ENC0022 Assessment Report – Spring 2015

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

1 INTRODUCTION

In Fall 2014, the English Department of Florida SouthWestern State College (FSW) outlined an initial plan for assessment in three courses: English for College Success (ENC0022), Composition I (ENC1101), and Composition II (ENC1102). For Spring 2015, assessment will include ENC0022 while both ENC1101 and ENC1102 will undergo further planning and discussions based on the results of the Fall 2014 assessment before implementing a new set of goals in Fall 2015. A baseline Student Learning Objective (SLO) for ENC0022 has been implemented based on the assessment results of Fall 2014 and 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, Coordinator of Academic Assessment, Academic Affairs (jfvangaalen@fsw.edu; x6965).

2 DESCRIPTIVE STATISTICS & LEARNING OBJECTIVES

ENC0022 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. Using this common rubric criterion as an assessment method and based on the results of the Fall 2014 assessment the English department has established a benchmark (SLO1) measuring the percentage of artifacts scoring a 2 or greater.

During the Spring 2015 semester, 128 total artifacts were recorded for ENC0022. The mean overall score for the 128 artifacts is 20.0/28, or 71.4% (Table 1). The Mechanics rubric dimension exhibits the lowest mean score (2.6). Additionally, just 7.0% of artifacts were scored at a 4. With the exception of Grammar, which shares a similar distribution of artifacts scored a 4 (8.6%), achievement at level 4 in other dimensions range from 22.7% to 35.9% (Figure 1).

	Introductory Paragraph	Support Paragraphs	Organization	Concluding Paragraph	Grammar	Mechanics	Research	Overall
mean	3.1	3.0	3.0	3.0	2.7	2.6	2.7	20.0
standard deviation	0.82	0.79	0.75	0.80	0.68	0.67	0.99	4.56
Rubric Dimension	%	%	%	%	%	%	%	
4	35.9	28.1	27.3	26.6	8.6	7.0	22.7	
3	42.2	47.7	51.6	46.1	53.1	50.0	37.5	
2	18.8	20.3	18.8	24.2	35.2	39.8	24.2	
1	3.1	3.1	2.3	3.1	3.1	3.1	14.8	
Benchmark Achievement	%	%	%	%	%	%	%	
3 or greater	78.1	75.8	78.9	72.7	61.7	57.0	60.2	
2 or greater	96.9	96.1	97.7	96.9	96.9	96.9	84.4	

Table 1. Basic descriptive statistics of Spring 2015 ENC0022 artifacts. Rubric dimensions are also shown with distribution of artifacts by rubric achievement level and by percentage scoring at benchmark levels (2 or greater & 3 or greater).

