Executive Summary

April 2022

This document describes the process to recommend performance level cut scores for the assessments associated with the Virginia Alternate Assessment Program (VAAP). A summary of the results is also provided.

Standard Setting Process and Results for the Virginia Alternate Assessment Program Assessments

The standard setting process involves several components and is designed to produce cut scores on a test that are used to classify students into performance levels. Performance Level Descriptors (PLDs) identify the specific knowledge and skills that students in each level should be able to demonstrate. The performance expectations defined by the PLDs are linked to the items on a test to determine cut scores that students must meet or exceed to be classified into each performance level.

The process used to obtain recommended performance level cut scores for the assessments is consistent with industry recognized standard setting best practices. An overview of the process and a summary of the results are presented in the following sections.

General Method

From Monday, April 11 through Friday, April 15, 2022, standard setting meetings were convened to recommend cut scores for the assessments associated with VAAP.

- Mathematics, Grades 3 – 8 and High School
- Reading, Grades 3 – 8 and High School
- Science, Grades 5, 8, and High School

The grade-level focused committees were composed of 7 to 9 panelists who were educators with relevant experience teaching students that are eligible to participate in VAAP. The standard setting committee participants were selected to provide content expertise and expertise in working with student from this special population during the meeting and to represent diverse state geographic regions, gender, ethnicity, educational experience, community size, and community socioeconomic status.

The Yes/No Angoff standard setting procedure was used. Participants were led through a standardized process in which they reviewed test items to consider how students in each performance level would answer each item. These judgments were used to obtain recommended cut scores for each performance level. The same standardized process was used by each grade-level committees to produce cut score recommendations for each assessment. Each committee used this process to recommend cut scores for all tests associated with the specific grade, starting with mathematics, followed by reading, and then, if needed, science.

The first part of each committee was an introduction to the purpose of the standard setting and process that would be used. They were told why new performance standards needed to be set and given an overview of the Yes/No Angoff procedure that would be used. The panelists then experienced the test, which allowed them to view the items as a student would and to consider the knowledge and skills needed to respond to each item. They reviewed the PLDs to gain a common understanding of the expectations for the performance levels and then narrowed the focus to key knowledge and skills at the borderlines separating the performance levels. They worked in small groups to create specific descriptions of the knowledge and skills expected of students who just barely enter a performance level.

After discussion and general agreement about the borderline descriptions, the participants were trained on the standard setting method and the judgment process to be applied during the remainder of the meeting. They reviewed each item and the borderline performance descriptions to answer the following question for each performance level: “Would a student performing at the borderline of the performance level likely get this item correct?” For the purposes of the standard setting, “likely” was defined as 2 out of 3 students at the borderline of the performance level correctly answering the item.

The panellists engaged in a practice judgment activity using sample items. They discussed the process and results to clarify their understanding of the judgment task. Then they completed three rounds of individual judgments. Following the first two rounds of judgments, the participants discussed their resulting individual cut score recommendations and the range of item judgments. The median of the group’s cut scores was the recommended cut score for the group for each round of ratings.

The cut score data from the standard setting committees are summarized in Tables ES1-ES3. The final round 3 median cut score recommendations for each performance level are highlighted.

Table ES1. Mathematics Cut Score Recommendations Summary Statistics

Grade	Statistic	Performance Level					
		Proficient			Advanced		
		R1	R2	R3	R1	R2	R3
Grade 3	N	9	9	9	9	9	9
	Minimum	10	10	7	22	18	19
	Median	17	13	12	28	24	23
	Maximum	26	26	22	30	29	25
Grade 4	N	8	8	8	8	8	8
	Minimum	12	10	11	24	21	21
	Median	22	18	17	29	30	25
	Maximum	29	30	17	30	30	28
Grade 5	N	8	8	8	8	8	8
	Minimum	8	11	12	17	22	21
	Median	14	13	13	24	24	24
	Maximum	19	14	14	29	24	24
Grade 6	N	8	8	7	8	8	7
	Minimum	7	5	10	7	5	18
	Median	14	12	11	24	22	21
	Maximum	23	15	12	28	21	22
Grade 7	N	8	8	8	8	8	8
	Minimum	16	11	10	22	19	20
	Median	22	14	14	30	28	23
	Maximum	25	30	25	30	30	26
Grade 8	N	8	8	8	8	8	8
	Minimum	4	3	3	15	16	16
	Median	12	14	13	23	24	25
	Maximum	16	18	17	27	29	26
High School	N	7	7	7	7	7	7
	Minimum	13	9	9	22	22	20
	Median	15	12	12	24	23	22
	Maximum	26	20	16	30	30	23

