Report on Options for Assessments Aligned with the Common Core State Standards

Submitted to the Michigan Legislature

December 1, 2013

MICHIGAN Department of Education

ACKNOWLEDGEMENTS

The completion of this report would not have been possible without the dedicated efforts of the Michigan Department of Education (MDE) staff listed below in alphabetical order.

Brandy Archer . . . . . Content Area Literacy Consultant, Office of Education Improvement and Innovation
Matt Ayotte . . . . . Online Assessment Specialist, Office of Systems, Psychometrics, and Measurement Research
Stephen Best . . . . . Assistant Director, Office of Education Improvement and Innovation
Phil Chase . . . . . Composition and Professional Development Manager, Office of Standards and Assessment
Deb Clemmons . . . . . Executive Director, School Reform Office
Doug Collier . . . . . Financial Manager, Office of Assessment Business Operations
Vince Dean . . . . . Director, Office of Standards and Assessment
Gregg Dionne . . . . . Supervisor, Curriculum and Instruction, Office of Education Improvement and Innovation
Jan Ellis . . . . . Communications and Process Specialist, Division of Accountability Services
Sue Fransted . . . . . Department Analyst, Office of Standards and Assessment
Linda Forward . . . . . Director, Office of Education Improvement and Innovation
Jill Griffin . . . . . Consultant, Urban Education, Grant Requirements, and Credit Recovery Office of Education Improvement and Innovation
Jim Griffiths . . . . . Test Adminstration Manager, Office of Standards and Assessment
Amy Henry . . . . . Consultant, Statewide System of Support, Office of Education Improvement and Innovation
Ruth Anne Hodges . . . . . Mathematics Consultant, Office of Education Improvement and Innovation
Linda Howley . . . . . Accessibility Specialist, Office of Standards and Assessment
Dave Judd . . . . . Director, Office of Systems, Psychometrics, and Measurement Research
Venessa Keesler . . . . . Deputy Superintendent, Division of Education Services
Pat King . . . . . Senior Project Management Specialist, Office of Systems, Psychometrics, and Measurement Research
Tom Korkoske . . . . . Financial Analyst, Office of Assessment Business Operations
Joseph Martineau . . . . . Deputy Superintendent, Division of Accountability Services
Kim Mathiot . . . . . Composition Editor, Office of Standards and Assessment
Andy Middlestead . . . . . Test Development Manager, Office of Standards and Assessment
Kim Mull . . . . . Graphic Designer, Office of Standards and Assessment
Jen Paul . . . . . English Learner Assessment Consultant, Office of Standards and Assessment
Karen Ruple . . . . . Statewide System of Support Manager Office of Education Improvement and Innovation
Paul Stemmer . . . . . National Assessment of Educational Progress State Coordinator Office of Standards and Assessment
Steve Viger . . . . . Measurement Research and Psychometrics Manager Office of Systems, Psychometrics, and Measurement Research
Shannon Vlassis . . . . . Graphic Designer, Office of Standards and Assessment
Kim Young . . . . . Interim Assessment Consultant, Secondary Level, Office of Standards and Assessment

In addition, MDE would like to thank Brian Gong, Executive Director of the Center for Assessment at the National Center for the Improvement of Educational Assessment, for providing an independent review of a draft of this report for the purpose of ensuring its clarity and usefulness.

Finally, MDE would like to express its sincere appreciation to all the service providers that responded to the abbreviated Request for Information (RFI) represented by the survey that forms the basis for this report's content.

TABLE OF CONTENTS

Introduction . . . . . 4 - 5

Summative Assessment. . . . . 6 - 21
- Content and Item Type Alignment . . . . . 6 - 7
- Transparency and Governance . . . . . 8 - 9
- Overall Design and Availability . . . . . 10 - 11
- Test Security . . . . . 12 - 13
- Scoring and Reporting . . . . . 14 - 15
- Cost - Standard Product. . . . . 16 - 17
- Constructed Response – Standard Product Cost Implications . . . . . 18 - 19

Interim Assessment . . . . . 20 - 33
- Content and Item Type Alignment . . . . . 20 - 21
- Transparency and Governance . . . . . 22 - 23
- Overall Design and Availability . . . . . 24 - 25
- Test Security . . . . . 26 - 27
- Scoring and Reporting . . . . . 28 - 29
- Cost - Standard Product. . . . . 30 - 31
- Constructed Response – Standard Product Cost Implications . . . . . 32 - 33

Accessibility. . . . . 34 - 35

Technical Requirements . . . . . 36 - 37

Formative Assessment Resources. . . . . 38 - 39

Local Implications . . . . . 40 - 41

Summary Conclusions and Recommendations . . . . . 42 - 43

References. 43

Appendices
- Appendix A - Survey
- Appendix B - Survey Responses
- Appendix C - Survey Questions Cross Reference Table

COMMON CORE Assessment Options Report

INTRODUCTION

Michigan students and educators need a rich, next-generation assessment system that is suitable for the numerous, high-stakes purposes toward which it will be applied. The solutions described in this report must be considered in light of how the test results will be used, and the fact that every school, educator and community will feel real consequences of their use, both intended and possibly unintended. Michigan's transition to new, online assessments that include multiple measures designed to capture student achievement and growth, is a powerful opportunity to improve the strength of our entire education system. This report represents an important source of information about the various options available to the state.

The Legislative Resolution

Both the House Concurrent Resolution 11 passed by the Michigan House of Representatives on September 26, 2013 and the substitute for House Concurrent Resolution 11 passed by the Senate on October 24, 2013 (subsequently adopted by the House on October 29, 2013) included a requirement for the State Board of Education and MDE to develop and submit a report on options for assessments fully aligned with the Common Core State Standards (CCSS). The report was to be completed and submitted to both chambers of the legislature by December 1, 2013 and be factual and unbiased. In addition, the final resolution expressed a preference for state assessments that are computer adaptive, provide real-time results, are able to be given twice per year, and assist in the evaluation of individual teachers. Other requirements included availability by the 2014-15 school year in grades 3 through 11.

In order to comply with the final resolution, the primary requirement for assessment solutions described in this report is that they be adequately aligned with Michigan's college- and career-ready standards, in this case the CCSS. Some aspects of alignment (e.g., coverage of the mathematics and reading standards) are relatively straightforward. Other facets are more challenging to capture and have far-reaching implications for categories such as cost. An example of this is the use of Constructed-Response (CR) items; test questions that require students to develop a short or long written response. These are often significantly better than other types of items for measuring complex skills such as research, problem solving or communicating reasoning, that are found in the CCSS. However, these types of items are often time-consuming for students to answer and are the most expensive and complicated to score. Because CR items have significant implications for a variety of categories presented in this report, references will be made to them in appropriate sections, and overall implications will be described in the summary conclusions and recommendations section.

MDE Request for Information Process

In order to complete this project by December 1, 2013, the decision was made to develop a survey covering the primary topics of concern and permit any vendor registered to do business in Michigan through the state's Buy4Michigan web system to respond. Development of the survey commenced immediately following the approval of the Senate resolution on October 24, when it was apparent that the final resolution was highly likely to require this report. Through the Buy4Michigan website, 185 entities are registered under the category of educational examination and testing services. The survey was posted to the site; each registered entity received a notification email indicating that an opportunity was available for them on October 30, and indicated that all replies were due in two weeks. Twelve service providers submitted responses, all of which were included in the report.

The survey questions were separated into three distinct categories, to capture information on the three primary types of assessment solutions that are essential elements of a balanced assessment system needed to support educational improvement. The goal of this was to learn what solutions were available or being developed for:

- **Summative purposes** (e.g., test like MEAP for high-stakes accountability)
- **Interim purposes** (e.g., tests administered multiple times throughout the year to measure student growth with relative frequency), and
- **Formative purposes** (e.g., resources to support real-time measurement of student learning).

It was also important to ask about these different classes to ensure that no service provider was excluded for having a viable solution for one product in any category, versus requiring that each vendor have something for all three categories. MDE is open to the idea that the strongest overall solution in the end may involve selecting the 'best in class' for each type, although this concept introduces substantial risk on aspects such as the comparability of test scores.

Responding to the extensive survey in two weeks was undoubtedly a challenging task for service providers, as the questions were detailed and covered a wide range of topics. MDE is appreciative that so many qualified teams made the time to submit complete responses. One service provider, the Northwest Evaluation Association (NWEA), which currently has products deployed in a number of Michigan schools, chose not to complete the survey and instead submitted a letter explaining some aspects of their assessments and why they elected to not complete the survey. Since they did not submit a full response, NWEA is not included in the report. However, as they were the only vendor to submit such a letter, and many Michigan stakeholders are familiar with what they have to offer, MDE felt it was appropriate to include their letter in Appendix B with the completed survey information from the other entities. The table below provides a summary of the report development schedule.

Report Development Milestones	
Senate passes resolution and survey development begins	October 24, 2013
Survey posted to Buy4Michigan website	October 30, 2013
Responses due from service providers	November 13, 2013
Report submitted to Legislature and State Board of Education	December 1, 2013

Organization and Composition of the Report

Once responses were received, MDE staff members needed to review them all, compile them by category, and assign ratings. In order to complete this task, teams of staff with relevant subject matter expertise were assigned to each category with explicit instructions on how to view the responses and assign ratings. It was determined that a 'Consumer Reports' type of display would be the most user-friendly. The tables displayed in the body of the report provide a snapshot of how each service provider completed the questions germane to each category. There are three important caveats about the ratings assigned:

- Due to the timeline, it was not possible to thoroughly evaluate the quality of evidence provided by each service provider. The highest rating is based on complete responses that included some evidence indicating they were likely to meet all requirements, the middle rating indicating unclear or partial meeting of requirements, etc. Therefore, development and rigorous vetting of scoring criteria could not be accommodated. Additionally, the decision was made to limit the number of rating categories to three, to help ensure that even if a longer timeline had been available and a more rigorous, fine-grained (e.g., 5 or 7 categories) scoring system developed, only minor changes in scoring would have likely resulted.
- Responses from service providers were not compared against each other, only against the content of the survey questions. Comparing responses across multiple survey questions related to each category would have required substantially more time in order to evaluate the quality of the response and accompanying evidence.
- It is important to remember that many of the solutions described in this report are under construction, so a true evaluation of their qualities will not be possible until after the first year of operational test administration.

