

2015-2016 Oral Communication Assessment Report- Results Summary

Executive Summary

Seventy-four previously recorded presentations were scored by two separate raters using a revised version of the AAC&U Oral Communication Rubric. Seven traits (Introduction, Organization, Language, Supporting Material, Voice Delivery, Nonverbal Delivery, and Conclusion) were rated on a five-point scale of 0=Unobserved, 1=Deficient, 2=Basic, 3=Proficient, and 4=Advanced. A total of 9% (7 presentations) of the 74 had a difference of more than 1.00 point requiring an additional rater, which was an improvement from the pilot assessment in which 17% (6) of the 35 presentations required an additional rater. Mean differences between raters' scores for the oral communication traits decreased from the pilot in 2014-2015 to the 2015-2016 assessment, as did the standard deviations of those means—showing that individual raters' scores became closer together, and the validity/reliability of the rater scores have improved due to better rater training. For 2015-2016, the raters were most consistent in the scoring of the Conclusion trait and least consistent on the scoring of the Supporting Material trait.

Twenty-three percent, or 67 of the 74 student presentations, achieved an Overall Total Average score for Oral Communication classified at the Proficient or Advanced performance levels; this was an improvement from the pilot assessment in which 17%, or 6 of the 35 student presentations, were classified at the Proficient or Advanced performance levels. The Oral Communication traits in which Washburn students were the strongest in 2015-2016 included Language, Voice Delivery, and Non-Verbal Delivery. The Conclusion, Introduction, Organization, and Supporting Material traits seem to be where Washburn students performed the lowest in 2015-2016. The percent of student presentations classified at the Proficient or Advanced performance levels increased from 2014-2015 to 2015-2016 for the Oral Communication traits of Non-Verbal Delivery, Supporting Material, Organization, and Introduction. However the Oral Communication traits of Language, Voice Delivery, and Conclusion experienced slight decreases from 2014-2015 to 2015-2016 in the percent of presentations classified at the Proficient or Advanced performance levels.

Key Findings for 2015-2016 Rater Differences

- 9% of presentations had a difference over 1 point- requiring a 3rd rater; 12% of the overall average scores from the two raters did not differ, 75% differed less than 1 point, 3% differed by 1 full point
- Raters were fairly consistent for the Conclusion trait with 96% of the ratings having score differences equal to 0 or 1; followed by the Introduction and Language traits at 89%, Non-Verbal Delivery at 88%, and the Organization and Voice Delivery traits at 85%
- Supporting Material seemed to be the trait in which raters were least consistent with 68% of the ratings differing by 0 or 1, and 32% of the Supporting Material trait ratings differing by 2-3 points

Key Findings for 2015-2016 Performance

- 23% of oral communication presentations achieved an overall total average score at the Proficient or Advanced performance levels
- 1% of the presentations attained an overall total average score at the Advanced performance level, 22% scored at the Proficient level, 70% at Basic, 7% at Deficient, and 0% at the Unobserved performance level

Key Findings for 2015-2016 Performance (continued)

- 78% attained average scores classified as Proficient or Advanced for the Language trait, 55% for Voice Delivery, 51% for Non-Verbal Delivery, 42% for Supporting Material, 36.5% for Introduction and for Organization, and 15% were Proficient/Advanced for the Conclusion trait
- The Language trait had the highest mean average score at 2.75, followed by Voice Delivery at 2.39 and Non-Verbal Delivery at 2.21; Organization rendered a mean score of 2.15, followed by Supporting Material at 2.09, and Introduction at 2.08; the Conclusion trait had the lowest mean at 1.39
- The Conclusion trait is where Washburn students performed the lowest; however, students also scored lower on Introduction, Supporting Material, and Organization relative to other traits

Key Findings for Pilot 2014-2015 and 2015-2016 Inter-Rater Reliability Comparison

- 35 presentations were collected/scored for the 2014-2015 pilot, but this increased to 74 for 2015-2016
- 9% of the presentations had a difference greater than 1 point and required a 3rd rater in 2015-2016; which was a decrease from 2014-2015 when 17% required an additional rater
- 12% of overall average scores from the two raters did not differ in 2015-2016, an increase from 11% of rater pairs not differing in 2014-2015
- Mean differences between raters for the overall average scores in 2014-2015 was 0.69 with a standard deviation of 0.445, and the mean difference between raters for 2015-2016 decreased by .024 to 0.45 with the standard deviation decreasing by 0.013 to 0.432
- Mean differences between raters' scores for the oral communication traits decreased from the pilot in 2014-2015 to the 2015-2016 assessment, as did the standard deviations of those means- showing that individual raters' scores became closer together, and the validity and reliability of the scores given by the individual raters have improved

Key Findings for Pilot 2014-2015 and 2015-2016 Performance Comparison

- 17% of the 35 student presentations in 2014-2015 attained a rounded overall average score classified at the Proficient or Advanced performance levels; this increased by almost 6 percentage points in 2015-2016 as 23% of the 74 overall average scores were categorized at the Proficient or Advanced performance levels
- The percent of rounded scores classified at the Proficient or Advanced performance levels increased from 2014-2015 to 2015-2016 for the Oral Communication traits of Non-Verbal Delivery (46% to 51%), Supporting Material (40% to 42%), Organization (29% to 36%), and Introduction (11% to 36%)
- The percent of rounded scores classified at the Proficient or Advanced performance levels decreased from 2014-2015 to 2015-2016 for the Oral Communication traits of Language (83% to 78%), Voice Delivery (57% to 55%), and Conclusion (17% to 15%)
- From 2014-2015 to 2015-2016 the mean scores increased for the Conclusion (1.30 to 1.39), Introduction (1.64 to 2.08), and Organization (2.14 to 2.15) traits; the mean for the overall average score remained the same at 2.15; and the mean scores decreased for the Language (2.87 to 2.75), Voice Delivery (2.48 to 2.39), Non-Verbal Delivery (2.30 to 2.21), and the Supporting Material (2.27 to 2.09) traits from the pilot in 2014-2015 to the 2015-2016 assessment
- The Conclusion, Introduction, and Organization are the lowest scoring traits in both years but are increasing; the Supporting Material traits scores are also relatively low but decreased between years

2015-2016 Oral Communication Assessment Report- Results Summary

Introduction: A total of 74 previously recorded presentations were scored by two separate raters using a revised version of the AAC&U Oral Communication Rubric. Seven traits (Introduction, Organization, Language, Supporting Material, Voice Delivery, Nonverbal Delivery, and Conclusion) were scored on a five-point scale with 0=Unobserved, 1=Deficient, 2=Basic, 3=Proficient, and 4=Advanced. When there were only two raters (their overall mean scores did not differ by more than 1.0) the mean for each trait and the overall total is the average of the two rater's scores. If the two raters differed by more than 1.0 for the overall total scores, then a third rater also scored the presentation. In cases of three raters, the mean for each trait and the overall total is the average of the three raters' scores.

Differences between Individual Raters: Approximately 12% of the overall average scores from the two raters did not differ, over 75% differed less than 1 point, and almost 3% differed by 1 full point. Therefore, 90% of the presentations did not have overall average scores that were disparate enough to require a 3rd reader, but 9% had a difference of 1.14 to 2.00 points (see Table 1) and an additional rater to score the presentations was necessary.

Table 1. Differences between Raters for Total Overall Average Scores

Score Difference	Frequency	Percent
0.00	9	12.16
0.14	21	28.40
0.29	8	10.81
0.43	11	14.86
0.57	6	8.10
0.71	8	10.81
0.86	2	2.70
1.00	2	2.70
1.14	1	1.35
1.29	1	1.35
1.43	2	2.70
1.57	2	2.70
2.00	1	1.35
Total	74	100.00

The mean difference between raters for the overall average scores was 0.45 with a standard deviation of 0.432, and differences ranged from 0 to 2.00 (see Table 2). For the individual traits scored, differences between raters ranged from 0 to 3 points. Supporting Material had the highest mean difference at 1.05; followed by the Organization, Voice Delivery, and Non-Verbal Delivery mean differences at 0.70. The mean difference between rater scores for Language was 0.66, for Conclusion was 0.64, and the trait was the smallest mean difference between raters was Introduction at 0.58.

