

English Assessment Report – Spring 2016

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

1 INTRODUCTION

Fall 2014 marked the beginning of a new assessment plan for the English Department of Florida SouthWestern State College (FSW) in three courses: ENC 0022 *Writing for College Success*, ENC 1101 *Composition I*, and ENC 1102 *Composition II*. For spring 2016, assessment will include ENC 0022 while both ENC 1101 and ENC 1102 undergo departmental discussions based on the results of fall 2015 assessment before data collection resumes during the fall term. The planned assessment practice for ENC 0022 continues in spring 2016 in which instructors use a common rubric with seven identified rubric dimensions using data collected from all course sections for ENC 0022 are assessed. Baselines set in place following fall 2014 analysis and discussion will serve as a correlative measure for supporting assessment driven instruction going forward (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 ENC 0022

2.1 LEARNING OBJECTIVES & DESCRIPTIVE STATISTICS

Using common rubric criterion as an assessment method, the FSW English faculty defined multiple areas of interest for evaluation based on core outcomes for the course. Those outcomes include:

- Plan and write paragraphs and essays reflecting styles and tones appropriate for their audience and use adequate support, coherence, and unity that demonstrate understanding of content for expository and persuasive purposes.
- Establish a substantive claim, link claims to relevant evidence, and acknowledge competing arguments, gather information needed, and accurately incorporate source material into their own writing to avoid plagiarism.
- Identify and correctly use proper conventions for sentence grammar and avoid illogical shifts in pronouns and verbs in their own writing and on tests.
- Identify and use proper conventions for spelling, capitalization, and punctuation in their own writing and on tests.
- Identify and correctly use the conventions of a variety of sentence structures and will be able to avoid sentence fragments, comma splices, and fused sentences in their own writing and on tests.
- Identify and write effective topic sentences and thesis statements that address task and audience and use logical structure, support, and transitional devices for expository and persuasive purposes.

2.1.1 Learning Objectives

ENC 0022 is scored using a rubric with seven dimensions: Introductory Paragraph, Support Paragraphs, Organization, Concluding Paragraph, Grammar, Mechanics, and Research. Each dimension is scored on a scale of 1 to 4 (1-Unacceptable, 2-Needs work, 3-Average, 4-Above average), with 0s if the baseline of 'Unacceptable' is not met. The English department has identified a target statistic for measurement purposes (SLO1) of measuring the percentage of artifacts scoring a 2 or greater.

For the spring 2016 assessment, 99 artifacts were collected for ENC 0022 from 10 of 10 course sections. The lowest scoring rubric dimension for percentage of artifacts scoring a 2 or greater is Research at 87%. This is consistent with fall 2015 results in which the Research dimension was also at 87%. All other dimensions exhibit percentage of 94% or higher (Table 1). For a visual comparison of scores by dimension, see Figure 1.

Rubric Score	Introductory Paragraph	Support Paragraphs	Organization	Concluding Paragraph	Grammar	Mechanics	Research
<i>Developing or higher</i>	94%	99%	100%	98%	95%	97%	87%
4	47%	43%	44%	35%	26%	21%	27%
3	36%	40%	39%	41%	48%	52%	37%
2	10%	16%	16%	22%	20%	23%	22%
1	6%	1%	0%	2%	5%	3%	12%
0	0%	0%	0%	0%	0%	0%	0%

Table 1. Percentage of student achievement level by rubric dimension (includes percentage of students scoring in developmental level or higher as per SLO) for ENC 0022.

