words, the analysis examined predictive validity between the assessments of interest. The results, shown in Table 3, indicate that spring DIBELS composite scores from Grades 1 and 2 were moderately strong predictors of Grade 3 outcomes on both the MCAP ELA/L overall scale scores and MAP-R RIT scores, with correlation coefficients between 0.70 and 0.74. In contrast, Lectura composite scores showed only a moderate relationship with the same Grade 3 outcome measures, with correlation coefficients between 0.51 and 0.60. Due to the moderate predictive relationship of Lectura with Grade 3 MCAP ELA/L and spring MAP-R, the data did not support proceeding with the more detailed threshold analyses planned for Research Question 2.

Table 3

Correlation between Spring DIBELS/Lectura and MCAP ELA/L and Spring MAP-R Scores

Assessment Grade 3 2024–2025	Correlation with Spring DIBELS Composite Scores		Correlation with Spring Lectura Composite Scores	
	Grade 2 2023–2024	Grade 1 2022–2023	Grade 2 2023–2024	Grade 1 2022–2023
MCAP ELA/L Scale Scores	.71 (N=10,621)	.72 (N=9,871)	.51 (N=573)	.59 (N=572)
Spring MAP-R RIT Scores	.74 (N=10,581)	.70 (N=9,825)	.54 (N=573)	.60 (N=572)

Note. N indicates the linking sample size, including students having both spring DIBELS/Lectura and MCAP ELA/L or spring MAP-R scores.

Research Question 2: What were the spring DIBELS composite score thresholds in Grades 1 and 2 that were associated with 75-percent and 80-percent probabilities to attain CCR benchmark score on Grade 3 MCAP ELA/L assessment and to earn Grade 3 spring MAP-R proficiency score established by the MCPS linking study?

Table 4 presents the predictive spring DIBELS composite scores thresholds in Grades 1 and 2. These thresholds served as indicators of a high probability (75% and 80%) of achieving reading proficiency on the Grade 3 MCAP ELA/L and spring MAP-R assessments. For the 75-percent probability level, the Grade 2 DIBELS composite score thresholds were identical for both outcome measures at a score of 472, while the Grade 1 DIBELS composite score thresholds were highly similar for the two outcome measures, with 489 for MCAP ELA/L and 490 for spring MAP-R. Similarly, at the 80-percent probability level, the Grade 2 DIBELS composite score thresholds remained consistent at 477 for both outcome assessments, and the Grade 1 DIBELS composite score thresholds again aligned closely at 496 for MCAP ELA/L and 497 for spring MAP-R.

Table 4

Primary-Grade Spring DIBELS Composite Score Threshold Associated with Predicted Probabilities of Being Proficient in Reading on MCAP ELA/L and Spring MAP-R in Grade 3

Spring DIBELS	Threshold at 75% Probability		Threshold at 80% Probability	
	Grade 3 MCAP ELA/L	Grade 3 Spring MAP-R	Grade 3 MCAP ELA/L	Grade 3 Spring MAP-R
Grade 2	472	472	477	477
Grade 1	489	490	496	497

Note. Including students with both of the DIBELS data in Grade 1 or Grade 2 and the MCAP ELA/L or MAP-R scores in Grade 3.

Research Question 3: How accurately did the established spring DIBELS composite score thresholds in Grades 1 and 2 predict attainment of reading proficiency on MCAP ELA/L and spring MAP-R in Grade 3? How did prediction accuracy vary by student race/ethnicity and service receipt?

The analysis of prediction accuracy for thresholds on the spring DIBELS composite scores in Grades 1 and 2 was conducted only for thresholds set at the 75-percent probability level for achieving reading proficiency on the MCAP ELA/L overall scale scores and the spring MAP-R RIT scores in Grade 3. The results are presented in Figures 1 to 8 and Tables 1 to 2 in the Appendix.

Spring DIBELS Composite Score Thresholds Prediction Accuracy for MCAP ELA/L

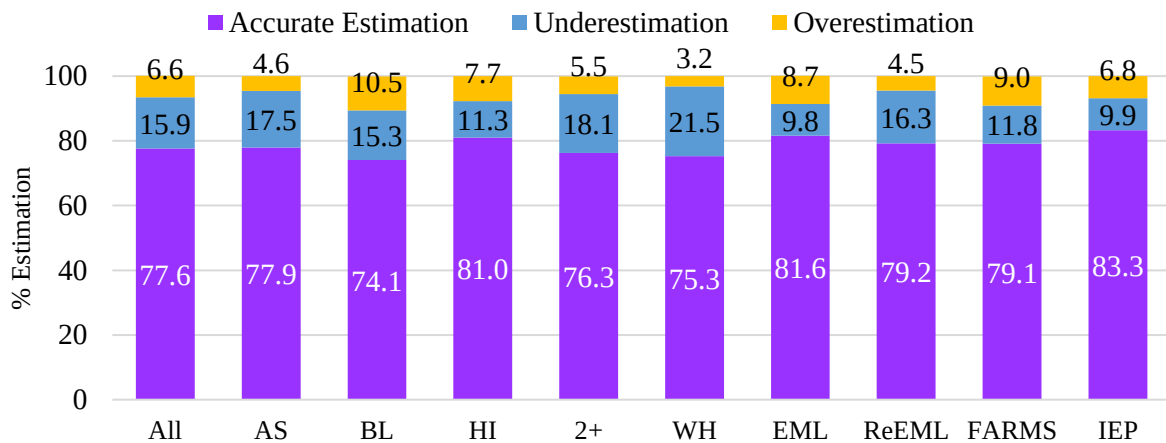
Figures 1 to 4 present results of prediction accuracy analyses for the Grade 2 spring DIBELS composite score threshold of 472 and the Grade 1 threshold of 489, respectively, which predicted reading outcomes measured by the Grade 3 MCAP ELA/L assessments (see the thresholds in Table 4). Table A1 in the Appendix provides the numbers of students in the linking samples and the prediction accuracy rates by student race/ethnicity and service receipt group.