Table 2. Descriptive Statistics for Rater Score Differences

Statistic (N=74)	Introduction Difference	Organization Difference	Language Difference	Supporting Material Difference	Voice Delivery Difference	Non-Verbal Delivery Difference	Conclusion Difference	Overall Total Difference
Mean	0.58	0.70	0.66	1.05	0.70	0.70	0.64	0.45
Std. Dev.	0.722	0.716	0.707	0.842	0.754	0.677	0.563	0.432
Min.	.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Max.	3.00	2.00	3.00	3.00	3.00	2.00	2.00	2.00

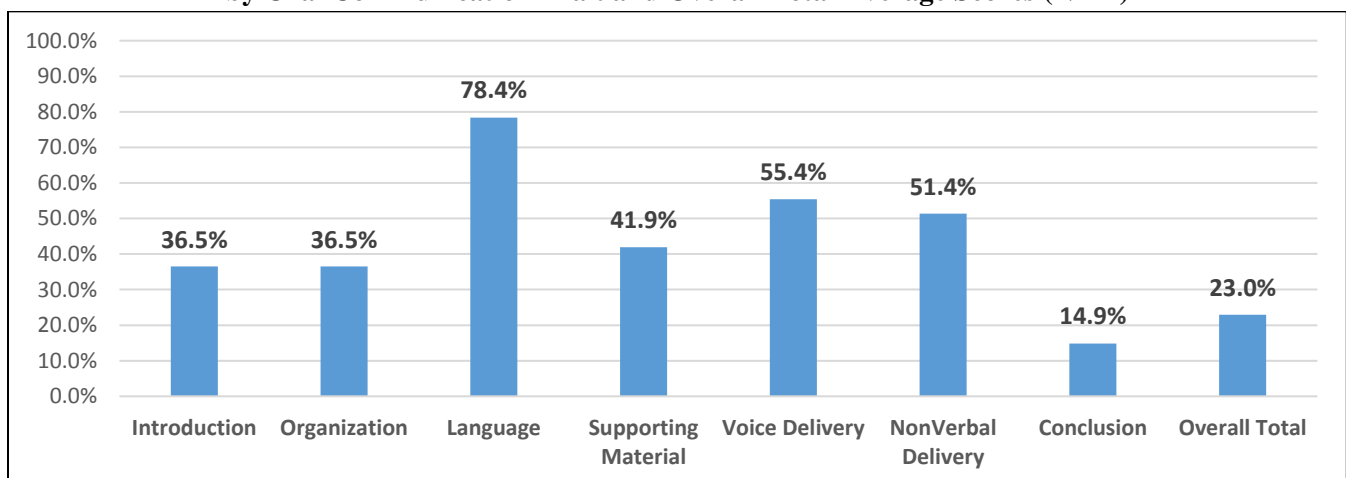
Frequencies and percentages of the difference scores for each trait on the Oral Communication rubric were computed; and the 0- and 1-point score differences, as well as the 2- and 3-point score differences, were combined to provide frequencies and percentages for the closer and more widely disparate scores by trait (see table below). Raters seemed to be most consistent for Conclusion scores with 96% of the ratings for that trait differing by 0-1 points. Raters were fairly consistent for the Introduction and Language traits with 89% of the ratings having score differences equal to 0 or 1; followed by Non-Verbal Delivery at 88%, then Organization and Voice Delivery at 85%. Supporting Material seemed to be the trait in which raters were least consistent with 68% of the ratings differing by 0 or 1, and 32% of the Supporting Material trait ratings differing by 2-3 points.

Table 3. Frequencies and Percentages of Score Differences between Raters by Trait

Rater Score Differences	Introduction		Organization		Language		Supporting Material		Voice Delivery		Non-Verbal Delivery		Conclusion	
	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc	Freq	Perc
0	40	54.1%	33	44.6%	34	45.9%	22	29.7%	34	45.9%	31	41.9%	30	40.5%
1	26	35.1%	30	40.5%	32	43.2%	28	37.8%	29	39.2%	34	45.9%	41	55.4%
0 or 1	66	89.2%	63	85.1%	66	89.2%	50	67.6%	63	85.1%	65	87.8%	71	95.9%
2	7	9.5%	11	14.9%	7	9.5%	22	29.7%	10	13.5%	9	12.2%	3	4.1%
3	1	1.4%	0	0.0%	1	1.4%	2	2.7%	1	1.4%	0	0.0%	0	0.0%
2 or 3	8	10.8%	11	14.9%	8	10.8%	24	32.4%	11	14.9%	9	12.2%	3	4.1%

Percent Scoring at Proficient/Advanced Performance Levels for Average Overall Total and Traits: Of the 74 student presentations, 23% achieved an overall total average score at the Proficient or Advanced performance levels (see Figure 1). Over 78% of the student presentations attained average scores at the Proficient or Advanced levels for the Language trait, 55% scored at Proficient or Advanced for Voice Delivery, 51% for Non-Verbal Delivery, and 42% scored at the Proficient or Advanced performance levels for the Supporting Material oral communication trait. For 36.5% of the student oral communication presentations, the Introduction and Organization traits both rendered average scores at the Proficient or Advanced performance levels, and 15% achieved average scores at the Proficient or Advanced levels for the Conclusion trait of the revised AAC&U Oral Communication Rubric.

Figure 1. Percent of Students Performing at the Proficient or Advanced Level by Oral Communication Trait and Overall Total Average Scores (N=74)



Descriptive Statistics for Average Overall Total and Trait Scores: The overall total scores for the Oral Communication presentations ranged from 0.50 to 3.79 with a mean or “average score” of 2.15, a median or “middle score” of 2.07, and a mode or “most frequent score” of 2.07 (see table 4). The consistency of the mean, median, and mode accompanied by the small standard deviation, or “average score distance” from the mean, of 0.513 indicate that most of the scores cluster together at the mean. Also, the score frequency sharply decreases towards the minimum and maximum values creating a more pointed/steep shape for the score distribution than that of the normal bell curve.

A larger standard deviation value signifies the scores are more spread out on the distribution, while a smaller standard deviation signifies that more of the scores are clustered at or near the mean, and the frequency of scores above or below the mean taper off towards the ends of the score distribution. If the mean is equal to the median, then the distribution is symmetric and has zero skewness. A positive skew occurs when the mean is greater than the median so the bulk of the score distribution is concentrated on the left of the mean (lower scores). A negative skew occurs when the mean is less than the median so the bulk of the score distribution is concentrated on the right of the mean (higher scores).

Table 4. Descriptive Statistics for Oral Communication Assessment Average Scores (N=74)

Statistic	Introduction	Organization	Language	Supporting Material	Voice Delivery	Non-Verbal Delivery	Conclusion	Overall Total
Mean	2.08	2.15	2.75	2.09	2.39	2.21	1.39	2.15
Median	2.00	2.00	2.83	2.00	2.50	2.50	1.50	2.07
Mode	2.00	2.00	3.00	2.00	2.00	2.50	1.50	2.07
Std. Dev.	0.691	0.622	0.691	0.646	0.663	0.688	0.816	0.513
Min.	0.33	0.00	0.00	0.00	0.00	1.00	0.00	0.50
Max.	4.00	4.00	4.00	4.00	4.00	4.00	3.00	3.79

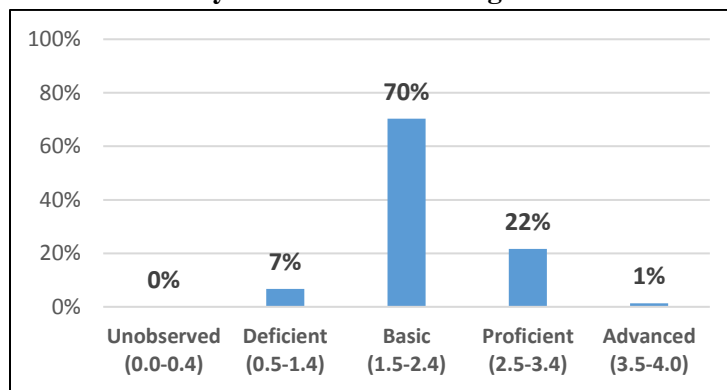
The Oral Communication AAC&U Rubric trait with the highest mean average score was Language at 2.75, followed by Voice Delivery at 2.39 and Non-Verbal Delivery at 2.21 (see Table 4). The median and mode for the average Language scores were a little higher than the mean, and the standard deviation among the scores was 0.691- meaning the Language scores were slightly negatively skewed more spread out on the continuum. The mode for Voice Delivery was lower than the mean, but the median was higher than the mean; thus, the Voice Delivery scores with a standard deviation of 0.663 were negatively skewed to a small extent but concentrated toward the mean. The standard deviation for the Non-Verbal Delivery scores was 0.688, and the mode and median were equal but slightly higher than the mean- indicating that the scores were faintly negatively skewed.