Table ES2. Reading Cut Score Recommendations Summary Statistics

Grade	Statistic	Performance Level					
		Proficient			Advanced		
		R1	R2	R3	R1	R2	R3
Grade 3	N	9	9	9	9	9	9
	Minimum	5	5	6	17	18	20
	Median	9	10	10	21	22	22
	Maximum	14	12	12	30	25	23
Grade 4	N	7	8	8	7	8	8
	Minimum	8	8	9	19	19	20
	Median	8	9	10	22	21	20
	Maximum	11	11	12	29	24	24
Grade 5	N	8	8	8	8	8	8
	Minimum	12	11	12	23	21	21
	Median	15	12	12	25	23	23
	Maximum	23	16	14	29	24	24
Grade 6	N	8	8	8	8	8	8
	Minimum	7	10	9	15	16	16
	Median	10	12	12	19	19	20
	Maximum	15	19	13	25	22	23
Grade 7	N	8	8	8	8	8	8
	Minimum	5	5	5	16	15	15
	Median	8	10	10	22	21	20
	Maximum	13	13	12	28	25	23
Grade 8	N	8	8	8	8	8	8
	Minimum	7	7	7	15	17	17
	Median	13	14	14	22	23	23
	Maximum	22	18	18	30	25	25
High School	N	7	7	7	7	7	7
	Minimum	10	7	9	28	17	17
	Median	24	12	11	30	21	20
	Maximum	30	28	14	30	30	23

Table ES3. Science Cut Score Recommendations Summary Statistics

Grade	Statistic	Performance Level					
		Proficient			Advanced		
		R1	R2	R3	R1	R2	R3
Grade 5	N	8	8	8	8	8	8
	Minimum	11	11	12	21	20	22
	Median	14	13	13	23	23	23
	Maximum	16	14	14	26	25	24
Grade 8	N	8	8	8	8	8	8
	Minimum	13	4	5	22	22	23
	Median	16	16	16	25	25	26
	Maximum	22	20	20	30	28	28
High School	N	7	7	7	7	7	7
	Minimum	11	11	10	20	19	20
	Median	14	12	11	24	22	21
	Maximum	29	17	15	30	30	23

After Round 3 of ratings, the participants from each standard setting committee completed an evaluation of the standard setting process and their confidence in their recommended cut scores. Overall, all participants understood the standard setting process and were confident about their recommendations.

After the standard setting committee, on Saturday, May 14, 2022, an articulation committee, composed of members from each grade-level committee, collectively reviewed the Round 3 recommendations from the individual committees. The articulation committee was held after the standard setting meeting to ensure that the impact data presented during the meeting was based on all students that took the assessment. This meeting was held virtually to ensure that individuals from the original standard setting committee were able to participate. The committee completed the articulation process for each assessment, starting with mathematics, followed by science, and then science.

This articulation committee reviewed the PLDs from each grade-level, discussed the Round 3 recommendations from each subject area, and reviewed the estimated impact data for each assessment. Impact data are the percentages of students classified into each performance level if the recommended Round 3 cut scores were to be adopted. Figures ES1-ES3 show the estimated impact based on the Round 3 recommendations for each assessment.

