Maryland State Department of Education
Assessment and Accountability Task Force
Minutes from Meeting 5
September 5, 2024

Welcome

Dr. Carey Wright welcomed all the attendees and thanked them again for participating in this important work. Chris Domaleski from the Center for Assessment shared updates since the last Assessment and Accountability Task Force meeting. These updates included the formation of two subcommittees on English language proficiency and post-secondary readiness.

Topic 1: Assessment Constraints and Requirements

Next, Scott Marion shared an overview of the constraints and requirements of state assessments. He emphasized that no assessment can do everything well. This means that the instructional utility of state assessments is challenging because test scores come at the end of the year and have the primary purpose of serving as an accountability and summative assessment role. Next, Scott briefly reviewed the purposes and uses assessments must support and the limits of what the summative assessment can provide.

Topic 2: Feedback on the Current System

The Task Force provided feedback on the current assessment system, discussing its strengths and weaknesses. Participants appreciated the ease of administration, flexibility in scheduling, quick access to student performance data, variety of question types, and accommodations for special education students. However, concerns were raised about the accuracy of questions, sufficiency of Algebra items, and trust in results. Issues included inconsistency in score distribution, difficulty interpreting scores due to non-released items, and insensitivity of score reporting to student growth at lower proficiency levels. Additional concerns involved the alignment of alternative assessments with MCAP results, the impact of writing assessments on result sharing, appropriateness of performance standards, alignment with district curricula, and clarity of scoring rubrics. The length of the test and its impact on instructional time, especially for students with accommodations, was also noted. The need for more parent-friendly score reports and timely reporting for system planning was emphasized.

Topic 3: Needs, Wants, and Credibility

Next, Task Force members discussed the needs of various constituent groups from the assessment system, with key themes of credibility and transparency emerging. Parents need clear, timely information on how state assessments relate to their child's overall academic performance, including actionable steps based on scores. Teachers require concise, impactful assessments that report student performance against standards. District administrators need valid, reliable performance data broken down by demographics and compared across similar districts. State education leaders need data that allows for comparisons between schools and districts, comparisons over time, and alignment with national assessments like NAEP. This includes disaggregated data by federal student groups, user-friendly access to information, transparent technical details, and research functions to address learning loss recovery and college/career readiness. During the group discussion, participants emphasized the need for clear communication on how state assessments fit with other state-administered tests and the importance of clear communication about student performance.

Topic 4: Adaptive vs. Fixed Form

Scott Marion then shared some of the strengths and limitations of adaptive versus fixed-form assessments. He reviewed two types of adaptive tests: item-level adaptive and multistage adaptive. Scott answered several clarifying questions about multi-stage adaptive testing in the follow-up discussion. The groups then discussed the pros and cons of each approach at separate tables and reported their thoughts to the

entire group. Opinions varied among different tables. After discussion, the consensus across groups was that the test's purpose should guide the decision but that either multistage adaptive or fixed form would be preferred for accountability and transparency.

Topic 5: Testing Time

Chris Domaleski reviewed constraints related to testing time and assessment length, including examination, preparation, and administrative support times. He shared the current state testing examination lengths for English language arts (3-5 hours) and Mathematics (2-4 hours). The group considered ways to reduce testing time, such as adaptive testing, replacing longer item types with shorter ones, and removing items. The possibility of through-course assessments was discussed, with concerns raised about interrupting instruction and the administrative burden on school districts.

Participants suggested how tests could be shortened with AI-scored technology-enhanced items (TEI). However, there were concerns about losing the depth of measurement. Additional concerns included the impact on instructional time for non-tested grades, the administrative burden when schools lack one-to-one digital devices, and the inefficiency of long testing blocks. Some participants voiced concerns about administrative burdens and suggested policy changes. The discussion also touched on the state statute limiting testing time to 2.2-2.3% of instructional time, including state, district, and local assessments. Ultimately, the task force emphasized the need for clarity on the purpose of assessments and prioritizing values to inform decisions about test length. Clear communication on the value of assessments was also highlighted as essential.

Topic 6: Summative/ System Considerations

Scott Marion shared with the task force the differences between formative, interim, and summative assessments. This presentation included both the purpose of each test type as well as the design features and the information each type provides. He also shared how other states have developed interim assessments collectively while others have developed their own benchmark assessments. Participants asked various questions about how summative and interim testing fits into the overall assessment system. The consensus of participants was that unless state interim assessments were replacing another assessment, districts would not welcome it. In addition, several participants shared that districts would not be willing to give up their interim assessments. Several questions were asked about if the state could simply purchase the most commonly used off the shelf interim. Other comments shared by participants included a need for trust, trend data, professional development supports, alignment, and coherence with interim assessments.

Topic 7: Timeline for New or Updated Assessments

Finally, Scott Marion reviewed the timeline for making changes to the state assessment and other considerations such as custom-made, consortium, or collection of states working with the same testing vendor. Participants asked questions about the impact of different design decisions on the timeline for development and the timeframe for receiving test scores on an annual basis.

Wrap Up and Next Steps

Chris wrapped up the meeting by summarizing key insights and reviewing action items. The meeting concluded with participants completing a meeting evaluation.

The next meeting will be held remotely on October 15, 2024.