The Organization trait rendered the next highest average score mean at 2.15, followed by Supporting Material at 2.09, and Introduction at 2.08 (see Table 4). The mode and median for the Organization average scores were equal and barely lower than the mean- showing a very slight positive skew in their distribution, but are clustered fairly close to the mean with a standard deviation of 0.622. The median and mode for the Supporting Materials trait average scores were identical, and nearly equal to the mean, with a somewhat small standard deviation of 0.646- indicating an almost normal distribution with most scores near the mean. The median and mode for the Introduction trait average scores were also identical, and nearly equal to the mean, but with a slightly larger standard deviation of 0.691. This suggests an almost normal distribution with the Introduction scores spread out more around the mean.

The Oral Communication trait with the lowest mean average score was the Conclusion trait at 1.39 with a rather large standard deviation of 0.816 (see Table 4). The median and mode scores were equal, and slightly higher than the mean for the Conclusion trait, which denotes a somewhat negatively skewed distribution where the scores are more scattered from the mean. Five of the seven Oral Communication traits assessed revealed average score values that ranged from a minimum of 0.00 (Unobserved) to a maximum of 4.00 (Advanced). However, there were no average scores for the Conclusion trait above 3.00 (Proficient); nor were there any average scores for Non-Verbal Delivery below 1.0 (Deficient).

Percent Scoring at Each Performance Level for Average Overall Total and Traits: One percent of the 74 student presentations attained an overall total average score at the Advanced performance level (see Figure 2/Figure 4), 22% scored at the Proficient level, 70% at Basic, 7% at Deficient, and 0% of the overall total average scores for Oral Communication scored at the Unobserved performance level.

Figure 2. Percent of Presentations by Overall Total Average Score and Performance Level (N=74)

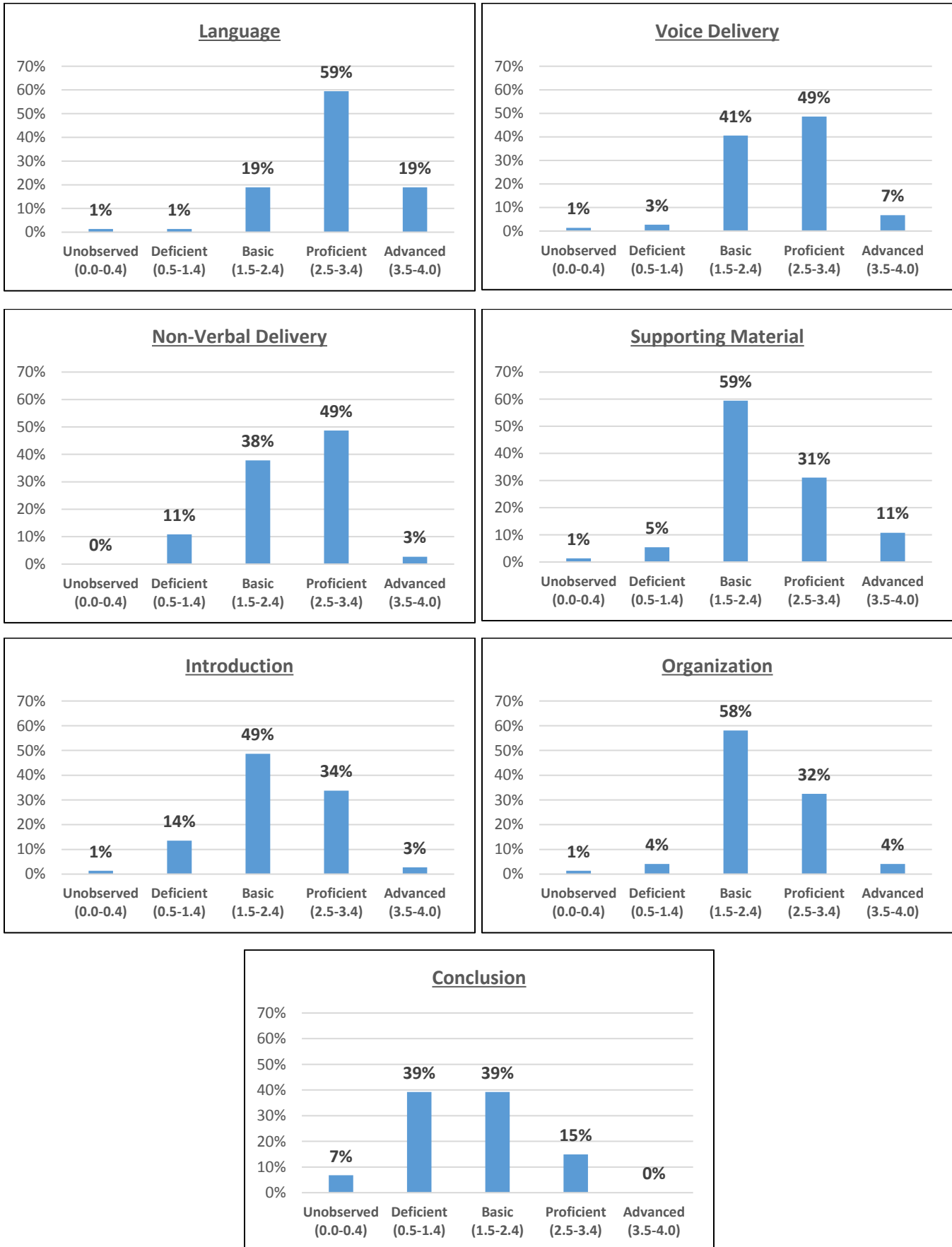


Nineteen percent of the oral communication presentations achieved an average Language trait score at the Advanced performance level, 59% scored at the Proficient level, 19% at the Basic level, 1% scored at the Deficient level, and 1% scored at the Unobserved performance level for the Language trait (see Figure 3/Figure 4). Seven percent of the presentations achieved an average Voice Delivery trait score at the Advanced performance level, 49% scored at the Proficient level, 41% at the Basic level, 3% scored at the Deficient level, and 1% scored at the Unobserved performance level for the Voice Delivery trait.

Three percent of the oral communication presentations achieved an average Non-Verbal Delivery trait score at the Advanced performance level, 49% scored at the Proficient level, 38% at the Basic level, 11% scored at the Deficient level, and 0% scored at the Unobserved performance level for the Non-Verbal Delivery trait (see Figure 3/Figure 4). Eleven percent of the presentations achieved an average Supporting Material trait score at the Advanced performance level, 31% scored at the Proficient level, 59% at the Basic level, 5% scored at the Deficient level, and 1% scored at the Unobserved performance level for the Supporting Material trait on the oral communication presentations.

