Foreign Language Assessment Report
Summer 2018

Author: Joseph F. van Gaalen, Ph.D., Asst. VP, Institutional Research, Assessment, & Effectiveness

1   INTRODUCTION

Florida SouthWestern's Foreign Language Department employs a common course assessment to measure student progress in course level objectives, a practice shown to be effective in establishing data driven instruction (Hall, 2010). Courses included in assessment are: FRE 1120 *Elementary French I*, FRE 1121 *Elementary French II*, SPN 1120 *Beginning Spanish I*, and SPN 1121 *Beginning Spanish II*. Through achievement of the courses students will acquire and demonstrate competency in speaking, writing, reading comprehension and listening comprehension in standard Spanish or French at the beginner's level. The assessment outcomes outlined below define the method of assessment for each course assessment as well as measure current Student Learning Outcomes (SLOs) and identify areas for future SLOs to be assessed. Additionally, the plan provides information on achievement levels of dual enrollment (concurrent) artifacts compared with traditional, as well as online artifacts compared with traditional artifacts as highlighted in the course level assessment plan.   This report provides achievement analysis for both summer 2018 as well as longitudinal studies, where applicable.

For additional detail or further analysis not provided in this report, please contact Dr. Joseph F. van Gaalen, Asst. VP, Institutional Research, Assessment, & Effectiveness (jfvangaalen@fsw.edu; x16965).

2   FRENCH

2.1   FRE 1120
Full implementation of the new assessment and/or modification for FRE 1120 will continue in the fall 2018 term.

2.2   FRE 1121
Full implementation of the new assessment and/or modification for FRE 1121 will continue in the fall 2018 term.

3   SPANISH

3.1   SPN 1120

3.1.1   Descriptive Statistics, Learning Outcomes, and Learning Objectives
During the summer 2018 semester 6 sections of SPN 1120 Beginning Spanish I were offered. Of those, artifacts from a common final were collected from 4 sections. One section did not report data. A second section reported incomplete data so it was not included in the study. A total of 83 students

were enrolled in SPN 1120. Of those, 55 artifacts were collected representing a sample size of 66% of the population.

Using a common course assessment, the FSW Spanish faculty defined three areas of interest for evaluation that apply to SPN 1120. The Student Learning Outcomes (SLOs) and their objectives or measures of success are:

➢ SLO 1: Students will be able to understand spoken Spanish. The faculty established measure of success for this SLO is for 80% of students to demonstrate competency with a score of 70% or better in the oral comprehension exam sections (Section I).
➢ SLO 2: Students will be able to understand written Spanish. The faculty established measure of success for this SLO is for 80% of students to demonstrate competency with a score of 70% or better in the reading comprehension exam sections (Section II and III).
➢ SLO 3: Students will be able to write effectively in the Spanish language. The faculty established measure of success for this SLO is for 80% of students to demonstrate competency with a score of 70% or better in the writing competency exam sections (Section IV and V).

The faculty established measure of success for SLO 1, 80% of students scoring 70% or higher in Section I, was nearly met as results exhibit 75% of artifacts score 70% or higher in the oral competency exam section (Section I) (Table 1). The faculty established measure of success for SLO 2, 80% of students scoring 70% or higher in Sections II and III, was partially met. Results exhibit 55% of artifacts scored 70% or higher in Section II and 85% of artifacts scored 70% or higher in Section III. The faculty established measure of success for SLO 3, 80% of students scoring 70% or higher in Sections IV and V, was partially met. Results exhibit 78% of artifacts scored 70% or higher in Section IV and 87% of artifacts scored 70% or higher in Section V. For a graphical representation of SLO achievement, see Figure 1.

n = 55	Section I (Oral)	Section II (Reading)	Section III (Reading)	Section IV (Written)	Section V (Written)	Combined (Overall)
Goal	*80% of artifacts scored ≥70% for all sections*					
% above 70%	75%	55%	85%	78%	87%	
Mean (as %)	79%	70%	85%	81%	82%	77%
Median (as %)	87%	75%	93%	87%	90%	79%
Section Score Max	15	60	15	15	20	125
Section Mean	11.9	42.1	12.8	12.1	16.3	95.3
Section Median	13	44	14	13	18	97.5

