

OnTarget Overview

Training Links in KB

Training Overview

The focus of this workshop will be to help educators explore the components of OnTarget. It will provide an overview of each of the components, including: Analyzing Local Assessments, Measure STAAR Growth, and Compare STAAR Assessments.

Analyze	
	Analyze Local Assessments Check each question on assessments for quality and validity. Upload Assessment Results File Upload Instructions

- TEA
 - The TEA Website
 - Technical Digest 2020-2021
 - o 2021 STAAR Mean P-Values
 - 2021 STAAR Score Distributions and Statistics
- Loading Exams Common Assessments
 - KB Article Loading Other Exams Common Assessment
 - o Common Assessment to OnDataSuite File Center Crosswalk
 - Common Assessment Template 080822
- Correlation Coefficient
 - Reliability
 - The concept of reliability is based on the idea that repeated administrations of the same test should generate consistent results. The degree to which results are consistent is assessed using a reliability coefficient. Reliability is a critical technical characteristic of any measurement instrument because unreliable scores cannot be interpreted in a meaningful way (TEA Technical Digest 2020-2021 p. 3-26).
 - Reliability coefficients based on one test administration are known as internal consistency measures because they measure the consistency with which students respond to the items within the test. As a general rule, reliability coefficients from 0.70 to 0.79 are considered adequate, those from 0.80 to 0.89 are considered good, and those at 0.90 or above are considered excellent (TEA Technical Digest 2020-2021 p. 3-26).



This graph demonstrates the reliability correlation coefficient for a particular STAAR assessment as released in the TEA in the Technical Digest for the most current available year, and the correlation coefficient as calculated by ODS for the assessment being reviewed.

- Assessment Analysis
 - P-Value
 - The p-value indicates the proportion of the total group of students answering a multiple-choice or gridded-response item correctly. An item's p-value shows how difficult the item was for the students who took the item. An item with a high p-value, such as 0.90 (meaning that 90 percent of students correctly answered the item), is a relatively easy item. An item with a low p-value, such as 0.30 (meaning that only 30 percent of students correctly answered the item), is a relatively easy item. An item with a low p-value, such as 0.30 (meaning that only 30 percent of students correctly answered the item), is a relatively difficult item (TEA Technical Digest 2020-2021 p. 3-14).



• Point- Biserial Correlation

The point-biserial correlation describes the relationship between a student's performance on a multiple-choice or gridded-response item (scored correct or incorrect) and performance on the assessment as a whole. A high point-biserial correlation indicates that students who answered the item correctly tended to score higher on the entire test than those who missed the item. In general, point-biserial correlations less than 0.20 indicate a potentially weaker-than-desired relationship. Note that the point-biserial correlation may be weak on items with very high or very low p-values. For example, if nearly all

students get an item correct (or incorrect), that item does not provide useful information for distinguishing between those students with higher performance from those students with lower performance on the entire test (TEA Technical Digest 2020-2021 p. 3-14).



• Item Analysis

Provides a general Overall Overview of the entire assessment at a glance and includes the Question Number, the Difficulty as calculated by the P-Value, The Question Quality as calculated by the Point-Biserial Correlation, and the Analyzed column which indicates if the item has been further analyzed for evidence of validity.

Item Analysis							
Question	Difficulty	Quality	Analyzed				
1	0.87	Fairly Good (0.28)	×				
2	0.96	Fairly Good (0.28)	×				
3	0.85	Good (0.37)	×				
4	0.71	Good (0.36)	×				
5	0.78	Fairly Good (0.27)	×				
6	0.88	Good (0.37)	×				
7	0.82	Fairly Good (0.29)	×				
8	0.84	Fairly Good (0.2)	×				
9	0.83	Good (0.31)	×				
10	0.73	Fairly Good (0.29)	×				
11	0.74	Good (0.3)	×				
12	0.84	Very Good (0.4)	×				
13	0.71	Good (0.34)	×				
14	0.76	Good (0.31)	×				
15	0.85	Good (0.3)	×				
Optimum Question Difficulty from .3 through .79							

- Statistical Details
 - Provides a comparison of the related STAAR test as released in the TEA in the Technical Digest for the most current available year.
 - Total Questions
 - Reliability Coefficient (Alpha)
 - Average Raw Score
 - Standard Deviation
 - Mean P-Value

