

Speech Assessment Report – Spring 2016

Author: Joseph F. van Gaalen, Ph.D., Director, Academic Assessment

1 INTRODUCTION

Florida SouthWestern's Speech Department has employed a common rubric used by all faculty as a means to evaluate an agreed upon series of student level outcomes. With a goal towards increasing student oral communication achievement, faculty have focused on a series of Student Learning Objectives (SLOs) using the rubric dimensions Introduction, Organization, Support, Oral Documentation, Language, NV-Vocal, NV-Physical, Presentation Media, Attire, and Conclusion, in a formative speech common assessment. Additional department goals for assessment include comparing results of SPC 1017, Fundamentals of Speech Communication, with that of SPC 2608, Introduction to Public Speaking, and comparisons by campus, dual enrollment/non-dual enrollment, and online/traditional, when applicable. These correlative measures will serve as support for instructive improvement (Cole et al., 2011; Elder and Paul, 2007).

For additional detail or further analysis not provided in this report, please contact Dr. Joseph F. van Gaalen, Director of Academic Assessment, Academic Affairs (jfvangaalen@fsw.edu; x16965).

2 LEARNING OBJECTIVES AND DESCRIPTIVE STATISTICS

Using common rubric criterion as an assessment method, in the 2014-15 academic year the FSW Speech faculty defined three areas of interest for evaluation that apply to both SPC 1017 and SPC 2608 and set goals appropriately. Results from that year showed that students met expectations (> 60% at achievement level). As a result, faculty determined that goals can be raised (see below for new goals).

The rubric dimensions are modeled on a 5-point scale where a score of 0 indicates "Insufficient", 1 indicates "Beginning" level, 2 indicates the "Developing" level, 3 indicates the "Accomplished" level, and 4 indicates the "Exemplary", or highest level. The SLOs and their measure of success are:

SLO1: Students will be able to prepare and deliver an effective speech introduction that gains the attention of the audience, establishes speaker credibility, relates the topic to the audience, identifies the topic and previews the main points of the speech. The faculty established measure of success for this SLO is a rating of "Developing" or higher for 70% of the students.

SLO2: Students will be able to provide effective speech closure by delivering a conclusion that signals the end, reviews purpose and main points and that ends smoothly and memorably. The faculty established measure of success for this SLO is a rating of "Developing" or higher for 70% of the students.

SLO3: Students will know how to avoid plagiarizing when speaking by incorporating an oral citation that includes appropriate information. The faculty established measure of success for this SLO is a rating of "Developing" or higher for 70% of the students.

2.1 SPC 1017

2.1.1 Learning Objectives

For the spring 2016 assessment, 548 artifacts were collected for SPC 1017 from 31 of 42 course sections. In some cases, rubric scores could either not be accessed or located. Faculty established measures of success for SLO1 was a rating of “Developing” or higher (≥ 2) in rubric dimension Introduction for 70% of the students. Spring 2016 artifacts successfully meet the goal as results exhibit 93% of artifacts scored level 2 or greater (Table 1). For SLO2, a rating of “Developing” or higher (≥ 2) in rubric criteria Conclusion for 70% of the students, spring 2016 artifacts successfully meet the goal as results exhibit 94% of artifacts scored level 2 or greater. For SLO3, a rating of “Developing” or higher (≥ 2) in rubric criteria Oral Documentation for 70% of the students, spring 2016 artifacts successfully meet the goal as results exhibit 82% of artifacts scored level 2 or greater. For a visual comparison of scores by dimension, see Section 2.3.

Rubric Score	Introduction	Organization	Support	Oral Documentation	Language	NV-Vocal	NV-Physical	Presentation Media	Attire	Conclusion
Developing or higher	93%	98%	98%	82%	99%	98%	96%	88%	98%	94%
4	33.9%	43.2%	88.1%	27.2%	49.1%	25.4%	67.0%	39.8%	77.6%	29.0%
3	34.1%	41.8%	10.2%	37.2%	45.3%	45.6%	28.6%	37.0%	17.0%	39.2%
2	25.0%	12.8%	0.0%	17.7%	4.9%	27.4%	0.0%	10.9%	3.6%	25.4%
1	6.4%	2.0%	1.1%	9.5%	0.2%	0.9%	3.8%	3.8%	0.9%	3.1%
0	0.4%	0.0%	0.0%	7.8%	0.0%	0.2%	0.0%	7.8%	0.4%	2.7%

Table 1. Percentage of student achievement level by rubric dimension (includes percentage of students scoring in developmental level or higher as per SLOs) for SPC 1017. Rubric dimensions identified in SLOs in blue.

Had a similar goal been set for the remaining rubric dimensions not focused in SLOs 1, 2, or 3, all would have been met (Table 1). Of these, the highest scoring dimension was Language at 99% while the lowest scoring dimension was Presentation Media at 88%. Organization, Support, NV-Vocal, NV-Physical, and Attire all exhibit greater than 96% achieving the goal.

Had the goal for achievement been set at “Accomplished” or (≥ 3) for 70% of the students all dimensions again would have been met with the exception of Introduction (68%), Oral Documentation (64%), and Conclusion (68%). The highest scoring dimension was Support at 96% while the lowest scoring dimension was Oral Documentation at 64%.

2.1.2 Descriptive Statistics & Longitudinal Data

Descriptive statistics for SPC 1017 artifacts can be found in Table 2. A histogram of artifact scores for all 548 artifacts is shown in Figure 1. To describe the behavior of the rubric dimensions based on overall achievement a color map, or binary raster image was created by calculating the mean scores for each dimension as a function of combined score (Figure 2). To create this image the rubric scores (4, 3, 2, 1, or 0) for each artifact was grouped based on combined raw rubric score (10 dimensions x maximum rubric level of 4 = 40 overall points). The color represents the mean rubric score achieved in each dimension based on the combined score as shown in the x-axis.

	<i>Introduction</i>	<i>Organization</i>	<i>Support</i>	<i>Oral Documentation</i>	<i>Language</i>	<i>NV-Vocal</i>	<i>NV-Physical</i>	<i>Presentation Media</i>	<i>Attire</i>	<i>Conclusion</i>
n	548	548	548	548	548	548	548	548	548	548
Max	10	10	15	10	10	10	15	5	5	10
Min	0	3	5	0	3	0	5	0	0	0
Mode	8	10	15	8	10	8	12	5	5	8
Mean	7.8	8.5	13.2	7.1	8.9	7.9	11.5	3.9	4.7	7.7
Standard deviation	2.04	1.58	2.18	2.89	1.20	1.56	2.44	1.38	0.64	2.15

Table 2. Descriptive statistics for SPC 1017 common course assessment.

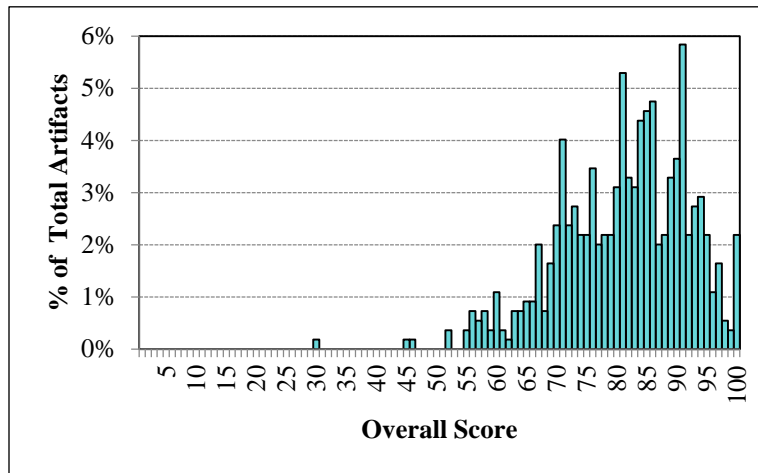


Figure 1. Overall score distribution for SPC 1017 artifacts (spring 2016 term).

