Comparing Early Literacy Screening Assessment Benchmarks in Massachusetts

Mariann Lemke, Dan Murphy, Aaron Soo Ping Chow, Hayley Spencer, Angela Zhang
Fall 2023

Background

Like many states, Massachusetts has intensified efforts in recent years to improve literacy outcomes, particularly in the early grades. Initiatives have focused on providing guidance and resources for implementing changes in curriculum, instruction, and additional screening and support for struggling students. In September 2022, the Board of Elementary and Secondary Education in Massachusetts amended state law to require that all elementary schools assess each student's reading abilities and early literacy skills a minimum of twice per year from kindergarten through at least grade 3. The goal of such screening is to identify students who may be at risk of reading difficulty and proactively intervene to ensure that all students have the foundational skills needed to be successful readers. Beginning with the 2020/21 school year, the Massachusetts Department of Elementary and Secondary Education (DESE) began collecting literacy screening assessment data from schools and districts participating in certain state grants. The goal of collecting the data is to analyze patterns of performance to inform improvement efforts—for example, knowing when students struggle or how they progress over time may help schools, districts, and the state target resources and supports.

Grantee schools and districts that provided literacy screener data to DESE selected their screening assessments from a list of state-approved commercially available literacy screener products. These assessments vary in how they were designed, how they are administered, and how they define risk of reading difficulty, making it difficult to compare and summarize performance for students in different schools and districts. However, grade 3 students with literacy screening assessment scores do take one common assessment: the Massachusetts Comprehensive Assessment

Data Included in Analysis

- *More than 4,000 students from the 2021/22 school year, including 28 districts and 63 schools*

- *Scores from multiple literacy screening assessments: DIBELS 8th Edition, FastBridge aReading, i-Ready Diagnostic, Lexia Rapid, MAP Growth, mClass, Star Reading*

System (MCAS) English Language Arts (ELA) assessment. MCAS therefore provides a mechanism to compare literacy benchmarks using a common metric. This issue brief describes what it means to be "at risk" of reading difficulty as defined by each screening assessment and then provides results of a study that links screener benchmark scores to the MCAS grade 3 scale.[1]

Findings

There is no common definition of "at risk"—screening assessments provide multiple benchmarks or performance levels defined in different ways.

The state's goal in approving early literacy screening assessments was to help schools and districts choose technically sound tools to identify students at risk of reading difficulties so that support services could be provided to them. Most of the approved screening assessments, however, do not simply identify risk using a single cut score, instead providing several benchmarks or risk levels (e.g., low risk, some risk, high risk; see Table 1). These levels differ across assessments in how they were determined and in what they represent. For example, the cut scores for each level for a screening assessment may have been determined based on data about students' later performance on that screening assessment or other reading assessments later in the year or in later grades. Alternatively, benchmarks may have been set based on normative data—selecting a percentile, such as the 40th percentile, to identify the lowest performing group of students.[2] Many vendors use both types of information to set their cut points. Benchmark levels are generally, although not always, based on composite scores derived from all the specific reading subtests administered at each grade level.

Table 1. Description of Risk Levels for Early Literacy Screening Assessments Included in Linking Analysis

Early literacy screening assessment	Benchmark or risk-level descriptions
DIBELS 8th Edition	Provides four levels to describe student performance: • Above Benchmark: 65th–99th percentile • At Benchmark: 55th–64th percentile • Below Benchmark: 30th–54th percentile* • Well Below Benchmark: below 30th percentile*
mCLASS	See description for DIBELS 8th Edition (mClass assessments are based on DIBELS 8th Edition and reported performance levels are the same).

---

1   Results described in this issue brief are based on grade 3 data only. The full set of grantee screening assessment data includes more than 35,000 student records. See *A First Look at Early Literacy Performance in Massachusetts: Results of Initial Analysis Based on State Grantee Literacy Screening Assessments* for details.

2   The populations on which norms were based may also differ across assessments.

| Early literacy screening assessment	Benchmark or risk-level descriptions
FastBridge aReading	Provides four levels to describe student performance: • Advanced/College Pathway: 71st–99th percentile • Low Risk: 40th–70th percentile • Some Risk: 15th–39th percentile* • High Risk: below 15th percentile*
i-Ready Diagnostic	Provides five levels to describe student performance: • Mid or Above Grade Level • Early On Grade Level • 1 Grade Level Below* • 2 Grade Levels Below* • 3 or More Grade Levels Below*
Lexia RAPID	Provides three levels to describe student performance: • High Likelihood of Success: 70th–99th percentile • Moderate Likelihood of Success: 31st–69th percentile* • Low Likelihood of Success: below 31st percentile*
MAP Growth	Provides two levels to describe student performance: • No intensive intervention • Intensive intervention* Note that MAP Growth also provides performance levels (Not Meeting, Partially Meeting, Meeting, Exceeding) designed to describe students as on or off track to meeting MCAS proficiency standards.
Star Reading	Provides four levels to describe student performance: • At/Above Benchmark: 40th–99th percentile • On Watch: 25th–39th percentile* • Intervention: 10th–24th percentile* • Urgent Intervention: below 10th percentile*

Note. The source is the authors' compilation based on documentation for each assessment and/or communication with assessment publishers.