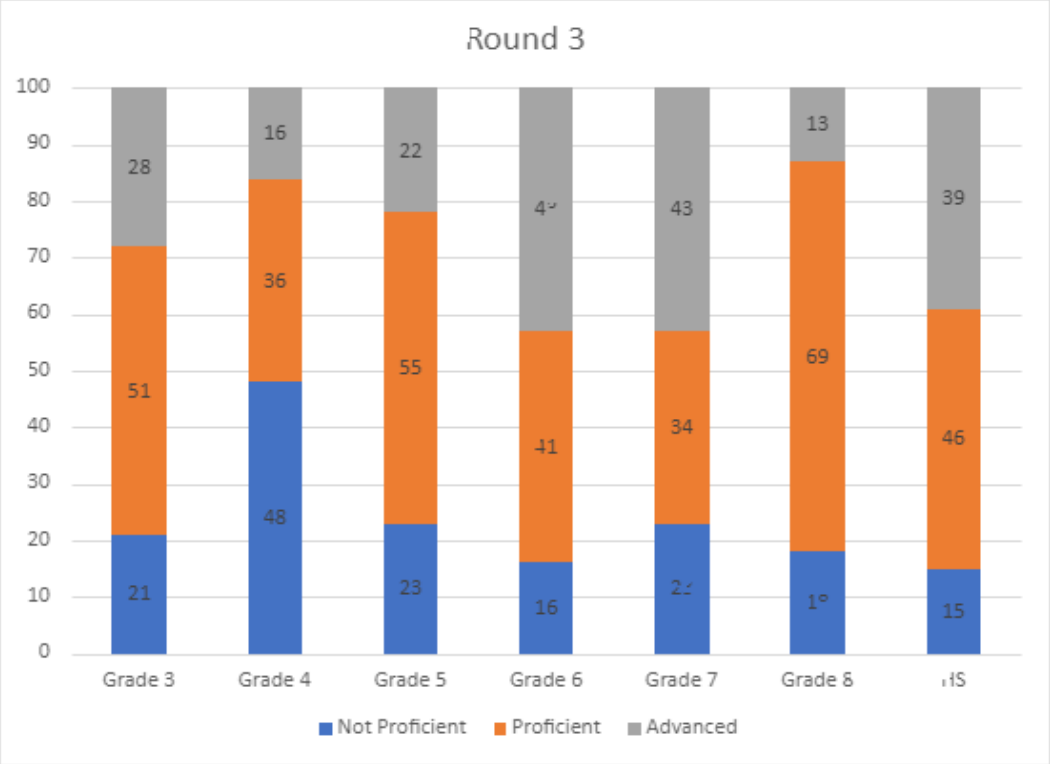

Figure ES1. Impact data from Round 3 for Mathematic by grade

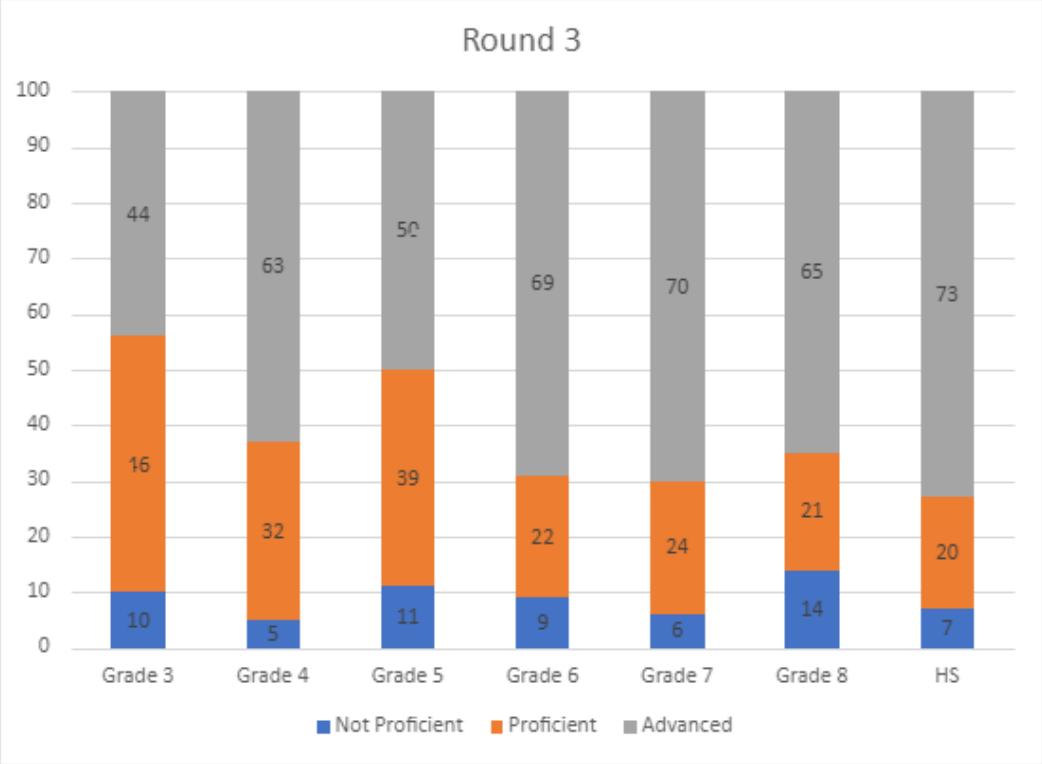

Figure ES2. Impact data from Round 3 for Reading by grade

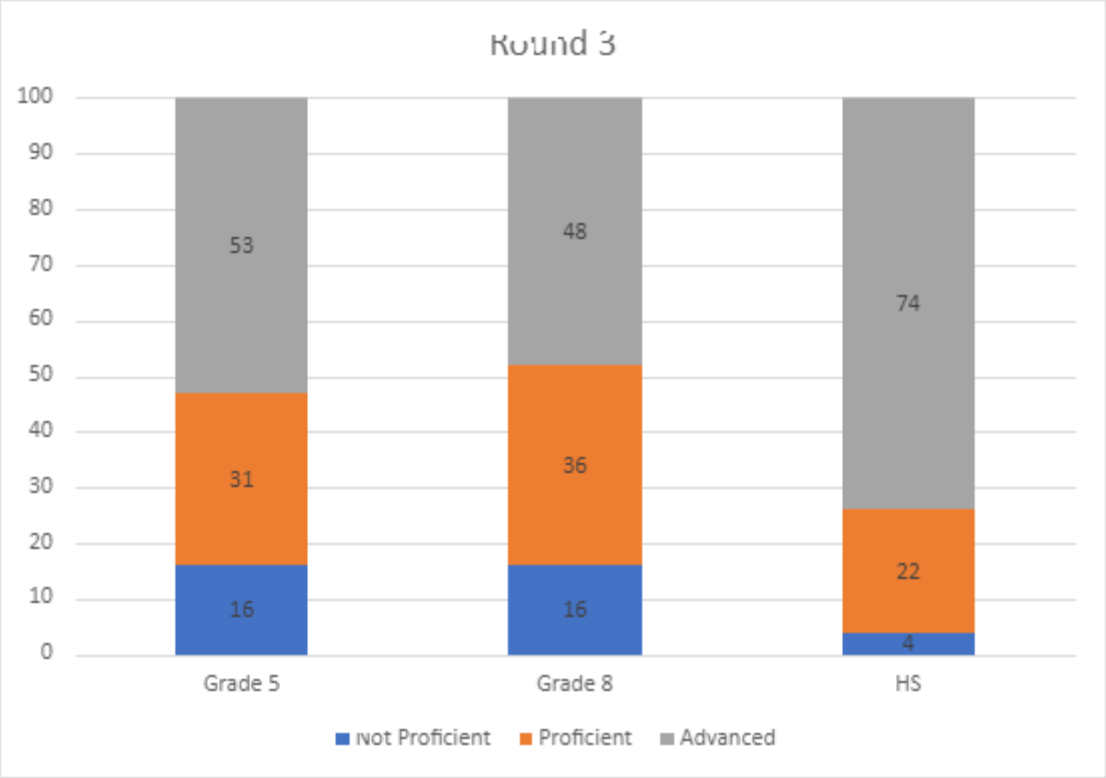

Figure ES3. Impact data from Round 3 for Science by grade

After reviewing the impact data associated with the Round 3 recommended cut scores, the articulation committee discussed whether any Round 3 recommended cut scores should be adjusted for consistency with expectations based on the scope and rigor of content covered across the assessments.

After group discussion, the articulation committee was permitted to 'smooth' the Round 3 results to align the recommendations with expectations and student performance across the grades. Each participant independently recommended cut scores for the "Proficient" and "Advanced" performance levels for each assessment informed by the group discussion and their expectations for student performance across grades. The panelist articulation judgment round was followed by a whole group discussion to determine the final recommended cut scores. Tables ES4-ES6 show the Round 3 raw score cuts and the articulation committee's recommended cut scores from each round for each assessment.