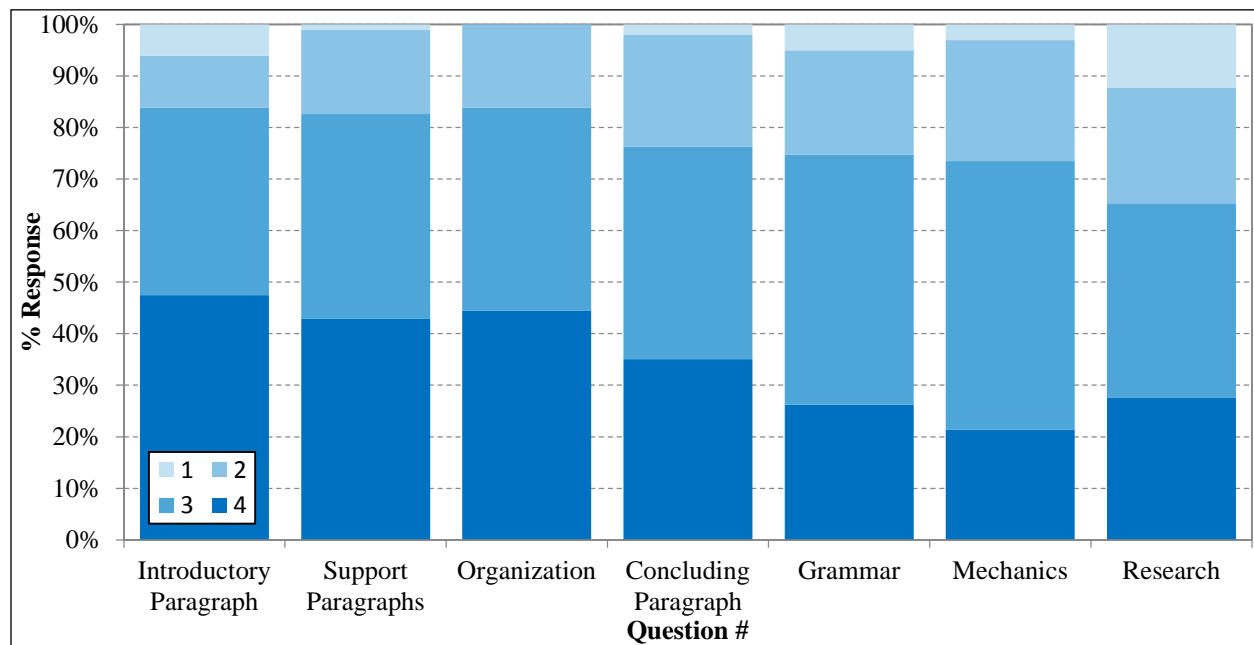


Figure 1. ENC 0022 distribution of rubric scores by dimension.

2.1.2 Descriptive Statistics & Longitudinal Studies

Descriptive statistics for ENC 0022 artifacts can be found in Table 2. A histogram of artifact scores for all 99 artifacts is shown in Figure 2. Distribution of artifact scores is bimodal centered on 21/28 and 25/28, and is moderately negatively skewed, meaning scores are shifted towards the higher range.

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 3). To create this image the rubric scores (4, 3, 2, 1, or 0) for each artifact was grouped based on combined raw rubric score (7 dimensions x maximum rubric level of 4 = 28 overall points). The color represents the mean rubric score achieved in each dimension based on the combined score as shown in the x-axis.

	Introductory Paragraph	Support Paragraphs	Organization	Concluding Paragraph	Grammar	Mechanics	Research	TOTAL
n	99	98	99	97	99	98	99	98
Max	4	4	4	4	4	4	5	28
Min	1	1	2	1	1	1	1	11
Median	3	3	3	3	3	3	3	22
Mode	4	4	4	3	3	3	3	25
Mean	3.3	3.2	3.3	3.1	3.0	2.9	2.8	21.5
Standard deviation	0.87	0.76	0.73	0.80	0.82	0.76	1.00	4.63
Skewness	-1.08	-0.59	-0.50	-0.42	-0.49	-0.30	-0.33	-0.62
Kurtosis	0.53	-0.55	-0.97	-0.68	-0.18	-0.21	-0.74	-0.43

Table 2. Descriptive statistics for ENC 0022 common course assessment.