*The asterisk with red-colored text indicates benchmarks used to indicate any level of risk of reading difficulties.

Screening assessment benchmarks that indicate any level of risk map to MCAS ELA scores that range from the middle of the Partially Meeting Expectations performance level to the beginning of the Meeting Expectations performance level.

Results of an equipercentile linking analysis show that most benchmark scores that identify students at any level of risk (for example, the score that divides students who are "At Benchmark" or "Above Benchmark" from students who are "Below Benchmark" or "Well Below Benchmark" in DIBELS 8th Edition terms) fall in the MCAS Partially Meeting Expectations performance level, though there is some variability (see Figure 1).
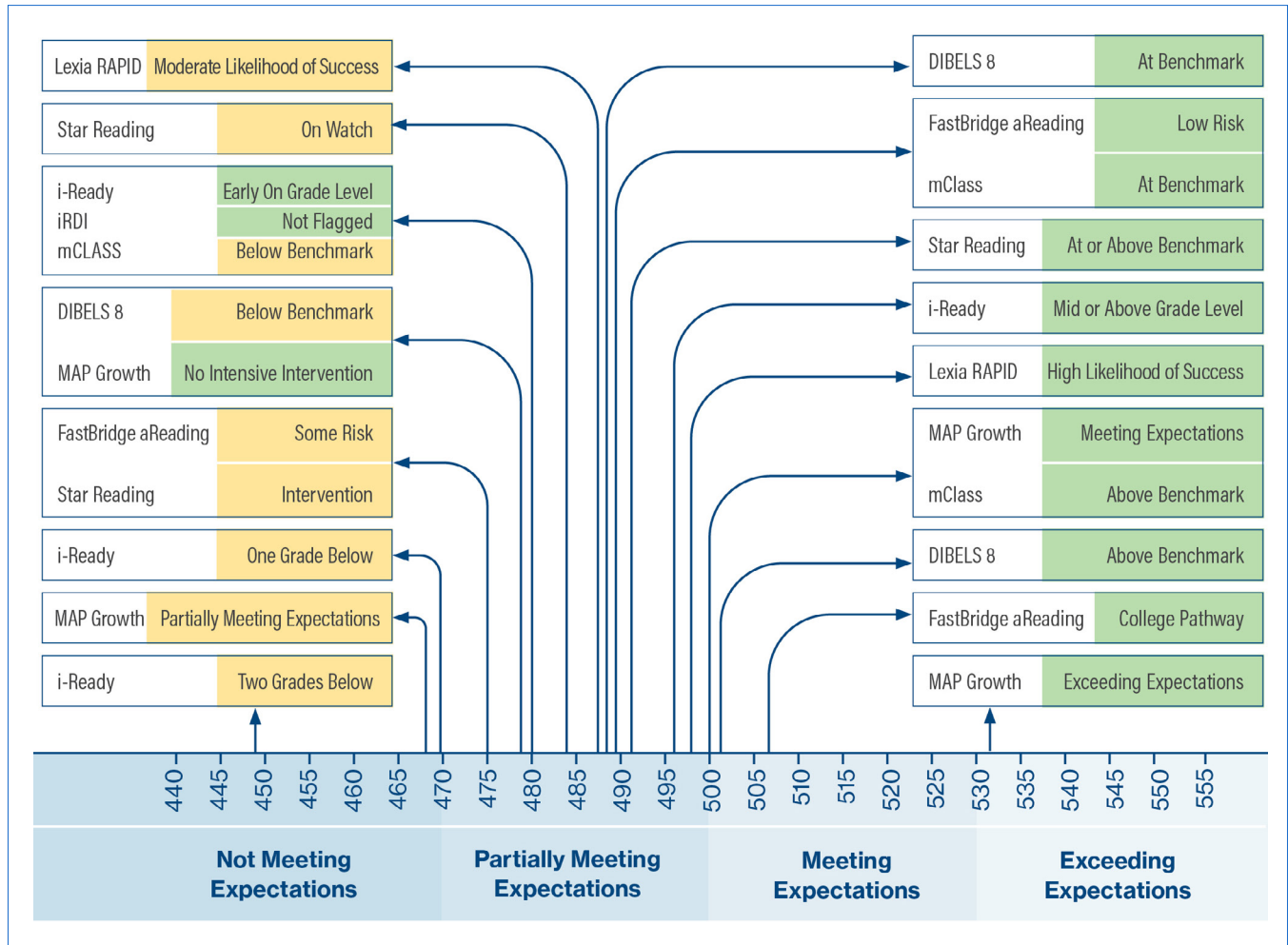
Figure 1. Grade 3 Screening Assessment Benchmark Scores That Distinguish Any Level of Risk From Little or No Risk on the MCAS Scale