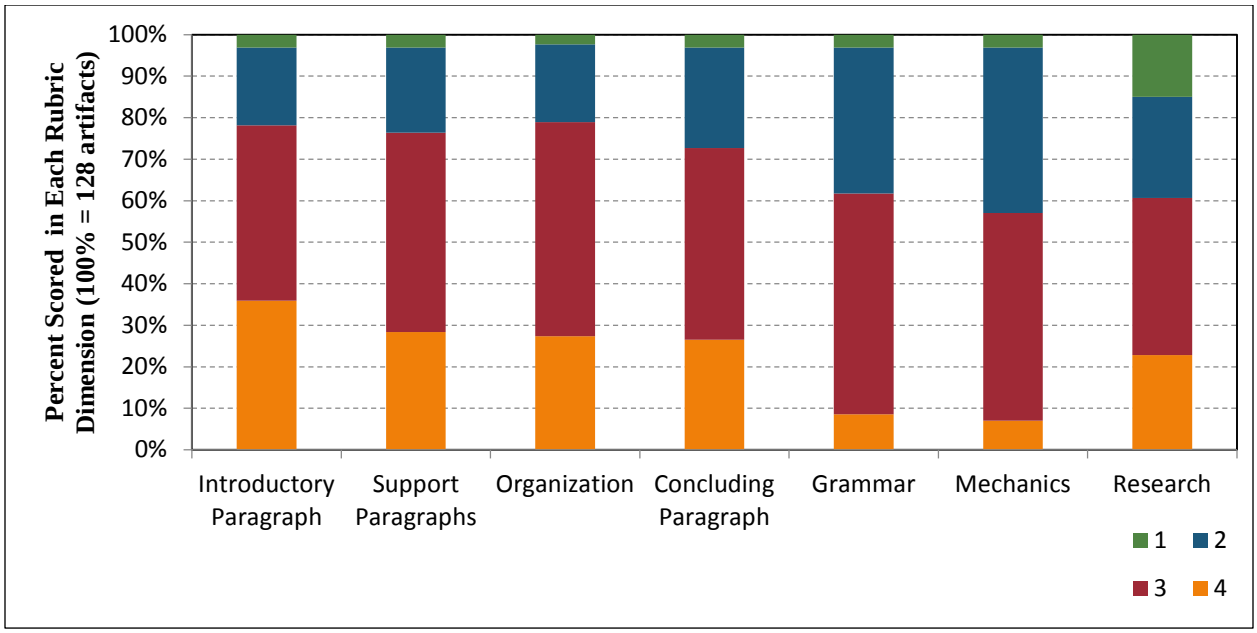

Figure 1. ENC0022 distribution of rubric scores by dimension.

The benchmark measurement, SLO1, exhibits achievement at 2 or greater ranging from 84.4% (Research) to 97.7% (Organization). Achievement at 3 or greater ranges from 57.0% (Mechanics) to 78.9% (Organization). While the Grammar, Mechanics, and Research dimensions achievement at level 4 vary by approximately 15% their achievement at level 3 or greater varies by less than 5%. A similar situation exists between Grammar and Mechanics and the remaining dimensions excluding Research. Here, Grammar and Mechanics exhibit achievement at 3 or greater at 10-20% lower than the other dimensions. At 2 or greater, this gap is reduced to 0-1%. These varied distributions speak to the typical achievement patterns in various dimensions. For a more thorough review of these patterns, see Section 3.2.

3 Exploratory Analysis & Significance Testing

3.1 Comparison by Site, Format, or Student Type

3.1.1 Dual Enrollment to non-Dual Enrollment Comparison

No dual enrollment sections of ENC0022 are offered nor do any dual enrollment students register for the course so no comparison studies were completed.

3.1.2 Online to Traditional Comparison

No online sections of ENC0022 are offered so no comparison studies were completed.

3.1.3 Full term to Mini-term Comparison

No 8-week mini-term sections were offered in Spring 2015 so no comparison studies were completed.

3.1.4 Comparison of Full-time and Part-time Faculty

During the Spring 2015 semester, 77 artifacts originate from courses taught by adjuncts while 51 artifacts originate from courses taught by full-time faculty. A comparison of the means for each rubric dimension and overall score was conducted. Each rubric dimension and the overall score was tested for significance using a Welch's t-test according to standard methods (Davis, 1973; McDonald, 2009; Wilkinson, 1999). The Introductory Paragraph, Support Paragraphs, Organization, and Research dimensions as well as the overall rubric score exhibit statistically significant differences in mean scores (see Table 2). Research is the only dimension in which the difference in the means is greater than 0.4. At 2.3, the full-time faculty exhibit average Research scores that are substantially lower than that of adjunct faculty (3.0).

df = 126^	**Introductory Paragraph**	**Support Paragraphs**	**Organization**	**Concluding Paragraph**	**Grammar**	**Mechanics**	**Research**	**Overall**
Adjunct	3.2	3.2	3.2	3.0	2.7	2.7	3.0	21.0
Full-time	2.9	2.8	2.8	2.8	2.6	2.5	2.3	18.7
Effect Size	-0.397	-0.525	-0.495	-0.256	-0.259	-0.257	-0.776	-0.496
p-value	***0.034****	***0.005***	***0.008***	0.183	0.159	0.198	***3.19x10^{-05}***	***0.004***

*Table 2. Mean scores by dimension and overall for both adjunct faculty and full-time faculty. Statistically significant results indicated in bold/italics. Positive effect sizes indicate a higher mean score for Full-time faculty artifacts. ^There are 126 degrees of freedom for all areas except Support Paragraphs and Research (125) and Overall (124). *Denote marginal significance as defined by Johnson (2013).*

Effect size was calculated using the Rosenthal and Rosnow (1991) for meta-analytical purposes to serve as a common thread across institutions (Lipsey and Wilson, 1993). The statistically significant results exhibit what Cohen (1988) would consider small to large effect sizes ranging from 0.26 to 0.77. In other words, non-overlap from adjunct artifacts to full-time artifacts ranges from approximately 17% in the case of the Grammar dimension to 47% in the case of the Research dimension.

