

STAAR Alternate Pilot Test
Technical Report
May 28, 2014

Contents

Executive Summary.....	3
Overview	5
Student Performance.....	5
Test Administrator Survey Summary	6
Test Coordinator Survey	13
Summary and Recommendations.....	15
Review period	15
Warm-up activities for students	15
Test length	15
Testing window	16
Use of accommodations	16
Test design	17
Access for visually impaired students.....	17

STAAR Alternate Pilot Test Report

Executive Summary

The STAAR Alternate Pilot test occurred during December 2013 and January 2014 for the purpose of gathering performance and survey data from students and test administrators regarding redesigned STAAR Alternate test items. The pilot test included 9 grades and subjects and participants included

- 152 districts,
- 1,562 students,
- 1,144 test administrators who responded to a survey, and
- 68 testing coordinators who responded to a survey.

Analysis of classical statistics and scoring were completed and items performed as expected.

- Items were ordered by difficulty (easy to hard).
- Item discrimination was strong indicating an ability to distinguish between test-takers' ability.
- Average scores indicated the test was not too difficult for the majority of participants.
- Scores were spread across the scale for all tests.

Recommendation: Although the test was not too difficult for the majority of participants it appears that students would not be adversely impacted if slightly more difficult items were added to the item bank.

The responses to the survey questions for test administrators who administered the pilot test were overwhelmingly positive. Test administrators responded to multiple-choice survey questions and were also given an opportunity to provide comments on a number of topics related to outstanding questions about the test design. The topics and recommendations resulting from the pilot analyses are provided below.

A review period for test administrators to familiarize themselves with test materials and provide accommodations, if necessary, has been considered and was included in the test administrator survey. Test administrators overwhelmingly indicated a need to review materials for the purpose of becoming familiar with scripts and presentation methods of items for the redesigned assessment.

Recommendation: Allow test administrators sufficient time to review the test materials in order to become familiar with and make accommodations to the assessment.

The September 2013 focus group indicated that warm-up activities for students might be necessary to allow students to feel comfortable with the testing format. Cognitive lab observations and pilot test analyses indicated a warm-up activity is probably not necessary.

Recommendation: Rather than creating student-centered warm-up activities, it may be more beneficial to provide resource material, including sample items, to test administrators regarding the presentation style of the assessment as well as to allow test administrators time to review the assessment materials prior to the administration.

The population of students for whom the alternate assessment is designed has not previously been exposed to a standardized, item-based assessment; therefore, there were concerns regarding the length of an item-based assessment that would be appropriate for this population. The pilot test did not indicate any concerns with the test length.

Recommendation: Continue to use the pilot test length as planning for the operational administration continues. More specifically, the operational test would consist of 24 total items, 20 of those being scored items and 4 as field test items. Continue to allow the test to be administered over multiple sessions depending on individual student needs.

The testing window has been a topic of discussion due to its likely significant change in length with the redesigned assessment. A testing window was determined with information about the appropriate test length, time to review and accommodate, as well as time to administer.

Recommendation: Overlap the new STAAR Alternate assessment testing window with the existing TELPAS window after the first operational year where it will be necessary to administer the assessment earlier than usual to meet reporting deadlines.

The use of accommodations is a critical need for the population of students assessed with an alternate assessment. Their needs are extremely diverse and a single test that is accessible for all students would be challenging and costly to create. It was clear that test administrators would need to provide accommodations and overwhelmingly desired more specific guidance than what was provided for the pilot on the types of accommodations that would be appropriate.

Recommendation: Provide test administrators with more specific guidance regarding allowable accommodations as well as sufficient time to prepare accommodations for their students.

Test administrators responded favorably to aspects of the test design such as organization of items by reporting category and difficulty as well as to the type of academic language used and content tested. Opinions were mixed regarding the presentation style and materials.

Recommendation: Consider providing tips about materials management strategies or providing options for how test administrators may document student performance during the administration. Consider providing printed student booklets and test administrator booklets as the standard testing materials. Consider providing item cards as an option that is available online for printing at the campus level (item cards will contain stimulus material and response options). Consider allowing some flexibility in gesturing as appropriate for the student as well as flexibility in how information is presented to students. This recommendation should not be considered an elimination of standardization, as it is not recommended that test administrators have the flexibility to change tested content, rather the way they communicate information (e.g., pointing, signing, tracing, rephrasing items into yes/no questions, replacing "Find" with an appropriate direction for the student such as "Show me", replacing "Say" in scripted directions with verbs that are more flexible such as "Communicate", etc.).