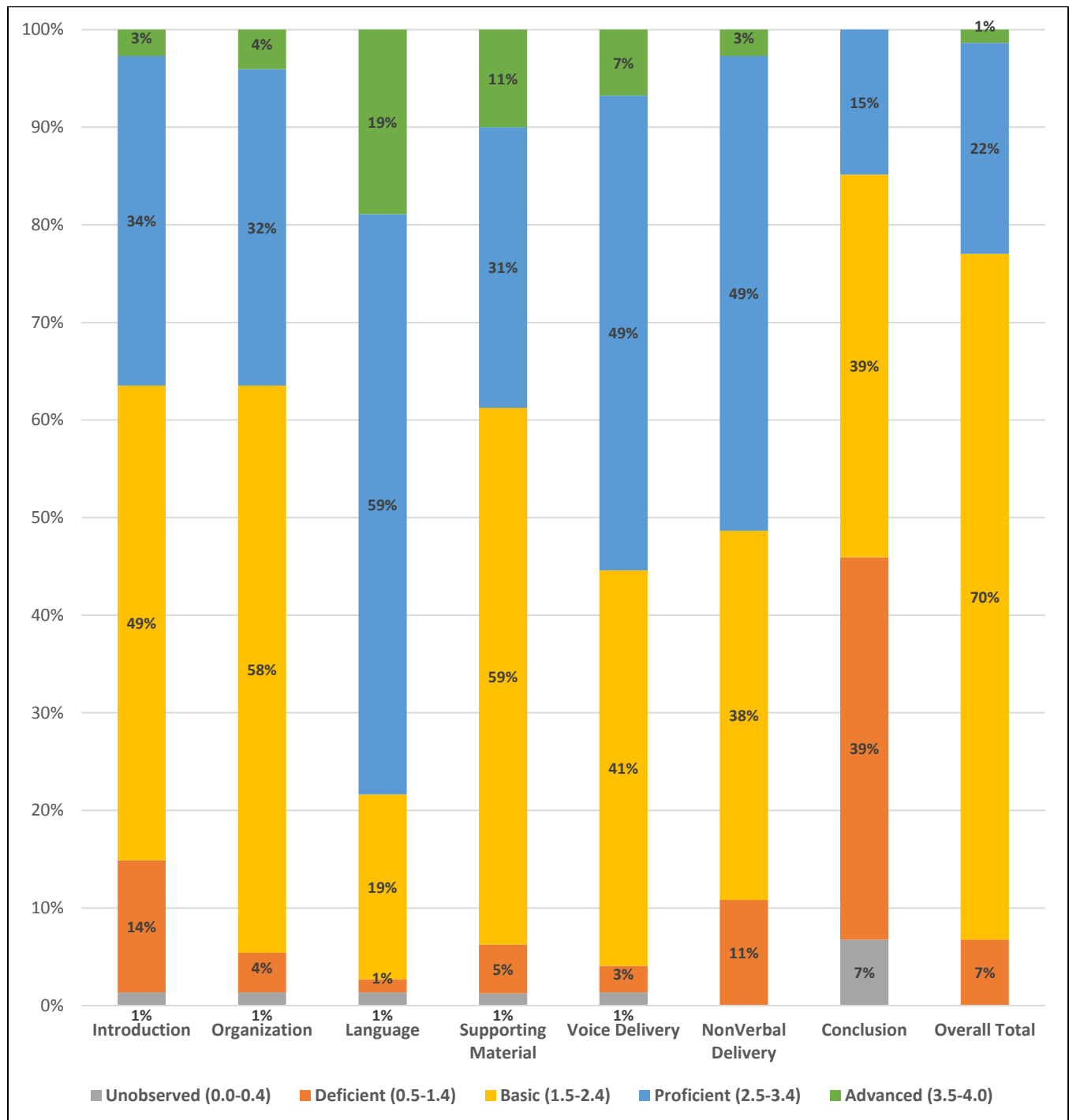
Three percent of the oral communication presentations achieved an average Introduction trait score at the Advanced performance level, 34% scored at the Proficient level, 49% at the Basic level, 14% scored at the Deficient level, and 1% scored at the Unobserved performance level for the Introduction trait (see Figure 3/Figure 4). Four percent of the presentations achieved an average Organization trait score at the Advanced performance level, 32% scored at the Proficient level, 58% at the Basic level, 4% scored at the Deficient level, and 1% scored at the Unobserved performance level for the Organization trait on the oral communication presentations.

Figure 3. Percent of Presentations by Average Trait Scores and Performance Levels (N=74)



Zero percent of the 74 student presentations assessed by at least two individual raters attained an average Conclusion trait score at the Advanced performance level, 15% scored at the Proficient level, 39% at the Basic level, 39% scored at the Deficient level, and 7% scored at the Unobserved performance level for Conclusion (see Figure 3/Figure 4). The Conclusion trait of the Oral Communication Assessment seems to be the trait in which Washburn students performed the lowest; however, students also scored lower on Introduction and Non-Verbal Delivery relative to the other oral communication traits assessed.

Figure 4. Comparison of Average Trait and Overall Total Scores by Percent of Student Presentations at each Performance Level (N=74)



Changes from Pilot 2014-2015 to 2015-2016 Oral Communication Results

Introduction: Considering that the Oral Communications assessment during 2014-2015 was a pilot project- the procedures for recording presentations to be rated were new for faculty, and it was the first time that the training for raters scoring using the revised AAC&U rubric was given- much had been learned about the process. The goals for the 2015-2016 Oral Communication Assessment were that it be based on better presentation recordings and that rater training concentrate on the AAC&U traits which were most disparate among raters during the pilot, and there would be more students participating in the assessment. The pilot included 35 presentations collected and scored, but this increased to 74 presentations for 2015-2016, the score differences between individual raters decreased and student scores for the average overall total and each oral communication trait increased.

Differences between Individual Raters: Approximately 12% of the overall average scores from the two raters did not differ in 2015-2016, compared to 11% in 2014-2015 (see Table 5). For 2015-2016 over 75% differed less than 1 point and almost 3% differed by 1 full point, compared to 57% differing less than 1 point and 14% differing by 1 point in 2014-2015. In 2015-2016, 90% of the presentations did not have overall average scores that were disparate enough to require a 3rd reader, and 9% had a difference of 1.14 to 2.00 points and an additional rater to score the presentations was necessary. In 2014-2015, 83% of the presentations did not have overall average scores more than 1 point disparate, and 17% had a difference of 1.14 to 1.72 points and an additional rater to score the presentations was necessary.

Table 5. Percent of Overall Total Average Differences between Raters by Difference Scores

Score Difference	2014-2015 (N=35)			2015-2016 (N=74)		
	Frequency	Percent	Cum. Percent	Frequency	Percent	Cum. Percent
0.00	4	11.4%	11.4%	9	12.1%	12.1%
0.14	1	2.9%	14.3%	21	28.4%	40.5%
0.28	2	5.7%	20.0%	0	0.0%	40.5%
0.29	1	2.9%	22.9%	8	10.8%	51.3%
0.43	4	11.4%	34.3%	11	14.8%	66.1%
0.57	3	8.5%	42.8%	6	8.1%	74.2%
0.58	2	5.7%	48.5%	0	0.0%	74.2%
0.71	2	5.7%	54.2%	8	10.8%	85.0%
0.72	3	8.5%	62.7%	0	0.0%	85.0%
0.86	2	5.7%	68.4%	2	2.7%	87.7%
1.00	5	14.3%	82.7%	2	2.7%	90.4%
1.14	2	5.7%	88.4%	1	1.4%	91.8%
1.28	1	2.9%	91.3%	0	0.0%	91.8%
1.29	0	0.0%	91.3%	1	1.4%	93.2%
1.43	1	2.9%	94.2%	2	2.7%	95.9%
1.57	0	0.0%	94.2%	2	2.7%	98.6%
1.58	1	2.9%	97.1%	0	0.0%	98.6%
1.72	1	2.9%	100.0%	0	0.0%	98.6%
2.00	0	0.0%	100.0%	1	1.4%	100.0%
Total	35	100.0%		74	100.0%	

The mean difference between raters for the overall average scores in 2014-2015 was 0.69 with a standard deviation of 0.445, and differences ranged from 0 to 1.72 (see Table 6). The mean difference between raters for the overall average scores for 2015-2016 decreased by .024 to 0.45, the standard deviation decreased by 0.013 to 0.432, and for the minimum score difference was the same at 0 but the maximum difference increased by 0.28 to 2.00.