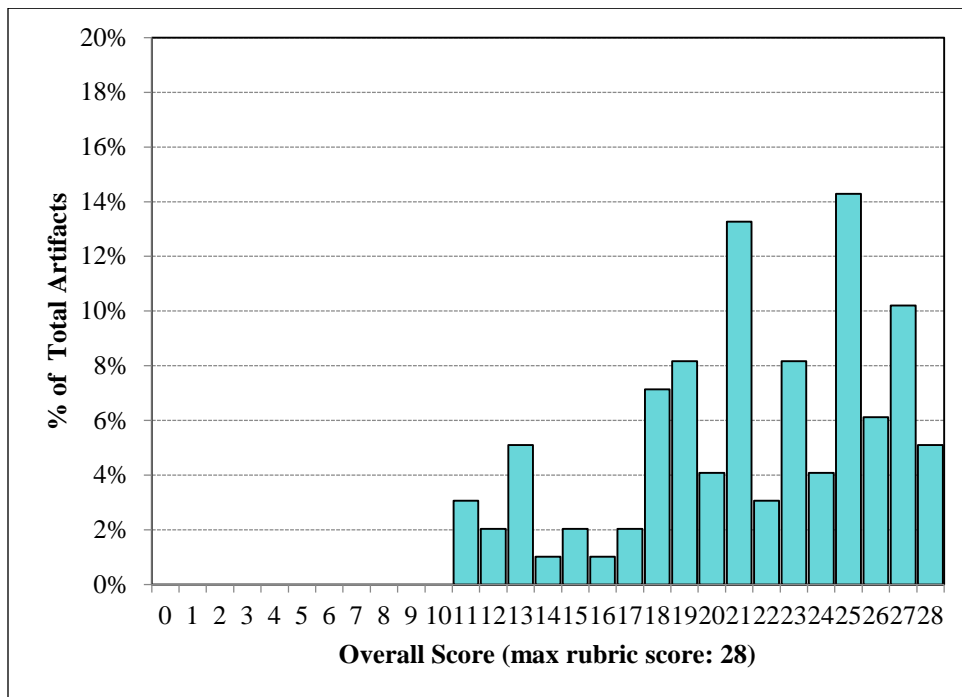


Figure 2. Overall score distribution for ENC 0022 artifacts (spring 2016 term).