Students with visual or hearing impairments are at a disadvantage when trying to access tested content that is primarily delivered through images and verbal cues.

Recommendation: Provide a list of allowable accommodations that includes replacing images with objects and adding texture to images and text. Consider adjustments to scripts during item development that allow access for as many students as possible (e.g., limit complex graphics, limit concepts that are not easily described, limit scripted terms related to visual access to content, allow flexibility in presentation instructions for the use of objects or other accommodations that allow for visual presentation of stimuli, etc.). Also consider the use of a supplement similar to specific Braille instructions that provides test administrators with guidance regarding how a script can be accommodated for students with visual impairments.

Overview

The STAAR Alternate Pilot testing window closed on January 31, 2014. Districts had until February 7, 2014, to submit online answer documents and complete surveys. A total of 152 districts participated in the pilot.

Assessment	Number of Answer Documents Submitted
Grade 4 Reading	176
Grade 4 Mathematics	197
Grade 5 Science	205
Grade 7 Reading	179
Grade 7 Writing	170
Grade 7 Mathematics	185
English I	146
Algebra I	149
U.S. History	155
Total	1562

Survey Name	Number of Surveys Submitted
Test Administrator Survey	1144
Coordinator Survey	68

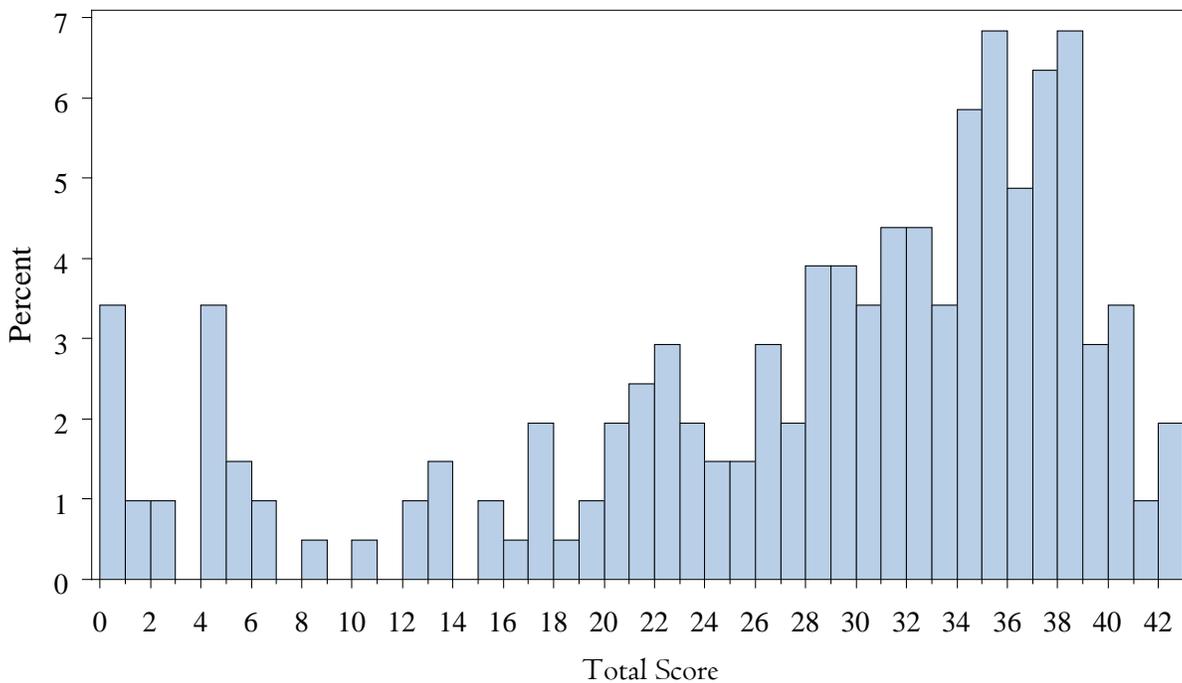
Student Performance

Overall, item and student performance were as expected. Analysis of classical statistics and scoring were completed and a summary of the results are presented below.