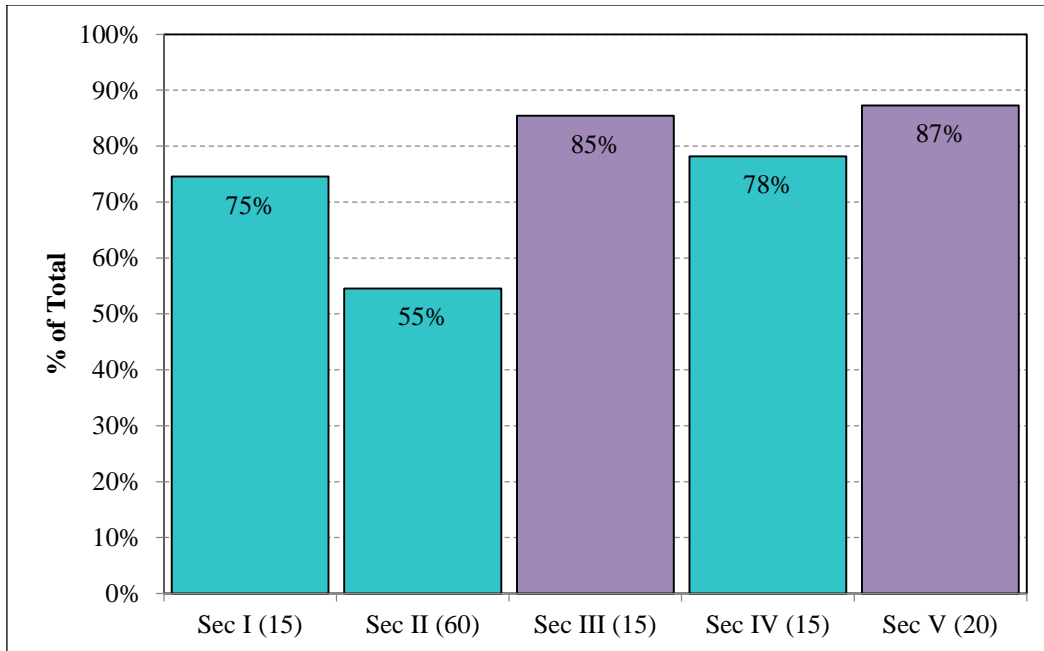Table 1. Percentage of student achievement level as per SLOs (SPN 1120).

Figure 1. SLO achievement for SPN 1120 by exam section (Percentage of student achievement level as per SLOs). Purple denotes having met objective.

3.1.2 Exploratory Analysis & Significance Testing

Multiple comparisons of artifact scores across varying formats, campuses, and student types were made to more thoroughly detail the assessment data. Each course was divided into appropriate subgroups such as by campus or enrollment status to perform the analysis. Where possible, additional methods of analysis were conducted to provide a broader picture of these comparisons.

3.1.2.1 Dual Enrollment (Concurrent) to Non-Dual Enrollment Comparison

No dual enrollment (concurrent) sections of the course were run during summer 2018 so no comparison study between dual enrollment and traditional courses could be completed.

3.1.2.2 Online to Traditional Comparison

During the spring 2018 semester, two online sections reported data. From those course sections, 23 total online artifacts were collected from SPN 1120 and 32 traditional artifacts were collected from SPN 1120. A comparison of basic statistics is provided in Table 2. Online artifact mean scores are 6.4 higher than traditional artifacts. Differences in the means were tested for significance using a Welch's t-test according to standard methods (Davis, 1973; McDonald, 2009; Wilkinson, 1999) and were found to not be statistically significantly different. Therefore, we cannot reject the null hypothesis that the differences in the means of the online and traditional artifacts are equal to 0, and we cannot conclude this with a 95% confidence that the differences in scores are not solely due to chance.