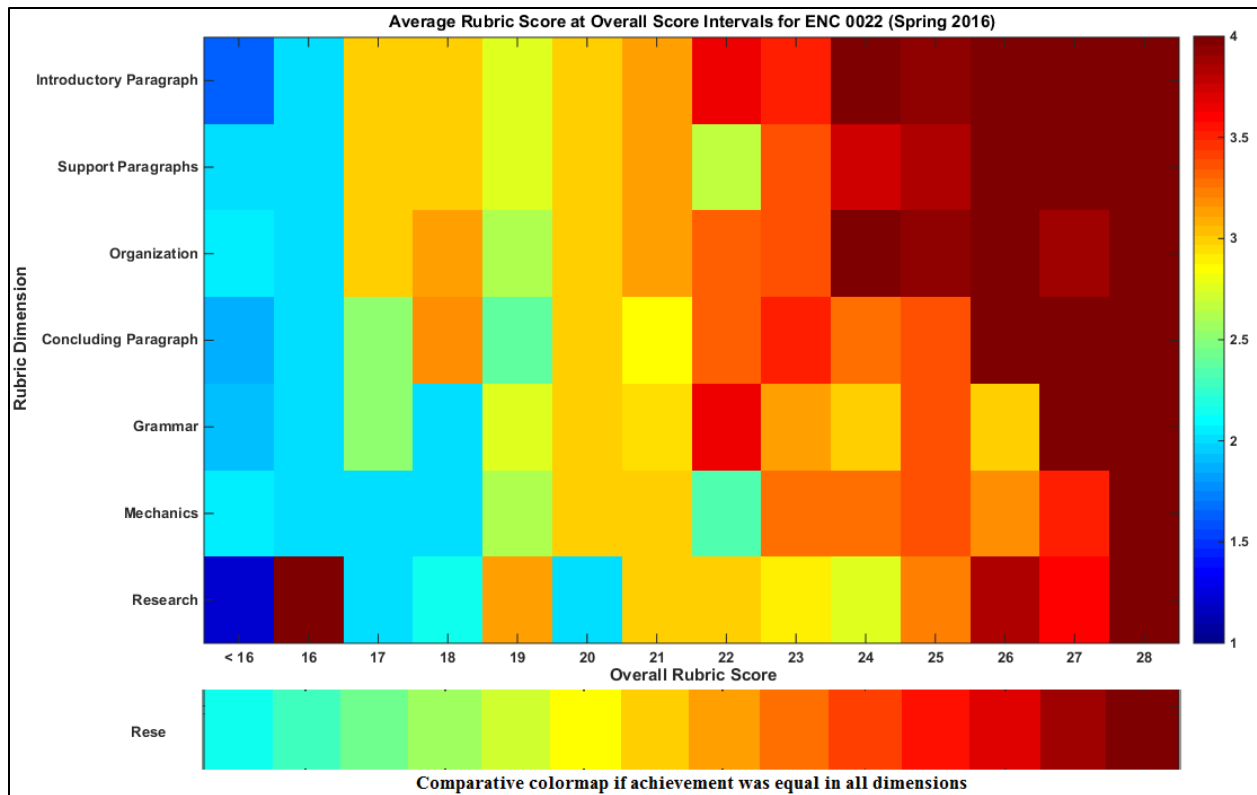


Figure 3. (Top) Colormap of mean scores for each rubric dimension (range: 0-4) based on overall rubric score (combined rubric score of all dimensions, max=28) for ENC 0022. (Bottom) Comparison rubric dimension if dimension score is the same as overall (i.e. artifact overall score is equally distributed 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 between 20/28 and 22/28 (approximately 75% overall score) all dimensions fair relatively equally (hot colors fairly evenly distributed). When overall rubric scores range above this, the Grammar and Mechanics dimensions lag somewhat behind all other dimensions. For example, at an overall score of 26/28, the Grammar and Mechanics dimensions exhibit average scores of 3.0/4 and 3.2/4, respectively, while other dimensions range from 3.8/4 to 4/4. From a student performance perspective, the average students tend to be equally strong in all dimensions while over achieving students tend to again lag in Grammar and Mechanics. In short, there is an upper limit to which even the best students do not attain for Grammar and Mechanics.

If we review spring 2016 colormap with reference to spring 2015, consistent patterns emerge (Figure 4; for comparison purposes fonts are small to allow for side-by-side imaging, please refer to Figure 3 for y-axis rubric dimension labels). Each term exhibits strong (near 4/4) rubric dimension scores for Introductory Paragraph, Support Paragraphs, Organization, and Concluding Paragraph being achieved when overall scores are 22/28 and higher. Each term also shows equally strong dimensions in all areas when overall scores are 21/28 and lower. And finally, each term shows Grammar and Mechanics lagging behind other dimensions when overall scores are 26/28 and higher.

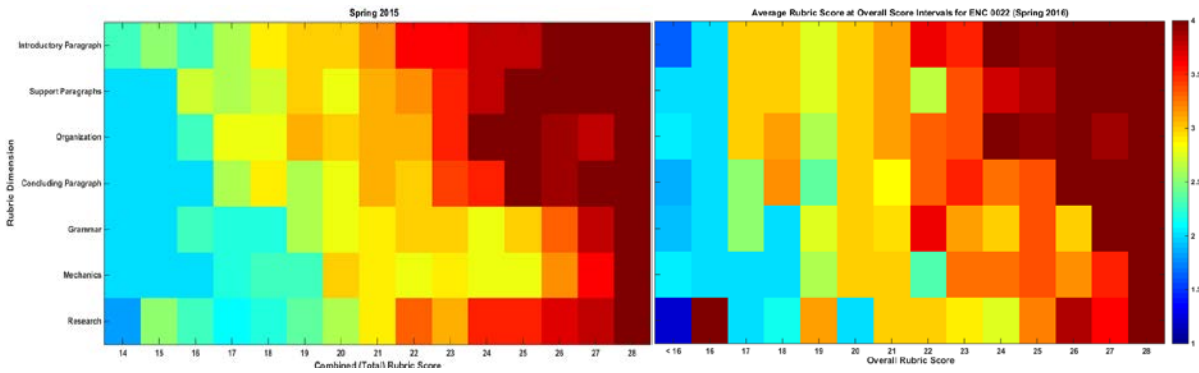


Figure 4. Side-by-side comparison of spring 2015 colormap (left) with spring 2016 colormap (right).