Grade 2 Prediction Accuracy: The prediction accuracy rate was 77.6 percent for Grade 2 spring DIBELS (Figure 1). Specifically, using the identified Grade 2 spring DIBELS composite score threshold of 472, the reading outcome on the Grade 3 MCAP ELA/L was accurately predicted, either to be proficient or not, for more than three-quarters of students in the linking sample. Of all students, however, there were a substantial number of students, 15.9 percent, were underestimated, meaning the students scored below the spring DIBELS composite score threshold of 472 in Grade 2 yet still achieved reading proficiency on the Grade 3 MCAP ELA/L. Conversely, a much smaller number of all students, 6.6 percent, were overestimated, meaning they met or exceeded the spring DIBELS composite score threshold of 472 in Grade 2 but did not attain the predicted reading outcome on the Grade 3 MCAP ELA/L.

Using the Grade 2 spring DIBELS composite score threshold of 472, the accuracy of predicting Grade 3 reading proficiency on the MCAP ELA/L varied across racial/ethnic and service receipt groups. The prediction accuracy rate ranged from 74.1 percent for students identified as Black or African American to 83.3 percent for those with an Individualized Education Program (IEP) (Figure 1). Table A1 in the Appendix provides the numbers of Grade 2 students in the linking sample and the prediction accuracy rates by student race/ethnicity and service receipt group.

Figure 1

Prediction Accuracy for Grade 2 Spring DIBELS Composite Score Threshold Predicting 75-Percent Probability of Being Proficient in Reading on Grade 3 MCAP ELA/L

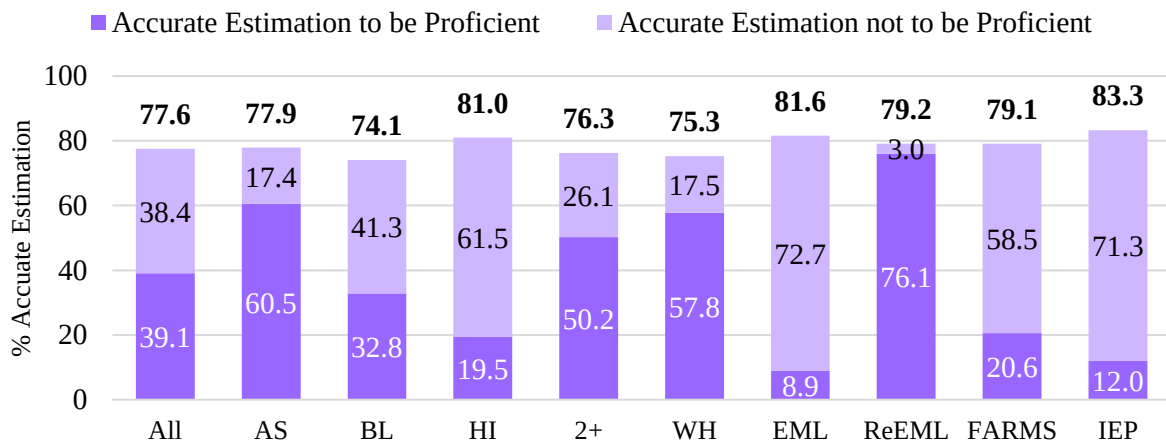


Notes. AS=Asian; BL=Black or African American; HI=Hispanic/Latino; 2+ =Two or More Races; WH=White; EML=Emergent Multilingual Learner; ReEM=Recently Exited Emergent Multilingual Learner; FARMS=Free and Reduced-price Meals System; IEP=Individualized Education Program. Students identified as American Indian or Alaskan Native and Native Hawaiian or Other Pacific Islander students are not reported separately due to small samples. The percentage of each category may not add up to total percentage due to rounding.

Grade 2 Accurate Prediction by Outcome: Of the students whose performance on the MCAP ELA/L was accurately predicted by the Grade 2 spring DIBELS composite score threshold of 472, performance was accurately predicted for 77.6 percent of all students (Figure 2). Of those with accurate prediction, 39.1 percent were correctly identified as likely to achieve reading proficiency on the spring MAP-R when the students were in Grade 3, while 38.4 percent were correctly foreseen not to attain proficiency (Figure 2). Figure 2 also presents accurate prediction rates by outcome across student groups. Accurate prediction for attainment of reading proficiency on Grade 3 MCAP ELA/L was more likely to occur for students identified as Asian, Two or More Races, White, and those who were identified as Recently-exited Emergent Multilingual Learner (ReEML) (ranging from 50.2% to 76.1%) than for other racial/ethnic and service receipt groups (ranging from 8.9% to 32.8%).