- Items were generally ordered as expected (item 1 within a cluster easiest, item 4 most difficult).
- Item discrimination was strong with item-total correlations generally around 0.60, and no negative values.
- Average scores were higher than 50% correct for all tests indicating that the tests were not too difficult for the majority of participating students.
- There was a reasonable spread of performance across the scale for all tests.

Summary of Test-Level Performance

Grade/Subject	N	Mean	SD	Min	Max (Possible)
4 mathematics	197	25.05	10.50	0	41 (41)
4 reading	176	29.42	10.10	0	41 (41)
5 science	205	27.47	11.37	0	42 (42)
7 mathematics	185	24.12	10.54	0	41 (41)
7 reading	179	25.66	10.49	0	41 (41)
7 writing	170	25.01	10.94	0	40 (40)
Algebra I	149	21.71	10.62	0	40 (40)
English I	146	25.03	10.37	0	40 (40)
US history	155	27.36	9.98	0	40 (40)



As indicated previously, average scores were higher than 50% of the test items answered correctly and the score distributions all reflect the negative skew (more scores at the high end of the score scale) that is shown in the sample histogram for grade 5 science above. Although cut scores for proficient performance have not yet been determined, it is possible with a distribution of scores like those seen during the pilot that proficiency rates could be high. Also, typical test performance shows students' scores improve over time implying proficiency rates might start relatively high and continue to grow over time. A recommendation coming from the pilot test analyses is that slightly more difficult items be added to the item bank with minimal concern that the test would become too difficult for students.

Test Administrator Survey Summary

Every test administrator was asked to enter student responses into an online data collection system. This data collection platform included survey questions in addition to the items the students were asked to complete. Responses to the survey questions for the pilot test are shown below.

The most test administrators who responded to the survey administered their student the grade 5 science pilot test., which corresponded to the highest frequency of student responses submitted across all pilot tests. The fewest students were administered the Algebra I assessment and the number of test administrator survey responses was correspondingly low as well.

Subject	Number of Test administrators	Percent
04 Math	155	<i>13.55</i>
04 Read	133	<i>11.63</i>
05 Science	176	<i>15.38</i>
07 Math	135	<i>11.80</i>
07 Read	109	<i>9.53</i>
07 Writing	113	<i>9.88</i>
Algebra 1	102	<i>8.92</i>
English 1	105	<i>9.18</i>
U.S. History	116	<i>10.14</i>
TOTAL	1144	<i>100.00</i>

Most of the pilot participants were classified as primarily taking Complexity Level 2 tasks on the current assessment. The breakdown of pilot participants is similar to what was seen in the 2012–2013 STAAR Alternate administration. The breakdown of complexity levels is shown below.

Complexity Level	Number of Students in Pilot	Percent in Pilot	Percent during 2013 Operational Administration
Level 1	321	28	29
Level 2	653	57	58
Level 3	170	15	12
TOTAL	1144	100	100

Below are the survey questions test administrators responded to upon completion of entry of the student response data into the online transcription form. The questions are provided along with the response options available to test administrators and the number and percent of responses in each response category. High frequency responses are shaded in green. Low frequency responses are shaded in red.

Almost all the test administrator participants indicated that they reviewed all test materials prior to administering the pilot test.

Q1. Did you review the test materials, including the STAAR Alternate Test Administrator Manual: Pilot Test, STAAR Alternate Test Administrator Instructions, and STAAR Alternate Student Booklet or Item Cards, prior to the administration?		
Yes	1141	<i>99.74</i>
No	3	<i>0.26</i>
TOTAL	1144	<i>100.00</i>

The majority of test administrators indicated that they reviewed materials for 46-60 minutes. Another large group of test administrators indicated that they spent more than 2 hours reviewing materials. The smallest number of test administrators indicated the shortest time period of review between 0 and 15 minutes.

Q2. How long did you review the test materials prior to the test administration?		
0-15 minutes	47	4.11
16-30 minutes	130	11.36
31-45 minutes	189	16.52
46-60 minutes	244	21.33
61-90 minutes	187	16.35
91-120 minutes	120	10.49
More than 2 hours	227	19.84
TOTAL	1144	100.00

About half the test administrators indicated that they gave the test in multiple sessions and half indicated that they gave the test in a single session. Many of the test administrator comments related to the open-ended portion of this question indicated that some students could complete the test in a single session but some students could not. The comments generally implied that test administrators want to be allowed to test in multiple sessions as appropriate for their students but not be required to test in multiple sessions.