A comparison of spring 2016 mean scores with past results is shown in Figure 5 below. Note that comparison from fall-to-spring is less useful as assessment reports across multiple course level and program level assessments at FSW typically exhibit substantial differences from fall to spring term and are most effectively interpreted when comparing fall-to-fall terms and spring-to-spring terms (see <http://www.fsw.edu/facultystaff/assessment/history> for examples).

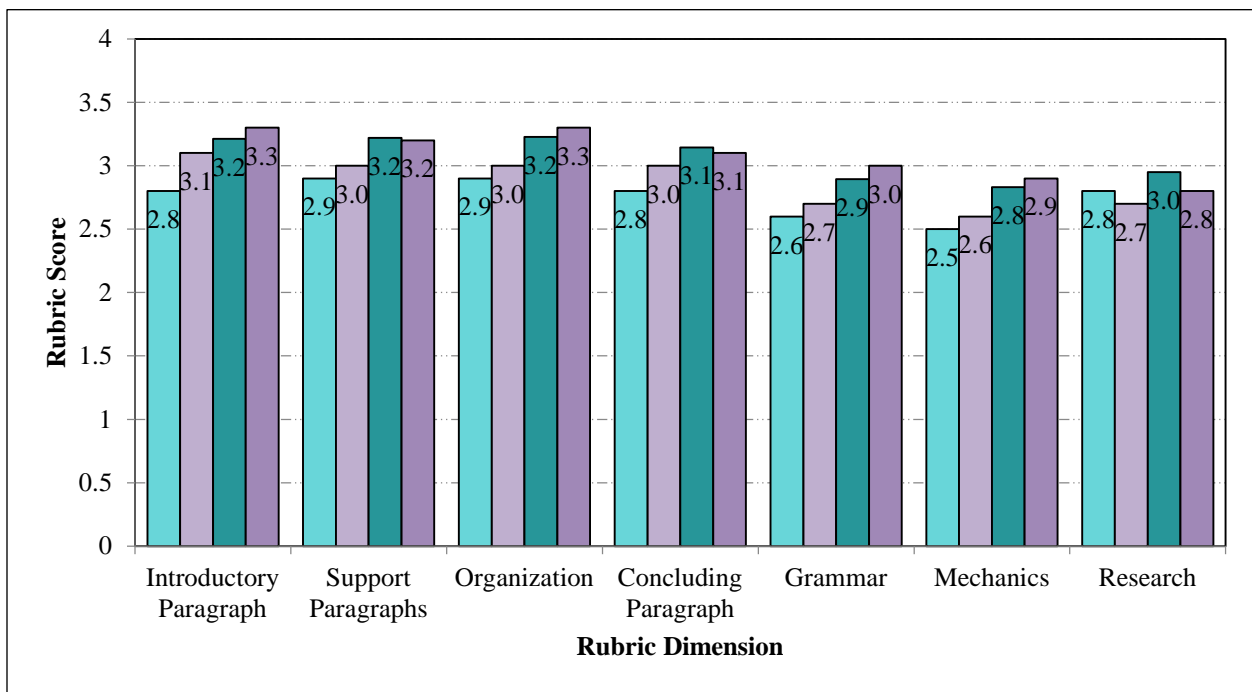


Figure 5. Comparison of mean scores for ENC 0022 through time beginning fall 2014 (light teal), spring 2015 (light purple), fall 2015 (dark teal), and spring 2016 (dark purple).

2.2 COMPARISONS BY SITE, FORMAT, AND STUDENT TYPE

2.2.1 Dual Enrollment to non-Dual Enrollment Comparison

ENC 0022 is not offered as a dual enrollment (offsite) course nor is it offered to dual enrollment students onsite and so no comparison study between dual enrollment artifacts and traditional artifacts can be made.

2.2.2 Online to Traditional Comparison

ENC 0022 is not offered as an online course and so no comparison study between online artifacts and traditional artifacts can be made.