Statistical Details

	State STAAR Test	My Test			
Total Questions	42	15			
Reliability Coefficient (Alpha)	0.91	0.7			
Average Raw Score	22.2	12.17			
Standard Deviation	9.37	2.6			
Mean P-Value	52.85	82.37			
STAAR Source: TEA, My Test Source: OnTarget					
"As a general rule, reliability coefficients from 0.70 to 0.79 are considered adequate, those from 0.80 to 0.89 are considered good, and those at 0.90 or above are considered excellent." — TEA Technical Digest 2018 - 2019					

- Student Performance By Demographics
 - Provides the Percentage Passing as denoted within the data file uploaded. ODS Does NOT determine passing, this is denoted by a standard set by the district and is included in the data file that is uploaded. ODS merely calculates the percentage based on the values identified by the data file uploaded.

Student Performance By Demographics								
###%TestedFailedPassedPassing								
All	266	42	224	84.21%				
Female	124	21	103	83.06%				
Male	142	21	121	85.21%				
Asian	24	1	23	95.83%				
Black or African American	22	6	16	72.73%				
Hispanic/Latino	96	20	76	79.17%				
Two or More Races	14	2	12	85.71%				
White	110	13	97	88.18%				

- Assessment Frequency Distribution
 - The number of students who obtained a particular raw score point on the assessment
- Standard Deviation
 - The standard deviation is another measure of statistical dispersion (variability or spread). It is an indicator of the degree of score variation around the mean.

- Mean / Average
 - The mean is a measure of central tendency. It is the average score for the assessment. It is computed by summing the scores of all students and dividing the sum by the total number of students (N).



- Question Analysis
 - Passage and Item Review Guidelines For STAAR, TEA's Curriculum Standards and Student Support and Student Assessment Divisions convene committees composed of Texas classroom teachers (including general education teachers, special education teachers, and bilingual and English as a second language [ESL] teachers), and curriculum specialists, to work with TEA staff in reviewing newly developed test items. TEA staff, along with Pearson and its subcontractors, train committee members on the proper procedures and criteria for reviewing newly developed items. Committee members judge each item for alignment, appropriateness, adequacy of student preparation, and any potential bias. Committee members discuss each test item and recommend whether the item should be field-tested as written or revised, recoded to a different TEKS or ELPS student expectation, or rejected (TEA Technical Digest 2020-2021 p. 2-6).
 - Table 2-1 shows the guidelines item-review committee members follow in their review (TEA Technical Digest 2020-2021 p. 2-6).

Table 2.1. Item-Review Guidelines

Passage and Item-Review Guidelines					
Reporting Category/Student Expectation Item Match	The item measures what it is supposed to assess.The item poses a clearly defined problem or task.				
Appropriateness (Interest Level)	 The item or passage is well written and clear. The point of view is relevant to students taking the test. The subject matter is of fairly wide interest to students at the grade being tested. The artwork is clear, correct, and appropriate. 				
Appropriateness (Format)	 The format is appropriate for the intended grade. The format is interesting for the student. The item is formatted so it is not unnecessarily difficult. 				
Appropriateness (Answer Choices)	 The answer choices are reasonably parallel in structure. The answer choices are worded clearly and concisely. The answer choices do not eliminate each other. There is only one correct answer. 				
Appropriateness (Difficulty of Distractors)	 The distractor is plausible. There is a rationale for each distractor. Each distractor is relevant to the knowledge and understanding being measured. Each distractor is at a difficulty level appropriate for both the objective and the intended grade. 				
Opportunity to Learn	The item is a good measure of the curriculum.The item is suitable for the grade or course.				
Freedom from Bias and Sensitivity Concerns	 The item or passage does not assume racial, class, or gender values or suggest such stereotypes. The item does not provide an advantage or disadvantage to any group of students because of their personal characteristics, such as race, gender, socioeconomic status, or religion. The item or passage avoids needless reference to topics that are extremely controversial or upsetting. The item or passage addresses sensitive topics in a careful, fair, and balanced way. The item fairly represents cultural, ethnic, social, and political diversity. 				