Figure 2

Accurate Prediction for Grade 2 Spring DIBELS Composite Score Threshold Predicting 75-Percent Probability of Being Proficient in Reading on Grade 3 MCAP ELA/L by Outcome

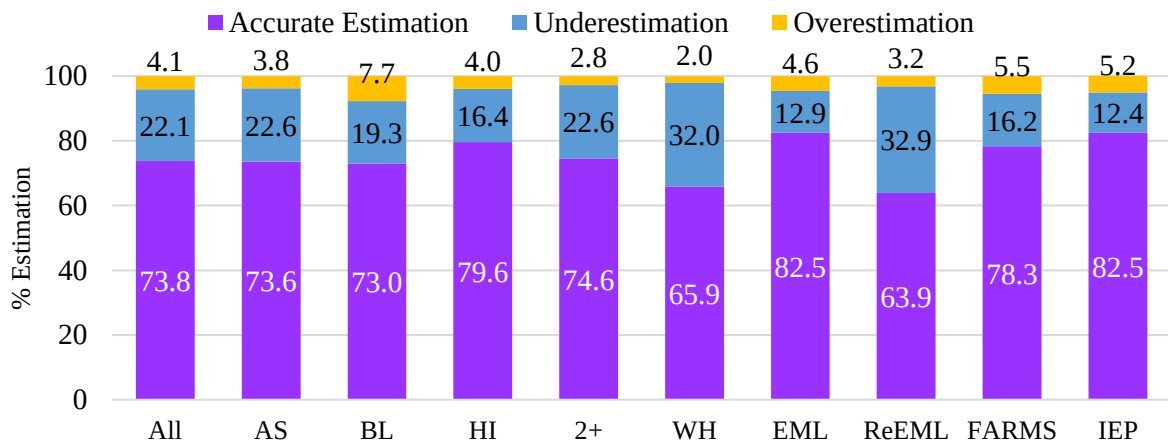


Grade 1 Prediction Accuracy: The prediction accuracy rate was 73.8 percent for Grade 1 DIBELS (Figure 3). Similarly, applying the Grade 1 spring DIBELS composite score threshold of 489 accurately predicted outcomes for nearly 74.0 percent of students in the linking sample. However, underestimation occurred for 22.1 percent of all students, referring to a substantial number of students who scored below the spring DIBELS composite score threshold of 489 in Grade 1 but achieved proficiency on the MCAP ELA/L in Grade 3. Conversely, a much smaller number of all students, 4.1 percent, were overestimated, meaning they met or exceeded the Grade 1 threshold of 489 yet did not attain reading proficiency on Grade 3 MCAP ELA/L.

When using the Grade 1 spring DIBELS composite score threshold of 489, prediction accuracy ranged from 63.9 percent for students who were identified as ReEML to 82.5 percent for those identified as Emergent Multilingual Learners (EML) and those with an IEP, as shown in Figure 3. Table A1 in the Appendix provides the numbers of Grade 1 students in the linking sample and the prediction accuracy rates by student race/ethnicity and service receipt group.

Figure 3

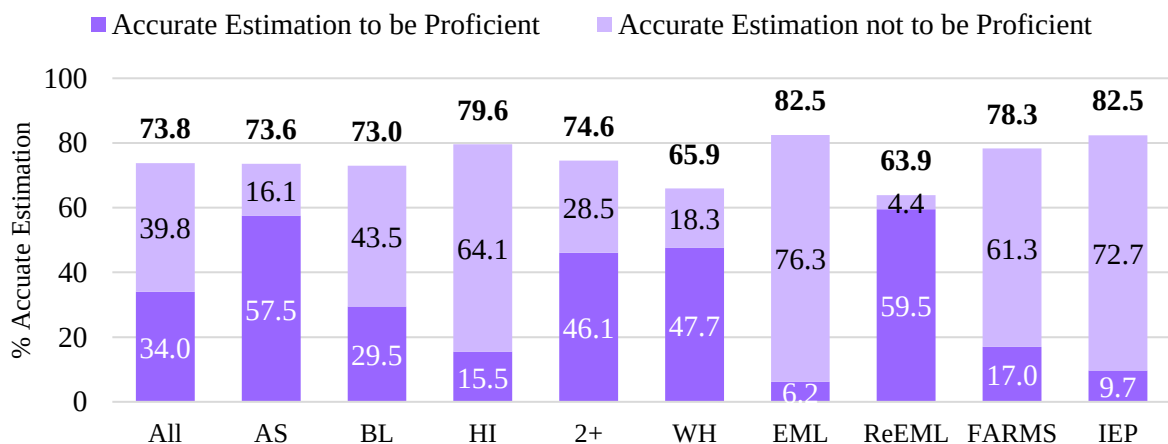
Prediction Accuracy for Grade 1 Spring DIBELS Composite Score Threshold Predicting 75-Percent Probability of Being Proficient in Reading on Grade 3 MCAP ELA/L



Grade 1 Accurate Prediction by Outcome: When applying the Grade 1 spring DIBELS composite score threshold of 489, performance was accurately predicted for 73.8 percent of all students, including 34.0 percent of those who were correctly projected to achieve proficiency on the Grade 3 MCAP ELA/L, and 39.8 percent accurately identified as not reaching proficiency (Figure 4). Figure 4 also presents accurate prediction rates by outcome across student groups. Similar to Grade 2, accurate prediction in Grade 1 for attainment of reading proficiency on Grade 3 MCAP ELA/L was more likely to occur for students who were identified as Asian, Two or More Races, White, and ReEML (ranging from 46.1% to 59.5%) than for other racial/ethnic and service receipt groups (ranging from 6.2% to 29.5%).