Based on these caveats, it is **essential** to recognize that this report alone is not sufficient to determine which assessments would truly be viable with regard to measuring the full breadth of career- and college-ready standards, interfacing with state systems, not adding additional burdens to local districts and schools, and cost effectiveness.

A conscious decision was made not to consolidate the ratings for each category into an overall Executive Summary. This process would have diluted the responses provided by each service provider by not properly accounting for the many areas where solutions are partially available or in various stages of development. Based on this, each category should be reviewed on its own merits and given equal weight. Additionally, in a number of cases, the survey responses required a 'yes' or 'no' response, but the opportunity to provide comments for the purpose of further clarification was made available. This introduced nuances, or possible opportunities for negotiation in areas such as control over data or opportunities to have Michigan educators involved with test question development, that could not be captured equitably in each section's table or narrative. The survey responses from each service provider are included in their entirety, unaltered, in Appendix B, if any readers of this report are interested in exploring the comments that accompanied some responses.

In addition to the specific items listed in the final resolution, four key documents guided the development of the survey questions and helped shape the lenses through which the responses were viewed by Department staff. Two of the documents, the *Standards for Educational and Psychological Testing* and the *Standards and Assessments Peer Review Guidance* have been important sources of requirements for technical quality and ensuring that all state standards and assessment systems meet criteria specified under the federal Elementary and Secondary Education Act. Recently, two other documents have been produced to guide the development of high quality, next-generation assessments and thoroughly define the requirements and responsibilities of clients (e.g., states) and service providers in all aspects of bringing large-scale assessment programs to operational status. Respectively, these are the *CCSSO Assessment Quality Principles* and the *Operational Best Practices for Statewide Large-Scale Assessment Programs-2013 Edition.*

Content & Item Type Alignment

SUMMATIVE

Service Provider	Content Alignment				Item Types		
	Content aligned to the CCSS	*Solution addresses all 5 content areas*	*Solution addresses all grade levels (G3-G11)*	*Qualifications for educators involved in alignment for content, diversity and special populations*	*Standard item types (multiple choice and constructed response) will be available*	*Diverse set of technology-enhanced item types will be available*	*Performance tasks/ assessments will be available*
ACT Aspire	◐	○	●	○	●	●	○
Amplify Education, Inc.	NR	NR	NR	NR	NR	NR	NR
College Board	○	○	○	○	●	○	○
CTB/McGraw-Hill	◐	●	●	○	●	●	●
Curriculum Associates LLC	◐	○	●	○	●	◐	○
Discovery Education Assessment	NR	NR	NR	NR	NR	NR	NR
Houghton Mifflin Harcourt/Riverside	○	○	●	○	●	◐	◐
Measured Progress	●	●	○	◐	NR	NR	●
PARCC	●	●	●	◐	●	●	●
Scantron	NR	NR	NR	NR	NR	NR	NR
Smarter Balanced	●	◐	●	◐	●	●	●
Triumph Learning	NR	NR	NR	NR	NR	NR	NR

INTRODUCTION

The Common Core State Standards are organized into five content areas: Mathematics, Reading, Writing, Listening and Speaking. They provide goals and benchmarks to ensure that students are achieving certain skills and knowledge by the end of each year. They were carefully written so that students leave high school with a deep understanding of the content and skills they need to be career- and college-ready. It is important, then, that the summative assessments accurately reflect the intended content emphasis and important understandings of each grade level, 3–8, and high school.

Multiple-choice and technology-enhanced item types are critical components of an assessment, and in order to truly assess the rigor and higher-order thinking skills required from the CCSS and career and college readiness, an assessment solution must offer a substantial number of constructed response items as well. Constructed response test questions are essential as they are the only item type that can truly measure certain areas of the CCSS such as writing, research, and problem solving skills. Please note a detailed report on constructed response items in the cost section of this report. The quantity of constructed response items will also be covered in both the cost and scoring and reporting sections of this report. Performance tasks provide insights into students' depth of knowledge on important content because they require students to engage in authentic problem solving and to persevere through the multiple steps of the task.

CONCLUSION

Of the 12 respondents, two of them, Measured Progress and PARCC, indicated that their solutions included all five subject areas for summative assessments and were able to demonstrate sufficient evidence that their solution is aligned with the CCSS. Smarter Balanced has all but speaking as part of their current solution.

In terms of item types, CTB McGraw-Hill, PARCC, and Smarter Balanced were able to demonstrate item types that included standard item types, technology enhanced item types, and performance tasks for all grade levels and content areas.

KEY:
- ● — Appears to fully meet requirements based on responses provided
- ◐ — Unclear if meets or appears to partially meet requirements based on responses provided
- ○ — Does not appear to meet requirements based on responses provided
- NR — No response or did not indicate having a summative product

Transparency & Governance

SUMMATIVE

Service Provider	Clear opportunities for Michigan educators to participate in...			Clear opportunities for MDE involvement in...			Clear evidence the State of Michigan retains sole and exclusive ownership of all student data
	Test question development processes	*Bias/sensitivity and accessibility reviews of test questions*	*Test question scoring processes*	*Test design*	*Test question scoring administration and reporting processes*	*Technical quality processes*	*Retains sole and exclusive ownership of all student data*
ACT Aspire	○	●	○	○	◐	○	○
Amplify Education, Inc.	NR	NR	NR	NR	NR	NR	NR
College Board	●	●	●	◐	◐	○	○
CTB/McGraw-Hill	◐	●	●	◐	◐	●	●
Curriculum Associates LLC	○	○	○	○	◐	○	○
Discovery Education Assessment	NR	NR	NR	NR	NR	NR	NR
Houghton Mifflin Harcourt/Riverside	●	●	●	●	●	●	●
Measured Progress	●	●	●	●	○	○	●
PARCC	◐	●	●	●	●	●	●
Scantron	NR	NR	NR	NR	NR	NR	NR
Smarter Balanced	●	●	●	●	●	●	●
Triumph Learning	NR	NR	NR	NR	NR	NR	NR

INTRODUCTION

It is essential that Michigan's educators and students have an assessment system that meets the unique needs of the state while providing opportunities for valid comparison with other states and large-scale assessment systems. This means that a balance must be found between customizability and compromise, with service providers (e.g., with off-the-shelf products) and other states (e.g., with multi-state consortia), in order to find the best solution for these two competing goals. Michigan is one of a few states that have significant experience with this challenge. Our current assessment programs include tests that are state-specific and completely customized (e.g., MEAP) and tests that are not customizable (e.g., the ACT, which is part of the Michigan Merit Examination) as they are administered across many states and therefore must be static for important reasons such as test security and comparability. Over the course of several months of testimony and debate around implementation of the Common Core State Standards, it was apparent that Michigan's ability to retain control over elements such as personally-identifiable student data was crucial. This section of the report includes ratings for responses to survey questions documenting opportunities for Michigan educators and MDE staff to have a direct and substantive influence in the development and operational implementation of the assessments.

The ratings in the table above were made in light of the high-stakes purposes that summative tests are designed to inform. These types of tests are the ones where comparability across schools, districts and/or states is paramount, and historically, data from them have formed the foundation for accountability systems. In light of this, it is essential that opportunities exist for Michigan educators to provide input in the development of the products used in the accountability system. This is very important with regard to demonstrating validity, especially when these instruments will be used for accountability metrics and evaluations. As administrators of these systems, it is also critical that MDE staff have a formal governance role to assure the results of the assessments are defensible.

If readers are interested in reviewing the specific survey questions and a particular service provider's response, Appendix B contains the necessary information.

CONCLUSION

As documented in the table above, it is evident that there is a limited number of options where the opportunity to strike a balance between a customizable solution for Michigan and a purely off-the-shelf product exists for summative assessments. Based on the responses to the survey questions on this topic, only the College Board, CTB/McGraw-Hill, Houghton Mifflin Harcourt/Riverside, Measured Progress, PARCC and Smarter Balanced appear to be developing solutions that permit robust opportunities for Michigan educator involvement in developing test questions. MDE input into essential areas of governance and design is only apparent with this same group of service providers, albeit with much more opportunity in the case of Houghton Mifflin Harcourt/Riverside, PARCC, and Smarter Balanced. The other service providers clearly indicated that significantly fewer opportunities existed for MDE input in these two areas.

It is also important to note that clear differences exist with regard to Michigan control of student data for these high-stakes summative tests. In light of that critical factor, only CTB/McGraw-Hill, Houghton Mifflin Harcourt/Riverside, PARCC, and Smarter Balanced would be recommended for further consideration based on the responses to this survey.

KEY:
● — Appears to fully meet requirements based on responses provided
◐ — Unclear if meets or appears to partially meet requirements based on responses provided
○ — Does not appear to meet requirements based on responses provided
NR — No response or did not indicate having a summative product

Overall Design & Availability

SUMMATIVE

INTRODUCTION

Michigan is committed to adopting online assessments, as well as providing paper-and-pencil test options while schools and districts continue to acquire and implement the technology required to administer online assessments. MDE believes strongly that the nature of computer-adaptive assessment, where each student receives a customized test event based on his or her performance, is the best solution for improving how student achievement and growth is measured. This is particularly true in the case of high-achieving students, and students that may be much lower than average due to factors such as disability, learning English, etc.

When reviewing the survey responses, MDE asked each respondent to note if their solution offered a computer-adaptive or computer-based test, as well as a paper-and-pencil option for administration. One key difference between computer-adaptive and computer-based assessments is that a computer-adaptive assessment scientifically selects items based on estimated student ability level, therefore a unique test is crafted for each student. A computer-based test is a fixed-form test (similar to paper-and-pencil solutions) where every student will see the same items. MDE also believes that offering a re-take opportunity for the summative assessments is a key component of our desired assessment system.