Table ES4. Cut Score Recommendations from Articulation Committee for Mathematics

Grade	Performance Level					
	Proficient			Advanced		
	Round 3	Judgment Round	Final Judgment	Round 3	Judgment Round	Final Judgment
3	12	12	12	23	23	23
4	17	12	12	25	24	24
5	13	13	13	24	24	24
6	11	11	11	21	23	23
7	14	14	14	23	24	24
8	13	13	13	25	24	24
High School	12	12	12	22	22	22

Table ES5. Cut Score Recommendations from Articulation Committee for Reading

Grade	Performance Level					
	Proficient			Advanced		
	Round 3	Judgment Round	Final Judgment	Round 3	Judgment Round	Final Judgment
3	10	10	10	22	22	22
4	10	10	10	20	20	20
5	12	12	12	23	23	23
6	12	12	12	20	21	21
7	10	11	11	20	21	21
8	14	14	14	23	23	23
High School	11	11	11	20	21	21

Table ES6. Cut Score Recommendations from Articulation Committee for Science

Grade	Performance Level					
	Proficient			Advanced		
	Round 3	Judgment Round	Final Judgment	Round 3	Judgment Round	Final Judgment
5	13	13	13	23	23	23
8	16	16	16	26	26	26
High School	11	14	14	21	24	24

Culminating Recommendations

Figures ES4-ES6 show the estimated impact based on the Articulation Committee's final cut score recommendations.

