

Apples, Oranges, and Inside Baseball: Key Components to Understanding 2017 STAAR Results

The Texas Education Agency's recent release of statewide 2017 STAAR results seems to have generated more heat than light about student achievement: Did performance improve, decline, or stay the same compared to prior years? In each of the three newspaper articles listed below, the reporters attempt to explain what they believe is a “drop” in STAAR scores when the 2017 results are compared to the 2016 results. Their work provides classic examples of how “apples-to-oranges” comparisons and the need to know technical details – “inside baseball” – affect accuracy and simplicity in reporting. Each of these issues is discussed in more detail below.

- July 7, 2017: Dallas Morning News - Texas students losing ground on STAAR tests <https://www.dallasnews.com/news/education/2017/07/07/texas-students-losing-ground-staar-tests>
- July 10, 2017: Drops in Texas STAAR scores raise questions — and testing angst. <http://www.mystatesman.com/news/drops-texas-staar-scores-raise-questions-and-testing-angst/60anViKo3ipmhKYTQKJ7QN/>
- July 13, 2017: Houston Chronicle - Newest STAAR exam results add up to difficult math problem for HISD, local districts <http://www.chron.com/news/education/amp/Fewer-Texas-Houston-students-pass-state-s-11286107.php>

Apples-to-Oranges

Most people are simply unaware that trying to compare the 2016 STAAR results to the 2017 results is like comparing “apples-to-oranges.” To understand that the two years’ results are not comparable requires one to know, remember, or even *find* information about the multiple changes made to the 2017 STAAR tests which render their results different from the prior year’s results. That is, one has to know a little “*inside baseball*.”¹

[Commissioner Morath’s January 30th blog](#) outlines a list of changes that TEA made to the STAAR program to address “*multiple issues on multiple fronts*.” The specific changes for 2017 that were listed in his blog are shown below, and provide evidence as to how the 2017 tests differ from the 2016 tests.

- Shortening the grade 4 and grade 7 writing assessment (from a two-day) to a one-day administration.
- Designing and implementing shorter test blueprints for the 2017 administration of STAAR in grades 3–8.
- Redesigning the three English end-of-course blueprints for the spring 2017 administration so that the reading sections no longer contain short answer questions.
- Incorporating changes to the Algebra I curriculum to correspond to the end-of-course assessment.
- **STAAR A and STAAR L were administered for the final time in December of 2016.** Beginning with the spring 2017 administration, the embedded supports that were available on these assessments are now individual embedded online accommodations on STAAR for eligible students. Information about these embedded supports and other accommodations, or Designated Supports, can be found on the [Accommodation Resources](#) webpage. (*emphasis added*)
- Accommodations policies were revised so that more of the decisions regarding the appropriateness of an accommodation for an individual student are made at the local level. **In most cases, accommodations eligibility widened significantly.** (*emphasis added*)

And that’s not all of the “inside baseball” one needs to know! In the past, TEA provided different paper versions of the STAAR tests. Regular STAAR, STAAR L, STAAR A, STAAR M (which was discontinued in 2015), STAAR Alternate (which was discontinued in 2015) and STAAR Alternate 2 were unique test versions, each

designed to assess the learning of specific groups of students. (For more information about unique STAAR test versions, please refer to TEA's Testing webpage at http://tea.texas.gov/Student_Testing_and_Accountability/Testing/)

TEA changed the testing program with the Spring 2017 STAAR administration. Gone are the different paper test versions; instead, students who in previous years took a different test version, now are required to log into a computer to take an online version of STAAR that contain a variety of accommodations, or designated supports, which are "*changes to materials or procedures that enable students to access learning and testing.*" TEA's [Accommodation Resources](#) webpage lists three main categories of accommodations and/or designated support options for students:

1. **Accessibility Features** – these are NEW features beginning with the Spring 2017 test administration. "*These are procedures and materials that are allowed for any student who needs them.*"
2. **Designated Supports** – which requires ARD, LPAC, and 504 committee approval before a student may access these supports
3. **Designated Supports Requiring TEA Approval** – before a student may access these supports

Therein lies part of the problem. TEA's press release **did not** mention that embedded support test results were included in the 2017 statewide STAAR summary reports – which is quite different from how results were reported in prior years. The agency **did not** include a footnote on the statewide summary reports to alert readers to this fact: the summary report titles look identical to prior year summary report titles. TEA **did not** publish a *separate* summary report showing the number of students who took an embedded support test and the percent who passed that test. Without a little inside baseball – knowing these substantive changes, or where to find the information about them – most people could not realize that comparing the 2016 to the 2017 STAAR results was largely inappropriate, just like comparing apples to oranges. (Appendix A shows the apples-to-oranges comparison.)