If the committee finds an item to be inappropriate after review and revision, it is removed from consideration for field testing. TEA field-tests the recommended items to collect student responses from representative samples across the state (TEA Technical Digest 2020-2021 p. 2-7).

Question Difficulty	0.58	Within Optimum Ra	nge		
Question Quality	0.31	Good			
er the following evidence of validity:					Action Taken
		\odot	Yes	No	Action Taken
		\odot	0	0	Adjust Question
			0	0	Adjust Responses
		۲	0	0	Clarify chart, graph, or picture
		\odot	0	0	Remove Item
		\odot			Notes
		\odot	Other not	es	
	Question Quality r the following evidence of validity:	Question Quality 0.31 If the following evidence of validity:	Question Quality 0.31 Good er the following evidence of validity: Image: Constraint of the second of	Question Quality 0.31 Good er the following evidence of validity: Image: Second	Question Quality 0.31 Good rt the following evidence of validity: Image: Control of Contro of Contro of Control of Control of Contro of Control of Contro

This table demonstrates the process outlined by ODS which follows the same process as outlined by TEA.

\otimes
and Skills (TEKS) student expectation (SE)
nd skill (identified by the cognitive verb) in the SE? (identify, describe, compare, analyze, etc.)
\odot
, ethnic, religion, socioeconomic, political, environmental, etc.)?
s?
ackgrounds so that students of one group do not have an unfair advantage over students of another group? (consider gender, rural/urban,
a a

Langua	ge and \	/ocabulary	\odot
Is the	item		
Yes	No		
0	0	Written using the appropriate verb tense throughout?	
0	0	Free of grammatical and spelling errors?	
0	0	Clear and concise, avoiding wordy, ambiguous, vague, irrelevant and unnecessary information and verbiage?	
0	0	Free of inappropriate colloquial and idiomatic language?	
0	0	In active voice rather than passive voice (unless passive voice is necessary or easier to understand)?	
0	0	Free of vocabulary and academic language that are not grade-level appropriate?	
0	0	Free of words with multiple meanings?	
0	0	Free of unnecessary or unclear pronouns?	
0	0	Free of complex lengthy clauses and sentences?	
0	0	Using consistent language when referring to the same object or concept?	

Does the item...

Yes	No	
0	0	Have a question, task, instructions, etc., that will be clear to students?
0	0	Use technology in a way that closely aligns to the item's content/skills SEs?
0	0	Avoid any clueing which may inappropriately influence a student's response to an item?
0	0	Have parallel structure so that the stem and answer choices make sense and and answer choices are similar in length, language, and structure?
0	0	Have a context that is clear, grade-level appropriate, and free from unnecessary complexity?

Answer Choices

Do the answer choices...

Yes No

- \odot $-\odot$ Include distractors in MC items that are plausible errors or misconceptions yet incorrect?
- \odot \odot Include distractors that are based on content that students for this grade level are expected to know?
- O O Avoid distractors that are too close to the correct answer that is likely to confuse or trick students who really do know the answer or can be considered outliers?
- O O Have only one correct answers (MC items) and is it marked as correct?

Does the item writer...

Yes No

O O Include rationales, where appropriate, that are complete (sufficiently explained Key/distractors) and written in acceptable style?

Visuals			\odot
Does t	he item		
Yes	No		
0	0	Only include art/table that provides support for the student to demonstrate proficiency of the standard?	
0	0	Include enough information for the item to be answered	
0	0	Include art/table that is legible, clear, and free from visual clutter?	

• KB Article Keywords

OnTarget	Common Assessment	Common Assessment Crosswalk
OnTarget Common Assessment File Format	Loading Common Assessments from Edugence	Loading Common Assessments from Eduphoria
Locally Developed Assessments	Combine	Combining
Local Assessments	Local Tests	

Additional Knowledge Base Articles

- How does OnTarget Determine is an Assessment is Valid And Reliable?
- Item Analysis
- <u>Reliability</u>
- <u>Validity</u>
- <u>A Locally Developed Assessment</u>
- Question Quality (point-biserial correlation)
- <u>Combining Local Assessments</u>



- Student Growth
 - Student Growth is a statistical analysis which utilizes the most current subject specific STAAR Raw Score mean and standard deviation to convert individual student raw scores into standardized z-scores. A standard score is a raw score expressed in terms of how many standard deviations it is away from the mean. This statistical process transforms the original raw scores into a common unit of measurement which allows OnTarget to compare variables which measure scores from different scales.
 - Statistical Details
 - Identifies average student growth by teacher as compared to the performance of the population in the state. By definition, the average standard score of the population in the state would be 0. Any number greater than 0 equals greater than average performance as compared to the population in the state. Conversely, any number less than 0 would equal less than average performance as compared to the population in the state.