Service Provider	Overall Test Design				Availability
	Solution will be available in a computer-adaptive modality	*Solution will be available in a computer-based modality*	*Solution will include a comparable paper-pencil option*	*Solution will offer a re-test option*	*Solution will be fully available (including all item types) for the 2014–2015 school year*
ACT Aspire	○	●	●	○	●
Amplify Education, Inc.	○	●	●	NR	○
College Board	○	○	●	NR	○
CTB/McGraw-Hill	●	●	●	○	●
Curriculum Associates LLC	●	○	○	●	◑
Discovery Education Assessment	○	●	●	NR	○
Houghton Mifflin Harcourt/Riverside	○	●	●	●	●
Measured Progress	○	●	●	●	○
PARCC	○	●	●	●	●
Scantron	◑	●	●	NR	●
Smarter Balanced	●	○	●	●	●
Triumph Learning	◑	●	●	NR	◑

All but one of the solutions presented clearly offered an online administration option; although only three, Curriculum Associates, CTB McGraw-Hill, and Smarter Balanced, offered an online computer-adaptive delivery of their assessment. Two of the three, CTB McGraw-Hill and Smarter Balanced, clearly offered a comparable paper-pencil alternative to their proposed computer-adaptive assessment. Two of the solutions presented were not clear in their offerings as they may have only had a computer-adaptive solution in certain grades or content areas.

Of the solutions presented, ACT Aspire and CTB McGraw-Hill did not offer a re-take option for their summative assessment.

Based on the information provided in the survey, many of the solutions would be fully available by the 2014-2015 school year as desired. ACT Aspire, CTB McGraw-Hill, Houghton-Mifflin Harcourt Riverside, PARCC, Scantron, and Smarter Balanced indicated they would have solutions ready within that timeframe.

CONCLUSION

Given the information provided, if Michigan desires to continue in the direction of adopting an online computer-adaptive assessment system with a comparable paper-pencil alternative and a re-take option, it appears that the Smarter Balanced solution would be the only one prepared to meet those requirements. If Michigan decides that a computer-adaptive solution is not indeed a requirement then the solutions from Houghton-Mifflin Harcourt Riverside and PARCC would also be suitable options.

KEY:
● — Appears to fully meet requirements based on responses provided
◑ — Unclear if meets or appears to partially meet requirements based on responses provided
○ — Does not appear to meet requirements based on responses provided
NR — No response or did not indicate having a summative product

INTRODUCTION

School accountability (including designation as priority or focus schools) and educator evaluation are high-stakes uses of Michigan's next generation assessment. Test results from the next generation assessment must therefore be valid for such uses. A key to maintaining validity is the assurance that student performance reflects the learning that students have experienced rather than advance familiarity with test questions or receiving inappropriate assistance in obtaining a high score.

Critical to assuring that student performance reflects student learning are two issues:

- Keeping test questions secure.
- Timely monitoring for security breaches and an ability to respond appropriately.

This section focuses on survey questions providing evidence regarding how well each solution is able to address these two issues.

The number of test forms available for administration to students is critical to keeping test questions secure. A minimum standard is having at least one additional test form to administer to students in the event of a security breach. Even better is to have many forms available such that multiple forms can be administered in the same classroom. In the optimal situation, each student would receive a unique test form, as is the case for Computer Adaptive Testing (CAT). Providers were identified as meeting this criterion if they meet the minimum standard of having at least one additional test form available in the event of a security breach.

Timely provision of security-related data to MDE is critical in being able to monitor for security breaches and respond appropriately. MDE will need to be provided with timely access to security-related data in order to perform forensic analyses on the data for potential security breaches. Timely analysis is needed to initiate and conduct investigations, and (if possible) provide for re-testing, before the testing window closes in the case of a security breach. Providers were asked whether MDE would be provided with timely access to security-related data for analysis.

CONCLUSION

Because maintaining test security is so integral to appropriate high-stakes use of Michigan's next generation assessments, MDE qualified only those that clearly indicated that at least one test form is available for use in the event of a breach of test security and clearly indicated that security-related data would be provided to MDE in a timely manner. The only service providers meeting this criteria based on the responses provided are CTB/McGraw-Hill, PARCC, and Smarter Balanced.

Service Provider	Assessment Integrity and Security	
	Multiple forms are used in operational testing with others available for emergency or misadministration.	*MDE will be provided timely and adequate information needed to monitor and investigate test administration, including student level data and psychometric data to perform forensic and security analyses*
ACT Aspire	○	○
Amplify Education, Inc.	NR	NR
College Board	◑	●
CTB/McGraw-Hill	●	●
Curriculum Associates LLC	●	◑
Discovery Education Assessment	NR	NR
Houghton Mifflin Harcourt/Riverside	○	●
Measured Progress	NR	NR
PARCC	●	●
Scantron	NR	NR
Smarter Balanced	●	●
Triumph Learning	NR	NR

KEY: ● — Appears to fully meet requirements based on responses provided
◑ — Unclear if meets or appears to partially meet requirements based on responses provided
○ — Does not appear to meet requirements based on responses provided
NR — No response or did not indicate having a summative product

Scoring & Reporting

SUMMATIVE

INTRODUCTION

Future scoring and reporting functions of state testing programs need to provide (1) faster, virtually instant, results back to the classroom; (2) more definition as to the depth of knowledge demonstrated by students on the content and standards being assessed; and (3) flexible testing reports that offer students, parents, teachers, and administrators the ability to access data specifically customized according to their individual learning, teaching, and/or evaluation needs.

To do this, we need systems designed to take the most efficient advantage of the data available.

Those who responded to the MDE request for information regarding their summative assessment offerings related to the Common Core State Standards were presented with a series of questions in two major areas regarding Scoring and Reporting. The two areas are Data Analysis Capabilities and Scoring, and Assessment Reporting.

In the area of Data Analysis Capabilities and Scoring, the focus was on vendor-provided products and data that would allow the MDE to run analyses verifying that vendor results were sufficient and accurate measures, as well as provide the MDE with additional opportunities for research and evaluation using the supplied data.

There was also emphasis on the amount of input the State would have into the design of student-level and aggregate data sets, statistical procedures, and scoring protocols. Having opportunities at the design level make it possible to assure the service provider is implementing processes that are the most current and efficient, with an aim to obtaining the highest degree of reliability.

In Assessment Reporting, the areas examined include vendor provisions for:

- reporting at a level sufficient to provide necessary information to educators, MDE, and satisfy federal and state requirements.
- reporting of assessment results that will be timely (i.e., significantly improved over results from current, paper-pencil tests). The immediacy with which reports can be obtained following testing is of constant concern to our stakeholders at all levels. It is critical that new systems take advantage of the opportunities made available by computer-delivered testing.
- assurance that MDE and schools/districts will be provided with all data underlying the reports and will have the capability to perform further analyses if desired. Many schools want and need the capability to examine data in ways that serve their unique populations. This also assures that data will be available as needed to those involved in efforts where improvement is a critical priority.
- parity for students who test with State-approved accommodations to the extent they will receive the same menus and types of score reports provided to students in the general population.

Service Provider	Data Analysis Capabilities and Scoring		Assessment Reporting			
	MDE will have sufficient information for verification and analysis done in-house, using vendor-provided products and data.	*MDE will have direct influence on student and aggregate level data structures, psychometric procedures, and scoring procedures and protocols.*	*Reporting will be at a level sufficient to provide necessary information to educators, MDE, and to satisfy federal and state requirements.*	*Reporting of assessment results will be timely (i.e., significantly improved over results from current, paper-pencil tests).*	*MDE and schools/ districts will be provided with all data underlying the reports and will have the capability to perform further analysis if desired.*	*Students who test with State-approved accommodations will receive the same menus and types of score reports provided to students in the general population.*
ACT Aspire	○	○	○	◐	○	●
Amplify Education, Inc.	NR	NR	NR	NR	NR	NR
College Board	○	○	●	◐	○	●
CTB/McGraw-Hill	●	●	●	●	●	●
Curriculum Associates LLC	○	○	○	●	○	●
Discovery Education Assessment	NR	NR	NR	NR	NR	NR
Houghton Mifflin Harcourt/Riverside	●	○	●	◐	○	●
Measured Progress	NR	NR	NR	NR	NR	NR
PARCC	●	◐	●	NR	●	●
Scantron	NR	NR	NR	NR	NR	NR
Smarter Balanced	●	●	●	●	●	●
Triumph Learning	NR	NR	NR	NR	NR	NR

CONCLUSION

The symbols displayed in the table on these pages provide a visual representation of how service providers offering a summative assessment product appear to meet the requirements for scoring and reporting. Based on responses provided, CTB/McGraw-Hill and the Smarter Balanced Assessment Consortium appear to fully meet requirements in all scoring and reporting categories.

KEY:
● — Appears to fully meet requirements based on responses provided
◐ — Unclear if meets or appears to partially meet requirements based on responses provided
○ — Does not appear to meet requirements based on responses provided
NR — No response or did not indicate having a summative product

INTRODUCTION

This table displays the average, per-student cost for the standard summative products offered by each service provider. While most offered thorough solutions for most of the desired grade span (3-11) indicated in the resolution, there was some degree of variability. Average cost was generated by taking the mean price for each modality (i.e., computer-based assessment/computer-adaptive assessment (CBA/CAT) or paper/pencil) across all grades. This table is provided as an informational snapshot, to which MDE staff did not attempt to assign ratings; therefore no conclusions are provided for this section. While these proposed costs give some idea as to which products are likely to be more or less expensive in a general sense, the information gathered by the survey is insufficient to determine an accurate cost model. As noted in the introduction to the report, that level of detailed information can only be produced by going through the full, formal state procurement process. The Grade Levels column of this section's table indicates that service providers reported having items of each type for only those grades. Additional notes about this are included in the Exceptions Column.

Summative Assessment Per Student Cost (Standard Product)

Service Provider	Average per student cost		Types of Test Questions Included					Exceptions
	CBA/CAT	*P & P*	*Grade Levels*	*Multiple Choice (ELA & Math)*	*Constructed Response (ELA & Math)*	*Technology Enhanced (ELA & Math)*	*Performance Assessment (ELA & Math)*	
ACT Aspire	22.00	28.00	Grades 3 - 10	●	◑	●	○	No constructed response test questions at Grade 9
Amplify Education, Inc.	NR	NR	NR	NR	NR	NR	NR	
College Board	NR	27.75	Grades 9 - 12	●	◑	○	○	No constructed response test questions in Mathematics; no constructed response test questions in ELA grades 9 and 10
CTB/McGraw-Hill	27.00	27.00	Grades 3 - 11	●	●	●	○	
Curriculum Associates LLC	11.00	NA	Grades 3 - 12	●	○	◑	○	No technology enhanced test questions in Grades 9-12
Discovery Education Assessment	NR	NR	NR	NR	NR	NR	NR	
Houghton Mifflin Harcourt/Riverside	20.00	25.00	Grades 3 - 12	●	●	●	◑	No performance assessments available for ELA
Measured Progress	NR	NR	NR	NR	NR	NR	NR	
PARCC	30.00	34.00	Grades 3 - 11	●	●	●	●	
Scantron	NR	NR	Grades 3 - 12	NR	NR	NR	NR	
Smarter Balanced	22.50	15.62	Grades 3 - 8, 11	●	●	●	●	
Triumph Learning	NR	NR	NR	NR	NR	NR	NR	

The Grade Levels column indicates that service providers reported having items of each type for only those grades. Additional notes about this are included in the Exceptions Column.