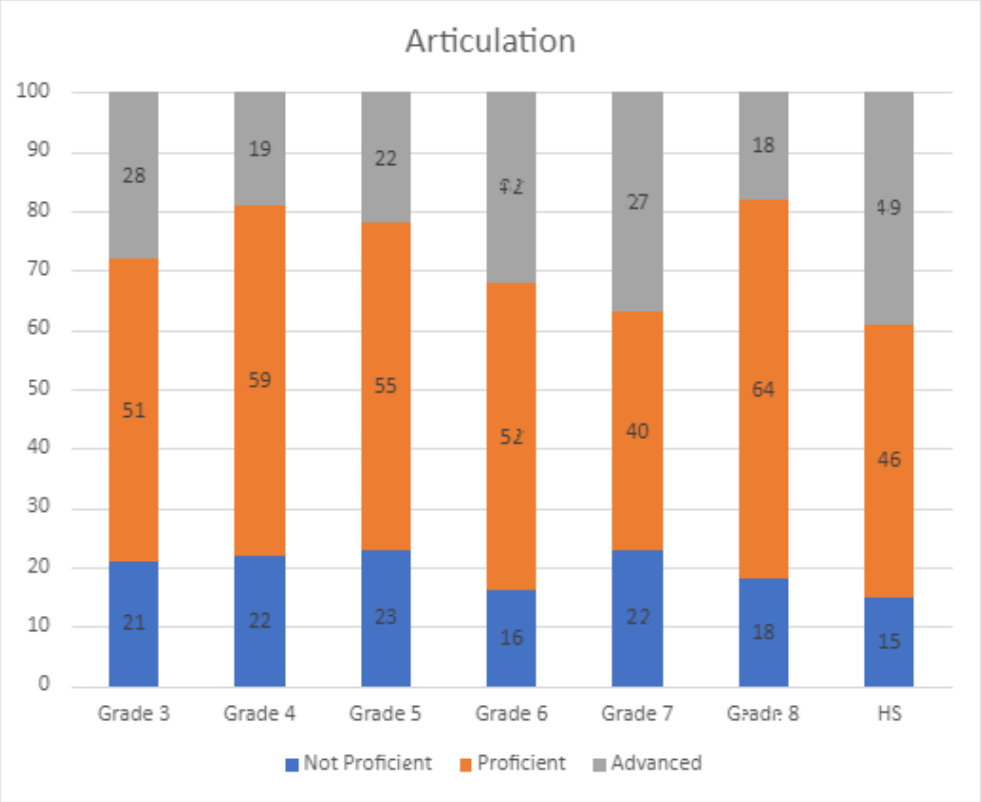

Figure ES4. Mathematics Impact Data by Assessment Following Articulation

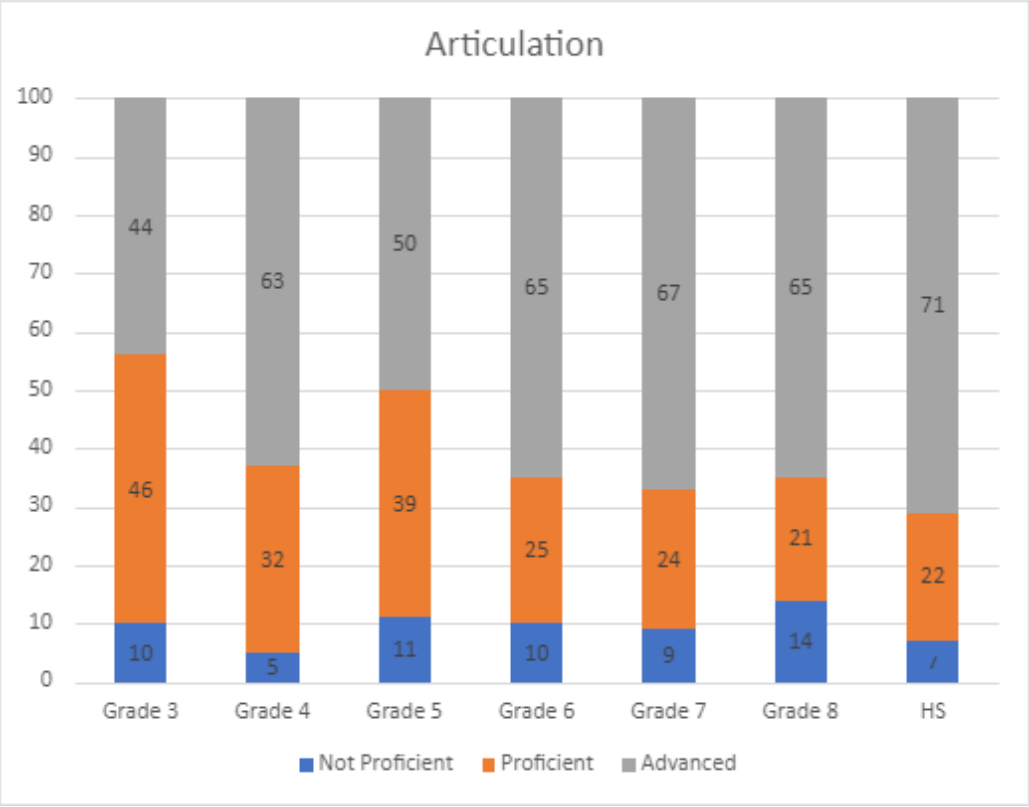