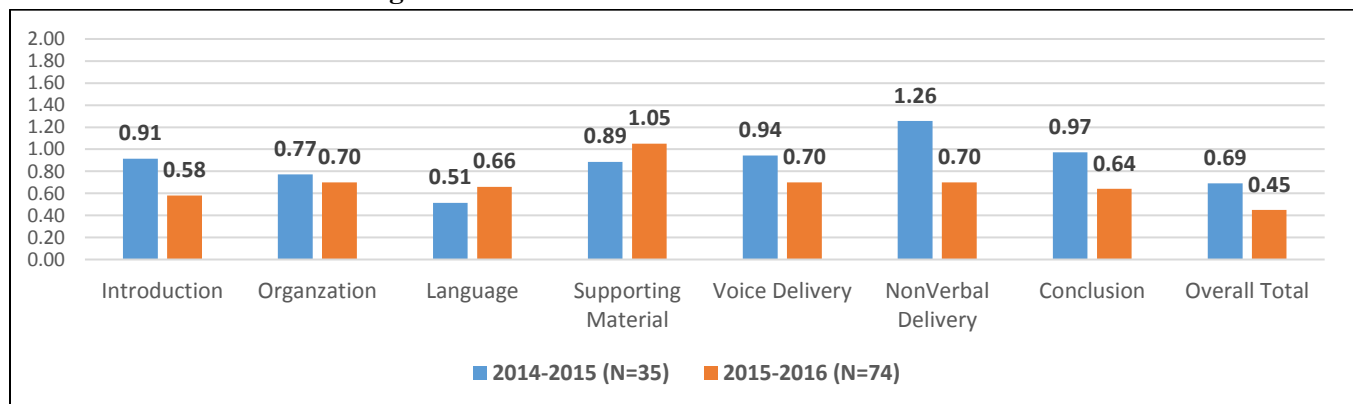
Table 6. Descriptive Statistics for Rater Score Differences

Statistic	Introduction		Organization		Language		Supportive Material		Voice Delivery		Non-Verbal Delivery		Conclusion		Overall Total	
	14-15	15-16	14-15	15-16	14-15	15-16	14-15	15-16	14-15	15-16	14-15	15-16	14-15	15-16	14-15	15-16
Mean	0.91	0.58	0.77	0.70	0.51	0.66	0.89	1.05	0.94	0.70	1.26	0.70	0.97	0.64	0.69	0.45
Std.Dev.	0.781	0.722	0.598	0.716	0.658	0.707	0.796	0.842	0.765	0.754	0.780	0.677	0.891	0.563	0.445	0.432
Min.	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Max.	3.00	3.00	2.00	2.00	2.00	3.00	3.00	3.00	3.00	3.00	3.00	2.00	3.00	2.00	1.72	2.00

For the oral communication rubric traits scored in 2014-2015 the differences between raters ranged from 0 to 2 points for Organization and Language, and ranged from 0 to 3 points for Introduction, Supportive Material, Voice Delivery, Non-Verbal Delivery, and Conclusion. For 2015-2016 the differences between raters ranged from 0 to 2 points for Organization, Non-Verbal Delivery, and Conclusion, and ranged from 0 to 3 points for Introduction, Language, Supportive Material, and Voice Delivery (see Table 6). The range of rater score differences between years was consistent for the traits of Introduction, Organization, Supportive Material, and Voice Delivery; the range of the rater score differences increased for Language, but decreased for the traits of Non-Verbal Delivery and Conclusion.

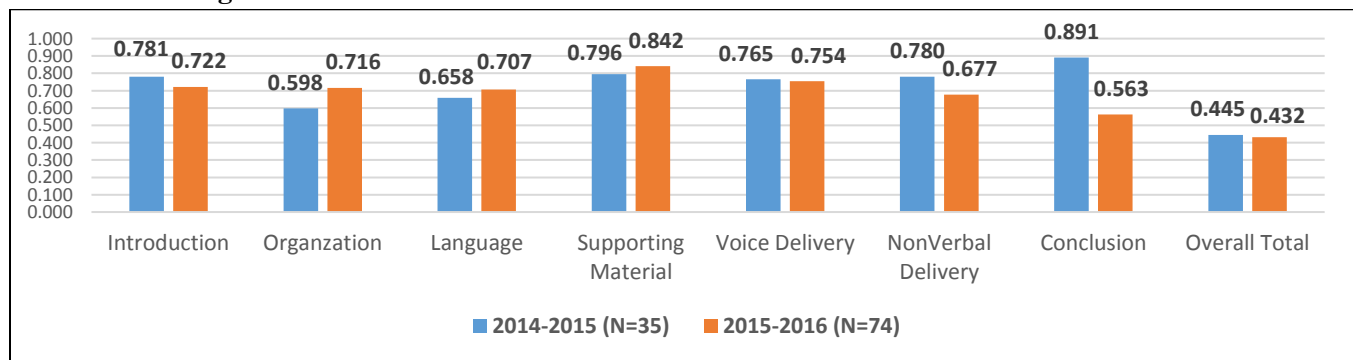
For the pilot assessment in 2014-2015, the mean differences between raters' scores for the AAC&U oral communication rubric traits ranged from 0.51-1.26, but most mean differences decreased for 2015-2016 for a range of 0.45-1.05 (see Table 6). In 2014-2015 the Non-Verbal Delivery trait had the highest mean difference at 1.26 but decreased to 0.70 in 2015-2016 (see Table 6/Figure 5). The Conclusion trait rater score mean difference for the pilot was 0.97 but decreased the next year to 0.64, the Voice Delivery mean difference decreased from 0.94 to 0.70, the Introduction trait decreased from 0.91 to 0.58, and the mean score difference between raters for Organization decreased slightly from 0.77 to 0.70. However, the mean difference between rater scores for Language in 2014-2015 was 0.51 and increased to 0.66 in 2015-2016, and the Supporting Materials trait mean difference between rater scores increased from 0.89 to 1.05- becoming the oral communication trait with the highest mean difference in 2015-2016.

Figure 5. Mean Differences between Raters' Scores



In 2014-2015, the standard deviation of the mean differences between raters' scores for the AAC&U oral communication rubric traits ranged from 0.445-0.891, but most of the standard deviations decreased for 2015-2016 for a range of 0.432-0.842 (see Table 6/Figure 6). For the pilot assessment, the largest standard deviation of the mean differences between raters' scores was 0.891 for the Conclusion trait which decreased to 0.563 in 2015-2016, the standard deviation for the Introduction trait decreased from 0.781 to 0.722, Non-Verbal decreased from 0.780 to 0.677, and the standard deviation for the Voice Delivery trait decreased slightly from 0.765 to 0.754. The standard deviation of the mean difference between rater's scores for the Language trait increased from 0.658 to 0.707, Organization increased from 0.598 to 0.716, and the standard deviation for the Supporting Material trait increased from 0.796 to 0.842- becoming the oral communication trait with the highest standard deviation of the mean differences in 2015-2016.

Figure 6. Standard Deviations of the Mean Differences between Raters' Scores



Overall, the majority of the mean differences between rater's scores for the AAC&U oral communication rubric traits and the average total score decreased from the pilot in 2014-2015 to the 2015-2016 assessment, as did the standard deviations of those means- showing that individual raters' scores became closer together, and the validity and reliability of the scores given by the individual raters have improved.