Additionally, a major driver of both cost and alignment is the number and type of constructed response items. Since the issues around these types of test questions are so pervasive, MDE staff determined it was necessary to display information about them in a separate table on pages 18-19.

KEY:
● — Appears to include this type of test question based on responses provided
◑ — Appears to include this type of question on some, but not all, subjects or grade levels. Please see the comment in the exception column
○ — Does not appear to include this type of test question based on responses provided
NR — No response or did not indicate having an summative product

Constructed Response

Standard Product Cost Implications

SUMMATIVE

INTRODUCTION

While multiple-choice and technology-enhanced test questions are types of items that are well-understood, easy to score, and comparatively cheap to produce, truly assessing the rigor and higher-order thinking skills required by career- and college-ready standards requires something more substantive. Any assessment solution that seeks to demonstrate the capability to measure and provide rich, student achievement and growth information on constructs such as writing, research, communicating reasoning and problem solving to the degree described in the CCSS, must offer test stimuli where students have the opportunity to do more than select 'a', 'b' or 'c'.

Examples of this idea include asking a student to summarize a reading passage in his or her own words, or write about the process he or she used to solve a math problem rather than just selecting the correct answer. Educators deserve strong strong data on how students are achieving and growing on these challenging topics. To attempt to learn more about what options are available now or in the near future to support this idea, the survey included questions specific to constructed-response items. Service providers were asked to list the number of constructed-response test questions that came with their standard product; numbers that are displayed in the following table.

Summative Assessment
Constructed Response (CR) Test Questions Included in Per Student Cost Estimate

Service Provider	Mathematics CR Test Questions				ELA CR Test Questions				Exceptions
	Hand Scored Short Answer	*Hand Scored Extended Response*	*AI Scored Short Answer* [1]	*AI Scored Extended Response* [1]	*Hand Scored Short Answer*	*Hand Scored Extended Response*	*AI Scored Short Answer* [1]	*AI Scored Extended Response* [1]	
ACT Aspire	0	4 -5*	1	0	0	1	1 -2	1	*No hand scored extended response test questions in Grade 9 mathematics
Amplify Education, Inc.	NR	NR	NR	NR	NR	NR	NR	NR	
College Board	0	0	0	0	NR	1**	NR	NR	**No hand scored extended response test questions in Grades 9-10 ELA
CTB/McGraw-Hill	4	1	0	0	3	1	0	0	
Curriculum Associates LLC	0	0	0	0	0	0	0	0	
Discovery Education Assessment	NR	NR	NR	NR	NR	NR	NR	NR	
Houghton Mifflin Harcourt/Riverside	3	1	0	0	3	2	0	0	
Measured Progress	NR	NR	NR	NR	NR	NR	NR	NR	
PARCC	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified	
Scantron	NR	NR	NR	NR	NR	NR	NR	NR	
Smarter Balanced	0	4 - 7	0	0	5	1	0	0	
Triumph Learning	NR	NR	NR	NR	NR	NR	NR	NR	

[1] Artificial Intelligence

CONCLUSION

ACT Aspire, CTB/McGraw-Hill, Houghton Mifflin Harcourt/Riverside and Smarter Balanced appear to include enough constructed-response items to measure student achievement deeply.

NOTES

As indicated in the text above and mentioned in other appropriate sections of this report, constructed-response test questions are considerably more expensive to score than other types of test questions and student responses. Therefore, the survey included an opportunity for service providers to indicate whether or not they were able to provide additional constructed-response items beyond what they offered in their standard package, and a corresponding pricing structure. However, the portion of the survey seeking to gather information on this augmented option functioned differently, depending on the method the service provider used to complete the survey. As a result, service providers interpreted the augmentation section differently and the information was not consistent or reliable. This was discovered as MDE staff began examining responses to this section and it was immediately evident that service providers interpreted this section in dramatically different ways. Therefore, the decision was made to not include information from the survey questions on augmented options (questions 70-73 in Appendix A).

Content & Item Type Alignment

INTERIM

INTRODUCTION

Interim assessments are given periodically throughout the school year. They provide information to educators about student learning and potential success with the summative assessments. The goal is to determine student achievement after instruction while there is still time to remediate areas in which students have done poorly. Michigan desires to have an interim assessment system that mirrors its summative counterpart and that uses an item pool that is independent from the summative assessment item pool.

Multiple-choice and technology-enhanced item types are critical components of an assessment, in order to truly assess the rigor and higher-order thinking skills required from the CCSS and career and college readiness standards an assessment solution must offer a substantial number of constructed-response items as well. Please note that a detailed report on constructed response items is included in the cost section of this report. The quantity of constructed-response items will also be covered in both the cost and scoring and reporting sections of this report. Performance tasks provide insights into students' depth of knowledge on important content because they require students to engage in authentic problem solving and to persevere through the multiple steps of the task. Comments on content alignment are based on survey responses only.

Service Provider	Content Alignment				Item Types		
	Content aligned to the CCSS	*Solution addresses all 5 content areas*	*Solution addresses all grade levels (G3-G11)*	*Qualifications for educators involved in alignment for content, diversity, and special populations*	*Standard item types (multiple choice and constructed response) will be available*	*Diverse set of technology-enhanced item types will be available*	*Performance tasks/ assessments will be available*
ACT Aspire	◐	○	●	○	○	○	○
Amplify Education, Inc.	◐	●	●	○	●	●	●
College Board	NR	NR	NR	NR	NR	NR	NR
CTB/McGraw-Hill	◐	●	●	○	●	●	●
Curriculum Associates LLC	◐	○	○	○	●	●	○
Discovery Education Assessment	◐	○	●	○	○	○	○
Houghton Mifflin Harcourt/Riverside	○	○	●	○	●	◐	○
Measured Progress	●	●	●	◐	◐	◐	◐
PARCC	●	●	●	◐	NR	NR	NR
Scantron	◐	○	●	○	◐	○	○
Smarter Balanced	●	◐	●	◐	●	●	●
Triumph Learning	○	●	●	NR	●	●	●

CONCLUSION

Amplify Education Inc., CTB Mcgraw-Hill, Measured Progress, PARCC, and Triumph report having all five content areas represented in interim assessments. Of these solutions, Measured Progress, and PARCC, demonstrated that their solutions were aligned to the CCSS through the survey. Five solutions (Amplify Education Inc., CTB McGraw-Hill, Measured Progress, Smarter Balanced, and Triumph) report the ability to offer standard item types, technology enhanced items, and performance tasks for the interim assessments they are building.

KEY:
- ● — Appears to fully meet requirements based on responses provided
- ◐ — Unclear if meets or appears to partially meet requirements based on responses provided
- ○ — Does not appear to meet requirements based on responses provided
- NR — No response or did not indicate having an interim product

Transparency & Governance

INTERIM

Service Provider	Clear opportunities for Michigan educators to participate in...			Clear opportunities for MDE involvement in...			Clear evidence the State of Michigan retains sole and exclusive ownership of all student data
	Test question development processes	*Bias/sensitivity and accessibility reviews of test questions*	*Test question scoring processes*	*Test design*	*Test question scoring administration and reporting processes*	*Technical quality processes*	*Retains sole and exclusive ownership of all student data*
ACT Aspire	○	●	○	○	◐	○	○
Amplify Education, Inc.	○	○	○	◐	◐	○	●
College Board	NR	NR	NR	NR	NR	NR	NR
CTB/McGraw-Hill	●	●	●	◐	◐	●	●
Curriculum Associates LLC	○	○	○	○	◐	○	○
Discovery Education Assessment	◐	●	○	○	◐	○	○
Houghton Mifflin Harcourt/Riverside	●	●	○	●	●	●	●
Measured Progress	●	●	●	●	○	○	●
PARCC	●	●	●	●	●	●	●
Scantron	●	●	●	●	●	○	NR
Smarter Balanced	●	●	●	●	●	●	●
Triumph Learning	●	●	●	NR	NR	NR	●

INTRODUCTION

It is essential that Michigan's educators and students have an assessment system that meets the unique needs of the state while providing opportunities for valid comparison with other states and large-scale assessment systems. This means that a balance must be found between customizability and compromise, with service providers (e.g., with off-the-shelf products) and other states (e.g., with multi-state consortia), in order to find the best solution for these two competing goals. Michigan is one of a few states that have significant experience with this challenge. Our current assessment programs include tests that are state-specific and completely customized (e.g., MEAP) and tests that are not customizable (e.g., the ACT, which is part of the Michigan Merit Examination) as they are administered across many states and therefore must be static for important reasons such as test security and comparability. Over the course of the months of testimony and debate around implementation of the Common Core State Standards, it was readily apparent that retaining Michigan control over elements such as personally-identifiable student data was crucial. This section of the report includes ratings for responses to survey questions documenting opportunities for Michigan educators and MDE staff to have a direct and substantive influence in the development and operational implementation of the assessments.

The ratings in this section's table above were made in light of the purposes that interim tests are typically designed to inform. For example, interim tests can be used to inform educator evaluations if all teachers in the same grade and subject administer them under the same conditions. Since it is likely that interim tests will be used for some accountability systems or purposes in Michigan schools, it is just as important that opportunities for involvement in test development and design be documented for them as for summative tests. If readers are interested in reviewing the specific survey questions and a particular service provider's response, Appendix B contains the necessary information.