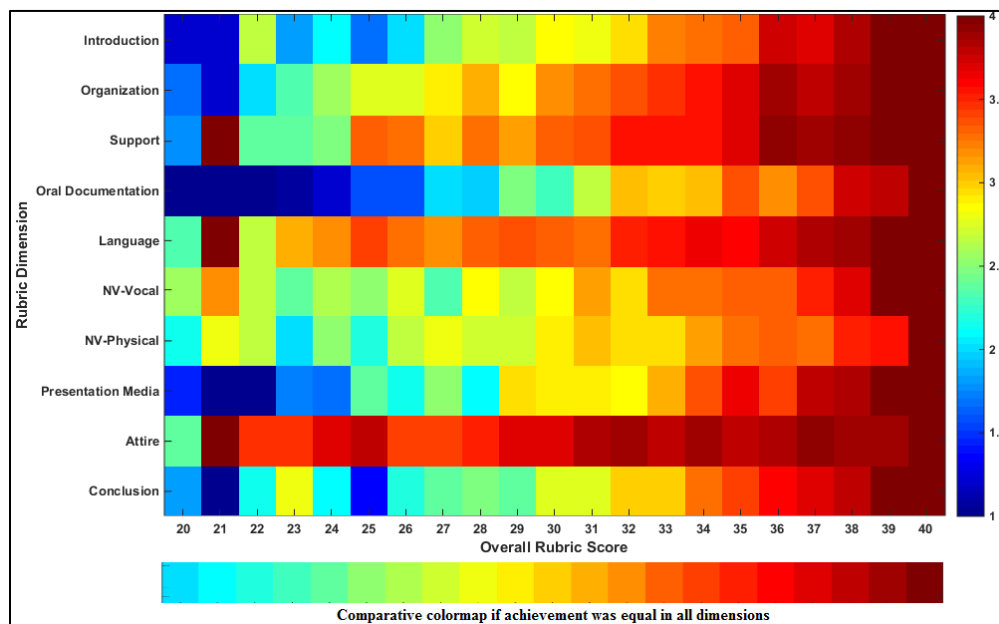


Figure 2. (Top) Colormap of mean scores for each rubric dimension (range: 0-4) based on overall rubric score (combined rubric score of all dimensions, max=40) for SPC 1017. (Bottom) Comparison rubric dimension if dimension score is the same as overall (i.e. artifact overall score is equally distributed across all sections). A rubric dimension with hotter colors (reds/yellows) means that dimension achievement exceeds the overall score and is an area of strength. An exam section with colder colors (blues/greens) means that section achievement is lower than the overall score and is therefore an area of weakness.

A review of the colormap in Figure 2 shows that at 34/40 and above (average rubric score of 3.4 or higher) all dimensions fair relatively equally (hot colors fairly evenly distributed). This is similar to fall 2015 data in which 33/40 and above exhibited relatively equal strength. When overall rubric scores range from 25-33, the Support, Language, and Attire dimensions exhibit strong scores even when the overall score is somewhat lower. For example, at an overall score of 27, those three dimensions exhibit average scores of 3.0/4, 3.2/4, and 3.4/4, well above the average rubric score at that level, 2.7/4. This is somewhat different from fall 2015 data in which Organization, Language, and Attire exhibited this behavior. When overall rubric scores range below 27, Oral Documentation and Presentation Median are weaker than the others. This is also different from fall data in which Support, Oral Documentation, Presentation Media and Conclusion were all somewhat weaker than the others. From a student performance perspective, over achieving students tend to be equally strong in all dimensions, moderate students tend to excel in Support, Language, and Attire, while beginning to lag in other dimensions, and under achieving students tend to be exceedingly lagging in Presentation Media and Conclusion.

A comparison of spring 2016 results with past results is shown in Figure 3 below. Results exhibit consistency among all rubric dimensions. Slight decreasing trends exist in both Support and NV-Vocal dimensions. However, at this point, these trends are minimal and not statistically significant.

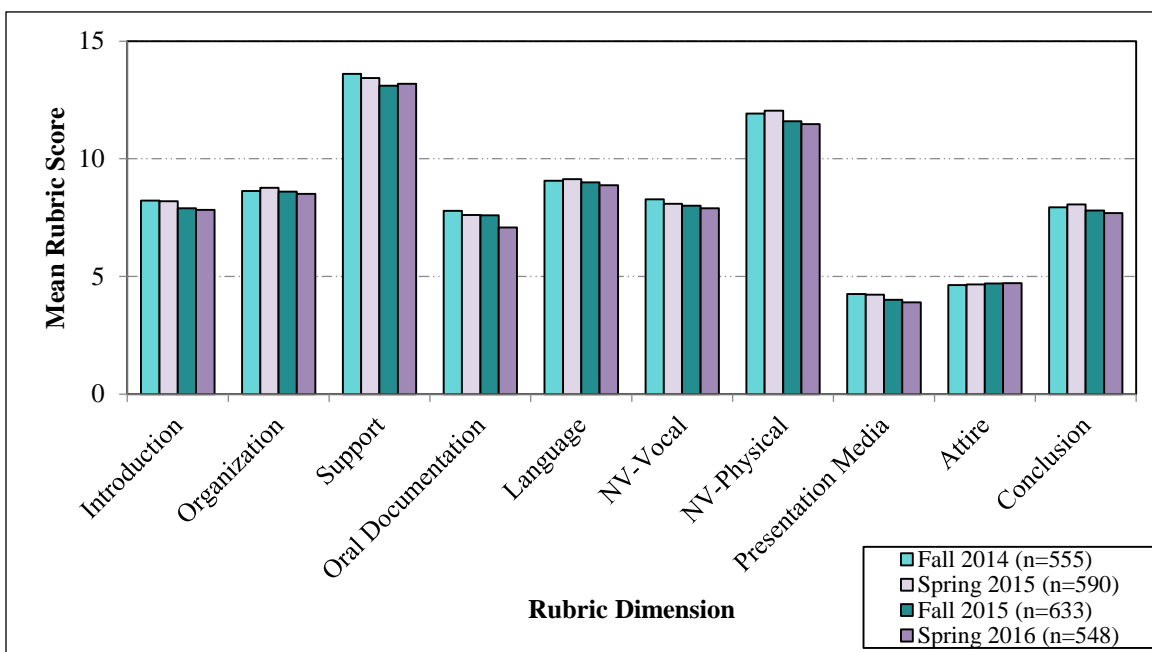


Figure 3. Comparison of mean scores for SPC 1017 through time beginning fall 2014 (teal) through spring 2015 (purple), fall 2015 (dark teal), and spring 2016 (dark purple).

2.2 SPC 2608

2.2.1 Learning Objectives

For the spring 2016 assessment, 222 artifacts were collected for SPC 2608 from 13 of 17 course sections. In some cases, rubric scores could either not be accessed or located. Faculty established measures of success for SLO1 was a rating of “Developing” or higher (≥ 2) in rubric dimension Introduction for 70% of the students. Spring 2016 artifacts successfully met the goal as results exhibit 99% of artifacts scored

level 2 or greater (Table 3). For SLO2, a rating of “Developing” or higher (≥ 2) in rubric criteria Conclusion for 70% of the students, spring 2016 artifacts successfully meet the goal as results exhibit 96% of artifacts scored level 2 or greater. For SLO3, a rating of “Developing” or higher (≥ 2) in rubric criteria Oral Documentation for 70% of the students, spring 2016 artifacts successfully meet the goal as results exhibit 90% of artifacts scored level 2 or greater.

Had a similar goal been set for the remaining rubric dimensions not focused in SLOs 1, 2, or 3, all would have been met (Table 3). Of these, the highest scoring dimension was Language at 99%, while the lowest scoring dimension was Presentation Media at 93%. All other rubric dimensions not included in the SLOs exhibit greater than 90% achieving the goal. For visual comparison of dimension scores, see Section 2.3.