Table 7. Percent of Presentations by Rounded Individual Rater Score Differences of Oral Communication Traits

Difference between Raters (Rounded)	Introduction		Organization		Language		Supporting Material		Voice Delivery		Nonverbal Delivery		Conclusion	
	2014-2015	2015-2016	2014-2015	2015-2016	2014-2015	2015-2016	2014-2015	2015-2016	2014-2015	2015-2016	2014-2015	2015-2016	2014-2015	2015-2016
0	31%	54%	31%	45%	57%	46%	34%	30%	29%	46%	17%	42%	31%	41%
1	49%	35%	60%	41%	34%	43%	46%	38%	51%	39%	43%	46%	49%	55%
0 or 1	80%	89%	91%	85%	91%	89%	80%	68%	80%	85%	60%	88%	80%	96%
2	17%	10%	9%	15%	9%	10%	17%	30%	17%	14%	37%	12%	11%	4%
3	3%	1%	0%	0%	0%	1%	3%	3%	3%	1%	3%	0%	9%	0%
2 or 3	20%	11%	9%	15%	9%	11%	20%	32%	20%	15%	40%	12%	20%	4%

Frequencies and percentages of the difference scores for each trait on the Oral Communication rubric were computed; and the 0- and 1-point score differences, as well as the 2- and 3-point score differences, were combined to provide frequencies and percentages for the closer and more widely disparate scores by trait (see Table 7). The rater scores for the traits of Language, Organization, and Supporting Materials became more disparate from 2014 to 2015 as the rater scores for the traits of Non-Verbal Delivery, Conclusion, Introduction, and Voice Delivery became closer or more consistent.

From 2014-2015 to 2015-2016, the percent of individual raters' scores differing by 0-1 points that increased substantially from 60% to 88% for the Non-Verbal Delivery trait of Oral Communication (see (Table 7), and a considerable increase from 80% to 96% was revealed for the Conclusion trait. The consistency between raters for the Introduction trait, as defined by a small score difference of 0-1 points, increased moderately from 80% to 89%, as did the rater consistency for the Voice Delivery trait increasing from 80% in 2014-2015 to 85% in 2015-2016. The percent of 0-1 point differences between individual raters' scores for the Language trait of the Oral Communication Assessment decreased from 91% in 2014-2015 to 89% in 2015-2016. The Organization trait showed a decreased in consistency between raters as the percent of smaller differences of 0-1 points between individual rater scores decreased from 91% to 85%, and the Supporting Material trait revealed a sizable decrease from 80% in 2014-2015 to 68% in 2015-2016. The trait of Supporting Material showed the largest decrease in percent of small, 0-1 point, differences between individual rater scores and rendered the lowest percent in 0-1 point differences between raters at 68%, while the percent of 2-3 point differences between individual rater scores reached 32% (see Table 7).

Descriptive Statistics for Average Overall Total and Trait Scores:

For the pilot in 2014-2015, the overall average total scores for the 35 oral communication presentations ranged from 1.29 to 3.72 with a mean of 2.15 and standard deviation of 0.538. In 2015-2016, the overall average total scores for the 74 oral communication presentations ranged from 0.50 to 3.79 with a mean of 2.15 and a standard deviation of 0.513 (see Table 8). The mean overall average score did not change from the pilot to 2015-2016 and the range of the overall average scores increased, but the standard deviation of scores from the mean decreased by 0.025.

Table 8. Descriptive Statistics Changes for Oral Communication Average Scores

Statistic	Introduction		Organization		Language		Supporting Material		Voice Delivery		Non-Verbal Delivery		Conclusion		Overall Total	
	2014-2015	2015-2016	2014-2015	2015-2016	2014-2015	2015-2016	2014-2015	2015-2016	2014-2015	2015-2016	2014-2015	2015-2016	2014-2015	2015-2016	2014-2015	2015-2016
2014-2015 N=35 2015-2016 N=74	2014-2015	2015-2016	2014-2015	2015-2016	2014-2015	2015-2016	2014-2015	2015-2016	2014-2015	2015-2016	2014-2015	2015-2016	2014-2015	2015-2016	2014-2015	2015-2016
Mean	1.64	2.08	2.14	2.15	2.87	2.75	2.27	2.09	2.48	2.39	2.30	2.21	1.30	1.39	2.15	2.15
Median	1.50	2.00	2.00	2.00	3.00	2.83	2.33	2.00	2.50	2.50	2.33	2.50	1.33	1.50	2.09	2.07
Mode	1.50	2.00	2.00	2.00	3.00	3.00	2.00	2.00	2.00	2.00	2.00	2.50	1.50	1.50	2.00	2.07
Std. Dev.	0.725	0.691	0.636	0.622	0.511	0.691	0.749	0.646	0.657	0.663	0.815	0.688	0.945	0.816	0.538	0.513
Minimum	0.33	0.33	1.00	0.00	1.50	0.00	1.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.29	0.50
Maximum	4.00	4.00	4.00	4.00	3.67	4.00	4.00	4.00	4.00	4.00	3.67	4.00	3.50	3.00	3.72	3.79

The descriptive statistics for the seven AAC&U oral communication rubric traits assessed revealed that the Language trait attained the highest mean score for both the pilot in 2014-2015 and in 2015-2016; however, the Language mean decreased by 0.12 from 2.87 to 2.75 and the standard deviation increased by 0.18 from 0.511 to 0.691 (see Table 8). The Voice Delivery trait obtained the second highest mean for both years, but it decreased slightly by 0.09 from 2.48 to 2.39 as the standard deviation increased slightly by 0.006 from 0.657 to 0.663. The trait with the third highest mean for both years of the oral communication assessment was Non-Verbal Delivery. The mean score for the Non-Verbal Delivery trait decreased slightly by 0.09 from 2.30 in 2014-2015 to 2.21 in 2015-2016, but the standard deviation also decreased by 0.127 from 0.815 to 0.688.

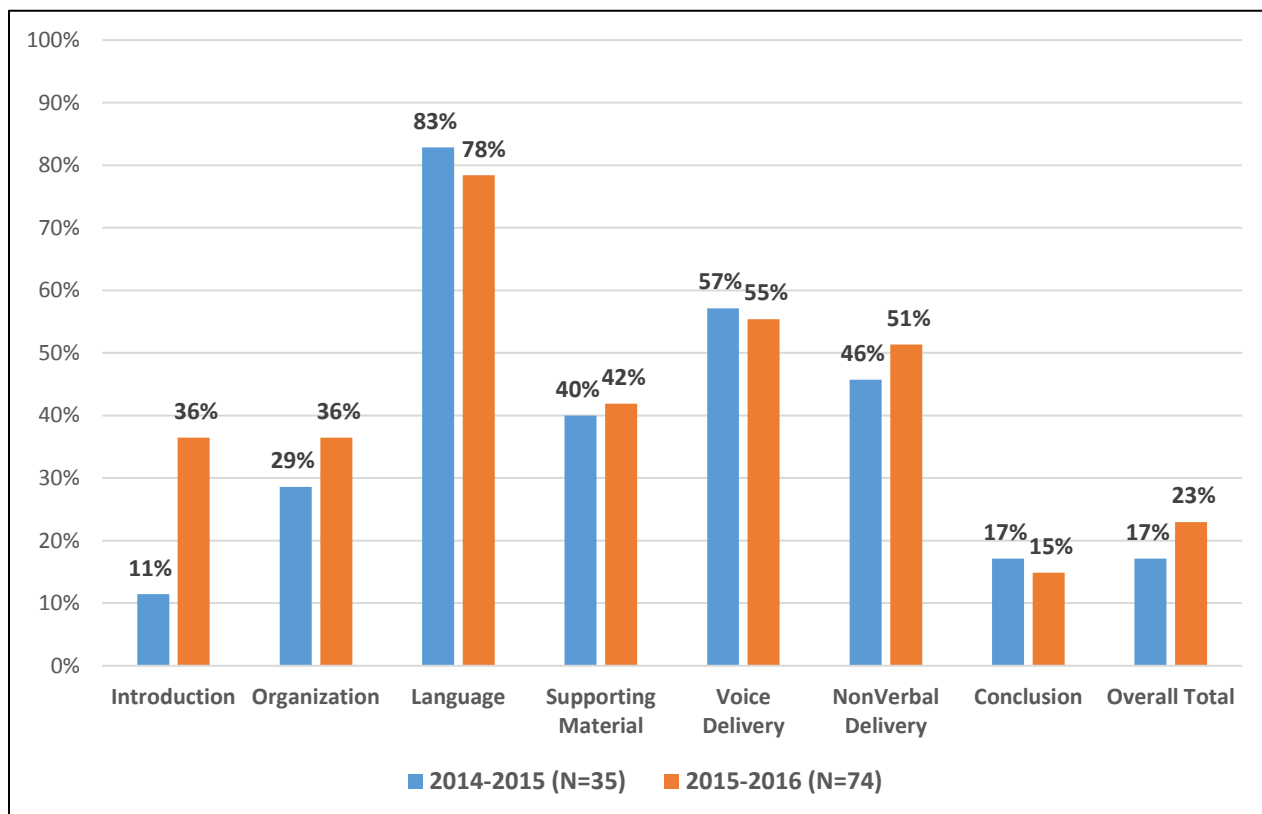
The Supporting Material oral communication trait achieved the fourth highest mean during the pilot in 2014-2015 but fifth in 2015-2016, and the Organization trait attained the fifth highest mean during the pilot but fourth in 2015-2016 (see Table 8). The mean for the Supporting Material trait decreased by 0.18 from 2.27 to 2.09 and the standard deviation decreased by 0.103 from 0.749 to 0.646. The mean for the Organization trait increased by 0.01 from 2.14 in 2014-2015 to 2.15 in 2015-2016, and the standard deviation decreased by 0.014 from 0.636 to 0.622.