CONCLUSION

In addition to more service providers indicating that they have an aligned interim solution compared to this category for summative assessments, it is clear that more opportunities exist for Michigan participation in test question development and governance activities. Based on the responses to the survey questions on this topic, CTB/McGraw-Hill, Houghton Mifflin Harcourt/Riverside, Measured Progress, PARCC, Scantron, Smarter Balanced, and Triumph Learning all provide substantial opportunities for Michigan educator involvement in developing test questions. Lack of MDE input into essential areas of governance and design eliminates Measured Progress and Triumph Learning from the list. Since all the remaining service providers except Scantron affirm Michigan's control over student data, CTB/McGraw-Hill, Houghton Mifflin Harcourt/Riverside, PARCC, and Smarter Balanced would be recommended for further consideration with regard to their interim assessment solutions.

It is also important to note that clear differences exist with regard to Michigan control of student data for these interim tests. In light of that critical factor, only CTB/McGraw-Hill, Houghton Mifflin Harcourt/Riverside, PARCC, and Smarter Balanced would be recommended for further consideration based on the responses to this survey.

KEY:
- ● — Appears to fully meet requirements based on responses provided
- ◐ — Unclear if meets or appears to partially meet requirements based on responses provided
- ○ — Does not appear to meet requirements based on responses provided
- NR — No response or did not indicate having an interim product

Overall Design & Availability

INTERIM

Service Provider	Overall Test Design				Availability
	Solution will be available in a computer adaptive modality	*Solution will be available in a computer based modality*	*Solution will include a comparable paper-pencil option.*	*Interim Solution(s) will have opportunity for multiple (at least twice per year) administrations.*	*Solution will be fully available (including all item types) for the 2014-2015 school year.*
ACT Aspire	○	●	●	●	●
Amplify Education, Inc.	○	●	●	●	○
College Board	○	○	●	NR	○
CTB/McGraw-Hill	●	●	●	●	●
Curriculum Associates LLC	●	○	○	◐	◐
Discovery Education Assessment	○	●	●	●	●
Houghton Mifflin Harcourt/Riverside	○	●	●	●	●
Measured Progress	○	●	●	●	●
PARCC	○	●	●	●	●
Scantron	◐	●	●	●	●
Smarter Balanced	●	○	●	●	●
Triumph Learning	◐	●	●	○	◐

INTRODUCTION

Michigan is committed to building an interim assessment system that is completely internet-based, as well as providing paper-and-pencil test options while schools and districts continue to acquire and implement the technology required to administer online assessments. MDE believes strongly that the nature of computer-adaptive assessment, where each student receives a customized test event based on his or her performance, is the best solution for improving how student achievement and growth is measured. This is particularly true in the case of high-achieving students, and students that may be much lower than average due to factors such as disability, learning English, etc.

Michigan also desires an interim assessment system that would provide a great amount of flexibility and applications for Michigan educators. This would require that an interim assessment is available to be given at least twice a year, which would allow it to be used as an end-of-course, or a mid-year checkpoint as examples for educators and their students.

When reviewing the survey responses, MDE asked each respondent to note if their solution offered a computer-adaptive or computer-based test, as well as a paper-and-pencil option for administration. One key difference between computer-adaptive and computer-based assessments is that a computer-adaptive assessment scientifically selects items based on estimated student ability level; therefore a unique test is crafted for each student. A computer-based test is a fixed-form test (similar to paper-and-pencil) where every student will see the same items.

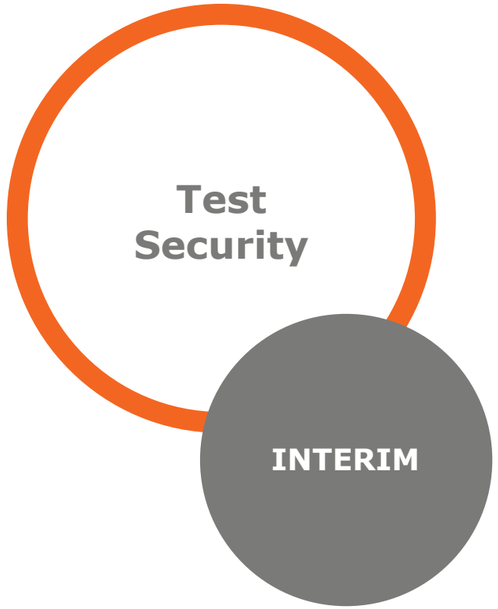
All but one of the solutions presented clearly offered an online administration option, although three, Curriculum Associates, CTB McGraw-Hill, and Smarter Balanced, offered an online computer-adaptive delivery of their assessment. Of those three, CTB McGraw-Hill and Smarter Balanced clearly offered multiple administration opportunities per year of their interim system.

Based on the information provided in the survey, many of the solutions would be fully available by the 2014-2015 school year as desired. ACT Aspire, CTB McGraw-Hill, Houghton-Mifflin Harcourt Riverside, PARCC, Scantron, and Smarter Balanced all would have solutions ready within that timeframe.

CONCLUSION

Given the information provided, if Michigan desires an interim assessment solution that could be administered in an online computer-adaptive system with a comparable paper-pencil alternative and offer multiple administrations per year, it appears that the solutions presented from CTB McGraw-Hill and Smarter Balanced would be prepared to meet those requirements. If Michigan decides that a computer-adaptive solution is not a requirement, then the solutions from ACT Aspire, Discovery Education, Houghton-Mifflin Harcourt Riverside, Measured Progress, PARCC and Scantron would also be suitable options.

KEY:
● — Appears to fully meet requirements based on responses provided
◐ — Unclear if meets or appears to partially meet requirements based on responses provided
○ — Does not appear to meet requirements based on responses provided
NR — No response or did not indicate having an interim product



INTRODUCTION

School accountability (including designation as priority or focus schools) and educator evaluation are high-stakes uses of Michigan's next generation assessment. Test results from the next generation assessment must therefore be valid for such uses. A key to maintaining validity is the assurance that student performance reflects the learning that students have experienced rather than advance familiarity with test questions or receiving inappropriate assistance in obtaining a high score.

Timely provision of security-related data to MDE is critical in being able to monitor for security breaches and respond appropriately. MDE will need to be provided with timely access to security-related data in order to perform forensic analyses on the data for potential security breaches. Timely analysis is needed to initiate and conduct investigations, and (if possible) provide for re-testing, before the testing window closes in the case of a security breach. Providers were asked whether MDE would be provided with timely access to security-related data for analysis.

CONCLUSION

Because maintaining test security is so integral to appropriate high-stakes use of Michigan's next generation assessments, for interim assessments MDE qualified only those that clearly indicated that security related data would be provided to MDE in a timely manner. The only service providers meeting this criteria based on responses provided were CTB/McGraw-Hill, Houghton Mifflin Harcourt/Riverside, Measured Progress, PARCC, and Smarter Balanced.

Service Provider	Assessment Integrity and Security
	MDE will be provided timely and adequate information needed to monitor and investigate test administration, including student level data and psychometric data to perform forensic and security analyses
ACT Aspire	○
Amplify Education, Inc.	◑
College Board	NR
CTB/McGraw-Hill	●
Curriculum Associates LLC	◑
Discovery Education Assessment	◑
Houghton Mifflin Harcourt/Riverside	●
Measured Progress	●
PARCC	●
Scantron	◑
Smarter Balanced	●
Triumph Learning	◑

KEY: ● — Appears to fully meet requirements based on responses provided
◑ — Unclear if meets or appears to partially meet requirements based on responses provided
○ — Does not appear to meet requirements based on responses provided
NR — No response or did not indicate having an interim product

Scoring & Reporting

INTERIM

INTRODUCTION

Future scoring and reporting functions of state testing programs need to provide (1) faster, virtually instant, results back to the classroom; (2) more definition as to the depth of knowledge demonstrated by students on the content and standards being assessed; and (3) flexible testing reports that offer students, parents, teachers, and administrators the ability to access data specifically customized according to their individual learning, teaching, and/or evaluation needs.

To do this, we need systems designed to take the most efficient advantage of the data available.

Those who responded to the MDE's request for information regarding their interim assessment offerings related to the Common Core State Standards were presented with a series of questions in two major areas regarding Scoring and Reporting. The two areas are Data Analysis Capabilities and Scoring, and Assessment Reporting.

In the area of Data Analysis Capabilities and Scoring, the focus was on vendor-provided products and data that would allow the MDE to run analyses verifying that the vendor results were sufficient and accurate measures, as well as provide the MDE with additional opportunities for research and evaluation using the supplied data.

There was also emphasis on the amount of input the State would have into the design of student-level and aggregate data sets, statistical procedures, and scoring protocols. Having opportunities at the design level make it possible to assure the service provider is implementing processes that are reliable, efficient, and valid for the intended purposes.

Service Provider	Data Analysis Capabilities and Scoring		Assessment Reporting			
	MDE will have sufficient information for verification and analysis done in-house, using vendor-provided products and data.	*MDE will have direct influence on student and aggregate level data structures, psychometric procedures, and scoring procedures and protocols.*	*Reporting will be at a level sufficient to provide necessary information to educators, MDE, and to satisfy federal and state requirements.*	*Reporting of assessment results will be timely (i.e., significantly improved over results from current, paper-pencil tests).*	*MDE and schools/ districts will be provided with all data underlying the reports and will have the capability to per-form further analysis if desired.*	*Students who test with State-approved accommodations will receive the same menus and types of score reports provided to students in the general population.*
ACT Aspire	○	○	○	◐	○	●
Amplify Education, Inc.	○	○	◐	◐	○	●
College Board	NR	NR	NR	NR	NR	NR
CTB/McGraw-Hill	●	●	●	●	●	●
Curriculum Associates LLC	○	○	○	●	○	●
Discovery Education Assessment	◐	○	○	●	○	●
Houghton Mifflin Harcourt/Riverside	●	○	●	◐	○	●
Measured Progress	○	○	○	◐	○	●
PARCC	●	◐	●	NR	●	●
Scantron	○	◐	○	●	●	●
Smarter Balanced	●	●	●	●	●	●
Triumph Learning	○	○	NR	◐	●	NR

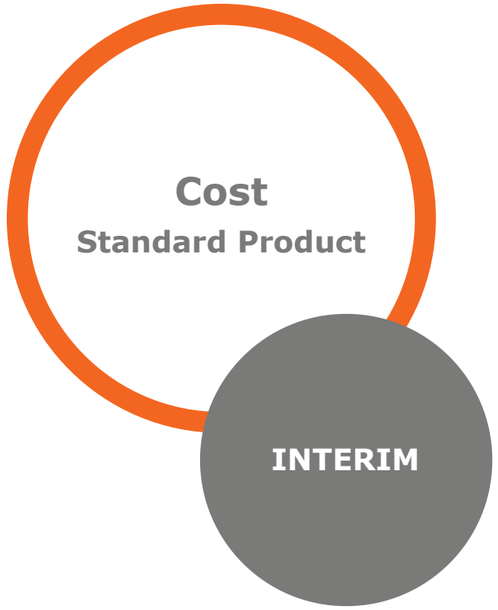
In Assessment Reporting, the areas examined include vendor provisions for:

- reporting at a level sufficient to provide necessary information to educators, MDE, and to satisfy federal and state requirements.
- reporting of assessment results that will be timely (i.e., significantly improved over results from current, paper-pencil tests). The immediacy with which reports can be obtained following testing is of constant concern to our stakeholders at all levels. It is critical that new systems take advantage of the opportunities made available by computer-delivered testing.
- assurance that MDE and schools/districts will be provided with all data underlying the reports and will have the capability to perform further analysis if desired. Many schools want and need the capability to examine data in ways that serve their unique populations. This also assures that data will be available as needed to those involved in efforts where improvement is a critical priority.
- parity for students who test with state-approved accommodations to the extent they will receive the same menus and types of score reports provided to students in the general population.