2.2.3 Comparison by Site/Campus

Of the 99 artifacts collected from ENC 0022, 6 originated from the Charlotte campus, 12 from the Collier campus, 2 from the Hendry-Glade Center, and 79 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 3.

	Introductory Paragraph	Support Paragraphs	Organization	Concluding Paragraph	Grammar	Mechanics	Research
Charlotte	4.0	3.8	3.7	3.3	3.2	2.8	3.8
Collier	2.7	2.8	2.9	2.6	2.7	2.7	3.1
Hendry-Glades	3.5	3.5	4.0	2.5	3.0	3.0	3.0
Thomas Edison (Lee)	3.3	3.3	3.3	3.2	3.0	3.0	2.7

Table 3. Comparison of mean scores by site for ENC 0022. Bold denotes highest mean score in that dimension among all sites.

Charlotte campus is consistently higher compared to other sites exhibiting the highest mean score in 5 of 7 dimensions. However, the sample size is limited (n=6). A plot comparing descriptive statistics of the combined (overall) scores by site is presented in Figure 6. There is extensive overlap between sites although limited data at both Charlotte and Hendry-Glades limit interpretation.

A one-way analysis of variance was used to compare means of the combined rubric scores at each site. Results of the ANOVA exhibit no statistically significant difference between sites (see Table 4). Therefore, we cannot reject the null hypothesis that the mean rubric scores at each site are equal to each other and we cannot conclude with a 95% confidence that the differences in scores are not solely due to chance. However, such low sample sizes make any analysis of variance results suspect (Brown and Forsythe, 1974).