Figure ES5. Reading Impact Data by Assessment Following Articulation

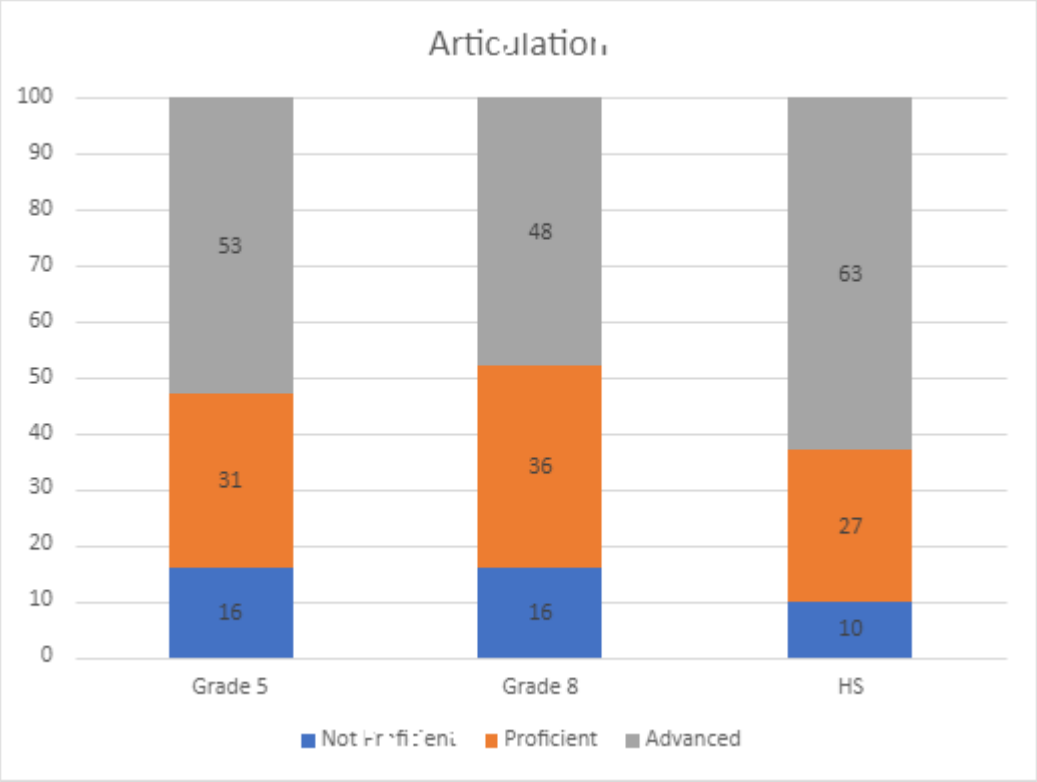

Figure ES6. Science Impact Data by Assessment Following Articulation