Introduction was the oral communication trait that achieved the sixth highest mean of the seven traits assessed traits for both years. The mean for Introduction during the pilot in 2014-2015 was 1.64 and increased by 0.44 to 2.08 in 2015-2016, and the standard deviation decreased by 0.034 from 0.725 to 0.691 (see Table 8). The trait with the lowest mean for both years was Conclusion. The Conclusion trait mean in 2014-2015 was 1.30 but increased by 0.09 to 1.39 in 2015-2016, and the standard deviation decreased by 0.129 from 0.945 to 0.816.

Percent Scoring at Proficient/Advanced Performance Levels for Average Overall Total and Traits:

In 2014-2015, 17% of the 35 student presentations assessed using the revised AAC&U Oral Communication rubric attained a rounded overall average score of 3.0 or above, and were classified as reaching the Proficient or Advanced performance levels (see Figure 7). The percent achieving a rounded overall mean score of 3.0 or higher in 2015-2016 had increased by almost 6 percentage points, and 23% of the 74 student presentations were categorized at the Proficient or Advanced performance levels.

Figure 7. Comparison of the Percent of Student Presentations Scoring at Proficient or Advanced Performance Levels by Trait and Overall Average



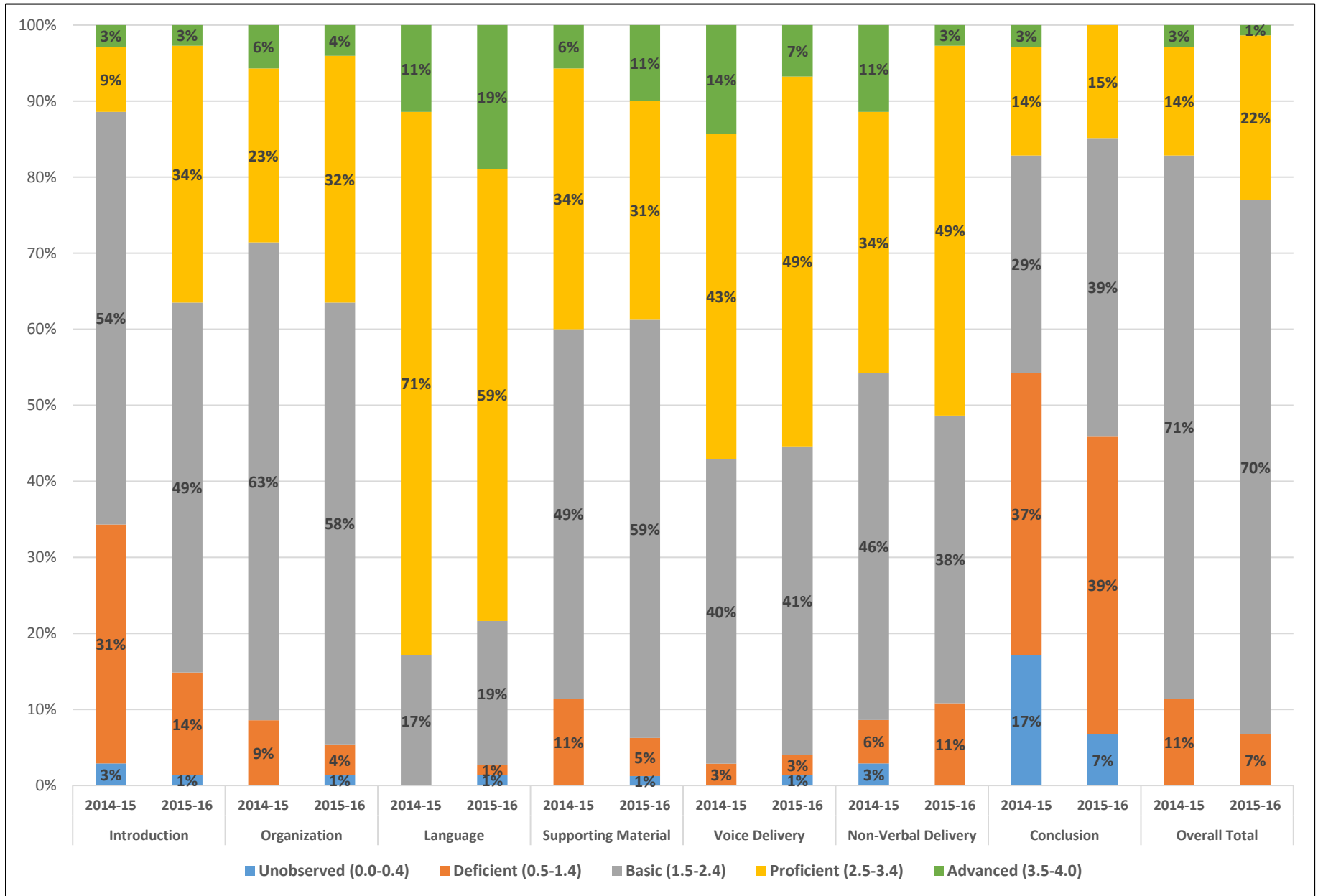
The Oral Communication traits that rendered the highest percent of scores classified at the Proficient or Advanced performance levels were Language, Voice Delivery, Non-Verbal Delivery, and Supporting Material. For the Language trait 83% of scores were classified at the Proficient or Advanced performance levels in 2014-2015, but a decrease of 5 percentage points led to 78% of the Language trait scores at 3.0 or higher (see Figure 7). Fifty-seven percent of the Voice Delivery trait scores in 2014-2015 were at the Proficient/Advanced levels with rounded scores of 3.0 or higher, and 55% of the Voice Delivery trait scores were classified as Proficient or Advanced in 2015-2016 after a 2 percentage point decrease. The Non-Verbal Delivery trait produced 46% of scores at the Proficient or Advanced performance levels for the pilot, and after a 5 percentage point increase for 2015-2016, 51% of the student presentations had rounded Non-Verbal Delivery trait scores of 3.0 or higher. The Supporting Material trait rendered 40% of scores categorized as Proficient or Advanced during the pilot, and increased by 2 percentage points for 42% of Supporting Material scores reaching Proficient or Advanced in 2015-2016.