CONCLUSION

The symbols displayed in the table on these pages provide a visual representation of how service providers offering an interim assessment product appear to meet the requirements for scoring and reporting. Based on responses provided, CTB/McGraw-Hill and the Smarter Balanced Assessment Consortium appear to fully meet requirements in all scoring and reporting categories.

KEY:
● — Appears to fully meet requirements based on responses provided
◐ — Unclear if meets or appears to partially meet requirements based on responses provided
○ — Does not appear to meet requirements based on responses provided
NR — No response or did not indicate having an interim product



INTRODUCTION

This table displays the average, per-student cost for the standard interim products offered by each service provider. While most offered thorough solutions for most of the desired grade span (3-11) indicated in the resolution, there was some degree of variability. Average cost was generated by taking the mean price for the computer-based assessment/computer-adaptive assessment (CBA/CAT) solution offered across all grades. This table is provided as an informational snapshot, to which MDE staff did not attempt to assign ratings. While these proposed costs give some idea as to which products are likely to be more or less expensive in a general sense, the information gathered by the survey is insufficient to determine an accurate cost model. As noted in the introduction to the report, that level of detailed information can only be produced by going through the full, formal state procurement process. The Grade Levels column of this section's table indicates that service providers reported having items of each type for only those grades. Additional notes about this are included in the Exceptions Column.

Interim Assessment Per Student Cost (Standard Product)							Exceptions
Service Provider	Average per student cost CBA/CAT	Types of Test Questions Included					
		GradeLevels	*Multiple Choice (ELA & Math)*	*Constructed Response (ELA & Math)*	*Technology Enhanced (ELA & Math)*	*Performance Assessment (ELA & Math)*	
ACT Aspire	7.00	Grades 3 - 12	●	○	◐	○	No technology enhanced test questions in mathematics; no technology enhanced test questions in ELA grades 9-10
Amplify Education, Inc.	4.25	Grades 3 - 12	●	●	●	●	
College Board	NR	NR	NR	NR	NR	NR	
CTB/McGraw-Hill	13.00	Grades 3 - 12	●	●	●	●	
Curriculum Associates LLC	11.00	Grades 3 - 8	●	○	●	○	
Discovery Education Assessment	8.00	Grades 3 - 11	●	○	○	○	
Houghton Mifflin Harcourt/Riverside	10.00	Grades 3 - 12	●	●	●	○	
Measured Progress	5.70	Grades 3 - 11	●	●	●	◐	No performance assessments in grades 9-11
PARCC	NR	NR	NR	NR	NR	NR	
Scantron	5.00	NR	●	●	○	○	
Smarter Balanced	4.80	Grades 3 - 11	●	●	●	●	
Triumph Learning	20.00	Grades 3 - 12	●	●	●	●	

The Grade Levels column indicates that service providers reported having items of each type for only those grades. Additional notes about this are included in the Exceptions Column.

Additionally, a major driver of both cost and alignment is the number and type of constructed response items. Since the issues around these types of test questions are so pervasive, MDE staff determined it was necessary to display information about them in a separate table on pages 32-33.

KEY: ● — Appears to fully meet requirements based on responses provided
◐ — Appears to include this type of question on some, but not all, subjects or grade levels. Please see the comment in the exception column
○ — Does not appear to meet requirements based on responses provided
NR — No response or did not indicate having an interim product

Constructed Response

Standard Product Cost Implications

INTERIM

INTRODUCTION

While multiple-choice and technology-enhanced test questions are types of items that are well-understood, easy to score, and comparatively cheap to produce, truly assessing the rigor and higher-order thinking skills required by career- and college-ready standards requires something more substantive. Any assessment solution that seeks to demonstrate the capability to measure and provide rich, student achievement and growth information on constructs such as writing, research, communicating reasoning and problem solving to the degree described in the CCSS, must offer test stimuli where students have the opportunity to do more than select 'a', 'b' or 'c'.

Examples of this idea include asking a student to summarize a reading passage in his or her own words, or write about the process he or she used to solve a math problem rather than just selecting the correct answer. MDE feels very strongly that educators deserve strong data on how students are achieving and growing on these challenging topics. To attempt to learn more about what options are available now or in the near future to support this idea, the survey included questions specific to constructed-response items. Service providers were asked to list the number of constructed-response test questions that came with their standard product; numbers that are displayed in the following table.

Interim Assessment
Constructed Response (CR) Test Questions Included in Per Student Cost Estimate

Service Provider	Mathematics CR Test Questions				ELA CR Test Questions				Exceptions
	Hand Scored Short Answer	*Hand Scored Extended Response*	*AI Scored Short Answer* [1]	*AI Scored Extended Response* [1]	*Hand Scored Short Answer*	*Hand Scored Extended Response*	*AI Scored Short Answer* [1]	*AI Scored Extended Response* [1]	
ACT Aspire	0	0	0	0	0	0	0	0	
Amplify Education, Inc.	66 - 171	16 - 59	0	0	5 - 21	19 - 50	0	0	
College Board	NR	NR	NR	NR	NR	NR	NR	NR	
CTB/McGraw-Hill	12 - 34	2 - 8*	2 - 20	0	12 - 19	8 - 10	2 - 20	0	*No hand scored extended response test questions for Grade 12 math
Curriculum Associates LLC	0	0	0	0	0	0	0	0	
Discovery Education Assessment	0	0	0	0	0	0	0	0	
Houghton Mifflin Harcourt/Riverside	0	0	3	1	0	0	3	1	
Measured Progress**	16	8	0	0	0	8	0	0	Information is for Grades 3-8 only; NR for High School
PARCC	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified	Not Specified	
Scantron	0	0	0	0	0	0	0	0	
Smarter Balanced	2***	4 - 5****	0	3***	5	1	0	0	***No short answer test questions in math grades 3-8; extended response test questions in math are hand scored for grades 3-8 and AI scored for grades 9-12
Triumph Learning**	60 - 70	35 - 75	0	0	60 - 70	35	0	0	Information is for Grades 3-8 only; NR for High School

[1] Artificial Intelligence

NOTES

As indicated in the text above and mentioned in other appropriate sections of this report constructed-response test questions are considerably more expensive to score than other types of test questions and student responses. Therefore, the survey included an opportunity for service providers to indicate whether or not they were able to provide additional constructed-response items beyond what they offered in their standard package, and a corresponding pricing structure. However, the portion of the survey seeking to gather information on this augmented option functioned differently, depending on the method the service provider used to complete the survey. As a result, service providers interpreted the augmentation section differently and the information was not consistent or reliable. This was discovered as MDE staff began examining responses to this section and it was immediately evident that service providers interpreted this section in dramatically different ways. Therefore, the decision was made to not include information from the survey questions on augmented options (questions 70-73 in Appendix A).

Accessibility

INTRODUCTION

Michigan is committed to the inclusion of ALL students, including students with disabilities (SWD) and English language learners (ELLs), in large-scale assessment and accountability systems. Assessment results should not be affected by disability, gender, ethnicity, or English language ability, and all students should have an opportunity to receive valid scores for summative and interim assessments. To ensure validity, assessments must promote an equitable opportunity for ELLs, SWDs, and general education students. The challenge of how to include all students in these assessments brings accessibility issues to the forefront. The purpose of the Accessibility Category is to ensure that all students have the supports and tools they require in order to fully access Michigan's assessment system. There are two types of accessibility features: Assessment Accommodations and Universal Tools. Assessment Accommodations are used to change the way students access a test without changing the content being assessed. In other words, accommodations equalize entry to the test without giving the student an unfair advantage, or altering the subject matter. For example, a blind student could access the test in Braille rather than print, and an English language learner may require test questions be translated into their primary language. Universal Tools can be used by any student who needs minor supports, such as a highlighter, magnifying device, or notepad. A series of questions aimed at determining the availability of these accessibility features for summative and interim assessments were included in the survey. Please refer to Appendices A and C.