Figure 4

Accurate Prediction for Grade 1 Spring DIBELS Composite Score Threshold Predicting 75-Percent Probability of Being Proficient in Reading on Grade 3 MCAP ELA/L by Outcome



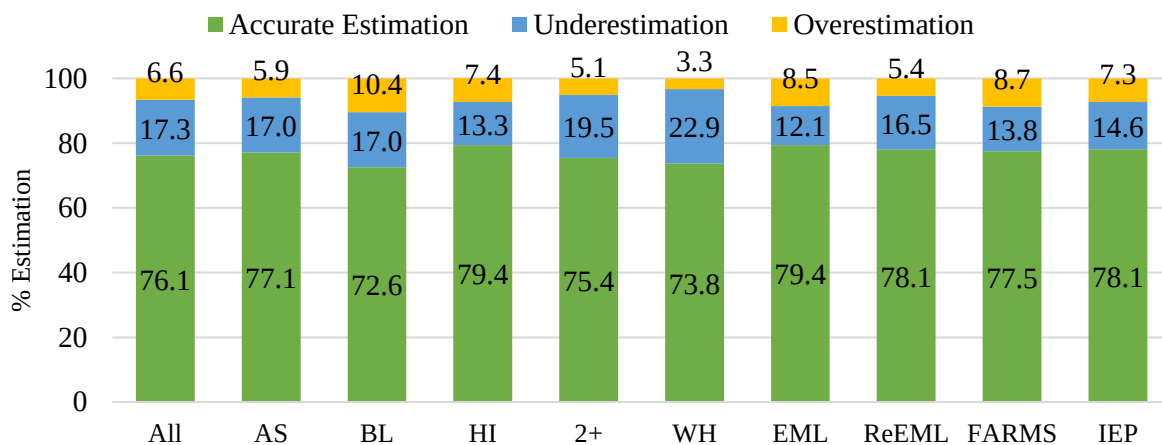
Spring DIBELS Composite Score Thresholds Prediction Accuracy for Spring MAP-R

Figures 5 to 8 present results of prediction accuracy analyses for the Grade 2 spring DIBELS composite score threshold of 472 and the Grade 1 threshold of 490, respectively, which predicted reading outcomes measured by the Grade 3 spring MAP-R assessments (see the thresholds in Table 4). The overall prediction accuracy rates and their variation across racial/ethnic groups and service receipt categories were similar to the rates obtained from the parallel threshold analysis for the Grade 3 MCAP ELA/L. Table A2 in the Appendix provides the numbers of students in the linking samples and the prediction accuracy rates by student race/ethnicity and service receipt group.

Grade 2 Prediction Accuracy: The prediction accuracy rate was 76.1 percent for Grade 2 spring DIBELS (Figure 5). Specifically, using the identified Grade 2 spring DIBELS composite score threshold of 472, the reading outcome on the Grade 3 spring MAP-R was accurately predicted, either to be proficient or not, for more than three-quarters of students in the linking sample. Of all students, however, there were a substantial number of students, 17.3 percent, were underestimated, meaning the students scored below the spring DIBELS composite score threshold of 472 in Grade 2 yet still achieved reading proficiency on the Grade 3 spring MAP-R. Conversely, a much smaller number of all students, 6.6 percent, were overestimated, meaning they met or exceeded the spring DIBELS composite score threshold of 472 in Grade 2 but did not attain the predicted reading outcome on the Grade 3 spring MAP-R.

Using the Grade 2 spring DIBELS composite score threshold of 472, the accuracy of predicting Grade 3 reading proficiency on the spring MAP-R varied across racial/ethnic and service receipt groups. The prediction accuracy rate ranged from 72.6 percent for students identified as Black or African American to 79.4 percent for those who were identified as Hispanic/Latino and EML (Figure 5). Table A2 in the Appendix provides the numbers of Grade 2 students in the linking sample and the prediction accuracy rates by student race/ethnicity and service receipt group.