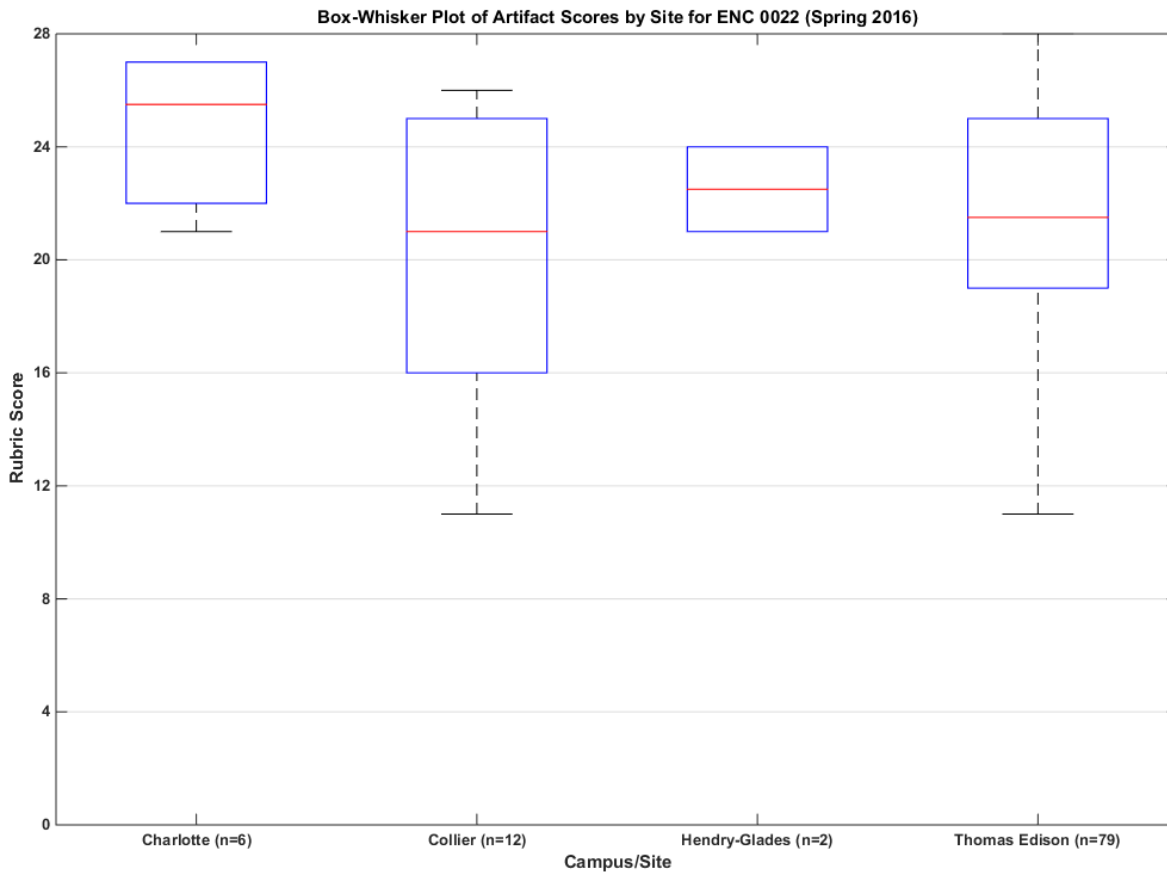


Figure 6. Box-Whisker plot of scores distributed by site for ENC 0022. Red line depicts median score. Upper and lower box boundaries indicate 75% quartile and 25% quartile (box represents central 50% of the scores). Vertical lines represent remaining scores outside central 50% that are not outliers. Red '+'s denote outliers.

Source of Variation	Sum of squared differences	df	Mean Squares	F _{obs}	p-value	F _{crit}
Between Sites	87.8	3	29.3	1.39	0.250	2.70
Within Sites	2018.4	96	21.0			
Total	2106.2	99				

Table 4. Results of one-way ANOVA of combined rubric scores at each site for ENC 0022.

2.2.4 Mini-term to Full-term Comparison

ENC 0022 was not offered as a mini-term course and so no comparison study between mini-term artifacts and full-term artifacts can be made.

3 ENC 1101

Course assessment for ENC 1101 follows a procedure of data collection in fall term only followed by departmental discussions in spring.

4 ENC 1102

Course assessment for ENC 1102 follows a procedure of data collection in fall term only followed by departmental discussions in spring.

5 CONCLUSIONS

FSW's English Department assessment plan includes three courses: ENC 0022 *Writing for College Success*, ENC 1101 *Composition I*, and ENC 1102 *Composition II*. For spring 2016, assessment will include ENC 0022 while both ENC 1101 and ENC 1102 undergo departmental discussions based on the results of fall 2015 assessment before data collection resumes during the fall term. The department has historically used a benchmark of percentage of students scoring 2 or higher in rubric dimensions as a means to measure achievement in the courses.

A drilldown of ENC 0022 results are as follows:

1. All seven rubric dimensions had > 87% achievement at level 2 or higher. The lowest dimension was Research while all other dimensions exceeded 94%.
2. Distribution of artifact scores is bimodal centered on 21/28 and 25/28, and is moderately negatively skewed, meaning scores are shifted towards the higher range.
3. In a study comparing rubric achievement based on overall score, average students tend to be equally strong in all dimensions while over achieving students tend to again lag in Grammar and Mechanics.
4. In a longitudinal study, results exhibit improvement across all areas.
5. No comparison of dual enrollment to traditional artifacts was completed because no dual enrollment sections of the course were offered.
6. No comparison of online to traditional artifacts was completed because no online sections of the course were offered.
7. In a cross-campus comparison, scores varied greatly across rubric dimensions. Charlotte campus is consistently higher compared to other sites exhibiting the highest mean score in 5 of 7 dimensions. There is extensive overlap between sites although limited data at both Charlotte and Hendry-Glades limit interpretation.
8. No comparison of mini-term artifacts and full-term artifacts was completed because no mini-term sections of the course were offered.

No drilldown of results for ENC 1101 is reported because the course follows a procedure of data collection in fall term only followed by departmental discussions in spring. Therefore, no results or analysis is reported here.

No drilldown of results for ENC 1102 is reported because the course follows a procedure of data collection in fall term only followed by departmental discussions in spring. Therefore, no results or analysis is reported here.

6 REFERENCES

- Brown, M.B., Forsythe, A.B. 1974. The small sample behavior of some statistics which test the equality of several means. *Technometrics*, 16(1), 129-132.
- 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.
- 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>.