Rubric Score	Introduction	Organization	Support	Oral Documentation	Language	NV-Vocal	NV-Physical	Presentation Media	Attire	Conclusion
Developing or higher	99%	98%	98%	90%	99%	99%	97%	93%	97%	96%
4	53.6%	52.3%	90.1%	51.4%	44.1%	29.7%	85.1%	57.2%	75.2%	50.5%
3	34.2%	36.0%	7.7%	25.2%	52.7%	53.2%	11.7%	27.9%	16.7%	33.3%
2	10.8%	9.9%	0.0%	13.1%	2.3%	15.8%	0.0%	7.7%	5.4%	12.6%
1	1.4%	1.8%	2.3%	5.9%	0.5%	1.4%	3.2%	2.7%	2.7%	2.7%
0	0.0%	0.0%	0.0%	4.5%	0.5%	0.0%	0.0%	4.5%	0.0%	0.9%

Table 3. Percentage of student achievement level by rubric dimension (includes percentage of students scoring in developmental level or higher as per SLOs) for SPC 2608. Rubric dimensions identified in SLOs in blue.

Had the goal for achievement been set at “Accomplished” or (≥ 3) for 70% of the students all dimensions again would have been met. The highest scoring dimension was Support at 98% while the lowest scoring dimension was NV-Vocal at 83%. All dimensions exhibit greater than 70% achieving the goal.

2.2.2 Descriptive Statistics & Longitudinal Data

Descriptive statistics for SPC 2608 artifacts can be found in Table 4. A histogram of artifact scores for all 222 artifacts is shown in Figure 4. To describe the behavior of the rubric dimensions based on overall achievement a color map, or binary raster image was created by calculating the mean scores for each dimension as a function of combined score (Figure 5). To create this image the rubric scores (4, 3, 2, 1, or 0) for each artifact was grouped based on combined raw rubric score (10 dimensions x maximum rubric level of 4 = 40 overall points). The color represents the mean rubric score achieved in each dimension based on the combined score as shown in the x-axis.

	Introduction	Organization	Support	Oral Documentation	Language	NV-Vocal	NV-Physical	Presentation Media	Attire	Conclusion
n	222	222	222	222	222	222	222	222	222	222
Max	10	10	15	10	10	10	15	5	5	10
Min	3	3	5	0	0	3	5	0	2	0
Mode	10	10	15	10	8	8	12	5	5	10
Mean	8.8	8.8	13.4	8.1	8.8	8.2	12.4	4.3	4.6	8.5
Standard deviation	1.52	1.54	2.28	2.63	1.29	1.46	2.31	1.19	0.71	1.87

Table 4. Descriptive statistics for SPC 2608 common course assessment.

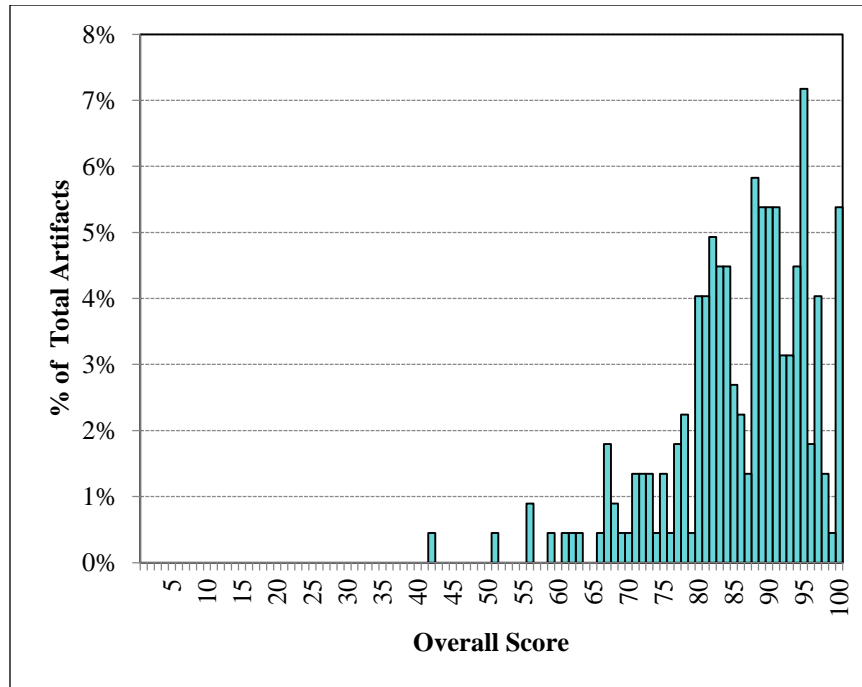


Figure 4. Overall score distribution for SPC 2608 artifacts (spring 2016 term).

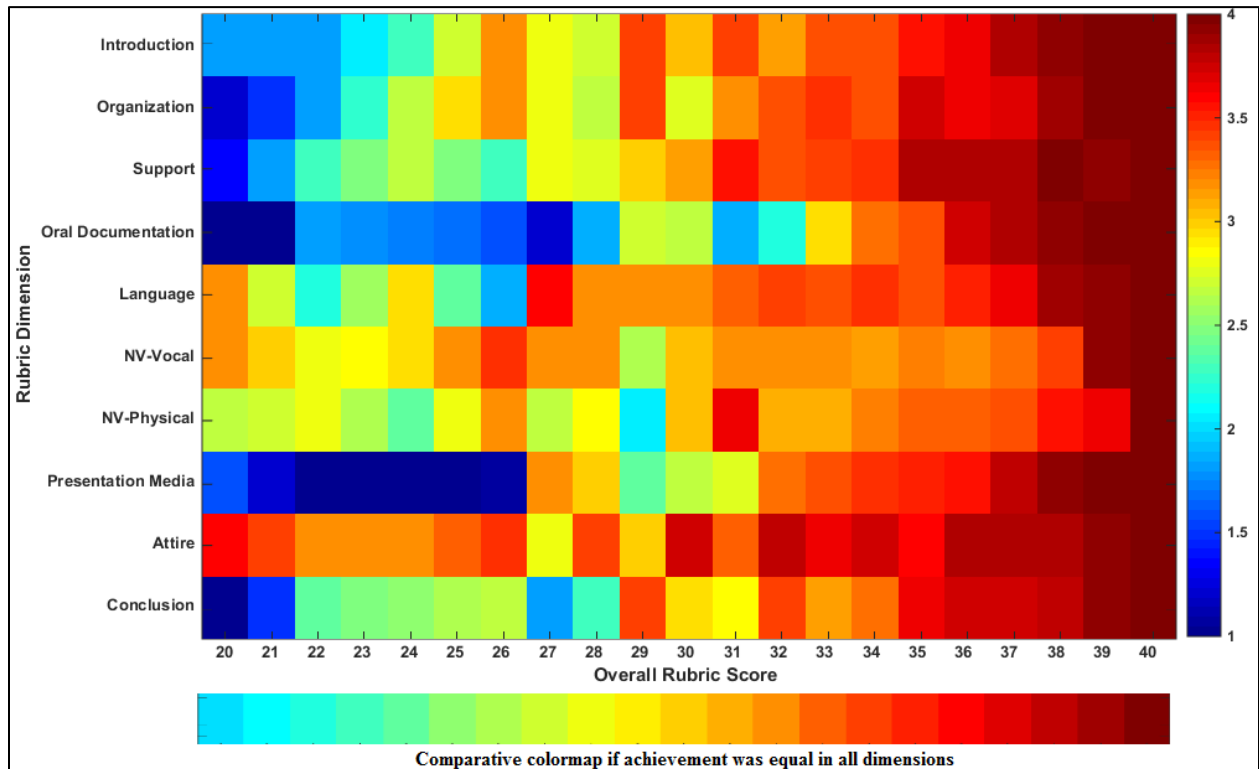


Figure 5. (Top) Colormap of mean scores for each rubric dimension (range: 0-4) based on overall rubric score (combined rubric score of all dimensions, max=40) for SPC 2608. (Bottom) Comparison rubric dimension if dimension score is the same as overall (i.e. artifact overall score is equally distributed across all sections). A rubric dimension with hotter colors (reds/yellows) means that dimension achievement exceeds the overall score and is an area of strength. An exam section with colder colors (blues/greens) means that section achievement is lower than the overall score and is therefore an area of weakness.