The Oral Communication traits that rendered the lowest percent of scores classified at the Proficient or Advanced performance levels were Conclusion, Introduction, and Organization. For the Introduction trait in 2014-2015, 11% of the student presentations were categorized as being at the Proficient or Advanced performance levels as they achieved a rounded average score of 3.0 or higher, and a 25 percentage point increase led to 36% being classified at the Proficient or Advanced performance levels in 2015-2016 (see Figure 7). Twenty-nine percent of the average Organization trait scores reached 3.0 or higher in 2014-2015- placing them in the Proficient or Advanced performance levels. In 2015-2016, an increase of 7 percentage points produced 36% of Organization trait scores that were categorized as Proficient or Advanced. Seventeen percent of the oral communication presentations in 2014-2015 attained Conclusion trait scores of 3.0 or higher being categorized at the Proficient or Advanced performance levels, and after a 2 percentage point decrease, 15% of the Conclusion trait scores were considered Proficient or Advanced in 2015-2016.

The largest increase in the percent of rounded traits scores considered to be at the Proficient or Advanced performance levels from 2014-2015 to 2015-2016 was observed for the Introduction trait with an increase 25 percentage points (see Figure 7). The Organization trait experienced an increase of 7 percentage points, Non-Verbal Delivery increased 5 percentage points, and the Supporting Material trait increased 2 percentage points from 2014-2015 to 2015-2016 in the percent of rounded trait scores classified at the Proficient or Advanced performance levels. The largest decrease in the percent of rounded traits scores considered to be at the Proficient or Advanced performance levels from 2014-2015 to 2015-2016 was observed for the Language trait with a decrease of 5 percentage points. The Voice Delivery and Conclusion traits each decreased 2 percentage points from 2014-2015 to 2015-2016 in the percent of rounded trait scores classified at the Proficient or Advanced performance levels.

The 2014-2015 and 2015-2016 percentages of student oral communication presentations scoring at the Unobserved, Deficient, Basic, Proficient, and Advanced performance levels for the average overall total and the seven separate oral communication trait scores are shown in Figure 8.

Figure 8. 2014-2015 and 2015-2016 Comparison of Average Trait and Overall Total Scores (Rounded) by Percent of Student Presentations at each Performance Level (2014-2015 N=35 and 2015-2016 N=74)



Conclusion:

In 2015-2016, 74 previously recorded student presentations were scored by two separate raters using a revised version of the AAC&U Oral Communication Rubric. Seven separate traits (Introduction, Organization, Language, Supporting Material, Voice Delivery, Nonverbal Delivery, and Conclusion) were rated on a five-point scale of 0=Unobserved, 1=Deficient, 2=Basic, 3=Proficient, and 4=Advanced. The trait scores were averaged to calculate an overall average score for each student presentation. Twelve percent of the overall average scores from the two raters did not differ, 75% differed less than 1 point, and 3% differed by 1 full point. A total of 9% of the 74 overall average rater scores had a difference of 1.14 to 2.00 points, and an additional rater was required to score the presentation. The individual raters were fairly consistent in their scores for the Introduction and Language traits with 89% of the ratings having score differences equal to 0 or 1; followed by Non-Verbal Delivery at 88%, Organization and Voice Delivery at 85% of raters scores having differences of 0 or 1. Supporting Material seemed to be the oral communication trait in which raters were least consistent with 68% of the scores differing by 0 or 1, and 32% of the Supporting Material trait rater scores differing by 2-3 points.

Twenty-three percent of the student presentations achieved an overall total average score that after rounding would be classified at the Proficient or Advanced performance levels. One percent of the presentations attained an overall total average score at the Advanced performance level, 22% scored at the Proficient level, 70% at Basic, 7% at Deficient, and 0% scored at the Unobserved performance level. Seventy-eight percent of student presentations attained average scores classified as Proficient or Advanced for the Language trait, 55% for Voice Delivery, 51% for Non-Verbal Delivery, 42% for Supporting Material, 36.5% for Introduction and Organization, and 15% were Proficient/Advanced for the Conclusion trait. The Language trait had the highest mean average score at 2.75, followed by Voice Delivery at 2.39 and Non-Verbal Delivery at 2.21; Organization rendered a mean score of 2.15, followed by Supporting Material at 2.09, and Introduction at 2.08; the Conclusion trait had the lowest mean at 1.39. The Conclusion trait seems to be where Washburn students performed the lowest; however, students also scored lower on the Introduction and Non-Verbal Delivery traits relative to the other oral communication traits assessed.

The changes in rater differences and student scores between the pilot in 2014-2015 and the assessment administered in 2015-2016 were examined. The number of student presentations scored increased from 35 to 74, the percentage of presentations in which the overall average scores from the two raters did not differ increased from 11% in 2014-2015 to 12% in 2015-2016, and the percent of presentations in which the overall average scores from the two raters differed more than one point decreased from 17% to 9%. The mean difference between raters for the overall average scores in 2014-2015 was 0.69 with a standard deviation of 0.445, and in 2015-2016 the mean difference decreased to 0.45 with the standard deviation decreasing to 0.432. The mean differences between raters' scores for the oral communication traits decreased from the pilot in 2014-2015 to the 2015-2016 assessment, as did the standard deviations of those means- showing that individual raters' scores became closer together, and the validity and reliability of the scores given by the individual raters have improved.

Seventeen percent of the 35 student presentations in 2014-2015 attained a rounded overall average score classified at the Proficient or Advanced performance levels, but this increased by almost 6 percentage points in 2015-2016, and 23% of the 74 overall average scores were categorized at the Proficient or Advanced performance levels. The percent of rounded scores classified at the Proficient or Advanced performance levels increased from 2014-2015 to 2015-2016 for the Oral Communication traits of Non-Verbal Delivery (46% to 51%), Supporting Material (40% to 42%), Organization (29% to 36%), and Introduction (11% to 36%). However, the percent of rounded scores classified at the Proficient or Advanced performance levels decreased from 2014-2015 to 2015-2016 for the Oral Communication traits of Language (83% to 78%), Voice Delivery (57% to 55%), and Conclusion (17% to 15%). From 2014-2015 to 2015-2016 the mean scores increased for the Conclusion (1.30 to 1.39), Introduction (1.64 to 2.08), and Organization (2.14 to 2.15) traits; the mean for the overall average score remained the same at 2.15; and the mean scores decreased for the Language (2.87 to 2.75), Voice Delivery (2.48 to 2.39), Non-Verbal Delivery (2.30 to 2.21), and the Supporting Material (2.27 to 2.09) traits.

The Oral Communication traits that rendered the highest percent of scores classified at the Proficient or Advanced performance levels were Language, Voice Delivery, Non-Verbal Delivery, and Supporting Material. The Oral Communication traits that rendered the lowest percent of scores classified at the Proficient or Advanced performance levels were Conclusion, Introduction, and Organization. In terms of what the assessment data suggests for written communication instruction in the future, perhaps more concentration should be placed upon the traits of Introduction, Organization, and Conclusion. The Introduction and Organization traits rendered only 36.5% of student papers reaching the Proficient or Advanced performance levels, and the Conclusion trait rendered 15% of the student papers scoring in the Proficient or Advanced performance levels. Improving these individual oral communication traits will improve the overall total average scores for the oral communication assessment.

The training held for the individual raters assessing the oral communication assessments seemed to be more effective in 2015-2016 than for the pilot assessment in 2014-2015. Ninety-one percent of the overall total average scores for student presentations provided by pairs of individual raters were not disparate enough to warrant an additional rater, but 9% of presentations did have overall total average scores that differed more than 1 point between rater pairs and an additional rater was necessary. Pairs of raters differed one point or less for the Conclusion oral communication trait on 96% of student presentations, followed by the Introduction and Language traits at 89%, and pairs of raters differed one point or less on the Non-Verbal Delivery trait for 88%. However, the Organization and Voice Delivery traits showed 85% of rater pairs with 0-1 point difference, and 15% with differences between raters of 2-3 points. The Supporting Material trait has the lowest inter-rater reliability with only 68% of rater pairs with 0-1 point differences, and 32% with 2-3 point differences. This inter-rater reliability could be improved by stressing the scoring of the Supporting Material, Organization, and Voice Delivery traits and providing extra examples during the rater training.