Based on the results of the categorization process, the following is a list of the responders in rank order, top-to-bottom, who had the most offerings meeting Michigan's accessibility requirements for their respective summative and/or interim assessments:

Service Provider	Accommodations for English language learners (ELLs)					Accommodations for Students with Disabilities (SWD)					Accessibility Tools	Translation Languages	Accommodations/ Reported Scores
	Embedded text-to-speech	*English Glossing*	*Foreign Language Glossing*	*Full translation of test questions into language other than English*	*Universal accommodations*	*Embedded text-to-speech*	*Embedded video in ASL (human)*	*Refreshable braille*	*Print-on-demand tactile graphics*	*Universal accommodations*	*Universally-provided accessibility tools*	*Minimally Spanish, Arabic*	*Official scores reported for State-approved accommodations*
ACT Aspire	●	○	○	●	●	●	◐	◐	◐	●	◐	○	●
Amplify Education, Inc.	○	○	○	●	●	○	○	○	○	●	◐	○	●
College Board	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	○	◐	●
CTB/McGraw-Hill	●	○	○	◐	●	●	○	○	○	●	◐	○	●
Curriculum Associates LLC	○	○	○	○	◐	○	○	○	○	◐	◐	○	●
Discovery Education Assessment	○	○	○	●	○	○	○	○	○	○	NR	○	●
Houghton Mifflin Harcourt/Riverside	○	○	○	○	●	○	○	○	○	●	◐	○	●
Measured Progress	●	○	○	○	●	●	○	○	○	●	◐	○	●
PARCC	●	●	●	●	●	●	●	●	●	●	◐	○	●
Scantron	○	○	○	●	○	○	○	○	○	◐	◐	○	●
Smarter Balanced	●	●	●	●	●	●	●	●	●	●	◐	●	●
Triumph Learning	○	○	○	○	●	○	○	○	●	●	◐	○	NR

Number of accessibility features meeting requirements	
Smarter Balanced	12
PARCC	11
ACT Aspire	6
CTB/McGraw-Hill	5
Measured Progress	5
Amplify Education, Inc.	4
Houghton Mifflin Harcourt/Riverside	3
Triumph Learning	3
Discovery Education Assessment	2
Scantron	2
College Board	1
Curriculum Associates LLC	1

CONCLUSION

Two of the respondents, Smarter Balanced and PARCC, provided sufficient evidence that their product meets all of Michigan's expectations for providing appropriate accommodations on their respective assessments for ELLs and SWDs. None of the respondents met all requirements for universally-provided tools. Smarter Balanced was the only respondent to report they currently provide the required languages for translation.

All respondents except for Triumph Learning indicated they meet Michigan's requirements for reporting valid scores for students using State-approved accommodations on their respective assessments.

KEY:
● — Appears to fully meet requirements based on responses provided
◐ — Unclear if meets or appears to partially meet requirements based on responses provided
○ — Does not appear to meet requirements based on responses provided
NR — No response

Technical Requirements

INTRODUCTION

Service providers were asked to indicate the device types and operating systems supported by their Computer Adaptive Testing solution. Service providers were also asked to provide the bandwidth requirement for each testing site. These factors have a significant effect on the level of school technology readiness as well as the overall cost to schools and districts.

All of the Service providers that responded, with the exception of Triumph Learning, indicated that their online testing system supports Windows XP/7 desktop and laptop testing devices. Since Windows XP is still widely used in Michigan schools, it is critical that the online testing system provide support for these devices.

All service providers that responded indicated that their online testing system supports Mac OSX desktop and laptop testing devices. According to MTRAx, a technology readiness survey tool, OSX devices are also common among Michigan schools. Therefore, it is critical that the online testing system provides support for these devices.

An increasing number of schools are adopting Chromebook devices for student instructional and assessment use. Seven of the responding service providers indicated that their online testing system supports Chromebook as a testing device. iPads are also widely used in Michigan schools for instruction and assessment. Six of the responding service providers indicated that their online testing system supports the iPad as a testing device.

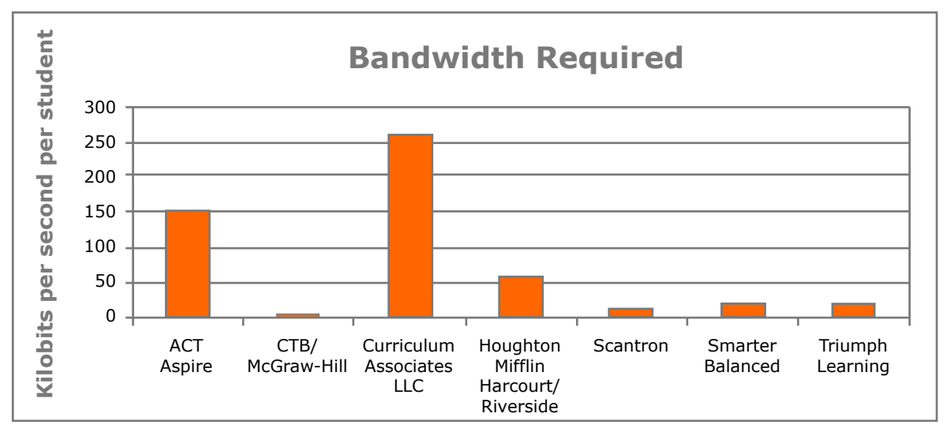
Service providers were asked if MDE would have a formal decision-making role with the ability to have a direct influence on the operating systems and technology platforms supported by their online testing system. Only PARCC and Smarter Balanced indicated MDE would have influence on the operating systems supported. Houghton Mifflin Harcourt/Riverside, PARCC, and Smarter Balanced indicated MDE would have influence on the technology platforms supported.

CONCLUSION

Many Michigan schools and districts have begun deployment of a variety of student-level mobile devices including Chromebooks and tablets. In many schools, these mobile devices are actually replacing the traditional computer lab configuration. Best practice calls for students to use the same device for both instruction and assessment. Therefore, the online testing system needs to support not only desktops and laptops but also Chromebooks and tablets (running iOS, Android, and Windows 8). Additionally, some schools have limited internet bandwidth available, which may limit the number of students that can test simultaneously.

Of the service providers that responded, Discovery Education Assessment and Smarter Balanced appear to meet the overall criteria regarding technical requirements.

Service Provider	The online testing system supports the use of						
	Windows XP/7 desktops/ laptops	*Windows 8 desktops/ laptops*	*Mac OS X desktops/ laptops*	*Chrome OS laptops (Chromebooks)*	*iOS tablets (iPads)*	*Android tablets*	*Windows 8 tablets*
ACT Aspire	●	●	●	●	●	●	○
Amplify Education, Inc.	●	○	●	●	●	○	○
College Board	NR	NR	NR	NR	NR	NR	NR
CTB/McGraw-Hill	●	●	●	○	○	○	○
Curriculum Associates LLC	●	●	●	●	○	○	●
Discovery Education Assessment	●	●	●	●	●	●	●
Houghton Mifflin Harcourt/Riverside	●	●	●	○	●	●	○
Measured Progress	NR	NR	NR	NR	NR	NR	NR
PARCC	NR	NR	NR	NR	NR	NR	NR
Scantron	●	○	○	●	○	○	○
Smarter Balanced	●	●	●	●	●	●	●
Triumph Learning	◐	●	●	●	●	●	○



*Only those service providers that responded with a bandwidth requirement are displayed.

KEY: ● — Appears to fully meet requirements based on responses provided
◐ — Unclear if meets or appears to partially meet requirements based on responses provided
○ — Does not appear to meet requirements based on responses provided
NR — No response

Formative Assessment Resources

INTRODUCTION

The formative assessment process differs from summative and interim assessments in many ways. Fundamental to understanding these differences is knowing how and when formative assessment is used.

In 2006, Michigan education representatives collaborated with other state education leaders, Council of Chief State School Officers (CCSSO), and national and international experts on formative assessment to develop a widely cited definition of formative assessment:

> "Formative assessment is a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievements of intended instructional outcomes." (CCSSO FAST SCASS 2006)

The importance of this definition is that it is compatible with research showing such practices to be an important driver of student learning gains. At the core of the formative assessment process is that it takes place during instruction to support student learning while learning is developing. This is a distinct difference from summative and interim assessment that are intended to assess students after an extended period of learning. Simply giving students an assessment in the classroom does not mean that the assessment is formative. Use of assessment evidence requires teachers to gain insights into individual student learning in relation to standards and to make instructional decisions and to use descriptive feedback to guide next steps. In addition, during the formative assessment process, student involvement is an essential component. Teachers seek ways to involve the student in "thinking about their thinking" (metacognition) to use learning evidence to close the gap and get closer to the intended learning target.

While formative assessment is not a new idea, teachers are not typically trained on it in-depth. Simply putting resources and tools into teacher hands is not sufficient. Sustained professional development is needed to apply sound formative assessment practices. This is why reviewing MDE staff included Professional Development as a rating category.

Service Provider	Compatible Definition	Online Availability	Variety of Classroom resources/tools/ strategies	Professional learning opportunities	Resources aligned to quality criteria	MI educator submission process	Cost Indicators
ACT Aspire	NR	○	NR	NR	NR	NR	NR
Amplify Education, Inc.	◑	◑	●	●	◑	○	No additional costs based on responses provided
College Board	NR	NR	NR	NR	NR	NR	NR
CTB/McGraw-Hill	●	●	●	◑	◑	●	Additional costs based on response provided
Curriculum Associates LLC	○	○	NR	NR	NR	NR	NR
Discovery Education Assessment	●	●	●	●	●	●	No additional costs based on responses provided
Houghton Mifflin Harcourt/Riverside	●	○	NR	NR	NR	NR	NR
Measured Progress	●	●	◑	●	●	◑	Additional costs based on response provided
PARCC	◑	●	◑	◑	◑	◑	NR
Scantron	●	●	●	●	◑	●	Additional costs based on response provided
Smarter Balanced	●	●	●	●	●	●	No additional costs based on responses provided
Triumph Learning	●	●	●	○	○	○	NR

CONCLUSION

Based on a review of survey responses, it appears that CTB/McGraw-Hill, Discovery, Measured Progress, Smarter Balanced, and Scantron may meet all or most of the stated requirements. Each indicates an online repository, a compatible definition of formative assessment, availability of classroom tools and professional learning resources, and opportunities for Michigan educators to submit additional resources. However, closer examination of resources and services that support educator understanding and use of the formative assessment process is encouraged.

KEY:
- ● — Appears to fully meet requirements based on responses provided
- ◑ — Unclear if meets or appears to partially meet requirements based on responses provided
- ○ — Does not appear to meet requirements based on responses provided
- NR — No response

Local Implications

INTRODUCTION

Many of the preceding sections focus on global aspects for how various products or solutions were designed or intended to function. The MDE believes a consideration that must be given equal weight is the set of implications for Michigan districts and schools that come with each solution. The opportunity to implement a new assessment system, especially in light of the shift from paper-pencil test to those delivered by computer, means that careful examination of several issues is important to determine if the transition will add or remove a significant amount of the burden that comes with secure, large-scale and high-stakes testing. In order to maintain the validity of test results, it is critical that standardized processes be in place and adhered to by test administrators so that the tests remain secure and the results uncompromised. What that principle in mind, four primary factors (Test Security, Test Design, Platform Availability and Bandwidth Requirements) are presented here in light of the potential they have to substantially increase or reduce burden on local districts and schools, depending on how they are implemented. In order to express the rationale for why these elements are so important, they are reiterated here, as opposed to only in the preceding sections where they originally appear.