3.1.5 Comparison by Campus/Site

Of the 128 artifacts collected from ENC0022, 11 originated from the Charlotte Campus, 30 from the Collier Campus, 5 from the Hendry-Glades Center, and 82 from the Thomas Edison (Lee) Campus. Mean scores vary by site with the Thomas Edison (Lee) Campus consistently exhibiting the lowest mean scores across all dimensions and overall score (Table 3). A plot comparing descriptive statistics of the scores by site is presented in Figure 2. While both Charlotte and Hendry-glades share the highest mean scores, both have low sample sizes with 11 for Charlotte Campus and just 5 for Hendry-Glades. Such low sample sizes make any analysis of variance results suspect and so no ANOVA was completed (Brown and Forsythe, 1974).

df = 3	**Introductory Paragraph**	**Support Paragraphs**	**Organization**	**Concluding Paragraph**	**Grammar**	**Mechanics**	**Research**	**Overall**
Charlotte	***3.7***	***3.4***	3.5	***3.2***	2.8	***3.0***	***3.3***	***22.9***
Collier	3.1	3.0	3.1	2.9	2.7	2.6	3.2	20.7
Hendry-Glades	3.4	3.2	***3.6***	3.0	***3.0***	2.8	3.0	22.0
Thomas Edison (Lee)	3.0	3.0	2.9	2.9	2.6	2.5	2.4	19.4

Table 3. Comparison of mean scores by site. Bold/italics denotes highest mean score in that dimension among all sites.

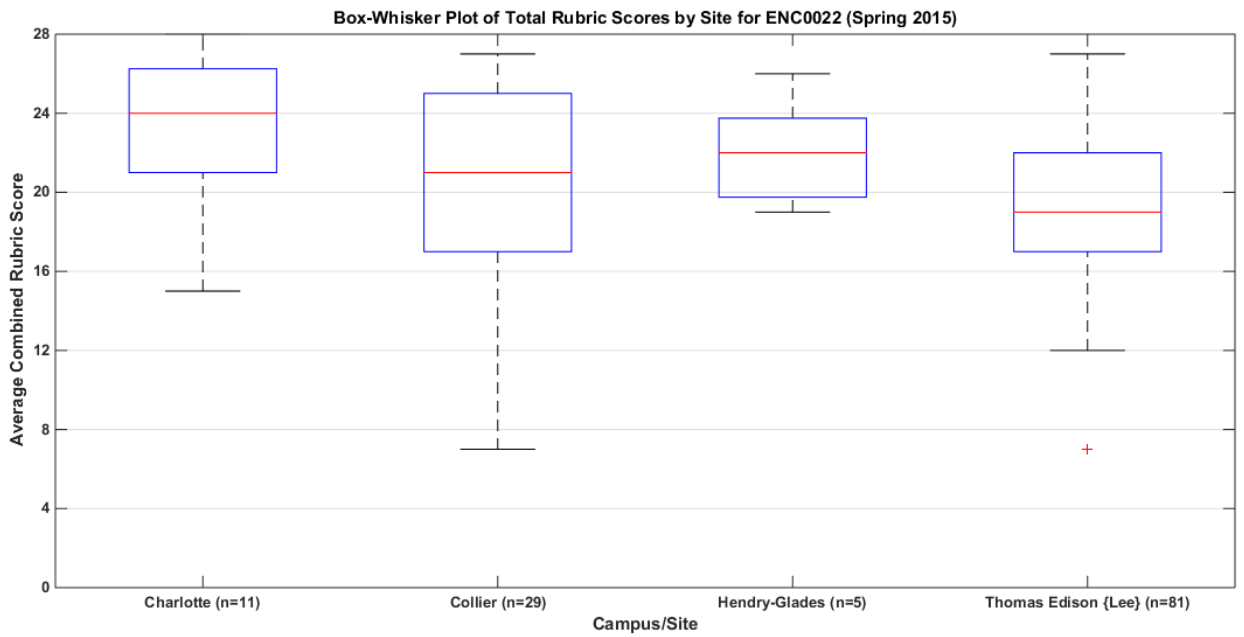

Figure 2. Box-Whisker plot of scores distributed by site for ENC0022. Red line depicts median score. Upper and lower box boundaries indicate 75% quartile and 25% quartile (box represents central 50% of the scores). Vertical lines represent remaining scores outside central 50% that are not outliers. Red '+'s denote outliers.