Teacher Average Growth Report 2022 - At Least 1 Student



This graph demonstrates the average student growth by teacher as compared to the performance of the population in the state as denoted through a normal distribution. The area on the y-axis denotes the number of students served by each teacher. The area on the x-axis denotes where that teacher's average student growth lies as represented by a standardized z-score.

- Growth Statistics
 - Represents the Total Students, mean, standard deviation, and standard error of the mean as represented by the number of students served in the district.

Growth Statistics						
	Total Students	Mean	Standard Deviation	Standard Error of the Mean		
Previous Year Scores for Same Students (2021)	12,512	0.43	1.01	0.01		
Current Filtered Test (2022)	12,512	0.83	0.93	0.01		

- Growth Breakdown by Levels
 - Represents the number of teachers and percentages that fall between each of the levels as represented by standardized z- scores. These levels can be adjusted by each district. If the levels are adjusted, the percentages will reflect the adjustment.

Print/Download

Growth Breakdown by Levels										
	Level 1 Level 2 Level 3 Level 4									
Level Range (Average Growth)	-2 to -1	-1 to -0.5	-0.5 to 0.5	0.5 to 1	1 to 2					
Count	0	3	209	144	5					
Percentage	0%	1%	58%	40%	1%					

• Teacher Average Growth Report

Provides an aggregate report of all of the teachers by subject area. The column include teacher name, campus, subjects taught, grades taught, total # of Students, GRowth of students, Total # of Eco Dis students, Total # of ESL Students, Total # of Sped students.

Show 10 ¢ entries								
Teacher Name	Campus(es)	1↓ Subject(s) Taught 1↓	Grade(s) Taught	Total # of Students	Growth of Students	Total # of Eco Dis Students ↑	Total # of ESL Students ↑↓	Total # of SPED Students
Acevedo, Sonia	Lando Calrissian Middle School	Reading	07	99	0.46	7	6	1
Allen, Carlo	Darth Vader Middle School	Reading	07	13	0.9	4	0	2
Anderson, Addison	Bail Organa Elementary School, Mon Mothma Elementary School	Math	04	13	0.36	1	5	2
Atkins, Carli	Anakin Skywalker Elementary School, Nien Nunb Elementary School	Reading	04	25	0.65	5	2	0
Atkinson, Rose	Watto Elementary School	Reading, Math	04	19	0.45	14	2	7
Ayers, Aubrey	Lela Oragana High School, Luke Skywalker High School	English II	09, 10, 11	82	-0.22	36	9	7

- Selecting a Teacher Provides a detailed look at a particular teacher's class roster.
- Growth Statistics

acher Average Growth Report 2022 - At Least 1 Student

Identifies average student growth by teacher as compared to the performance of the population and compares it to the Previous year's scores. This table provides the following columns: Current Filtered Test, as compared to the previous year scores for the same students and includes Total students, Mean, Standard Deviation, Average Growth, Total students with negative growth and Total Students with Positive Growth.

GTOWTH REPORT Reading - for Talon Karrde Elementary School - Grade: 05 - Ayers, Kenna										
Growth Statistics										
	Total Students	Mean	Standard Deviation	Average Growth	Total Students with Negative Growth	Total Students with Positive Growth				
Current Filtered Test (2022)	41	0.07	0.8	0.73	4	25				
Previous Year Scores for Same Students (2021)	41	-0.66	1.11	0.75	4	30				

• Growth Report

Provides individual student information and denotes the previous year scores as denoted by a z-score, selected year score as denoted by a z-score, and the difference between the two scores also denoted by a z-score.



