

UT researchers pour cold water on STAAR concerns



University of Texas researchers evaluated the State of Texas Assessments of Academic Readiness and delivered their findings to the Texas Legislature on Monday. [FILE PHOTO]

By Julie Chang

University of Texas researchers found that the state's standardized tests are not too advanced for the grade levels of the students taking the tests, [according to a study released Monday](#).

The findings blunt criticism of the State of Texas Assessments of Academic Readiness, used to determine whether public school students in some grades can advance to the next grade level and for evaluations of school districts and campuses.

[State lawmakers required an audit of the STAAR](#) this year, amid results from other studies that had concluded the tests were too difficult for children to read and thus were not reliable testing instruments. House Bill 3, the overhaul of the school finance system approved earlier this year, included a provision to evaluate whether STAAR passages and questions are written at an appropriate grade level for students to understand, and whether the tests align with the state's curriculum standards.

The results were delivered to the Legislature on Monday.

Researchers at UT's Meadows Center for Preventing Educational Risk found that between 93% and 100% of questions in each of five subjects — 17 tests in total across third through eighth grades in the 2018-19 school year — aligned with curriculum standards, also called the Texas Essential Knowledge and Skills. Researchers could not determine whether questions on those tests were readable because of a lack of reliable evaluation tools. But, they determined that most passages in the reading and writing tests were readable — between 86% and 97% depending on the methodology used.

"TEA is pleased with the report's findings which show our assessments overwhelmingly align with the Texas Essential Knowledge and Skills (TEKS), with readability appropriate to students' grade level," the Texas Education Agency said in a statement. "The agency thrives on a culture of continuous improvement and studies like these provide us with additional benchmarks that we can effectively learn from in an effort to better serve Texas students."

Previous studies

Critics of the STAAR had called on the state to produce a meaningful assessment of the test because so much rides on how students perform on it. STAAR scores determine whether fifth and eighth graders can advance to the next grade level as well as whether high school students can graduate. How the state evaluates schools and school districts annually also heavily depends on student performance on the STAAR.

Lawmakers asked for the study to be included in legislation after researchers from Texas A&M University-Commerce in March found that the ease with which passages in the reading tests can be read and understood was above grade level for students in third through eighth grades. A study by the same university in 2012 and another by University of Mary Hardin-Baylor researchers in 2016 also found the STAAR to be too difficult.

"What is clear from this report is that the state of Texas is asking the right questions. ... The question that may never be answered is, 'Are the STAAR questions and passages too hard for students to read?'" said Emily Sass, an education policy director for conservative think tank Texas Public Policy Foundation. "As noted in this report, determining readability is a far from exact method of measure and results can vary widely by metric chosen and passage style."

Texas Education Agency officials at the time defended the STAAR, saying the Legislature had required an independent evaluation of the STAAR that was completed in 2016. It found that tests yielded "valid and reliable assessment scores." The agency also regularly invites teachers to comb through the STAAR for problems.

The Legislature will receive a second study that will evaluate the 2019-20 STAAR by Feb. 1.

Readability

Researchers tested readability of the STAAR based on three indices: word and sentence length as well as difficulty; the structure of the text; and the vocabulary or the likelihood of a text to contain familiar words.

They said they could not apply their methodology to measuring the readability of question items because question items have too few words. Researchers noted they were not told to evaluate the difficulty of the questions, only the readability.

“We approached the evaluation of the readability of items with a great deal of caution,” researchers said in the report.

Researchers concluded a passage was readable if it fell within or below grade level on two of the three indices. Using one method, they found 30 of the 35 reading and writing STAAR passages evaluated fell within or below the grade level. Using another method that researchers said was better suited for measuring informational texts, all but one of the passages were considered readable.

“This study confirms that the STAAR test measures our state standards. Now, we should all turn our focus to ensuring every Texas student can demonstrate mastery of these important skills each year,” said Will Fullerton, executive director of Texas Aspires, which has supported the test.