Red Apples-to-Green Apples

Unfortunately, it is impossible to prepare the most appropriate comparison between the two years' results with the current, publicly available data. To create a true "apples-to-apples", year-over-year (YOY) comparison, the agency must release the embedded support test results separately so that recalculations of the most accurate YOY comparisons can be made.

However, MCA can (and did) create a "red apples-to-green apples" YOY comparison by aggregating the 2016 STAAR, STAAR L, and STAAR A statewide test results and then calculating the one-year change relative to the 2017 STAAR with embedded support test results. MCA finds that if this type of modified comparison were used (which we acknowledge is still not the most statistically accurate comparison), the headlines would be very different from what was reported using the apples-to-oranges comparison. When comparing "red apples-to-green apples":

- Reading for all grade levels, except grade 7, declined slightly. 7th grade reading improved by three percentage points. And instead of a 7 percentage point drop in 4th grade reading and a 6 percentage point drop in 8th grade reading as was reported, scores dropped 5 and 3 percentage points, respectively.
- Math scores increased at all grade levels. The increases ranged from a one percentage point gain in 7th grade to five percentage point gains in both grades 5 and 8.
- First-time EOC testers improved on 3 out of the 5 subjects, including Algebra I, which increased by five percentage points. Performance in the other 2 subjects (Biology and English II) remained relatively flat.

(Appendix A shows the red apples-to-green apples comparison.)

What else is important to know?

It was surprising to see that the number of students testing in 4th grade this year increased dramatically from the number tested in 2016. More than 13,000 additional students tested in 4th grade reading (and math and writing) in 2017 than were tested in 2016 (as shown in Appendix B).

What we **do not** know is how many students took an embedded test and how they performed as a group. We don't know who / which students took an embedded test. We don't know if there is an over-representation from one particular student group or not. And we also don't know how or even if this significant increase in the number of students has impacted teaching, learning and student performance on STAAR.

Until TEA publishes the full complement of facts about the 2017 embedded support test results, opinions will continue to vary as to what the YOY change in STAAR performance mean on a statewide basis. Cypress-Fairbanks ISD, and all other districts that are treating the 2017 STAAR with embedded support tests as new baseline results (as per the Houston Chronicle article), provide some much-needed light on the very heated debate about student performance in 2017. Others may wish to emulate these exemplars. In addition, districts would be wise to prepare for a much larger group of 5th grade students in SY 2017-18, who may need a variety of new designated supports and accelerated instruction to be successful in the upcoming school year.

¹ The phrase "**inside baseball**" is a term used mainly in the United States. It usually refers to a detail-oriented approach to the minutiae of a subject, which in turn requires such a specific knowledge about what is being discussed that the nuances are not understood or appreciated by outsiders. [https://en.wikipedia.org/wiki/Inside_baseball_\(metaphor\)](https://en.wikipedia.org/wiki/Inside_baseball_(metaphor))

STATEWIDE STAAR RESULTS – 2016 COMPARED TO 2017

Apples to Oranges2016 excludes STAAR A. 2017 includes embedded support tests

STAAR Reading	Spring 2016	Spring 2017	1-year Diff.
Grade 3	74%	72%	-2
Grade 4	77%	70%	-7
Grade 5*	75%	71%	-4
Grade 6	71%	67%	-4
Grade 7	72%	72%	0
Grade 8*	82%	76%	-6

Source: TEA Spring 2016 and 2017 Statewide Summary Reports.
Excludes Spanish results. *1st administration results only.

Red Apples to Green Apples2016 includes STAAR A. 2017 includes embedded support tests

STAAR Reading	Spring 2016	Spring 2017	1-year Diff.
Grade 3	73%	72%	-1
Grade 4	75%	70%	-5
Grade 5*	73%	71%	-2
Grade 6	68%	67%	-1
Grade 7	69%	72%	+3
Grade 8*	79%	76%	-3

Source: TEA Spring 2016 and 2017 Statewide Summary Reports.
Excludes Spanish results. *1st administration results only.

2016 excludes STAAR L and A. 2017 includes embedded support tests

STAAR Math	Spring 2016	Spring 2017	1-year Diff.
Grade 3	76%	76%	0
Grade 4	74%	75%	+1
Grade 5	79%	81%	+2
Grade 6	74%	75%	+1
Grade 7	71%	68%	-3
Grade 8	73%	74%	+1

Source: TEA Spring 2016 and 2017 Statewide Summary Reports.
Excludes Spanish results. *1st administration results only.