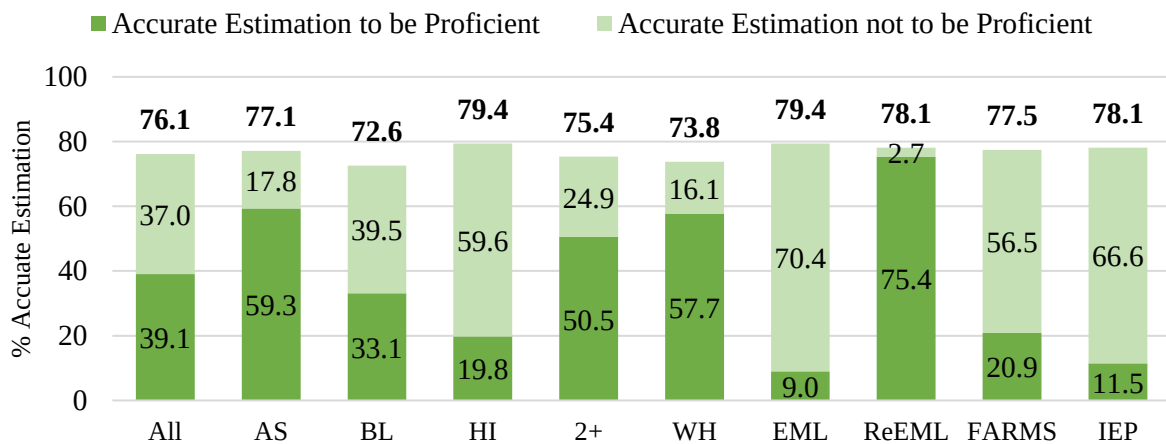
Figure 5
Prediction Accuracy for Grade 2 Spring DIBELS Composite Score Threshold Predicting 75-Percent Probability of Being Proficient in Reading on Grade 3 Spring MAP-R



Grade 2 Accurate Prediction by Outcome: Of the students whose performance on spring MAP-R was accurately predicted by the Grade 2 spring DIBELS composite score threshold of 472, performance was accurately predicted for 76.1 percent of all students (Figure 6). Of those with accurate prediction, 39.1 percent were correctly identified as likely to achieve reading proficiency on the spring MAP-R when the students were in Grade 3, while 37.0 percent were correctly foreseen not to attain proficiency. Figure 6 also presents accurate prediction rates by outcome across student groups. Accurate prediction for attainment of reading proficiency on Grade 3 spring MAP-R was more likely to occur for students who were identified as Asian, Two or More Races, White, and ReEML (ranging from 50.5% to 75.4%) than for other racial/ethnic and service receipt groups (ranging from 9.0% to 33.1%).

Figure 6

Accurate Prediction for Grade 2 Spring DIBELS Composite Score Threshold Predicting 75-Percent Probability of Being Proficient in Reading on Grade 3 Spring MAP-R by Outcome

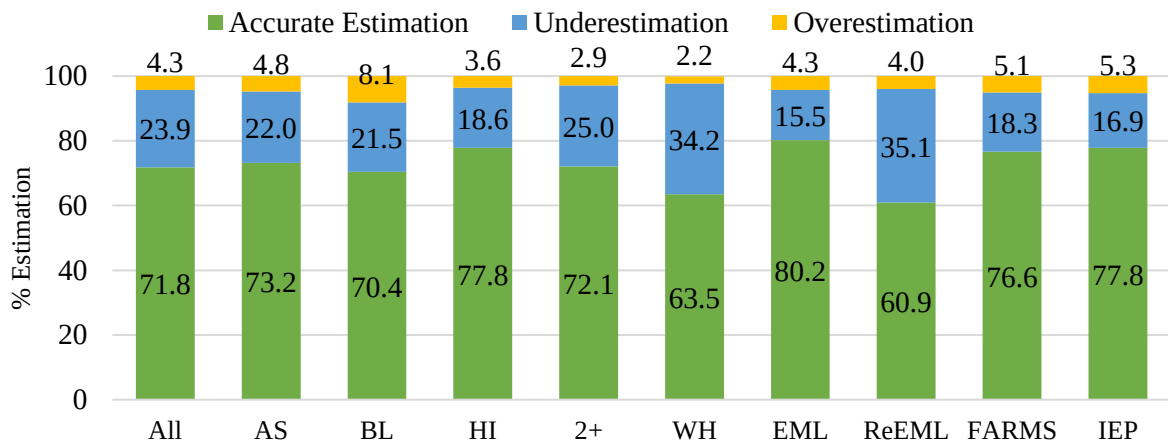


Grade 1 Prediction Accuracy: The prediction accuracy rate was 71.8 percent for Grade 1 DIBELS (Figure 7). Similarly, applying the Grade 1 spring DIBELS composite score threshold of 490 accurately predicted outcomes for nearly 72.0 percent of students in the linking sample. However, underestimation occurred for 23.9 percent of all students, referring to a substantial number of students who scored below the spring DIBELS composite score threshold of 490 in Grade 1 but achieved proficiency on the spring MAP-R in Grade 3. Conversely, a much smaller number of all students, 4.3 percent, were overestimated, meaning they met or exceeded the Grade 1 threshold of 490 yet did not attain reading proficiency on Grade 3 spring MAP-R.

When using the Grade 1 spring DIBELS composite score threshold of 490, prediction accuracy ranged from 63.5 percent for students identified as White to 80.2 percent for those identified as EML as shown in Figure 7. Table A2 in the Appendix provides the numbers of Grade 1 students in the linking sample and the prediction accuracy rates by student race/ethnicity and service receipt group.