Definitions

- Growth: (z-Score current year) (z-Score previous year)
- Average Growth: (z-Score current year) (z-Score previous year) / (Total # of Students)
- Positive Growth: Number of Students with an increase in the standard score form the previous year to the current year.
- Negative Growth: Number of Students with a decrease in the standard score form the previous year to the current year.
- Mean: The average calculated by summing all the values in a data set and dividing by the number of cases.
- Standard Score: A raw score expressed in terms of how many standard deviations it falls away from the mean (also known as z-score).
- Standard Deviation: Tells the average distance by which average scores deviate from the mean.
- z-Score: A raw score expressed in terms of how many standard deviations it falls away from the mean (also known as standard score).

Additional Knowledge Base Articles

- District/Campus Teacher Average Growth
- Student Growth By Teacher

Compare



Compare STAAR Assessments

Compare your assessment between two STAAR test, between two different teachers, or compare several different test across multiple years.

Compare Students	Compare Subjects
<u>Two Year Comparison</u> Subject to Subject Comparison	<u>Multi Year Comparison</u> <u>Multi Subject Comparison</u>

Average Standard Scores for STAAR Tests 2021 vs 2022

Print/Download



• Compare STAAR Assessments

- Average Standard Scores for STAAAR utilizes a statistical analysis where the most current subject specific STAAR Raw Score mean and standard deviation are used to convert individual student raw scores into standardized z-scores. The subject Specific areas aggregates all of the corresponding z-scores for the district, campus or grade level.
- By definition, the average standard score of the population in the state would be 0. Any number greater than 0 equals greater than average performance as compared to the population in the state. Conversely, any number less than 0 would equal less than average performance as compared to the population in the state.
- The gray shaded areas denote the prior year's overall z-score performance.



- Compare Students
 - Two-Year Comparison allows the user to compare the same students across <u>two different</u> <u>years of STAAR on the same subject area assessment</u> (ex. 3rd grade 2020 Reading to 4th grade 2021 Reading). Each dot located on the coordinate plane represents a student's raw score for each year in the same subject. The scale on the plane represents the z-score location as a standard score. To be included in this report, a student is required to have at least two scores, one for each year included in the analysis.



The scale on the plane represents the z-score location as a standard score. The z-score calculation used the mean and standard deviation of the population (Total Group STAAR Mean and standard deviation score for each subject as identified by the TEA on the Mean P-value and Internal Consistency Report). The center section identifies the students whose z-score were "average" or within 1 standard deviation in both assessments. The top right identifies students who were at least 1 standard deviation *below* average on both assessments. The top left identifies students who were at least 1 standard deviation *below* average on both assessments. The top left identifies students who were at least 1 standard deviation below average on both assessments. The top left identifies students who were at least 1 standard deviation above average on one assessment but below average on the second assessment. Average Standard Scores for STAAR utilizes a statistical analysis where the most current subject.

- Subject to Subject Comparison
 - Subject to subject comparison allows the user to compare the same students across <u>two different subject areas of STAAR on the same year</u> (ex. 3rd grade 2021 Reading to 3rd grade 2021 Math). Each dot located on the coordinate plane represents a student's raw score for that year on two different subject areas. The scale on the plane represents the z-score location as a standard score. To be included in this report, a student is required to have at least two scores, one for each subject included in the analysis.



The scale on the plane represents the z-score location as a standard score. The z-score calculation used the mean and standard deviation of the population (Total Group STAAR Mean and standard deviation score for each subject as identified by the TEA on the Mean P-value and Internal Consistency Report). The center section identifies the students whose z-score were "average" or within 1 standard deviation in both assessments. The top right identifies students who were at least 1 standard deviation *below* average on both assessments. The top left identifies students who were at least 1 standard deviation *below* average on both assessments. The top left identifies students who were at least 1 standard deviation below average on both assessments. The top left identifies students who were at least 1 standard deviation above average on one assessment but below average on the second assessment. Average Standard Scores for STAAR utilizes a statistical analysis where the most current subject.



- Compare Subjects
 - Multi Year Comparison is intended to show a trend of a particular subject area over several years. The state average (Z-score = 0) is identified in red, the sample mean is plotted in green, and the standard deviation (the spread of the scores) are plotted in yellow.