3.2 Data Distribution & Longitudinal Study

3.2.1 Data Distribution

Results from Section 2 briefly described the distribution in scores among rubric dimension. Varied distributions exhibited achievement gaps between dimensions at 2 or greater compared with 3 or greater. To further explore this aspect, a color map or binary raster image was created by calculating the average scores for each dimension for a given overall (total) rubric score (Figure 3).

The most effective way to read the colormap is to associate relationships of the colors based on overall scores. For example, an overall score of 21 evenly distributed across all seven rubric dimensions means each dimension would be scored a 3. The dimensions in Figure 3 above depict colors of dark yellow to light orange for the corresponding combined score of 21. When interpreted with the color bar on the right, these colors correspond to a rubric score range from 2.9 to 3.1. By comparison, an overall score of 25 evenly distributed would yield an average across each dimension of 3.6, or orange. The dimensions in Figure 3 above depict colors of yellow to dark red, corresponding to rubric score ranges from 2.7 to 4.0.

From combined rubric scores ≥ 24, the Mechanics dimension and to a lesser extent the Grammar and Research dimensions exhibit average scores that lag the other five dimensions. The Mechanics dimension exhibits an average of 3.5 when the overall rubric score is 27/28. That is, when students score a 27/28, it is most likely the Mechanics dimension which is scored a 3/4 while all others are 4/4.

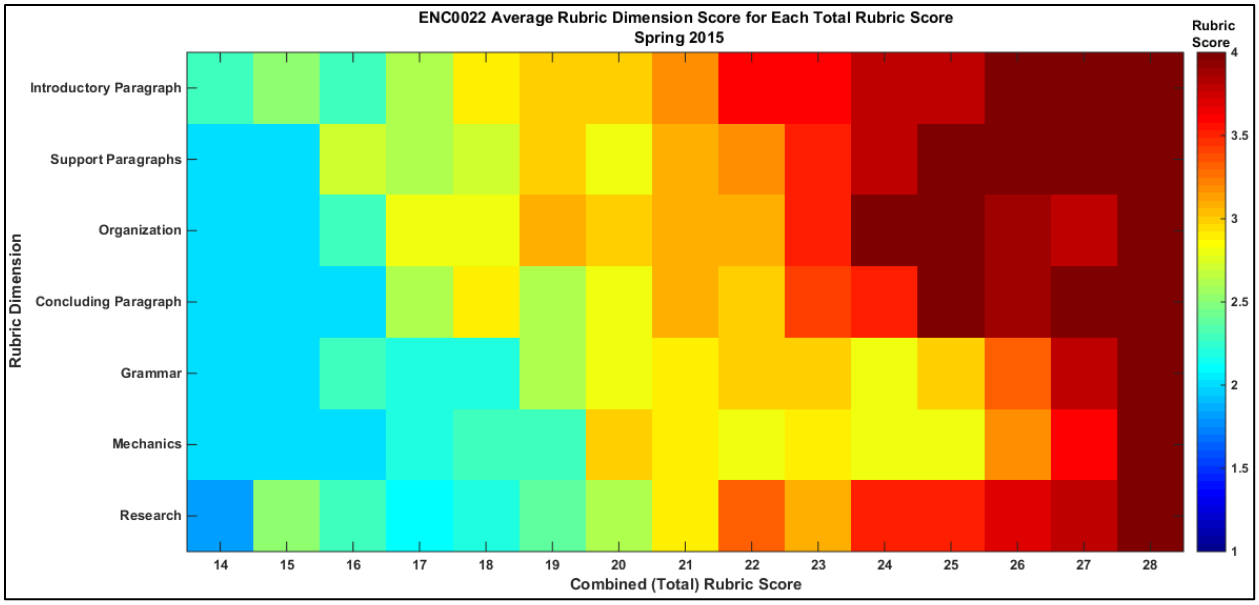

Figure 3. Colormap of mean scores for each rubric dimension for each combined (total) rubric score for ENC0022.

From combined rubric scores of 17-23, it is the Grammar, Mechanics, and Research dimensions which lag compared with the other four dimensions. At the lower end of the overall scores (< 17) mean scores exhibit fairly even distribution across all dimensions. In short, Grammar, Mechanics, and Research score similar to other dimensions when scores are below 17 and lag other dimensions in overall score range of ≥ 17-23, with Grammar substantially so above 23.