A review of the colormap in Figure 5 shows that at 34/40 and above (average rubric score of 3.4 or higher) all dimensions fair relatively equally (hot colors fairly evenly distributed), which is quite similar to colormap results of fall 2015. When overall rubric scores range below 29, Oral Documentation and Conclusion are somewhat weaker than the others, which is also in keeping with fall 2015 data. From a student performance perspective, strong students tend to be equally strong in all dimensions. Meanwhile, weaker students tend to be exceedingly lagging in Oral Documentation and Conclusion.

A comparison of spring 2016 results with past results is shown in Figure 6 below. Results exhibit consistency among all rubric dimensions. At this point, only the increase in Support in fall 2015 data can be considered substantial and is statistically significantly different.

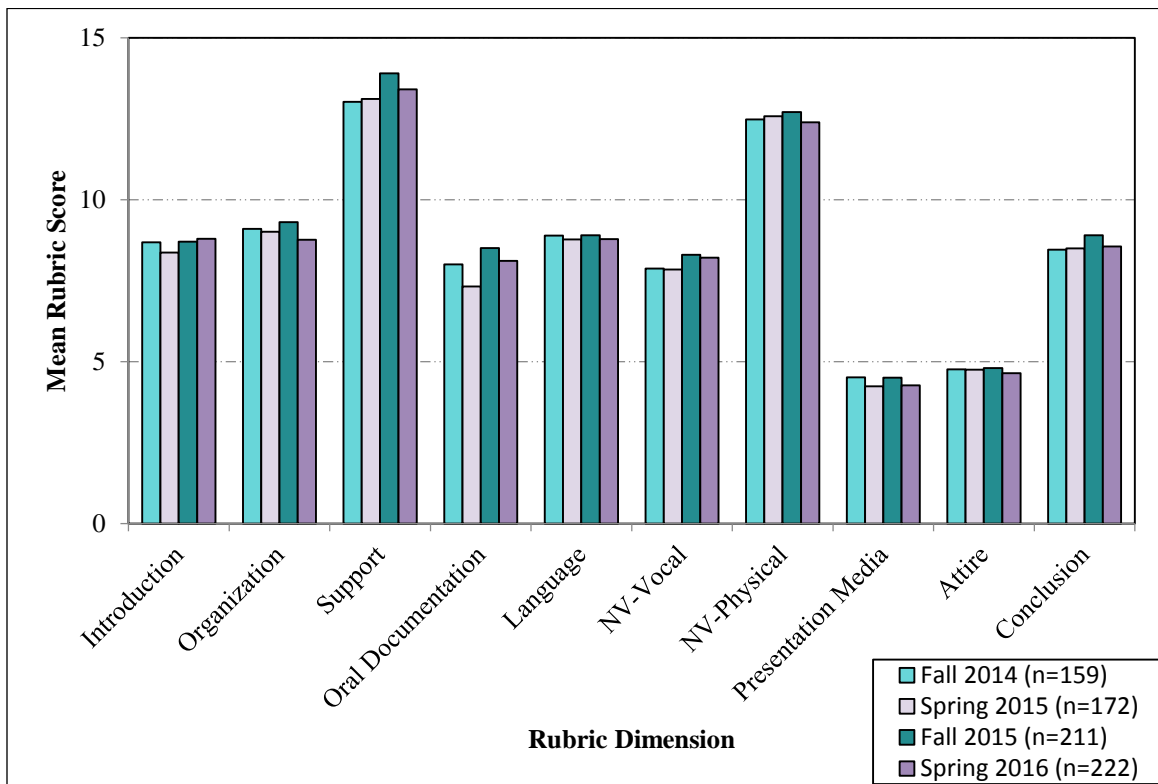


Figure 6. Comparison of mean scores for SPC 2608 through time beginning fall 2014 (teal) through spring 2015 (purple), fall 2015 (dark teal), and spring 2016 (dark purple).

2.3 COMPARISON OF SPC 1017 AND SPC 2608

Achievement levels for both SPC 1017 and SPC 2608 were plotted together for comparison (Figure 7). Achievement at the exemplary level is greater in SPC 2608 for all rubric dimensions except Language and Attire. By comparison, in fall 2015, 9/10 dimensions exhibited higher exemplary levels in SPC 2608. And spring 2015 exhibited 7/10 dimensions higher for SPC 2608 as compared to SPC 1017. The lowest achievement at the exemplary level for both SCP 1017 and 2608 is Vocal, at 25% and 30%, respectively.

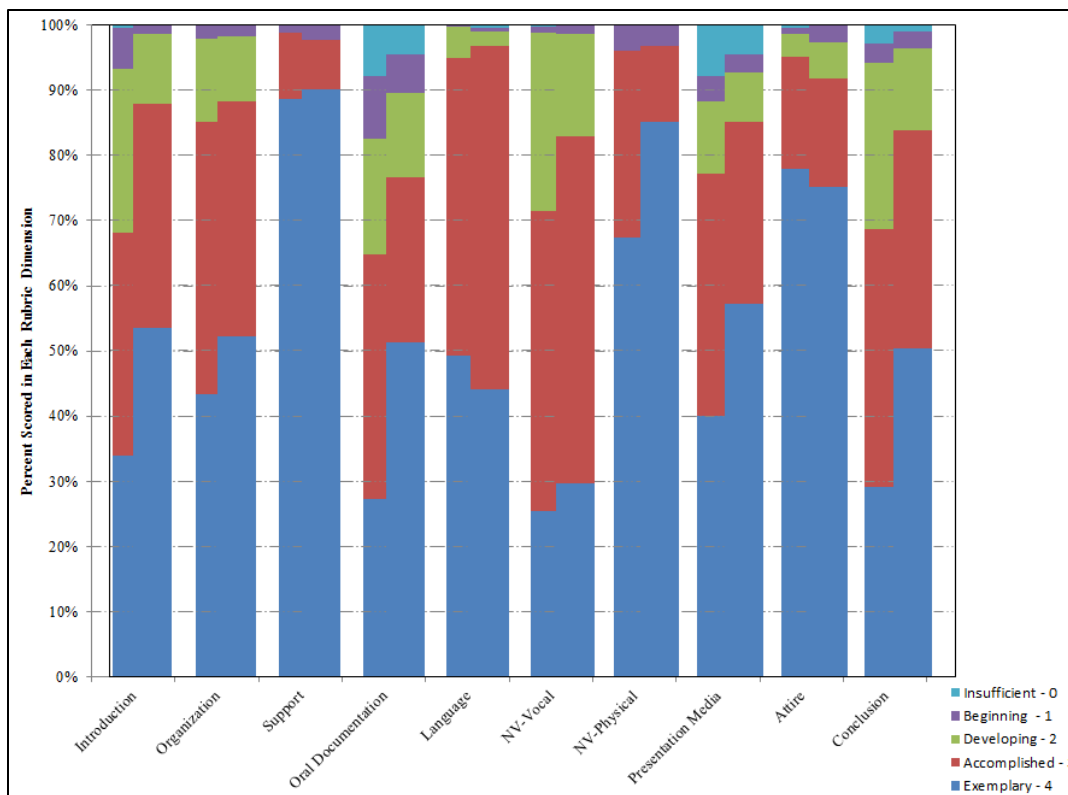


Figure 7. Comparison of Rubric achievement by dimension. Bar graphs are split for each dimension with left-side representing SPC 1017 and right-side representing SPC 2608.

A comparison of mean scores by rubric dimension is provided in Table 5. Mean scores are higher for SPC 2608 in all rubric dimensions except Language and Attire (see Figure 8 for graphical representation). By comparison, in fall 2015, SPC 2608 was higher in 9/10 dimensions. For spring 2015, SPC 2608 was higher in 5/10 dimensions. Differences in the means for all dimensions and combined score were tested for significance using a Welch's t-test according to standard methods (Davis, 1973; McDonald, 2009; Wilkinson, 1999). Introduction, Organization, Oral Documentation, NV-Vocal, NV-Physical, Presentation Media, Conclusion, and the overall score were found to be statistically significantly different. Therefore, in these cases, we must reject the null hypothesis that the differences in the means of the artifacts of the two courses for these dimensions are equal to 0, and we can conclude with a 95% confidence that the differences in scores are not solely due to chance. However, based on the work of Johnson (2013), there is a 17-25% chance that the marginally significant result (denoted in Table 5 as *) may be false positives (i.e. Type I errors).