Q3. Did you administer the test in multiple sessions?		
Yes	562	49.13
No	582	50.87
TOTAL	1144	100.00

The majority of test administrators indicated that it took between 6 and 10 minutes to administer a cluster of 4 items. Seventy-five percent of test administrators indicated that a single cluster took 15 minutes or less.

Q4. What was the average amount of time it took to administer each cluster?		
0-5 minutes	244	21.33
6-10 minutes	383	33.48
11-15 minutes	234	20.45
16-20 minutes	134	11.71
More than 20 minutes	149	13.02
TOTAL	1144	100.00

Almost 50% of test administrators indicated that it took between 46 minutes and an hour and a half to administer all items on the test. Almost 20% of test administrators indicated that it took between 31 and 45 minutes whereas 24% indicated that it took more than an hour and a half to administer all items.

Q5. How long did it take for you to administer the entire test (regardless of whether it took multiple sessions)?		
30 minutes or less	91	7.95
31-45 minutes	225	19.67
46-60 minutes	289	25.26
61-90 minutes	263	22.99
91-120 minutes	149	13.02
More than 2 hours	127	11.10
TOTAL	1144	100.00

Eighty-four percent of test administrators indicated that the test materials were easy to use. If the test administrators indicated difficulty with test materials, they were asked to describe their challenges. Some test administrators indicated that sturdier materials for the student booklet/cards would be preferred. Some test administrators also indicated that it was difficult to manage the number of materials, particularly with students who needed physical assistance to stay on task. A few test administrators indicated the need to spread apart materials, which could be solved using photocopies or

cards, whereas other test administrators indicated that cards were too large making management of materials more difficult. A number of test administrators pointed out the difficulty in presenting and scoring an item when the item in the test administrator instructions spanned two pages. Finally, some test administrators indicated the need to paraphrase or change the script in order to make the presentation appropriate for students with certain disabilities (e.g., visual impairments). Again, most test administrators felt the materials were easy to use and comments were only solicited from test administrators who indicated that the materials were not easy to use.

Q 6. Were the STAAR Alternate Student Booklet or Item Cards and STAAR Alternate Test Administrator Instructions easy to use during testing?		
Yes	963	84.18
No	181	15.82
TOTAL	1144	100.00

Seventy-one percent of test administrators indicated that they had some challenges with entering data into the online data collection system, a free internet application available for use but with limited flexibility in question and response formats. It is important to note that the online system used during the pilot will not be used during the operational administration.

Q7. Did you encounter any challenges with the Administrator Materials (i.e., STAAR Alternate Student Booklet or Item Cards, STAAR Alternate Test Administrator Instructions, STAAR Alternate Answer Document, STAAR Alternate Online Data Collection system).		
No	80	6.99
Yes, with the STAAR Alternate Test Administrator Instructions	153	13.37
Yes, with the STAAR Alternate Student Booklet or Item Cards	21	1.84
Yes, with the STAAR Alternate Online Data Collection system	811	70.89
Yes, with multiple components of the administration materials	68	5.94
Yes, with the STAAR Alternate Answer Document	11	0.96
TOTAL	1144	100.00

Almost 60% of test administrators indicated that their students were able to engage with the images provided. Eighty-two percent of test administrators said their students could engage with the images either as they were presented or with accommodations. Test administrators' comments for students who could not engage with images described a range of concerns with the primary concern being that students who would encounter challenges with the images would be students who would need concrete objects regardless of how graphics were presented. Test administrators also mentioned that many students would benefit from having graphics that are printed in color. Most of the comments; however, could be addressed via allowable accommodations made to the items.

8. Was the student able to engage with the images provided in the STAAR Alternate Student Booklet (e.g., size, detail, layout, etc.)?		
Yes, as presented	670	58.57
Yes, with accommodations	279	24.39
No	195	17.05
TOTAL	1144	100.00

A majority of test administrators indicated that the font size in the student booklet was appropriate.

Q9. Was the font size in the STAAR Alternate Student Booklet appropriate for your student?		
Yes	1034	90.38
No	110	9.62
TOTAL	1144	100.00

Of the test administrators who indicated that this question was applicable, 55% of them indicated that the assessment could be appropriately accommodated for students who are blind or visually impaired.