3.2.2 Longitudinal Study

Figure 4 shows the comparison of each rubric dimension achievement percentages from Fall 2014 to Spring 2015. The Introductory Paragraph, Support Paragraphs, Organization, Concluding Paragraph, and Mechanics rubric dimensions exhibit increased achievement at both level 3 and 4. It should be noted that a comparison of achievement from term-to-term as opposed to year-to-year isn't necessarily a one-to-one comparison at FSW. Assessment reports across multiple course level and program level assessments support this and should be taken under consideration upon drawing any relevant conclusions (see http://www.fsw.edu/facultystaff/assessment/history for further details).

In comparing mean rubric score from term-to-term there is an increase in all dimensions except Research, which exhibits a decline from 2.8 to 2.7. The largest increases are in the Introductory paragraph dimension and Concluding Paragraph dimension, at +0.3 and +0.2. All other increases are +0.1 or less.

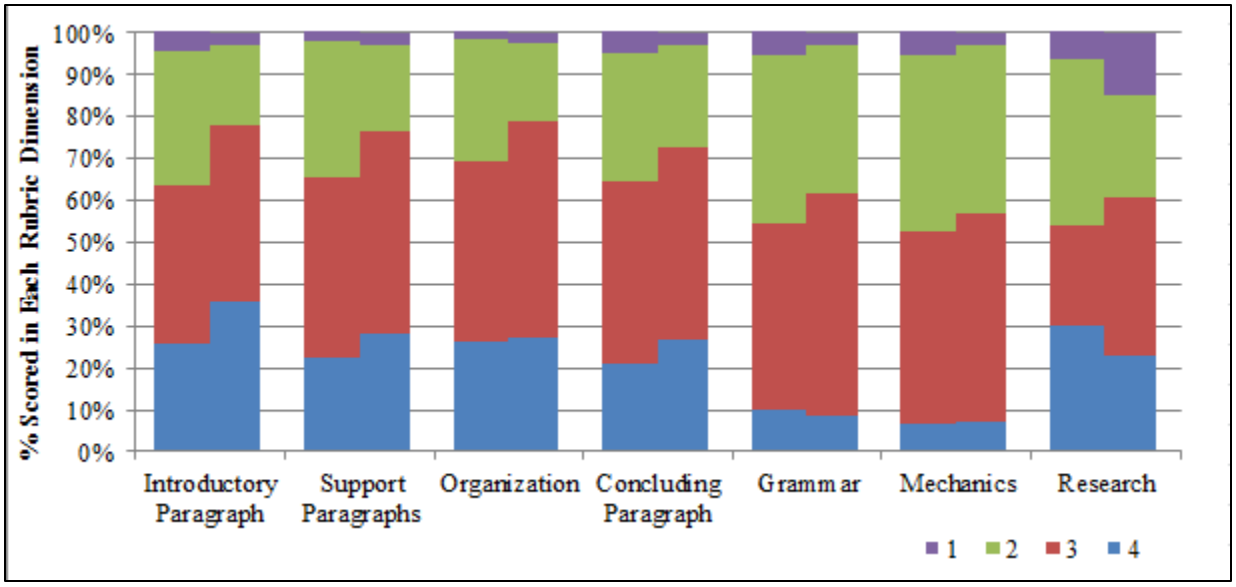

Figure 4. Distribution of rubric scores by dimension for both Fall 2014 (left half of bar graph on each dimension) and Spring 2015 (right half of bar graph on each dimension).