Figure 7

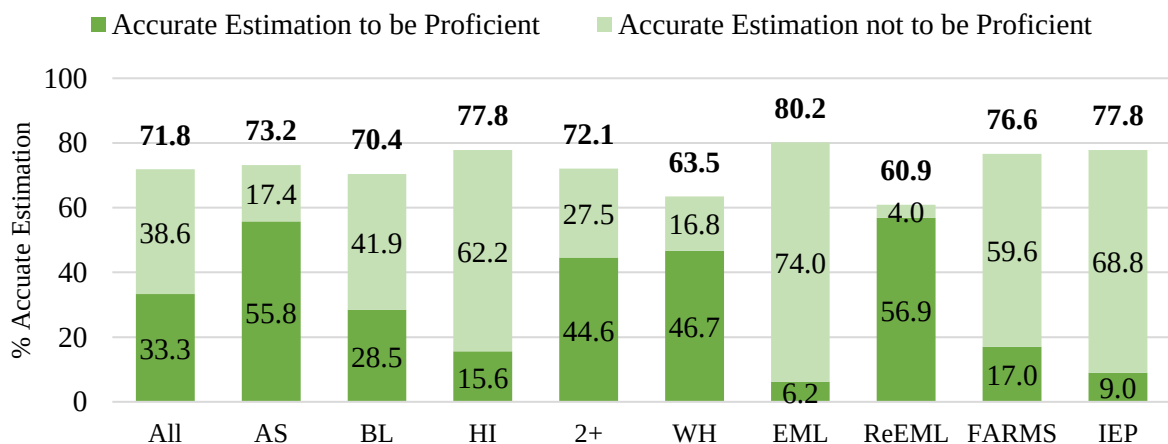
Prediction Accuracy for Grade 1 Spring DIBELS Composite Score Threshold Predicting 75-Percent Probability of Being Proficient in Reading on Grade 3 Spring MAP-R



Grade 1 Accurate Prediction by Outcome: When applying the Grade 1 spring DIBELS composite score threshold of 490, performance was accurately predicted for 71.8 percent of all students, including 33.3 percent of those who were correctly projected to achieve proficiency on the Grade 3 spring MAP-R, and 38.6 percent accurately identified as not reaching proficiency (Figure 8). Figure 8 also presents accurate prediction rates by outcome across student groups. Similar to Grade 2, accurate prediction in Grade 1 for attainment of reading proficiency on Grade 3 spring MAP-R was more likely to occur for students who were identified as Asian, Two or More Races, White, and ReEML (ranging from 44.6% to 56.9%) than for other racial/ethnic and service receipt groups (ranging from 6.2% to 28.5%).

Figure 8

Accurate Prediction for Grade 1 Spring DIBELS Composite Score Threshold Predicting 75-Percent Probability of Being Proficient in Reading on Grade 3 Spring MAP-R by Outcome



Conclusions and Implications

Discussion of Results

The magnitude of the correlation coefficients, ranging from .70 to .74, observed between spring DIBELS composite scores in Grades 1 and 2 and MCAP ELA/L overall scale scores and spring MAP-R RIT scores in Grade 3 provided predictive validity evidence for the threshold estimate analysis. The thresholds identified for the spring DIBELS composite scores in Grades 1 and 2, which were associated with the high probability (75% and 80%) of achieving reading proficiency as measured by the MCAP ELA/L and spring MAP-R, were identical or highly similar for the two outcome measures. These consistent results corroborated across outcome measures and probability levels validated the spring DIBELS composite score threshold estimates for Grade 3 reading proficiency. The consistency also strengthened the practical applicability of the established thresholds for early prediction of reading outcomes. The alignment between the threshold estimates for MCAP ELA/L and spring MAP-R may be attributed to the rigorous analytical methodology employed in this study and to the very strong concurrent relationship between the two assessments (Wang, 2023).

Among students whose performance was inaccurately predicted, underestimation was more common than overestimation, indicating that the spring DIBELS composite score thresholds set at a 75-percent probability level were reasonably rigorous. In practical terms, these thresholds were set high enough to maintain a relatively low rate of overestimating reading proficiency on the Grade 3 MCAP ELA/L and spring MAP-R assessments.

The consistent results of spring DIBELS composite score threshold estimates for Grade 3 MCAP ELA/L and spring MAP-R made it possible to recommend single spring DIBELS composite score thresholds in Grade 1 and Grade 2, respectively. The final spring DIBELS composite score thresholds were then determined based on analyses identifying scores associated with a 75-percent probability of achieving reading proficiency by the end of Grade 3. In practice, the following thresholds were recommended for predicting Grade 3 reading outcomes:

- For Grade 2, a single spring DIBELS composite score threshold of 472 was recommended to use, as it was identical for both the MCAP ELA/L and spring MAP-R outcomes (see Table 4).
- For Grade 1, the spring DIBELS composite score thresholds were 489 for MCAP ELA/L and 490 for MAP-R; the more conservative score of 489 was recommended for use across both outcomes (see Table 4).

The recommended spring DIBELS composite score thresholds indicated that a student who met or exceeded the scores in spring DIBELS at a primary grade would have a 75-percent probability to become reading proficient on the MCAP ELA/L and spring MAP-R in Grade 3. Otherwise, the student would be indicated in need of additional support or intervention to improve his/her ability in reading. The thresholds were estimated based on group performance which did not exactly equal to estimates for individual students. In other words, an expected score for a particular student can be slightly higher or lower than the threshold due to estimation error.