Q10. For students who are blind or visually impaired, could the assessment be appropriately accommodated?			
	Number of respondents	Percent based on all respondents	Percent based on respondents for which question was applicable
Yes	182	15.91	54.98
No	149	13.02	45.01
N/A	813	71.07	100.00
TOTAL	1144/331	100.00	

Almost 81% of test administrators indicated that it took 30 minutes or less to make any accommodations they felt were necessary for their students.

Q11. If you determined it necessary to accommodate any items, how long did it take to prepare the accommodations for the student?			
0-15 minutes		805	70.37
16-30 minutes		121	10.58
31-45 minutes		69	6.03
46-60 minutes		51	4.46
61-90 minutes		41	3.58
91-120 minutes		18	1.57
More than 2 hours		39	3.41
TOTAL		1144	100.00

A small majority of test administrators indicated that the presentation instructions were easy to understand for their students. For test administrators who responded that the language was not understandable, the two most frequent reasons given were that the word "FIND" was confusing to students and that the script/item was "too wordy." A few test administrators also pointed out that the difficulty of items was too high for the student to whom he/she was administering the pilot.

Q12. Was the student able to understand the language provided in the Presentation Instructions?			
Yes		665	58.13
No		479	41.87
TOTAL		1144	100.00

Seventy-eight percent of test administrators indicated that they use the academic language in the assessment during classroom instruction.

Q13. Has the student been exposed to the academic language used in the assessment?			
Yes		896	78.32
No		248	21.68
TOTAL		1144	100.00

Similar to the academic language question above, 81% of test administrators indicated that they have taught or plan to teach the academic skills addressed on the pilot test.

Q14. Are the academics addressed in the questions the type of skills taught or planned to be taught during the school year?			
Yes		936	81.82
No		208	18.18
TOTAL		1144	100.00

A large majority of test administrators thought ordering the items by difficulty within a cluster was helpful to their students.

Q15. Was having the tested skill organized by difficulty within a cluster of 4 questions helpful to your student?		
Yes	930	<i>81.29</i>
No	214	<i>18.71</i>
TOTAL	1144	100.00

Approximately 66% of test administrators indicated that the presentation instructions were appropriate for their students. However, those test administrators who disagreed pointed out confusion with the word “FIND” and difficulty with the complexity and amount of language in the item. Test administrators also mentioned their desire to repeat part or all of the item as necessary to engage the student in the task.

Q16. Were the Presentation Instructions appropriate for the student?		
Yes	752	<i>65.73</i>
No	392	<i>34.27</i>
TOTAL	1144	100.00

Over 90% of test administrators indicated that the presentation instructions were easy to follow.

Q17. Were the Presentation Instructions easy to follow?		
Yes	1042	<i>91.08</i>
No	102	<i>8.92</i>
TOTAL	1144	100.00

Almost 80% of test administrators indicated that actions such as point, sweep, and say were helpful to their students.

Q18. Were the actions in the Presentation Instructions (e.g., point to, sweep, and say) helpful to the student?		
Yes	910	<i>79.55</i>
No	234	<i>20.45</i>
TOTAL	1144	100.00

Over 96% of test administrators felt the scoring instructions were clear.

Q19. Were the scoring instructions clear?		
Yes	1103	<i>96.42</i>
No	41	<i>3.58</i>
TOTAL	1144	100.00

A majority of test administrators indicated that students need a second presentation of the stimulus in order to access the item content.

Q20. Would the content be accessible if you were not able to provide a second presentation of the stimulus?		
Yes	566	<i>49.48</i>
No	578	<i>50.52</i>
TOTAL	1144	100.00

For test administrators who did provide a second presentation of the stimulus, a majority of them indicated that the support provided for the second presentation was helpful.

Q21. If the student needed a second presentation of the stimulus after answering the item incorrectly, was the teacher assistance provided in the Presentation Instructions helpful to the student (e.g., replication, modeling, etc.)?		
Yes	832	<i>72.73</i>
No	237	<i>20.72</i>
N/A	75	<i>6.56</i>
TOTAL	1144	<i>100.00</i>

A majority of test administrators indicated that their students would be able to engage in an online version of the test. Test administrators who responded that an online version of the test would not be suitable for their students indicated the difficulty their students have in focusing on a single object (computer screen) or the idea that online instruction was uncommon for this population so an online test would be unfamiliar.