Effect size was calculated using a method devised by Rosenthal and Rosnow (1991) for meta-analytical purposes in potential comparisons with other institutions (Lipsey and Wilson, 1993). The statistically significant results exhibit what Cohen (1988) would consider small-to-medium effect sizes ranging from 0.09 to 0.52 (Table 5). In other words, non-overlap from SPC 1017 artifacts to SPC 2608 artifacts range from approximately 7% to 34%.

	<i>Introduction</i>	<i>Organization</i>	<i>Support</i>	<i>Oral Documentation</i>	<i>Language</i>	<i>NV-Vocal</i>	<i>NV-Physical</i>	<i>Presentation Media</i>	<i>Attire</i>	<i>Conclusion</i>	<i>Combined Score</i>
<i>Rubric Max</i>	10	10	15	10	10	10	15	5	5	10	100
SPC 1017 Mean	7.8	8.5	13.2	7.1	8.9	7.9	11.5	3.9	4.7	7.7	81.1
SPC 2608 Mean	8.8	8.8	13.4	8.1	8.8	8.2	12.4	4.3	4.6	8.5	85.9
<i>Effect size</i>	0.52	0.14	0.09	0.34	-0.07	0.19	0.35	0.26	-0.09	0.39	0.42
<i>p-value</i>	1.91×10^{-12}	0.045*	0.213	2.43×10^{-6}	0.346	0.007	1.42×10^{-6}	2.82×10^{-4}	0.219	6.25×10^{-8}	8.66×10^{-9}

Table 5. Comparison of mean scores for SPC 1017 and SPC 2608. Bold denote statistically significant difference. Rubric dimensions identified in SLOs in blue. Positive effect sizes indicate a higher mean score for SPC 2608 artifacts. *Denote marginal significance as defined by Johnson (2013).

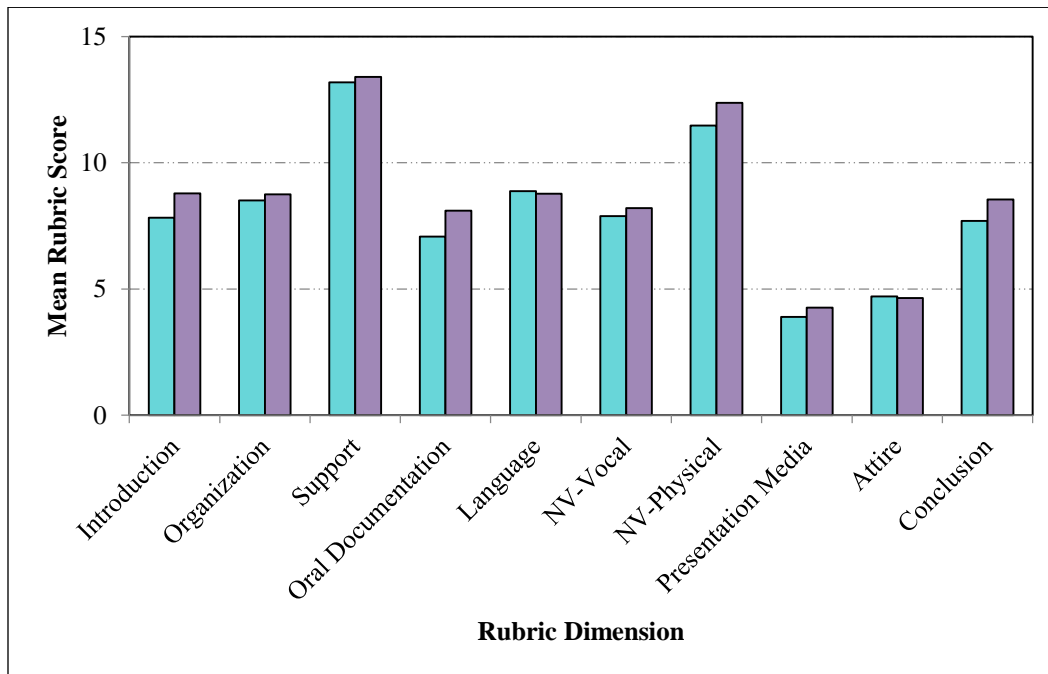


Figure 8. Comparison of mean scores for SPC 1017 (teal) and SPC 2608 (purple) for spring 2016 term.

3 EXPLORATORY ANALYSIS AND SIGNIFICANCE TESTING

Multiple comparisons of artifact scores across varying formats, campuses, and student types were made in order to add depth to the distribution of the artifacts by achievement level. Each course was divided into the appropriate subgroups to perform the analysis. Where possible, additional methods of analysis were conducted to provide a broader picture of these comparisons.

3.1 SPC 1017

3.1.1 Dual Enrollment to non-Dual Enrollment Comparison

No dual enrollment sections of SPC 1017 were offered during spring 2016 so no comparison study could be completed.

3.1.2 Online to Traditional Comparison

During the spring 2016 semester, 146 total online artifacts and 402 traditional artifacts were collected from SPC 1017 course sections. Before reviewing results it is important to note that in some cases scores did not include a Support Dimension score.

A comparison of mean scores by rubric dimension is provided in Table 6 and a graphical representation is provided in Figure 9. Mean scores are higher for online courses in all dimensions except Presentation Media. Differences in the means for all dimensions and overall score were tested for significance using a Welch’s t-test according to standard methods (Davis, 1973; McDonald, 2009; Wilkinson, 1999). Of these, Introduction, NV-Vocal, and Presentation media are statistically significantly different. Therefore we must reject the null hypothesis that the differences in the means of the artifacts of the two courses are equal to 0 for these dimensions, and we can conclude with a 95% confidence that the differences in scores are not solely due to chance. However, based on the work of Johnson (2013), there is a 17-25% chance that the marginally significant result (denoted in Table 6 as *) may be false positives (i.e. Type I errors).

Effect size was calculated using a method devised by Rosenthal and Rosnow (1991) for meta-analytical purposes in potential comparisons with other institutions (Lipsey and Wilson, 1993). The statistically significant results exhibit what Cohen (1988) would consider small-to-medium effect sizes ranging from 0.04 to 0.46 (Table 6). In other words, non-overlap from online artifacts to traditional artifacts range from approximately 3% to 30%.

	<i>Introduction</i>	<i>Organization</i>	<i>Support</i>	<i>Oral Documentation</i>	<i>Language</i>	<i>NV-Vocal</i>	<i>NV-Physical</i>	<i>Presentation Media</i>	<i>Attire</i>	<i>Conclusion</i>	<i>Combined Score</i>
<i>Rubric Max</i>	10	10	15	10	10	10	15	5	5	10	100
Online mean	8.2	8.7	13.4	7.4	9.0	8.4	11.6	3.2	4.7	8.0	82.7
Traditional mean	7.7	8.4	13.1	7.0	8.8	7.7	11.4	4.1	4.7	7.6	80.6
Effect Size	-0.21	-0.12	-0.11	-0.12	-0.13	-0.39	-0.07	0.46	0.04	-0.15	-0.15
p-value	0.027*	0.169	0.217	0.885	0.126	7.66x10⁻⁶	0.422	1.72x10⁻⁹	0.499	0.152	0.078

Table 6. Comparison of mean scores for Online and Traditional for SPC 1017. Bold denote statistically significant difference. Rubric dimensions identified in SLOs in blue. Positive effect sizes indicate a higher mean score for Traditional artifacts. *Denote marginal significance as defined by Johnson (2013).