Discussion of Prediction Accuracy

The analysis revealed moderately strong predictive relationships between spring DIBELS composite scores in Grades 1 and 2 and the subsequent Grade 3 MCAP ELA/L overall scale scores and spring MAP-R RIT scores. While these relationships were meaningful, they helped explain the moderate prediction accuracy rates (ranging from 71.8% to 77.6%) and the correspondingly high inaccuracy rates (22.5% to 28.2%). A notable portion of this inaccuracy stemmed from underestimation, where a significant number of students scored below the spring DIBELS composite score thresholds yet still achieved reading proficiency in Grade 3. Specifically, approximately 16.0-to-17.0 percent of students in Grade 2 and 22.0-to-24.0 percent of students in Grade 1 were not projected to meet proficiency standards but actually did so. This level of underestimation may limit confidence in using these thresholds alone to monitor student progress and achievement in reading. It would be more beneficial to provide support to students based on more comprehensive performance data.

The limited prediction accuracy of the identified DIBELS composite score thresholds may reflect potential differences between the assessments of interest in terms of test purposes and contents. This can be clarified by examining the differences between the assessments in terms of test purpose, content, etc. DIBELS, administered in primary grades, measures foundational skills like word decoding. However, as noted in recent Amplify research (Aldeman, 2025), mastering decoding does not guarantee reading comprehension achievement in Grade 3 or above, which depends heavily on background knowledge and meaning-making. In contrast, district and state assessments like the Grade 3 MCAP ELA/L and MAP-R are designed to measure higher-order, content-relevant skills (e.g., identifying main ideas and author's message), but do not evaluate much foundational decoding. This fundamental gap in what is measured, foundational skills versus content-based comprehension, likely explains both the moderately strong correlations between the assessments of interest and therefore the moderate predictive accuracy of DIBELS in primary grades for reading outcome in later grades.

Application of the study

- Predictive thresholds were established for the spring DIBELS composite scores earned in Grades 1 and 2, corresponding to the high probabilities (75% and 80%) of reading proficiency on the Grade 3 MCAP ELA/L and spring MAP-R criteria. The 75-percent probability thresholds, 472 for Grade 2 DIBELS and 489 for Grade 1 DIBELS, were recommended for use due to their relatively high rates of accurate prediction to the 80-percent thresholds.
- Evidence obtained from the study supported the use of spring DIBELS composite scores in Grades 1 and 2 to predict reading outcomes on Grade 3 MCAP ELA/L and spring MAP-R assessments. As a starting point, school administrators and teachers should be provided with support in understanding how to interpret and use the predictive data. The study result gives a clear guide to identify students in Grades 1 and 2 who are on track to meet grade-level reading proficiency expectations, as well as those scoring below the identified spring DIBELS composite score thresholds in Grades 1 and 2 who may benefit from early support and target reading intervention.

- Given the moderate predictive accuracy of the assessments, the identified spring DIBELS composite score thresholds for Grades 1 and 2 should be used in conjunction with other measures, such as course performance, classroom observation of motivation, and test skills, when making predictions about Grade 3 reading outcomes on MCAP ELA/L and spring MAP-R.
- Under the forthcoming Maryland State Board of Education PreK-3 literacy policy, Grade 3 students who do not demonstrate reading proficiency may be considered for grade retention with parent/guardian consent. To align with these state mandates, this study focused on the early identification of literacy performance in the primary grades. By establishing reliable predictive thresholds on the early reading scores, this research facilitates targeted early intervention for at-risk students in order to reduce the likelihood of retention.

References

- Aldeman, Chad (2025). *Are kids making progress in reading? It all depends on how you measure it.* Amplify Education Company. <https://www.the74million.org/article/are-kids-making-progress-in-reading-it-all-depends-on-how-you-measure-it/>
- Amplify (2020). *What is DIBELS?* Amplify Education Company. <https://dibels.uoregon.edu/>
- Dorans, N. J., Moses, T. P. & Eignor, D. R. (2010). *Principles and practices of test score equating.* Educational Testing Services, Princeton: New Jersey.
- Maryland State Board of Education (2024). *Comprehensive PreK-3 Literacy Policy.* Maryland State Board of Education. <https://marylandpublicschools.org/programs/Documents/ELA/PreK-Literacy-Policy-Fall-Final-a.pdf>.
- Wang, Helen, (2023). *Linking MAP and MCAP assessments for Grades 3–8 students.* Montgomery County Public Schools, Rockville, MD. [MCPS MAP-MCAP Linking Study](#).