Q22. Would the student you tested be able to engage with an online version of the test?		
Yes	601	<i>52.53</i>
No	543	<i>47.47</i>
TOTAL	1144	<i>100.00</i>

Sixty percent of test administrators indicated that their students would have a better chance demonstrating their proficiency in an online delivery platform than with a paper test.

Q23. Would students be better able to demonstrate proficiency in an online version of the test where there are more options for color, magnification, etc.?		
Yes	688	<i>60.14</i>
No	456	<i>39.86</i>
TOTAL	1144	<i>100.00</i>

Despite indicating higher engagement and potentially better access to content, most test administrators wanted the primary version of the test to be a paper delivery with online delivery as an option.

Q24. Which testing format would you prefer?		
Online as the primary delivery method with a paper option	307	<i>26.84</i>
Paper as the primary delivery method with an online option	501	<i>43.79</i>
Paper	245	<i>21.42</i>
Online	91	<i>7.95</i>
TOTAL	1144	<i>100.00</i>

Test Coordinator Survey

Test coordinators in districts that volunteered to participate in the pilot test were asked to respond to a short survey regarding administration procedures for the new STAAR Alternate assessment. The window for completing the test coordinator survey was January 17 to February 7, 2014. Out of 152 participating districts, 68 district coordinators completed a survey for a return rate of 45%.

Almost 60% of respondents indicated that their district would test fewer than 50 students across all grades where the new STAAR Alternate assessment will be administered.

Q1. How many students will your district test on the new STAAR Alternate Assessment?		
Less than 100 (< 100) across all grades	8	11.76
Less than 50 (< 50) across all grades	39	57.35
Less than 500 (< 500) across all grades	7	10.29
Greater than 500 (> 500) across all grades	3	4.41
N/A Campus Coordinator	11	16.18
TOTAL	68	100.00

Over 80% of respondents felt the testing window could reasonably be between 4 and 8 weeks.

Q2. What would be a reasonable length of testing window for the new STAAR Alternate assessment?		
2 weeks	10	14.71
4 weeks	23	33.82
6 weeks	11	16.18
8 weeks	21	30.88
Other: 10 weeks (2) Need flexibility (1)	3	4.41
TOTAL	68	100.00

Although the majority of district coordinators indicated that they would not divide the general testing window into subject-specific windows, there was still a large number of responses indicating that district coordinators would divide the window if given the opportunity. The proposal for dividing the testing window into subject-specific periods of testing was originally suggested by participants in the September 2013 focus group.

Q3. Would your district or campus divide the testing window into subject-specific windows (e.g., first week for reading, second week for mathematics, etc.)?		
Yes	29	42.65
No	38	55.88
Blank	1	1.47
TOTAL	68	100.00

Seventy-five percent of respondents indicated that 10 days would be sufficient to administer a single subject-specific assessment to most students.

Q4. A 10-day per test testing window is appropriate for the majority of students.		
Yes	51	75.00
No	17	25.00
Blank	0	0.00
TOTAL	68	100.00

Responding district coordinators overwhelmingly agreed that test administrators should be able to review testing materials prior to the administration.

Q5. Test administrators should be able to check out testing materials prior to the administration to prepare for the administration.		
Agree	67	98.53
Disagree	1	1.47
Blank	0	0.00
TOTAL	68	100.00

Most respondents agreed that maintaining test security for the new assessment will be no different than test security for the general assessment.

Q6. Maintaining test security of the new STAAR Alternate assessment is no different than maintaining the security of the general education assessment.		
Agree	58	85.29
Disagree	10	14.71
Blank	0	0.00
TOTAL	68	100.00

Most respondents also agreed that item cards for the new assessment would be secured and maintained in the same way as materials for the general assessment.

Q7. Maintaining the security and returning the item cards for the new STAAR Alternate assessment will be no different than returning general assessment test materials.		
Agree	58	85.29
Disagree	9	13.24
Blank	1	1.47
TOTAL	68	100.00

Most respondents agreed that the amount of materials used during the pilot seemed appropriate for the new assessment.