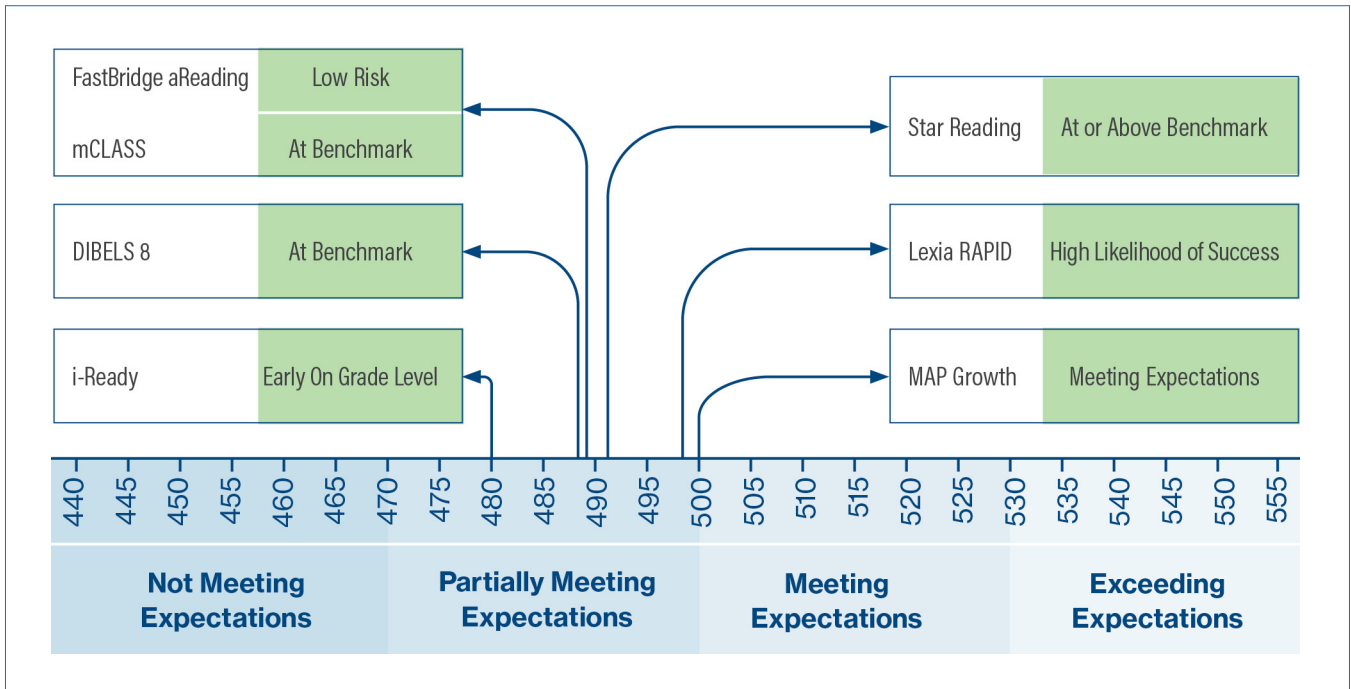
Mapping all benchmark scores to the MCAS scale illustrates a wide range of what "at risk" and benchmark-level performance can mean.

As described earlier, most of the approved screening assessments provide multiple benchmark scores that classify students into different levels of risk or performance. Figure 1 shows how benchmark scores that classify students into "at risk" or not compare; Figure 2 locates all screening assessment benchmarks on the MCAS scale and shows they cover a wide range of the scale from Not Meeting Expectations to Exceeding Expectations.

Figure 2. All Grade 3 Screening Assessment Benchmark Scores on the MCAS Scale

Literacy screening assessment benchmark scores that identify students at risk successfully predict which students will not meet MCAS expectations but not necessarily who will meet them.