FACTOR DESCRIPTIONS

Test Security Deploying a large number of comparable forms can assure that few students see a particular set of test questions, significantly reducing the potential for cheating. Computer adaptive testing (CAT) takes this further in that each student sees a unique test form matched to his or her performance, dramatically reducing the opportunities for cheating. If cheating occurs, identifying the extent, and providing additional testing opportunities places a significant burden on local districts and schools affected by the compromised test questions. A large number of test forms can help to reduce this risk for schools. CAT can substantially mitigate this risk. As described in the Overall Design & Availability and Test Administration & Security sections, providers responded to questions regarding number of forms and use of CAT. Providers meeting thresholds for multiple forms and/or CAT were: CTB/McGraw-Hill, Curriculum Associates, PARCC, and Smarter Balanced.

Test Design The design (e.g., CAT vs. fixed-form) of tests delivered via computer has another substantial implication with regard to test administration. In order to maintain test security, because students see the same set of test questions, fixed-form testing requires that all students be tested on the same day (or the same small set-of-days). This scenario would require every student to have a suitable device. This student to device ratio is a major cost driver for local districts and schools in moving from paper-and-pencil testing to online testing. Because each student taking a CAT test sees a unique test form, CAT allows for a long testing window, in turn making it possible for local districts and schools to move testing online even without one-to-one student-to-device ratios. Districts that will not be ready, even for this low bar of technology readiness, will need to have a paper-and-pencil option available. As described in the overall design & availability section, the two providers with CAT solutions with a paper & pencil option are CTB/McGraw-Hill and Smarter Balanced. While Scantron and Triumph Learning noted similar solutions it was unclear if they would meet Michigan's needs as they only noted CAT solutions at certain grades.

Platform Availability In order to take advantage of technology purchases already made by local schools and districts, the solution adopted for Michigan must support the widest possible array of computing devices. The fewer platforms that are supported, the fewer the number of students that will be able to take the tests online, or the more new devices local schools and districts will need to purchase to make the move to online testing. As described in the technical requirements section, the providers indicating adequate availability on a wide variety of platforms include: Discovery Education Assessment, Smarter Balanced, and ACT Aspire.

Service Provider	Test Security	Test Design	Platform Availability	Bandwidth Requirements
ACT Aspire	○	○	◐	○
Amplify Education, Inc.	○	○	○	○
College Board	○	○	○	○
CTB/McGraw-Hill	●	●	○	●
Curriculum Associates LLC	●	○	○	○
Discovery Education Assessment	○	○	●	○
Houghton Mifflin Harcourt/Riverside	○	○	○	○
Measured Progress	○	○	○	○
PARCC	●	○	○	○
Scantron	○	○	○	●
Smarter Balanced	●	●	●	●
Triumph Learning	○	○	○	●

Bandwidth Requirements To maximize the number of students who can take assessments online without significant costs put toward increased bandwidth, the solution provided must require minimal bandwidth. As described in the technical requirements section, providers responded to a question about the bandwidth required for each student taking a test. The MDE review team qualified only those that were reasonably near the lowest requirement listed by any provider. MDE leniently qualified, based on our experience, providers requiring less than 50kbps per student. The providers meeting this threshold were: CTB/McGraw-Hill, Scantron, Smarter Balanced, and Triumph Learning.

CONCLUSION

In these four primary areas driving local implications, ACT Aspire, Curriculum Associates, Discovery Education Assessment, PARCC, Scantron and Triumph met one threshold, CTB/McGraw-Hill met three thresholds, and Smarter Balanced met all four.

KEY: ● — Appears to fully meet requirements based on responses provided
◐ — Unclear if meets or appears to partially meet requirements based on responses provided
○ — Does not appear to meet requirements based on responses provided
NR — No response

SUMMARY CONCLUSIONS

This report on options for assessments aligned with the Common Core State Standards (CCSS) contains a substantial amount of information on the current status of a number of potentially viable solutions. Each element (i.e., Summative, Interim and Formative) required for a balanced, rigorous and fair system of measuring student achievement and growth currently exists or will be operational in the near future. However, since many components of the solutions presented for consideration are not yet fully operational, and none of the solutions currently provides all three components, a definitive recommendation for a full-service system is difficult.

Additionally, assessments used to inform the state accountability system are subject to review by the U.S. Education Department, as part of the No Child Left Behind Act of 2001. This review requires the state to demonstrate how well the tests match the purposes to which they are being applied. In order to do this, Michigan needs to be able to have complete information on all aspects of the development, administration, scoring and reporting of the assessment. Therefore, multiple survey questions that formed the basis of the content of this report attempted to capture the degree to which Michigan can participate in or have opportunities to thoroughly understand aspects of the assessments proposed by service providers.

As Michigan moves forward with new levels of accountability for districts, schools and for teachers, the Department believes strongly that Michigan educators and assessment experts must have opportunities to inform the design of the tests. This includes how test questions will be developed and scored, results will be reported and how the technical adequacy will be documented. MDE must have access to sufficient documentation to permit staff with content and assessment expertise to evaluate the quality of processes used to develop and implement each aspect of the system.

As important as ensuring LEAs have access to high-quality, secure summative (once-yearly) assessments is the need to provide high-quality interim (pre-post or more often) assessments and formative resources and tools (to provide professional learning to educators regarding gathering and using data to inform day-to-day instruction). LEAs currently procure interim assessments and formative assessment resources individually or in small groups (e.g., across Intermediate School Districts). This small-scale procurement is costly and creates significant challenges with regard to comparability. It will be much more cost-effective for the state to provide interim assessments and formative assessment resources online to LEAs, freeing up local resources and helping to ensure comparability across the state. This is essential as Michigan moves forward with implementing reforms such as educator evaluations. Another key factor is whether a provider's solution will increase or decrease the burden on local districts. Multiple questions addressed these issues as well.

Finally, a major driver is cost. As noted in this report's introduction, the cost information captured in this report only serves as a limited benchmark for off-the-shelf products. The only way to truly determine specific and detailed costs, at the student level or otherwise is to complete the full state procurement process. As all providers were within a reasonable ballpark on prices, all were identified as meeting this threshold. The full state procurement process has been completed recently for all aspects of test development, and is in the final stages of being completed for test administration. This process (from issuing an RFP through signing contracts with successful service providers) currently takes approximately eighteen months. The contracts currently in place or that are being finalized are scheduled to expire after the spring of 2016.

MDE had been proceeding with implementation of the CCSS and participating in the development of Smarter Balanced for three years, with the aim of ensuring at least one viable option for an assessment system aligned to CCSS is available to the state. At the time that these contracts were being prepared, Smarter Balanced was the only viable option available to the state, and as this report demonstrates, it remains the only viable option that can satisfy all of the multiple needs for test security, student data privacy, a Michigan governance role, Michigan educator involvement, minimizing local burdens, cost effectiveness, Michigan access to all data to allow for verification, and so on. Because Smarter Balanced was designed primarily by state assessment directors who understand these needs, this should not be a surprising result.

RECOMMENDATIONS

The state procurement process is lengthy in great part because there are appropriate protections built into the system. It also takes significant time, once a contract is signed, for vendors to get systems in place to serve the needs of Michigan students, schools, districts, and the state. Because of these time constraints, adopting a different solution at this time will result in not having an assessment for the 2014-15 school year, and would likely result in not having an assessment for the 2015-16 school year, putting MDE in violation of both state and federal law. This is the case even with continuing forward with MEAP, where development has been ceased to avoid unnecessary costs.

As the current contracts expire after the spring of 2016, it presently takes approximately eighteen months to complete the formal state procurement process, and it takes time for a new contractor to put systems in place, MDE recommends developing and issuing a new RFP in late 2014 that incorporates information from this report. Contracts put in place from that RFP process will be geared toward delivering summative, interim and formative solutions beginning with the 2016-17 school year. By issuing a new RFP in the fall of 2014, we hope that more providers will be able to put forward a product that can meet Michigan's needs at that time. MDE is agnostic regarding what solution is ultimately chosen for the 2016-17 school year and beyond, as long as it meets Michigan's needs.

FINAL NOTE

One potential avenue for assessing student achievement against the Common Core that (due to the timeframe) could not be produced for this report is that MDE has the capability to develop a customized, high-quality assessment in-house. Technology solutions such as MDE's Item Banking System, Secure Site, and existing service provider systems are already in place. MDE is currently developing a suite of interim assessments across grades 3-12 in science and social studies, to support reform efforts such as educator evaluations. This path gives the state complete control over such things as alignment to standards and test administration procedures, and makes sense in those content areas as there are no multi-state consortia or test companies that have developed tests specifically to measure Michigan's science and social studies content.

MDE has not been pursuing this path for English language arts and Mathematics, as the in-house approach does not permit the state to take advantage of the resources available from a consortium of states working on rich solutions to measure the same content (i.e., the Common Core State Standards). For example, the MDE-developed science and social studies will have a limited number of item types (e.g., constructed-response) and resources to support professional development. Without substantial new funding, MDE would not be able to develop its own assessment with the rich item types necessary to adequately measure the level of knowledge and skill described by the CCSS. The economies of scale provided by working with a consortium have allowed Michigan to avoid those substantial new funding needs.

REFERENCES

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). Standards for Educational and Psychological Testing. Washington, D.C.: American Educational Research Association.

The Council of Chief State School Officers (2013). CCSSO Assessment Quality Principles. Washington, DC: Author.

The Council of Chief State School Officers & The Association of Test Publishers (2013). Operational Best Practices for Statewide Large-Scale Assessment Programs-2013 Edition. Washington, DC: Author.

U.S. Department of Education. (2009). Standards and assessments peer review guidance: Information and examples for meeting requirements of the No Child Left Behind Act of 2001. [Revised December 21, 2007 to include Modified academic achievement standards. Revised with technical edits, January 12, 2009] Washington, DC: Author.

MICHIGAN Department of Education

Phone: 1-877-560-8378
Website: www.michigan.gov/baa
Email: baa@michigan.gov