Q8. The amount of materials needed for the new STAAR Alternate assessment (test administrator manual, student booklet, test administrator instructions booklet, and item cards) is right for the assessment.		
Agree	63	92.65
Disagree	4	5.88
Blank	1	1.47
TOTAL	68	100.00

There was an even split among respondents as to whether one test administrator booklet per student or one for every five students was appropriate. If an answer document is used, it is likely that a single test administrator booklet per grade/subject would be sufficient for an administration.

Q9. The ideal number of test administrator books per subject for a campus is:		
1 per student	34	50.00
1 for every 5 students	34	50.00
Blank	0	0.00
TOTAL	68	100.00

Summary and Recommendations

Generally, test administrators provided positive feedback about their experiences with the pilot test. The positive feedback included:

- the organization of the content into clusters, ordered within a cluster by difficulty;
- the academic language used and content tested;
- the font size;
- detail of graphics;
- paper presentation; and
- presentation and scoring instructions.

There were a number of outstanding questions awaiting pilot test results. Below are a list of topics where pilot test results provide insight into potential recommendations.

Review period

Almost all pilot survey respondents indicated that they had reviewed the test materials prior to the administration and 68% of respondents indicated that they spent more than 45 minutes reviewing the materials. There were a number of comments related to the test administrator's unease with handling the materials, gesturing, and following the script. It is likely that test administrators would feel more comfortable administering the test once they were given sufficient training to become familiar with the presentation style of the new assessment as well as ample time to review the actual test materials in order to present the tested content fluidly and with ease using a tone, cadence, and demeanor that is familiar to the student.

Recommendation: Allow test administrators sufficient time to review the test materials in order to become familiar with and provide accommodations to the assessment.

Warm-up activities for students

Although warm-up activities for students were discussed in the early stages of the assessment design primarily based on feedback from the September 2013 focus group, there were no specific questions posed to test administrators regarding activities like this during the pilot test. Cognitive lab observations and interviews, supported by pilot test results, indicate that test administrator familiarity with the test format and content assessed is more important for preparing students to complete the assessment than warm-up activities. Several test administrators indicated that the item style with a stimulus and answer choices was not the most typical instructional style, but with knowledge of the test design test administrators could adapt some instructional methods to make the format more familiar to students.

Recommendation: Rather than creating student-centered warm-up activities, it may be more beneficial to provide resource material, including sample items, to test administrators regarding the presentation style of the assessment as well as to allow test administrators time to review the assessment materials prior to the administration.

Test length

The pilot test consisted of 24 items administered per tested grade/subject. The operational design was planned to mirror this test length with 20 of the 24 items serving as scored, operational items and 4 of the 24 serving as non-scored, field-test items. There was a query, prior to the pilot, as to the appropriateness of the test length because the current version of the STAAR Alternate assessment consists of only 4 tasks, with 3 predetermined criteria per task. It appears that 24 items was not overly burdensome with the qualification that only one test was administered per student during the pilot administration. Operationally, in grade 8, a test administrator would be required to administer as many as 4 tests with 24 items each to meet testing requirements. Several test administrators commented that there would be

instances where 24 items would be too burdensome for their students but allowances to take frequent breaks would mitigate any need to shorten the test for that reason. Many test administrators participating in the pilot survey indicated that the test length was appropriate and their student did not need multiple sessions to complete the assessment.

Recommendation: Continue to use the pilot test length as planning for the operational administration continues. More specifically, the operational test would consist of 24 total items, 20 of those being scored items and 4 as field test items. Continue to allow the test to be administered over multiple sessions depending on individual student needs.

Testing window

There are a number of outstanding questions regarding the testing window including the length of the window, the division of the window into subject-specific components, the length of the preview period, and the period within the academic year in which the window should fall. Information about the appropriateness of the test length also relates to the length of the testing window. Considering that most test administrators indicated that 24 items was appropriate, particularly if given flexibility to administer the test over multiple sessions, the length of the window needs to accommodate the amount of time necessary to complete a maximum of 4 tests in a single grade/subject. Seventy-six percent of test administrators indicated that it took 1.5 hours or less to administer the pilot test. One can estimate that it would take 6 hours (at a maximum of 90 minutes per test) to administer all tests in grade 8 to a single student. Although we do not have a direct comparison for the current STAAR Alternate because the same exact question was not asked of test administrators, a question regarding the amount of time needed to prepare for the current administration was asked during a survey of test administrators in 2012–2013 and 67% of teachers indicated it took more than 4 hours to prepare an assessment per student and per subject area. The expectation is that it would take significantly less time to prepare for the administration of the redesigned STAAR Alternate because all materials a teacher needs are provided. The only preparations teachers would need to make would be any accommodations to items, such as enlarging images through the use of photocopies. For test administrators who did provide accommodations, over 80% of them indicated that it took 30 minutes or less to make accommodations for one test and one student.