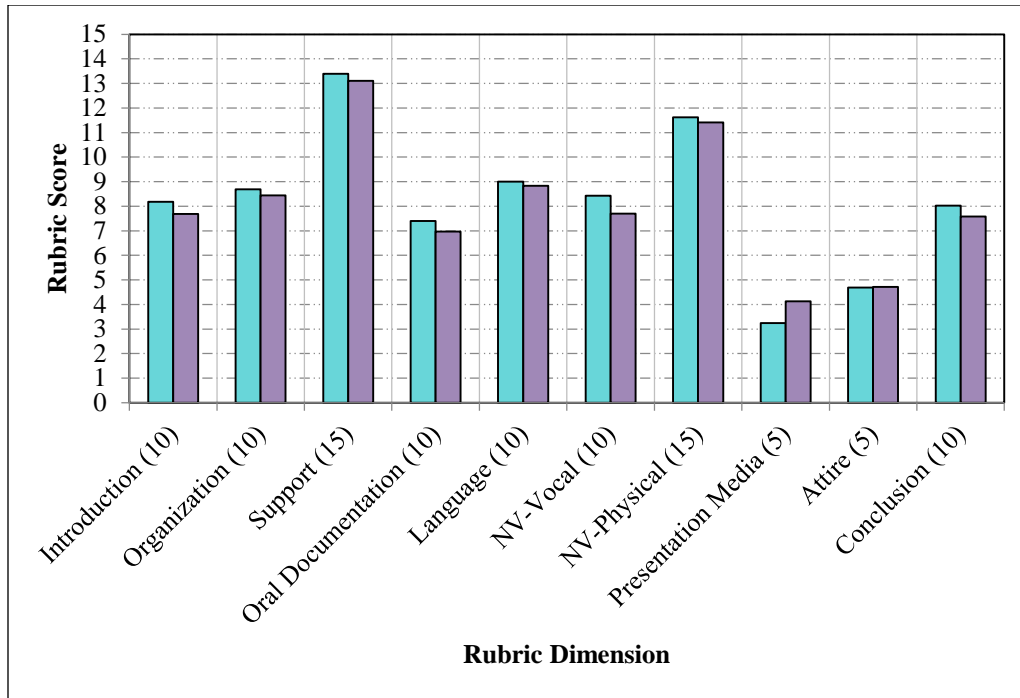


Figure 9. Comparison of mean scores for online (teal) and traditional (purple) scores for SPC 1017 (spring 2016 term).

3.1.3 Comparison by Campus/Site

Of the 548 artifacts collected from SPC 1017, 73 originated from the Charlotte campus, 115 from the Collier campus, 146 from FSW Online, and 214 from the Thomas Edison (Lee) campus. Scores by rubric dimension varied greatly across campuses. A comparison of mean scores by rubric dimension is provided in Table 7.

	<i>Introduction</i>	<i>Organization</i>	<i>Support</i>	<i>Oral Documentation</i>	<i>Language</i>	<i>NV-Vocal</i>	<i>NV-Physical</i>	<i>Presentation Media</i>	<i>Attire</i>	<i>Conclusion</i>	<i>Combined Score</i>
<i>Rubric Max</i>	10	10	15	10	10	10	15	5	5	10	100
Charlotte	8.8	8.9	13.9	7.8	9.0	8.4	12.2	4.9	4.8	8.7	87.4
Collier	7.2	8.6	14.7	7.3	9.6	7.2	10.4	4.0	4.8	7.5	81.2
FSW Online	8.2	8.7	13.4	7.4	9.0	8.4	11.6	3.2	4.7	8.0	82.7
Thomas Edison (Lee)	7.6	8.2	12.0	6.5	8.4	7.7	11.7	4.0	4.7	7.2	77.9

Table 7. Comparison of mean scores by site for SPC 1017. Bold denotes highest mean score in that dimension among all sites. Rubric dimensions identified in SLOs in blue.

Charlotte campus exhibits higher scores in 8/10 dimensions (up from 5/10 in fall data) as well as overall score while the remaining 2/10 highest scores occur with Collier. A plot comparing descriptive statistics of the combined (overall) scores by site is presented in Figure 10. There is extensive overlap between sites. However, Charlotte exhibits a distribution on the upper end of that overlap with progressively negatively shifted modes (central tendencies) for Collier, FSW Online, and eventually, Thomas Edison. For example, over 35% of Charlotte's scores range from 93 or higher. Meanwhile, in that same range, Thomas Edison exhibits only 8%.

A one-way analysis of variance was used to compare means of the combined rubric scores at each site. Results of the ANOVA exhibit a statistically significant difference between sites (see Table 8) likely caused by the differences between Charlotte, the highest performing site, and Thomas Edison, the lowest. Therefore, we can reject the null hypothesis that the mean combined rubric scores at each site are equal to each other and we can conclude with a 95% confidence that the differences in scores are not solely due to chance.

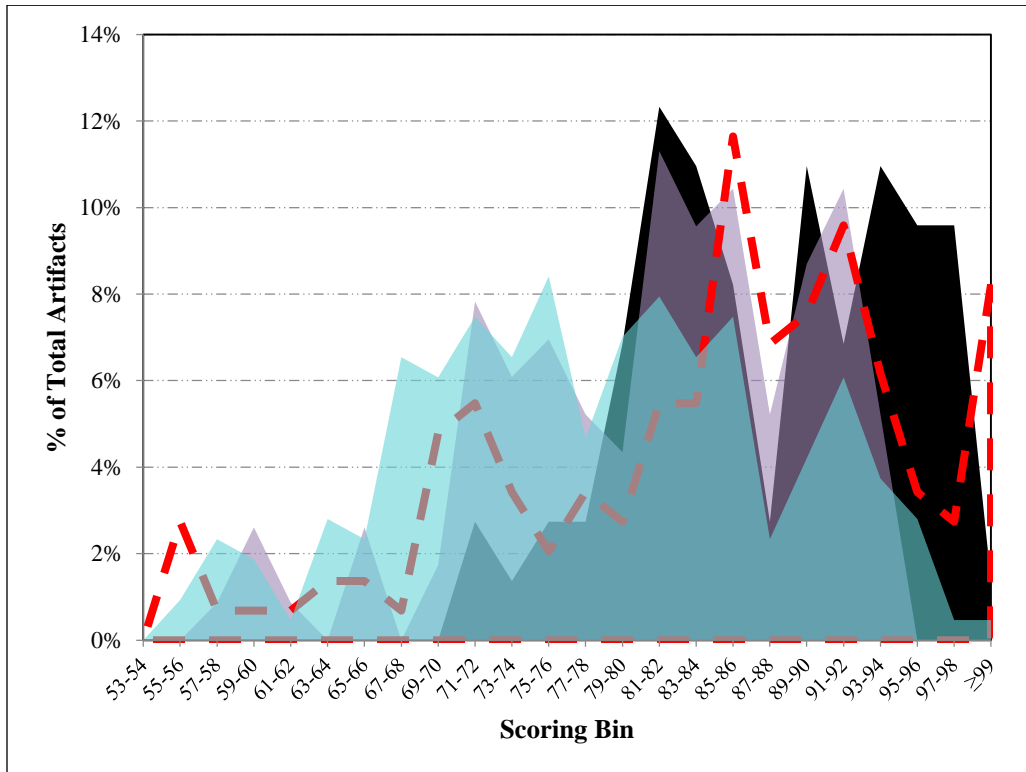


Figure 10. Comparison of artifact score distribution for Charlotte (black), Collier (purple), FSW Online (outlined in red), and Thomas Edison (teal).

Source of Variation	Sum of squared differences	df	Mean Squares	F _{obs}	p-value	F _{crit}
Between Sites	5414.8	3	1804.9	17.24	1.04x10 ⁻¹⁰	2.62
Within Sites	56,940.8	544	104.7			
Total	62,355.6	547				

Table 8. Results of one-way ANOVA of combined rubric scores at each site for SPC 1017.

3.2 SPC 2608

3.2.1 Dual Enrollment to non-Dual Enrollment Comparison

While there was one offsite (dual enrollment) course section offered during spring 2016, the section did not report data. As a result, no comparison study could be completed.