Appendix

Table A1

Prediction Accuracy for Grade 2 and Grade 1 Spring DIBELS Composite Score Thresholds Predicting 75-Percent Probability of Being Proficient in Reading on Grade 3 MCAP ELA/L

MCAP ELA/L		N Students	% Accurate Estimation	Accurate Estimation		% Under-estimation	% Over-estimation
				% Accurately Predicted to be Proficient in Reading	% Accurately Predicted Not to be Proficient in Reading		
Grade 2	All	10,621	77.6	39.1	38.4	15.9	6.6
	Asian	1,457	77.9	60.5	17.4	17.5	4.6
	Black or African American	2,088	74.1	32.8	41.3	15.3	10.5
	Hispanic/Latino	3,781	81.0	19.5	61.5	11.3	7.7
	Two or More Races	651	76.3	50.2	26.1	18.1	5.5
	White	2,613	75.3	57.8	17.5	21.5	3.2
	EML	2,920	81.6	8.9	72.7	9.8	8.7
	ReEML	264	79.2	76.1	3.0	16.3	4.5
	FARMS	4,677	79.1	20.6	58.5	11.8	9.0
	IEP	1,475	83.3	12.0	71.3	9.9	6.8
Grade 1	All	9,871	73.8	34.0	39.8	22.1	4.1
	Asian	1,335	73.6	57.5	16.1	22.6	3.8
	Black or African American	1,921	73.0	29.5	43.5	19.3	7.7
	Hispanic/Latino	3,497	79.6	15.5	64.1	16.4	4.0
	Two or More Races	614	74.6	46.1	28.5	22.6	2.8
	White	2,475	65.9	47.7	18.3	32.0	2.0
	EML	2,577	82.5	6.2	76.3	12.9	4.6
	ReEML	252	63.9	59.5	4.4	32.9	3.2
	FARMS	4,314	78.3	17.0	61.3	16.2	5.5
	IEP	1,416	82.5	9.7	72.7	12.4	5.2

Notes. EML=Emergent Multilingual Learner; ReEML=Recently Exited Emergent Multilingual Learner; FARMS=Free and Reduced-price Meals System; IEP=Individualized Education Program. Students identified as American Indian or Alaskan Native and Native Hawaiian or Other Pacific Islander students are not reported separately due to small samples. Including students with both of the DIBELS data in Grade 2/Grade 1 and the MCAP ELA/L data in Grade 3. Accurate estimation means a student met or did not meet the proficiency standard on MCAP ELA/L in Grade 3, as predicted by spring DIBELS in Grade 2/Grade 1. Underestimation means a student was observed proficient on MCAP ELA/L in Grade 3 while scoring below the established spring DIBELS composite score threshold in Grade 2/Grade 1. Overestimation means a student was not observed proficient on MCAP ELA/L in Grade 3 while scoring above the identified spring DIBELS composite score threshold in Grade 2/Grade 1. The percentage of each category may not add up to total percentage due to rounding.

Table A2

Prediction Accuracy for Grade 2 and Grade 1 Spring DIBELS Composite Score Thresholds Predicting 75-Percent Probability of Being Proficient in Reading on Grade 3 Spring MAP-R

Spring MAP-R		N Students	% Accurate Estimation	Accurate Estimation		% Under- estimation	% Over- estimation
				% Accurately Predicted to be Proficient in Reading	% Accurately Predicted Not to be Proficient in Reading		
Grade 2	All	10,581	76.1	39.1	37.0	17.3	6.6
	Asian	1,448	77.1	59.3	17.8	17.0	5.9
	Black or African American	2,079	72.6	33.1	39.5	17.0	10.4
	Hispanic/Latino	3,768	79.4	19.8	59.6	13.3	7.4
	Two or More Races	650	75.4	50.5	24.9	19.5	5.1
	White	2,606	73.8	57.7	16.1	22.9	3.3
	EML	2,908	79.4	9.0	70.4	12.1	8.5
	ReEML	260	78.1	75.4	2.7	16.5	5.4
	FARMS	4,655	77.5	20.9	56.5	13.8	8.7
	IEP	1,460	78.1	11.5	66.6	14.6	7.3
Grade 1	All	9,825	71.8	33.3	38.6	23.9	4.3
	Asian	1,325	73.2	55.8	17.4	22.0	4.8
	Black or African American	1,913	70.4	28.5	41.9	21.5	8.1
	Hispanic/Latino	3,478	77.8	15.6	62.2	18.6	3.6
	Two or More Races	612	72.1	44.6	27.5	25.0	2.9
	White	2,469	63.5	46.7	16.8	34.2	2.2
	EML	2,557	80.2	6.2	74.0	15.5	4.3
	ReEML	248	60.9	56.9	4.0	35.1	4.0
	FARMS	4,287	76.6	17.0	59.6	18.3	5.1
	IEP	1,400	77.8	9.0	68.8	16.9	5.3

Notes. EML=Emergent Multilingual Learner; ReEML=Recently Exited Emergent Multilingual Learner; FARMS=Free and Reduced-price Meals System; IEP=Individualized Education Program. Students identified as American Indian or Alaskan Native and Native Hawaiian or Other Pacific Islander students are not reported separately due to small samples. Including students with both of the DIBELS data in Grade 2/Grade 1 and the MAP-R data in Grade 3. Accurate estimation means a student met or did not meet the proficiency standard on spring MAP-R in Grade 3, as predicted by spring DIBELS in Grade 2/Grade 1. Underestimation means a student was observed proficient on spring MAP-R in Grade 3 while scoring below the established spring DIBELS composite score threshold in Grade 2/Grade 1. Overestimation means a student was not observed proficient on spring MAP-R in Grade 3 while scoring above the identified spring DIBELS composite score threshold in Grade 2/Grade 1. The percentage of each category may not add up to total percentage due to rounding.