House Public Education Committee

Testimony on State Assessment System

Dr. Mike Waldrip, Superintendent

Recommendations

Utilizing Performance-Based Assessments

Frisco ISD is one of the highest performing school districts in the state. Frisco ISD is one of only two districts with over 50,000 students to receive an "A" in the state's accountability system and it is the highest performing district with over 50,000 students, having received a 96. Yet, this system is based on standardized test scores and is failing Texas students. Current assessments do not allow for students who are poor test-takers, who have a non-educational emotional or social struggle on the day of the exam, or for students who may be missing a single, but related piece of knowledge. As a result, many students score below their actual mastery level on standardized testing. However, the STAAR also fails to fully assess student performance and can lead to a student scoring above their mastery level on standardized tests and masking a problem that needs to be addressed.

In order to address the concerns related to standardized testing, performance-based assessments should be utilized, *in addition* to STAAR testing.

Performance-based assessments allow educators to measure growth in real-time, allowing for proactive adjustments in instruction throughout the year as opposed to the STAAR which only allows for a reactive approach. These assessments provide additional data points and a more holistic review of the student's mastery of content. For example, the Developmental Reading Assessment (DRA)[1] is a proven diagnostic reading assessment that establishes each student's reading level and gives the teacher a focus for future instruction. DRA measures specific skills related to the three components of reading: reading engagement, oral reading fluency, and comprehension. The results provide teachers with important information about each student's reading behaviors, which allows them to make sound decisions to help all learners become proficient, motivated, independent readers. Unlike the STAAR,

[1] The Developmental Reading Assessment is an assessment tool created by Pearson Education.

DRA allows teachers to monitor student reading progress in real time in order to provide students with immediate feedback and respond to their needs through targeted reading instruction.

The DRA also provides teachers with information about a student's Reading Engagement, which the STAAR test does not. Research shows that good readers have developed the following strategies, skills, and habits, which the DRA measures:

- enjoy reading, read often, and read a wide variety of genres to meet multiple purposes.
- successfully select texts that match their reading level, interests, and purposes.
- are confident about their ability to read, are aware of their strengths as reader, and are goal-directed.

The DRA provides teachers with information about a student's Oral Reading Fluency, which the STAAR test does not. Research shows that good readers have developed the following strategies, skills, and habits, which the DRA measures:

- read appropriately leveled texts with a high level of accuracy, monitor meaning, and use fix-up strategies to quickly self-correct miscues that interfere with meaning.
- read quickly and smoothly.
- read in longer, meaningful phrases with effective expression.

The DRA provides teachers with information about a student's Comprehension, which the STAAR test only provides if a student is able to read the text. For example, a student on a 2nd grade reading level may not be able to read a 3rd grade reading level passage. As a result, the STAAR test would indicate the student's comprehension is below grade level, when in reality, the issue may be reading fluency, not comprehension. The data provided by STAAR selected response questions only allows teachers a narrow view of a student's reading abilities. Research shows that good readers have developed the following strategies, skills, and habits, which the DRA measures:

- review texts, making predictions about what is likely to happen or identifying topics and information that may be included.
- ask themselves questions prior to and during the reading of a text.
- use text features and graphic organizers.
- comprehend what they read and are able to use their own language and key vocabulary from the text to identify and organize important information into a written summary.
- understand what is explicitly stated in the text.
- interpret what they read by making inferences and connections.
- support their responses with evidence from the text and personal examples.

- are aware of the strategies they use to construct and monitor meaning while reading.

The DRA is only one example of a performance-based assessment. However, performance-based assessments, such as the portfolio assessment method, can be implemented across subjects.

Reduced Reliance on STAAR to Measure Growth

Currently, the STAAR is administered in every grade from 3rd through 8th grade. The results of STAAR are used to measure student achievement and growth. However, STAAR has limitations related to measuring student growth. Although STAAR, or some form of standardized testing, is necessary to comply with federal law, and the data it provides is important, STAAR is not required to be used to measure student growth. Texas should reduce its reliance on STAAR for this metric and turn to performance-based assessments instead.

Texas should reduce the reliance on STAAR to measure growth and instead utilize performance-based assessments.

Student growth is arguably more important than absolute achievement. Unfortunately, one of the limitations of STAAR is that it cannot accurately measure student growth over a school year. Because each STAAR administration is given at the end of every year, the baseline for student growth is inaccurate. Research shows that many students regress over the summer. As a result, the current assessment system expects teachers to make up the gap caused by the summer break by assuming a student is at the same level at the beginning of a grade as they were at the end of the previous grade. Performance-based assessments can be more easily administered throughout the year and therefore provide a better baseline to measure growth.

Reduced Impact of Cross-Subject Testing

On the 2018 3rd Grade Math STAAR test, every single question required reading to solve. Although accommodations are available for students who have been identified as having some sort of reading difficulties, many students who are simply behind receive no such accommodations. As a result, a 3rd grader who is either unable to read the questions on the Math STAAR, or who is too slow to read all the questions in the time allotted, may significantly underperform his or her actual ability. Indeed, all STAAR exams require students to read questions, yet accommodations are not available for students who are simply behind in reading ability. As a result, their reading deficiencies cause underperformance in other subjects.

To reduce the impact of cross-subject testing, STAAR administrations should be modified to allow for oral administration of assessments to any student performing below grade level in reading.

To accurately determine the achievement of a student in any given subject, the student must be able to understand a question. Allowing for oral administration to students who perform below grade level in

reading ensures that student scores on other subjects are an accurate representation of their mastery of those subjects, rather than skewed as a result of their reading deficiencies.

Reduced STAAR Administration Time

Because of the stakes attached to the STAAR exam, many schools drastically alter operations during STAAR administration to prevent any distractions and allow for teachers to proctor the STAAR. Because of the length of the STAAR administration, this disrupts the day of students in non-testing grades. In addition, many students, particularly those with reading deficiencies, simply do not have the stamina to complete such a long exam.

To reduce the impact of STAAR administration on non-testing students and to decrease the likelihood of fatigue for testing-students, the STAAR should be shortened or broken up into smaller chunks to allow for administration of each STAAR in a single class period.

Prolonged testing administration can cause fatigue, particularly for those students who have to struggle simply to read and understand the questions and students who struggle to sit still for long stretches of time. Often these students are not eligible for accommodations, or the accommodations do not truly address the problem. As a result, students are more likely to give up make mistakes as a result of being tired. This can also lead to underperformance for students. Making STAAR exams short enough to be administered during a single class period would mitigate these issues. Additionally, implementing the performance-based assessments discussed above would replace any loss of data resulting from shorter STAAR tests.

Conclusion

Relying solely on high-stakes, standardized testing to measure student achievement and growth subjects students to undue stress and frustration. These tests, used alone, fail to accurately and adequately measure how Texas students are performing and in what areas instruction should be focused. STAAR places the emphasis on a single data point, rather than viewing students holistically over many data points. Utilizing performance-based assessments and making changes to the way STAAR is administered can correct many of these problems and provide a more accurate assessment of how well Texas students are being educated.