Using a regression-based approach to link screening assessment benchmarks to the MCAS ELA scale suggests that early literacy screening assessments can be useful predictors of MCAS performance. Estimates of association, including correlations, classification accuracy, and area under the receiver operating characteristic (ROC) curve, indicate strong to very strong associations between the tests.

About 8 percent of students identified as at any level of risk on a literacy screening assessment at end of year (EOY) in grade 3 met or exceeded MCAS ELA standards, while about 60 percent who met screening assessment benchmarks met or exceeded expectations on MCAS (see Figure 3). Thus, about 40 percent of students who met literacy screening assessment benchmarks did not meet or only partially met MCAS expectations. These "false positive" cases are likely the result of differences in how screening assessments and MCAS define performance levels and the content and skills they assess. The MCAS "Meeting Expectations" level reflects a high performance expectation, with only about half of students across the state achieving it in 2021. Understanding that "on track" performance on screening assessments does not

necessarily indicate strong likelihood of scoring above 500 or "Meeting Expectations" on MCAS may be especially important for schools in light of accountability requirements that begin in grade 3.

Figure 3a. Percentage of Grade 3 Students At Risk on Literacy Screening Assessments by MCAS Performance Level
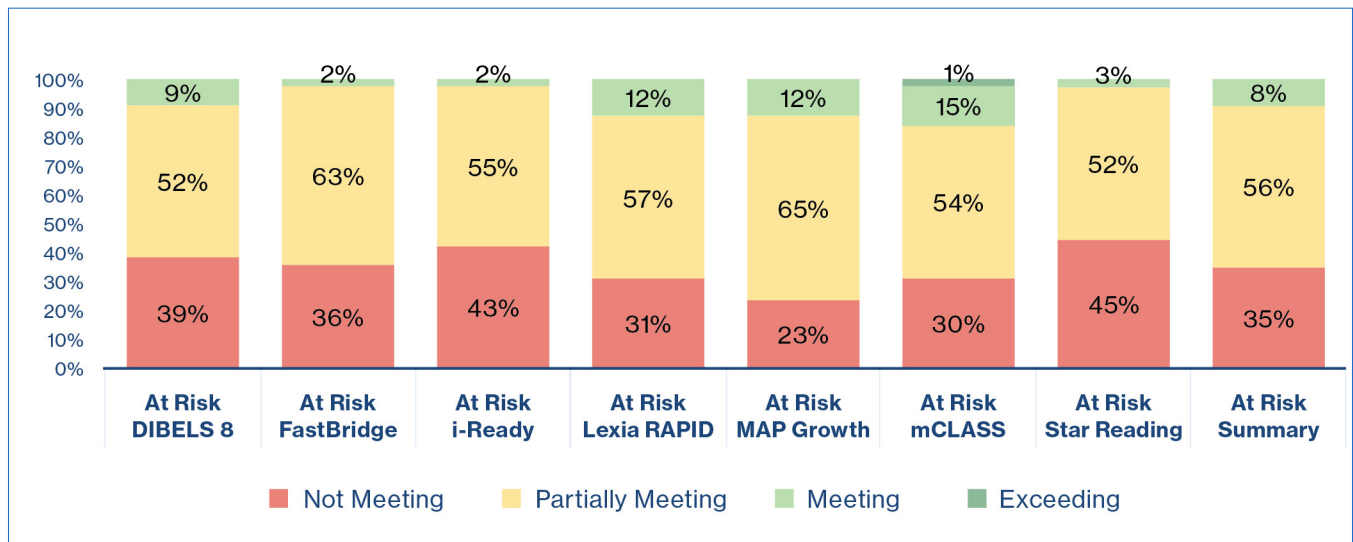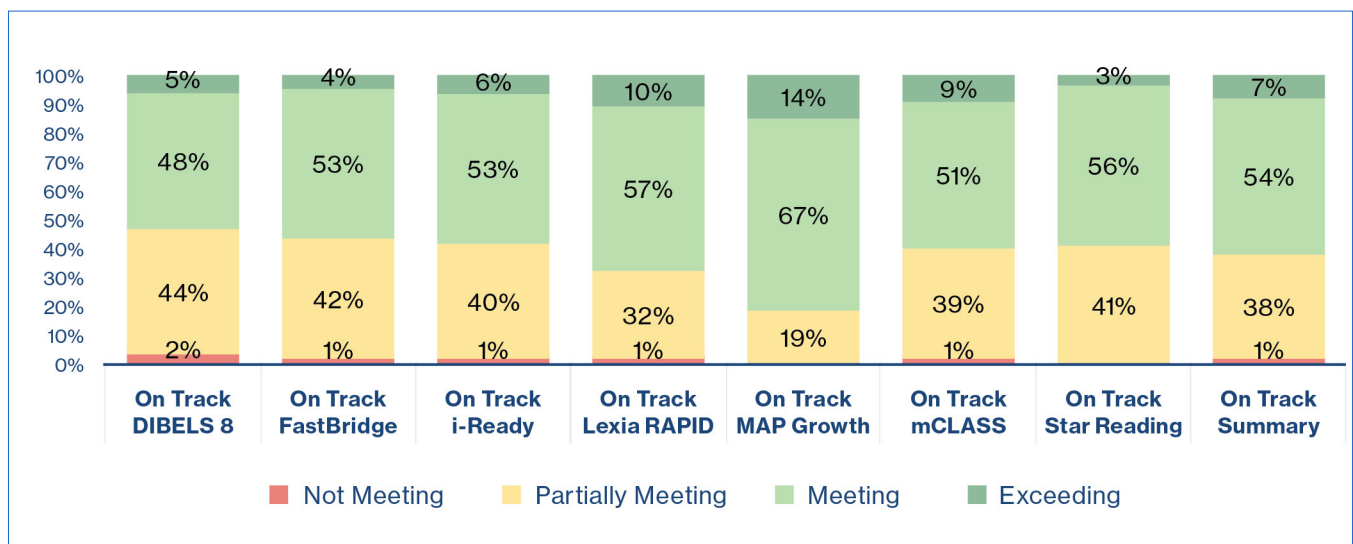