3.2.2 Online to Traditional Comparison

During the spring 2016 semester, 56 total online artifacts and 166 traditional artifacts were collected from SPC 2608 course sections. A comparison of mean scores by rubric dimension is provided in Table 9

and a graphical representation is provided in Figure 11. Mean scores are higher for online courses in all dimensions. Differences in the means for all dimensions and overall score were tested for significance using a Welch's t-test according to standard methods (Davis, 1973; McDonald, 2009; Wilkinson, 1999). Of these, all but the Oral Documentation and the Presentation Media dimension mean scores are statistically significantly different. Therefore we must reject the null hypothesis that the differences in the means of the artifacts of the two courses are equal to 0 for these dimensions, and we can conclude with a 95% confidence that the differences in scores are not solely due to chance. It should be noted that there is a 17-25% chance that the marginally significant results between the means of the NV-Physical and Conclusion dimensions may be a Type I error according to Johnson (2013).

Effect size was calculated using a method devised by Rosenthal and Rosnow (1991) for meta-analytical purposes in potential comparisons with other institutions (Lipsey and Wilson, 1993). The statistically significant results exhibit a wide range of effect sizes from 0.09 to 0.94 (Table 9). In other words, non-overlap from online artifacts to traditional artifacts range from approximately 7% to 53%.

	<i>Introduction</i>	<i>Organization</i>	<i>Support</i>	<i>Oral Documentation</i>	<i>Language</i>	<i>NV-Vocal</i>	<i>NV-Physical</i>	<i>Presentation Media</i>	<i>Attire</i>	<i>Conclusion</i>	<i>Combined Score</i>
<i>Rubric Max</i>	<i>10</i>	<i>10</i>	<i>15</i>	<i>10</i>	<i>10</i>	<i>10</i>	<i>15</i>	<i>5</i>	<i>5</i>	<i>10</i>	<i>100</i>
Online mean	9.4	9.3	14.6	8.8	9.4	8.6	12.9	4.4	4.8	9.2	91.4
Traditional mean	8.6	8.6	13.0	7.9	8.6	8.1	12.2	4.2	4.6	8.3	84.1
Effect Size	-0.54	-0.45	-0.94	-0.30	-0.64	-0.31	-0.29	-0.09	-0.38	-0.46	-0.77
p-value	1.01x10⁻⁴	0.001	3.40x10⁻¹¹	0.157	8.56x10⁻⁶	0.026*	0.037*	0.874	0.005	0.002	8.43x10⁻⁸

Table 9. Comparison of mean scores for Online and Traditional for SPC 2608. Bold denote statistically significant difference. Rubric dimensions identified in SLOs in blue. Positive effect sizes indicate a higher mean score for Traditional artifacts. *Denote marginal significance as defined by Johnson (2013).

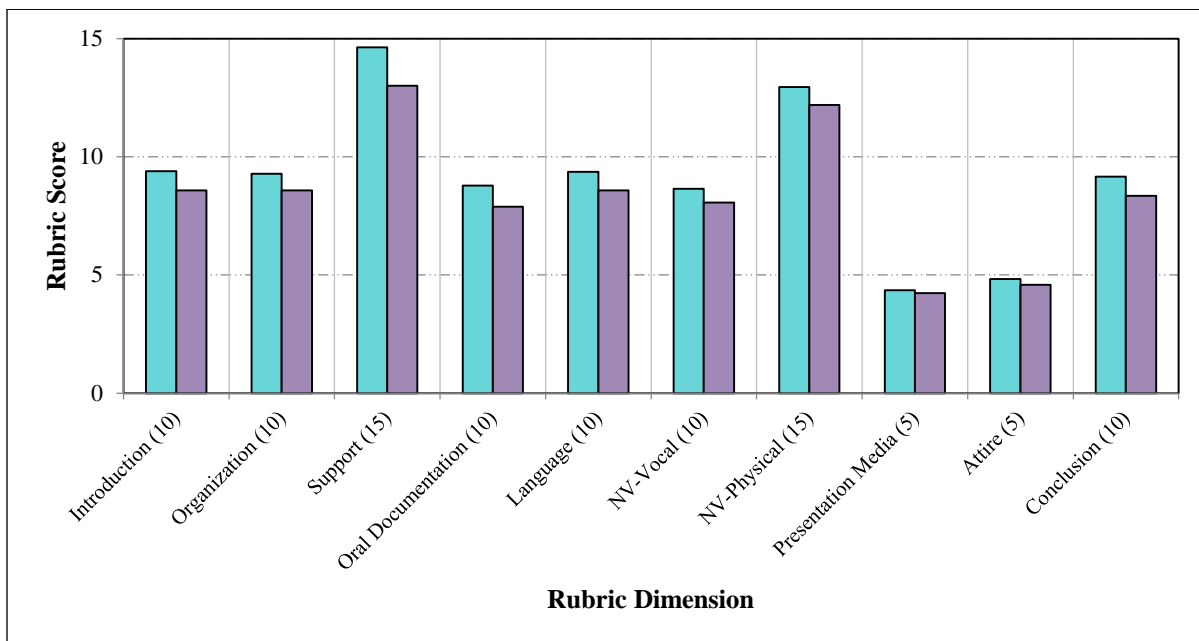


Figure 11. Comparison of mean scores for online (teal) and traditional (purple) scores for SPC 2608 (spring 2016 term).

3.2.3 Comparison by Campus/Site

Of the 222 artifacts collected from SPC 2608, 41 originated from the Charlotte campus, 56 from FSW Online, 12 from Hendry-Glades Center, and 113 from the Thomas Edison (Lee) campus. Scores by rubric dimension are consistently higher for FSW Online artifacts in all areas except Introduction, Oral Documentation, and Presentation Media. A comparison of mean scores by rubric dimension is provided in Table 10. A plot comparing descriptive statistics of the combined scores by site is presented in Figure 12. Charlotte and Thomas Edison sites exhibit very similar distributions. The FSW Online and Hendry Glades Center, however, exhibit opposite ends of the spectrum and are substantially different from both the other sites and each other. A one-way analysis of variance was used to compare means of the combined rubric scores at each site. Results of the ANOVA exhibit a statistically significant difference between sites largely as a result of the shifts exhibited by both FSW Online and Hendry Glades (see Table 11). Therefore, we can reject the null hypothesis that the mean combined rubric scores at each site are equal to each other and we can conclude with a 95% confidence that the differences in scores are not solely due to chance.

	<i>Introduction</i>	<i>Organization</i>	<i>Support</i>	<i>Oral Documentation</i>	<i>Language</i>	<i>NV-Vocal</i>	<i>NV-Physical</i>	<i>Presentation Media</i>	<i>Attire</i>	<i>Conclusion</i>	<i>Combined Score</i>
<i>Rubric Max</i>	10	10	15	10	10	10	15	5	5	10	100
Charlotte	9.5	8.6	13.9	8.1	9.2	8.1	12.2	4.8	4.7	8.7	87.8
FSW Online	9.4	9.3	14.6	8.8	9.4	8.6	12.9	4.4	4.8	9.2	91.4
Hendry-Glades	7.5	8.0	13.0	9.3	8.0	7.5	12.3	4.2	4.5	7.6	81.8
Thomas Edison (Lee)	8.4	8.6	12.7	7.6	8.4	8.1	12.2	4.0	4.5	8.3	82.9

Table 10. Comparison of mean scores by site for SPC 2608. Bold denotes highest mean score in that dimension among all sites. Rubric dimensions identified in SLOs in blue.