Consideration also needs to be given to the length of the review period for test administrators to familiarize themselves with the assessment and make any accommodations necessary to items. Over 99% of respondents said they reviewed materials and 80% of respondents indicated that they reviewed materials for 2 hours or less (53% reviewed for an hour or less). It is likely that test administrators would spend the most time reviewing the test when they administer it the first time, to their first student, and that they might spend less time reviewing that test for the remainder of the students in their class.

Considering estimates of the time needed to review, accommodate, and administer the assessment it is recommended that the testing window itself be two weeks. However materials could be made available to a district two to three weeks prior to the opening of the window to allow time for teachers to review and accommodate any materials necessary.

Recommendation: Overlap the new STAAR Alternate assessment testing window with the existing TELPAS window after the first operational year where it will be necessary to administer the assessment earlier than usual to meet reporting deadlines.

Use of accommodations

Feedback was overwhelmingly positive regarding the test design, the test administrator scripts, the academic content assessed, and language used; however, when test administrators provided comments relative to questions about the test design those comments most often related to issues with accommodating materials as appropriate for individual students. Comments were related to the following:

- More time to preview materials for accommodation purposes
- More guidance regarding the types of accommodations allowed
- More time to actually provide accommodations

Test administrators indicated that they were unclear as to the accommodations allowed until they reached the online data-capturing platform where a list of possible accommodations was provided. Several indicated that they would prefer to have the list prior to the administration. It should be noted that some test administrators indicated that, for certain students, providing the necessary accommodations would be time consuming and burdensome in a manner similar to preparations for the existing alternate assessment.

Recommendation: Provide test administrators with specific guidance regarding allowable accommodations as well as ample time to prepare accommodations for their students.

Test design

Test administrators responded positively to questions about the test design including the organization of items by reporting category and difficulty, academic language, presentation instructions, and test materials. The most common feedback related to difficulty managing the materials (test administrator booklet, student booklet/ item cards, and answer document) while managing challenging student behaviors. Comments were mixed regarding the student booklet versus item cards. Other comments related to test administrators' unfamiliarity with the presentation style and gestures outlined in the script as well as use of the term "Find" in the question portion of each item.

Recommendation: Consider providing tips about materials management strategies or providing options for how test administrators may document student performance during the administration. Consider providing printed student booklets and test administrator booklets as the standard testing materials. Consider providing item cards as an option that is available online for printing at the campus level (item cards will contain stimulus material and response options). Consider allowing some flexibility in gesturing as appropriate for the student as well as flexibility in how information is presented to students. This recommendation should not be considered an elimination of standardization, as it is not recommended that test administrators have the flexibility to change tested content, rather the way they communicate information (e.g., pointing, signing, tracing, rephrasing items into yes/no questions, replacing "Find" with an appropriate direction for the student such as "Show me", replacing "Say" in scripted directions with verbs that are more flexible such as "Communicate", etc.).

Access for visually impaired students

Only 30% of test administrators indicated that the question about accommodating items for students with visual impairments was applicable for their students. Of those who responded it was applicable, approximately 55% of respondents said the test used during the pilot could be accommodated for students with visual impairment whereas 45% indicated that it could not be accommodated. Most test administrators who responded to this question with comments indicated that students need concrete objects to manipulate in order to access the content. A large number of test administrators also indicated that raised lines and textures or Braille would be beneficial for these students. There was no definitive response to this question based on the test administrator survey results.

Recommendation: Provide a list of allowable accommodations that includes replacing images with objects and adding texture to images and text. Consider adjustments to scripts during item development that allow access for as many students as possible (e.g., limit complex graphics, limit concepts that are not easily described, limit scripted terms related to visual access to content, etc.). Also consider the use of a supplement similar to specific Braille instructions that provides test administrators with guidance regarding how a script can be accommodated for students with visual impairments.