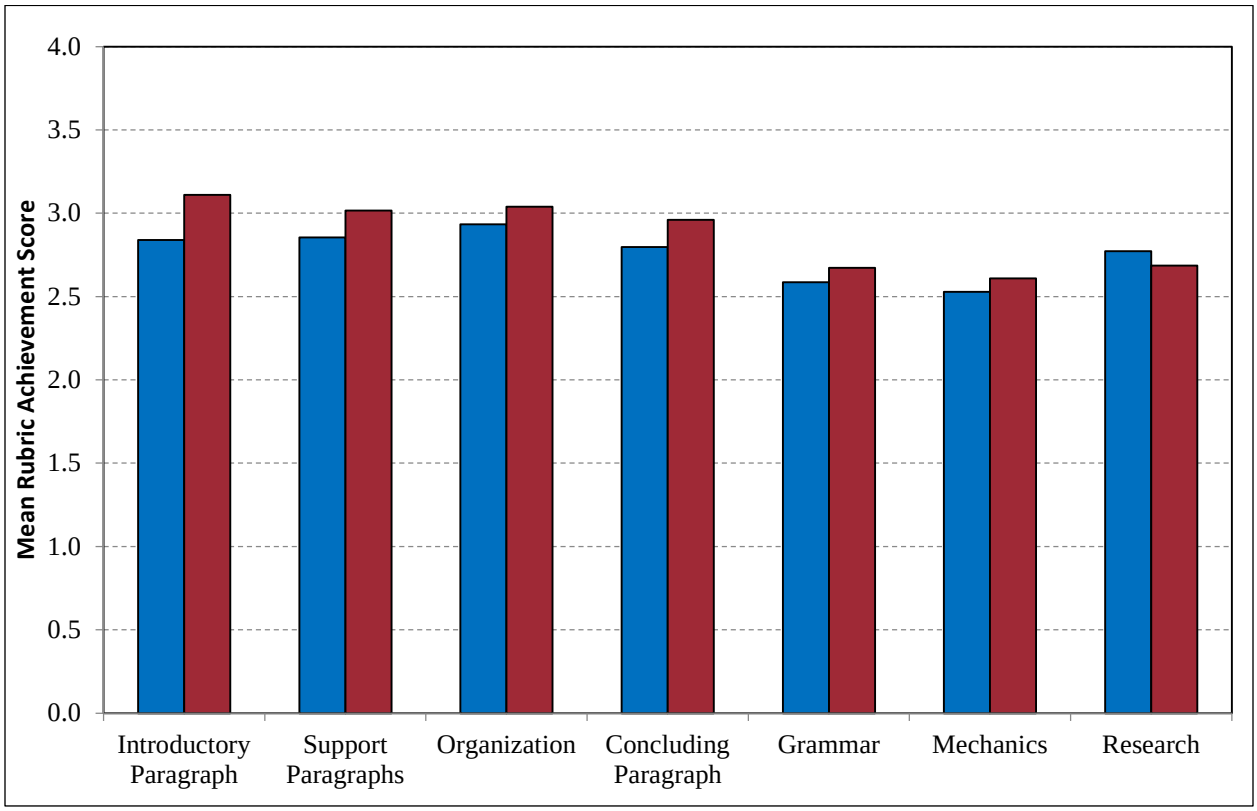

Figure 5. Mean scores by rubric dimension for Fall 2014 (blue) and Spring 2015 (red).

4 CONCLUSIONS

The objective of Spring 2015 assessment for the FSW English Department was to assess the ENC0022 English for College Success course using the new Student Learning Objective (SLO) while both ENC1101 Composition I and ENC1102 Composition II undergo further development using new learning objectives in Fall 2015. Using the same common rubric criterion as Fall 2014, the results of the Fall 2014 assessment resulted in the establishment by the English department of a benchmark (SLO1) measuring the percentage of artifacts scoring a 2 or greater.

A drilldown of ENC0022 results are as follows:

1. All seven rubric dimensions have > 80% achievement at level 2 or higher. The lowest dimension, Research, exhibits achievement of 84.4% at 2 or higher.
2. All rubric dimensions except for Mechanics exhibit have > 60% of achievement at level 3 or higher. The Mechanics dimension exhibits a rate of 57.0% at level 3 or higher.
3. No dual enrollment sections of ENC0022 are offered nor do any dual enrollment students register for the course so no comparison studies were completed.
4. No online sections of ENC0022 are offered so no comparison studies were completed.
5. No 8-week mini-term sections were offered in Spring 2015 so no comparison studies were completed.
6. In a comparison of full-time faculty to adjunct faculty, there was a statistically significantly higher mean score for adjunct faculty artifacts in all rubric dimensions except Grammar and Mechanics. In the case of the Research dimension, full-time faculty exhibit average scores that are substantially lower than that of adjunct faculty at 2.3 compared with 3.0, respectively.
7. In a cross-campus comparison, both the Charlotte campus and Hendry-Glades center exhibit consistently higher mean rubric scores compared with the other two sites, although low sample size limits validity of the comparison.
8. In a study comparing average rubric dimension score according to overall score, the Grammar, Mechanics, and Research score similar to other dimensions when scores are below 17/28 and lag other dimensions in overall score range of ≥ 17-23, with Grammar substantially so above 23.
9. In a longitudinal study, mean rubric scores increased from Fall 2014 to Spring 2015 in all dimensions except Research.

5 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.

Cohen, J. 1988. Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Earlbaum 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.