- The scale on the plane represents the z-score location as a standard score. The z-score calculation used the mean and standard deviation of the population (Total Group STAAR Mean and standard deviation score for each subject as identified by the TEA on the Mean P-value and Internal Consistency Report).
- Statistical Details
 - Provides information specifically related to the sample for that particular academic year and includes the total student count, the average standard score, and the standard deviation of scores. The standard deviation is represented in the graph with yellow dotted lines. This allows the user to see how one standard deviation or about 68% of the population scored during that year.

Mathematics May 2022 - All Campuses - STAAR Test - Grade 03 - English - All Admin Modes - Ethnicity All - Special Population None		Mathematics May 2021 - All Campuses - STAAR Test - Grade 03 - English - All Admin Modes - Ethnicity All - Special Population None		Mathematics May 2019 - All Campuses - STAAR Test - Grade 03 - English - All Admin Modes - Ethnicity All - Special Population None	
Total Student Count	1,270	Total Student Count	1,129	Total Student Count	1,052
Average Standard Score	0.79	Average Standard Score	0.21	Average Standard Score	0.61
Standard Deviation of Standard Scores	0.91	Standard Deviation of Standard Scores	1.13	Standard Deviation of Standard Scores	1.09

*Students may be excluded because they scored a unique standard score not synchronous with standard scores from other tests, or because they were outliers beyond the scope of the graph.

- Statistical Details
 - Provides the number of students for each academic year scoring within each of the z-score ranges.

Negative	Range	Positive	Negative	Range	Positive	Negative	Range	Positive
140	0 through 0.5	233	186	0 through 0.5	165	159	0 through 0.5	181
73	0.51 through 1.0	322	200	0.51 through 1.0	199	116	0.51 through 1.0	231
13	1.1 through 1.5	211	62	1.1 through 1.5	72	31	1.1 through 1.5	80
7	1.51 through 2.0	169	56	1.51 through 2.0	137	12	1.51 through 2.0	181
0	2.1 through 2.5	69	9	2.1 through 2.5	0	2	2.1 through 2.5	0
0	2.51 through 2.9	0	0	2.51 through 2.9	0	0	2.51 through 2.9	0
0	3.0 and above/below	33	3	3.0 and above/below	40	1	3.0 and above/below	58
233	Totals	1,037	516	Totals	613	321	Totals	731

 Multi Subject Comparison is intended to show a trend of several subject areas during the same academic year. The state average (Z-score = 0) is identified in red, the sample mean is plotted in green, and the standard deviation (the spread of the scores) are plotted in yellow.



- The scale on the plane represents the z-score location as a standard score. The z-score calculation used the mean and standard deviation of the population (Total Group STAAR Mean and standard deviation score for each subject as identified by the TEA on the Mean P-value and Internal Consistency Report).
- Statistical Details
 - Provides information specifically related to the sample for that particular academic year and includes the total student count, the average standard score, and the standard deviation of scores. The standard deviation is represented in the graph with yellow

dotted lines. This allows the user to see how one standard deviation or about 68% of the population scored during that year. Not all students would have taken all tests for this particular report therefore the sample size for each subject area may differ.



- Statistical Details
 - Provides the number of students for each academic year scoring within each of the z-score ranges.

Negative	Range	Positive	Negative	Range	Positive	Negative	Range	Positive
133	0 through 0.5	215	164	0 through 0.5	293	110	0 through 0.5	161
75	0.51 through 1.0	308	96	0.51 through 1.0	271	49	0.51 through 1.0	284
21	1.1 through 1.5	133	16	1.1 through 1.5	167	18	1.1 through 1.5	229
10	1.51 through 2.0	253	9	1.51 through 2.0	162	5	1.51 through 2.0	262
1	2.1 through 2.5	0	1	2.1 through 2.5	0	3	2.1 through 2.5	86
0	2.51 through 2.9	0	0	2.51 through 2.9	0	0	2.51 through 2.9	0
0	3.0 and above/below	58	2	3.0 and above/below	52	0	3.0 and above/below	25
240	Totals	967	288	Totals	945	185	Totals	1,047

Additional Knowledge Base Articles

- <u>District/Campus Teacher Average Growth</u>
- Student Growth By Teacher