2016 includes STAAR L and A. 2017 includes embedded support tests

STAAR Math	Spring 2016	Spring 2017	1-year Diff.
Grade 3	74%	76%	+2
Grade 4	72%	75%	+3
Grade 5*	76%	81%	+5
Grade 6	71%	75%	+4
Grade 7	67%	68%	+1
Grade 8*	69%	74%	+5

Source: TEA Spring 2016 and 2017 Statewide Summary Reports.
Excludes Spanish results. *1st administration results only.

Apples to Oranges2016 excludes STAAR A. 2017 includes embedded support tests

STAAR Writing	Spring 2016	Spring 2017	1-year Diff.
Grade 4	69%	63%	-6
Grade 7	70%	68%	-2

Source: TEA Spring 2016 and 2017 Statewide Summary Reports.
Excludes Spanish results.2016 excludes STAAR L and A. 2017 includes embedded support tests

STAAR Science	Spring 2016	Spring 2017	1-year Diff.
Grade 5	75%	73%	-2
Grade 8	76%	74%	-2

Source: TEA Spring 2016 and 2017 Statewide Summary Reports.
Excludes Spanish results.2016 excludes STAAR L and A. 2017 includes embedded support tests

STAAR Soc.Studies	Spring 2016	Spring 2017	1-year Diff.
Grade 8	65%	62%	-3

Source: TEA Spring 2016 and 2017 Statewide Summary Reports.

Apples to Oranges - EOCs2016 excludes STAAR L and A. 2017 includes embedded support tests

1st-Time Tested	Spring 2016	Spring 2017	1-year Difference
Algebra I	85%	87%	+2
Biology	92%	88%	-4
English I	71%	70%	-1
English II	74%	71%	-3
U. S. History	95%	93%	-2

Source: TEA Spring 2016 and 2017 Statewide Summary Reports.

Red Apples to Green Apples2016 includes STAAR A. 2017 includes embedded support tests

STAAR Writing	Spring 2016	Spring 2017	1-year Diff.
Grade 4	68%	63%	-5
Grade 7	67%	68%	+1

Source: TEA Spring 2016 and 2017 Statewide Summary Reports.
Excludes Spanish results.2016 includes STAAR L and A. 2017 includes embedded support tests

STAAR Science	Spring 2016	Spring 2017	1-year Diff.
Grade 5	73%	73%	0
Grade 8	73%	74%	+1

Source: TEA Spring 2016 and 2017 Statewide Summary Reports.
Excludes Spanish results.2016 includes STAAR L and A. 2017 includes embedded support tests

STAAR Soc. Studies	Spring 2016	Spring 2017	1-year Diff.
Grade 8	61%	62%	+1

Source: TEA Spring 2016 and 2017 Statewide Summary Reports.

Red Apples to Green Apples - EOCs2016 includes STAAR L and A. 2017 includes embedded support tests

1st Time Tested	Spring 2016	Spring 2017	1-year
Algebra I	82%	87%	+5
Biology	89%	88%	-1
English I	68%	70%	+2
English II	72%	71%	-1
U. S. History	92%	93%	+1

Source: TEA Spring 2016 and 2017 Statewide Summary Reports.

Research Question: How many students tested in 2016 and 2017?									
STAAR 3-5	Spring 2016 Number Tested			Spring 2017 Number Tested			2016 to 2017 YOY Change		
	Grade 3	Grade 4	Grade 5	Grade 3	Grade 4	Grade 5	Grade 3	Grade 4	Grade 5
Reading	366,635	366,872	374,671	370,790	380,063	379,532	4,155	13,191	4,861
Math	385,912	381,929	383,082	389,765	395,332	387,610	3,853	13,403	4,528
Writing		365,319			378,591			13,272	
Science			380,807			385,853			5,046
STAAR 6-8	Spring 2016 Number Tested			Spring 2017 Number Tested			2016 to 2017 YOY Change		
	Grade 6	Grade 7	Grade 8	Grade 6	Grade 7	Grade 8	Grade 6	Grade 7	Grade 8
Reading	384,564	382,860	375,534	391,613	389,226	380,566	7,049	6,366	5,032
Math	378,151	354,546	320,150	384,610	355,878	324,154	6,459	1,332	4,004
Writing		382,533			389,357			6,824	
Science			377,881			382,710			4,829
SS			377,898			383,327			5,429
STAAR EOCs	Spring 2016 Number Tested	Spring 2017 Number Tested	2016 to 2017 YOY Change						
Algebra I	381,153	386,756	5,603						
Biology	376,665	382,319	5,654						
English I	382,739	383,641	902						
English II	360,733	368,728	7,995						
U.S. History	347,096	342,620	(4,476)						