Chart showing percentage breakdown of At Risk students by assessment:

	At Risk DIBELS 8	At Risk FastBridge	At Risk i-Ready	At Risk Lexia RAPID	At Risk MAP Growth	At Risk mCLASS	At Risk Star Reading	At Risk Summary
Meeting/Exceeding (top)	9%	2%	2%	12%	12%	1% / 15%	3%	8%
Partially Meeting	52%	63%	55%	57%	65%	54%	52%	56%
Not Meeting	39%	36%	43%	31%	23%	30%	45%	35%

Legend: Not Meeting | Partially Meeting | Meeting | Exceeding

Figure 3b. Percentage of Grade 3 Students Not At Risk on Literacy Screening Assessments by MCAS Performance Level

Chart showing percentage breakdown of On Track students by assessment:

	On Track DIBELS 8	On Track FastBridge	On Track i-Ready	On Track Lexia RAPID	On Track MAP Growth	On Track mCLASS	On Track Star Reading	On Track Summary
Exceeding	5%	4%	6%	10%	14%	9%	3%	7%
Meeting	48%	53%	53%	57%	67%	51%	56%	54%
Partially Meeting	44%	42%	40%	32%	19%	39%	41%	38%
Not Meeting	2%	1%	1%	1%		1%		1%

Legend: Not Meeting | Partially Meeting | Meeting | Exceeding

Implications for Policy and Practice

Although data included in the analysis are somewhat limited, the results suggest that differences in benchmark scores and their relationship to MCAS may have potential implications for policy and practice.

- **Differences among benchmark scores that define risk within and across assessments mean that schools and districts must carefully consider their goals and resources in choosing how to use them to identify and address student needs.** Screening assessment benchmark scores link to a wide range on the MCAS scale and provide a wide range of benchmark performance options. Schools and districts must consider what they mean by "risk" and how that corresponds to available benchmark scores so that they can identify students who need support in a way that matches their goals and available resources and instructional supports. If the score that defines risk is too low, then students who may need additional support will not receive it. If the cut score is too high, then resources will be used unnecessarily to provide support for students who do not need it. The state may also wish to consider defining a common level to indicate risk across screening assessments.

- **Goals and local definitions of risk could be framed in MCAS or other terms.** If educators are interested primarily in designing instruction and intensive interventions geared toward students who are most in need of support, then benchmarks located near the Partially Meeting Expectations cut score (or even lower) could be useful at-risk indicators. On the other hand, if educators are interested in identifying students who need support to meet MCAS expectations, then screener assessment benchmarks located closer to the MCAS Meeting Expectations cut score (e.g., DIBELS 8th Edition or mClass Above Benchmark or i-Ready Mid or Above Grade Level) would be more useful indicators. Screening assessments could be used to identify students who need support to meet MCAS ELA goals. However, such definitions of risk would be more stringent and likely identify a larger proportion of the student population than was previously identified in younger grades.

WestEd
WestEd.org

WestEd, a nonpartisan research, development, and service agency, works with education and other communities to promote excellence, achieve equity, and improve learning for children, youth, and adults. WestEd has more than a dozen offices nationwide. More information about WestEd is available at WestEd.org.