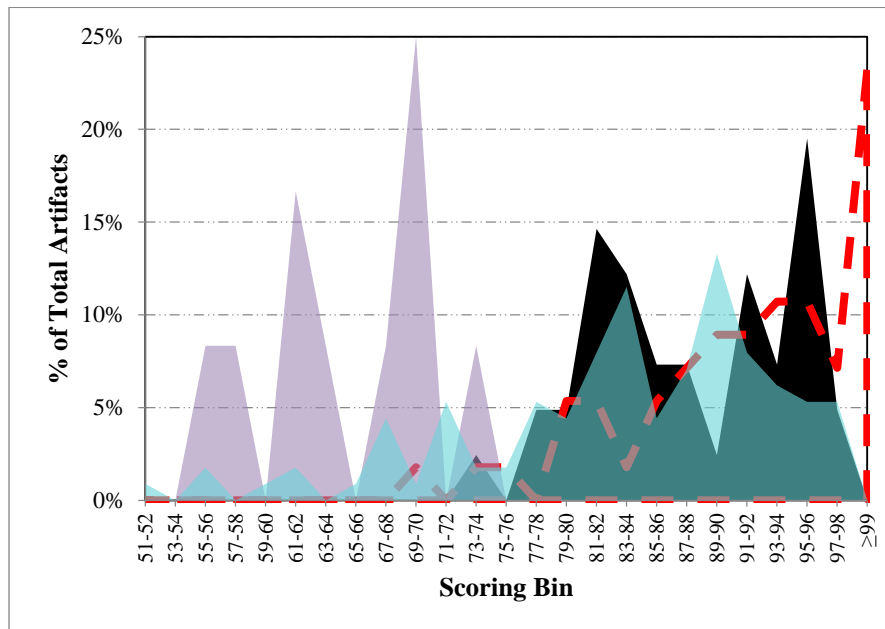


Figure 12. Comparison of artifact score distribution for Charlotte (black), FSW Online (outlined in red), Hendry Glades (purple), and Thomas Edison (teal).

Source of Variation	Sum of squared differences	df	Mean Squares	F _{obs}	p-value	F _{crit}
Between Sites	3003.8	3	1001.3	11.48	5.10x10 ⁻⁷	2.65
Within Sites	19,010.0	218	87.2			
Total	22,013.8	221				

Table 11. Results of one-way ANOVA of combined rubric scores at each site for SPC 2608.

4 CONCLUSIONS

FSW's Speech Department employed a common rubric used by all faculty as a means to evaluate an agreed upon series of student level outcomes for SPC 1017 and SPC 2608. Faculty goals in assessment included tracking rubric implementation, Student Learning Objectives (SLOs) to include Introduction, Oral Documentation, and Conclusion, and comparisons between dual enrollment and non-dual enrollment students, online and traditional students, and by site.

A drilldown of SPC 1017 results are as follows:

1. SLO 1 – Achievement of 70% at the developing (2) level or higher for SLO1: Introduction was met with achievement at 93%.
2. SLO 2 – Achievement of 70% at the developing (2) level or higher for SLO2: Conclusion was met with achievement at 94%.
3. SLO 3 – Achievement of 70% at the developing (2) level or higher for SLO3: Documentation was met with achievement at 82%.
4. Had a similar goal been set for the remaining rubric dimensions not focused in SLOs 1, 2, or 3, all would have been met. Had the goal for achievement been set at “Accomplished” or (≥ 3) for 70% of the students all dimensions again would have been met with the exception of Introduction (68%), Oral Documentation (64%), and Conclusion (68%).
5. Distribution of artifact scores was centered on 86/100 and is moderately negatively skewed, meaning scores are shifted towards the higher range.
6. In a study comparing rubric achievement based on overall score, over achieving students tend to be equally strong in all dimensions, moderate students tend to excel in Support, Language, and Attire, while beginning to lag in other dimensions, and under achieving students tend to be exceedingly lagging in Presentation Media and Conclusion.
7. In a longitudinal study, results exhibit consistency among all rubric dimensions. Slight decreasing trends exist in both Support and NV-Vocal dimensions. However, at this point, these trends are minimal and not statistically significant.
8. No comparison between dual enrollment sections and traditional sections could be made because no dual enrollment sections were offered during spring 2016.
9. In a comparison of online to traditional artifacts mean scores are higher for online courses in all dimensions except Presentation Media. Of these, Introduction, NV-Vocal, and Presentation media are statistically significantly different.
10. In a cross-campus comparison, there is extensive overlap between sites. However, Charlotte exhibits a distribution on the upper end of that overlap with progressively negatively shifted modes (central tendencies) for Collier, FSW Online, and eventually, Thomas Edison. Results of the ANOVA exhibit a statistically significant difference between sites.

A drilldown of SPC 2608 results are as follows:

1. SLO 1 – Achievement of 70% at the developing (2) level or higher for SLO1: Introduction was met with achievement at 99%.
2. SLO 2 – Achievement of 70% at the developing (2) level or higher for SLO2: Conclusion was met with achievement at 96%.
3. SLO 3 – Achievement of 70% at the developing (2) level or higher for SLO3: Documentation was met with achievement at 90%.
4. Had a similar goal been set for the remaining rubric dimensions not focused in SLOs 1, 2, or 3, all would have been met. Had the goal for achievement been set at “Accomplished” or (≥ 3) for 70% of the students all dimensions again would have been met.
5. Distribution of artifact scores is centered on 88/100 and has a large negative skew, meaning scores are substantially shifted towards the higher range.
6. In a study comparing rubric achievement based on overall score, strong students tend to be equally strong in all dimensions. Meanwhile, weaker students tend to be exceedingly lagging in Oral Documentation and Conclusion.
7. In a longitudinal study, results exhibit consistency among all rubric dimensions. At this point, only the increase in Support in fall 2015 data can be considered substantial and is statistically significantly different.
8. No comparison between dual enrollment sections and traditional sections could be made because no dual enrollment sections were offered during spring 2016.
9. In a comparison of online to traditional artifacts mean scores are higher for online courses in all dimensions. Of these, all but the Oral Documentation and the Presentation Media dimension mean scores are statistically significantly different.
10. In a cross-campus comparison, scores by rubric dimension are consistently higher for FSW Online artifacts in all areas except Introduction, Oral Documentation, and Presentation Media. Results of the ANOVA exhibit a statistically significant difference between sites largely as a result of the shifts exhibited by both FSW Online and Hendry Glades.

A drilldown of a comparison study between SPC 1017 and SPC 2608 results are as follows:

1. Achievement at the exemplary level is greater in SPC 2608 for all rubric dimensions except Language and Attire. By comparison, fall 2015 exhibited 9 of 10 dimensions higher for SPC 2608, and in spring 2015, 7 of 10 dimensions higher for SPC 2608 as compared to SPC 1017.
2. Mean scores are higher for SPC 2608 in all rubric dimensions except Language and Attire. Introduction, Organization, Oral Documentation, NV-Vocal, NV-Physical, Presentation Media, Conclusion, and the overall score were found to be statistically significantly different.

5 REFERENCES

- Cohen, J. 1988. Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates, Hillsdale, NJ.
- Cole, R., Haimson, J., Perez-Johnson, I., and May, H. 2011. Variability in Pretest-Posttest Correlation Coefficients by Student Achievement Level. NCEE Reference Report 2011-4033. Washington, DC: National Center for Education Evaluation and Regional Assistance, U.S. Department of Education.

- Davis, J.C. 1973. *Statistics and Data Analysis in Geology*. John Wiley & Sons, New York, New York, 564 pp.
- Elder, L, and Paul, R. 2007. Consequential Validity: Using Assessment to Drive Instruction. In: *Foundation For Critical Thinking*. Retrieved from <http://www.criticalthinking.org/pages/consequential-validity-using-assessment-to-drive-instruction/790>.
- Johnson, V. 2013. Revised Standards for Statistical Evidence. *Proceedings of the National Academy of Science*, 110(48), 19313-19317.
- Lipsey, M.W. and Wilson, D.B. 1993. The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist*, 48, 1181-1209.
- McDonald, J.H. 2009. *Handbook of Biological Statistics (2nd ed.)*. Sparky House Publishing, Baltimore, Maryland.
- Rosenthal, R. and Rosnow, R.L. 1991. *Essentials of behavioral research: Methods and data analysis (2nd ed.)*. McGraw Hill, New York, NY.
- Wilkinson, L. 1999. APA Task Force on Statistical Inference. *Statistical Methods in Psychology Journals: Guidelines and Explanations*. *American Psychologist* 54 (8), 594–604.