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 results exhibit what Cohen (1988) would consider a medium effect size. In other words, non-overlap score distribution from online artifacts to traditional artifacts is approximately 22%. For a graphical representation of this see Figure 2.

df = 53	
Online mean	79.9%
Online standard deviation	21.33
Traditional mean	73.5%
Traditional standard deviation	16.44
Effect size	-0.33
p-value	0.234

Table 2. Comparison of mean scores (as %) for online and traditional artifacts. Positive effect sizes indicate a higher mean score for traditional artifacts.
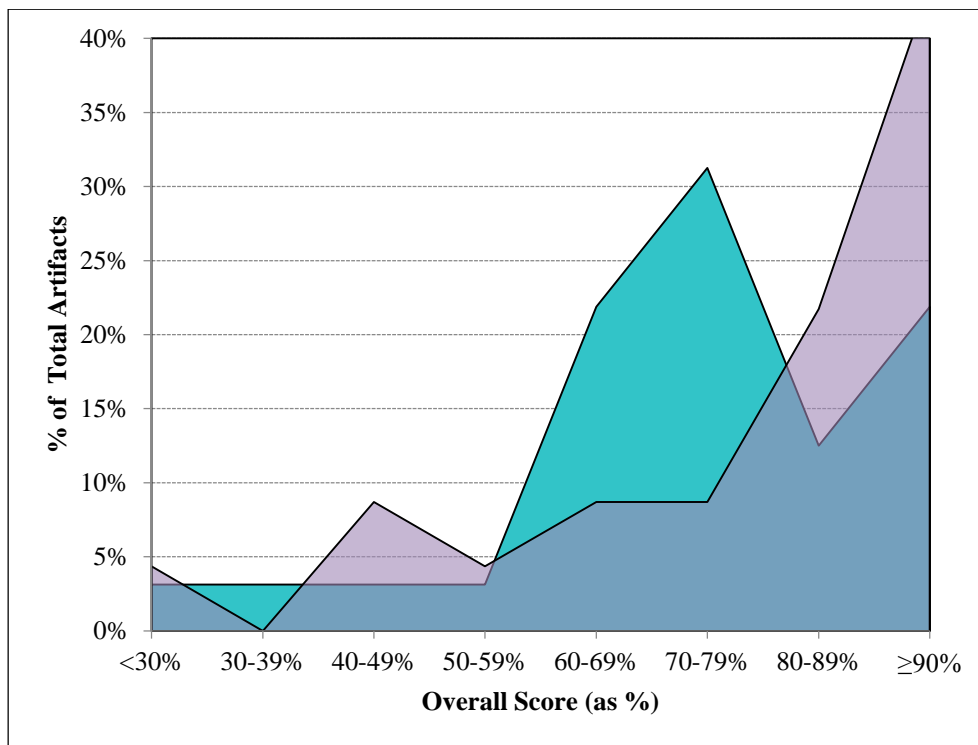

Figure 2. Score distribution for online (purple) and traditional (aqua) artifacts of SPN 1120.

3.1.2.3   Comparison by Campus/Site

Of the 55 artifacts collected from SPN 1120, 6 originated from the Charlotte campus, 13 from the Collier campus, 23 from FSW Online, and 13 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.

	Section I	Section II	Section III	Section IV	Section V	Combined Score
Rubric Max	*15*	*60*	*15*	*15*	*20*	*125*
Charlotte	10.1	37.3	10.5	12.4	14.0	84.3
Collier	12.0	**44.5**	**14.5**	**13.8**	**18.0**	**102.7**
FSW Online	**13.2**	43.9	13.1	12.5	17.3	99.9
Thomas Edison (Lee)	10.4	38.8	11.7	9.6	14.0	84.6

Table 3. Comparison of mean scores by site for SPN 1120. Bold denotes highest mean score in that dimension among all sites. Rubric dimensions identified in SLOs in blue.

The Collier campus exhibits the highest mean score in four of five sections (II, III, IV, and V) and in the overall. FSW Online exhibits the highest mean score in Section I (Figure 3).
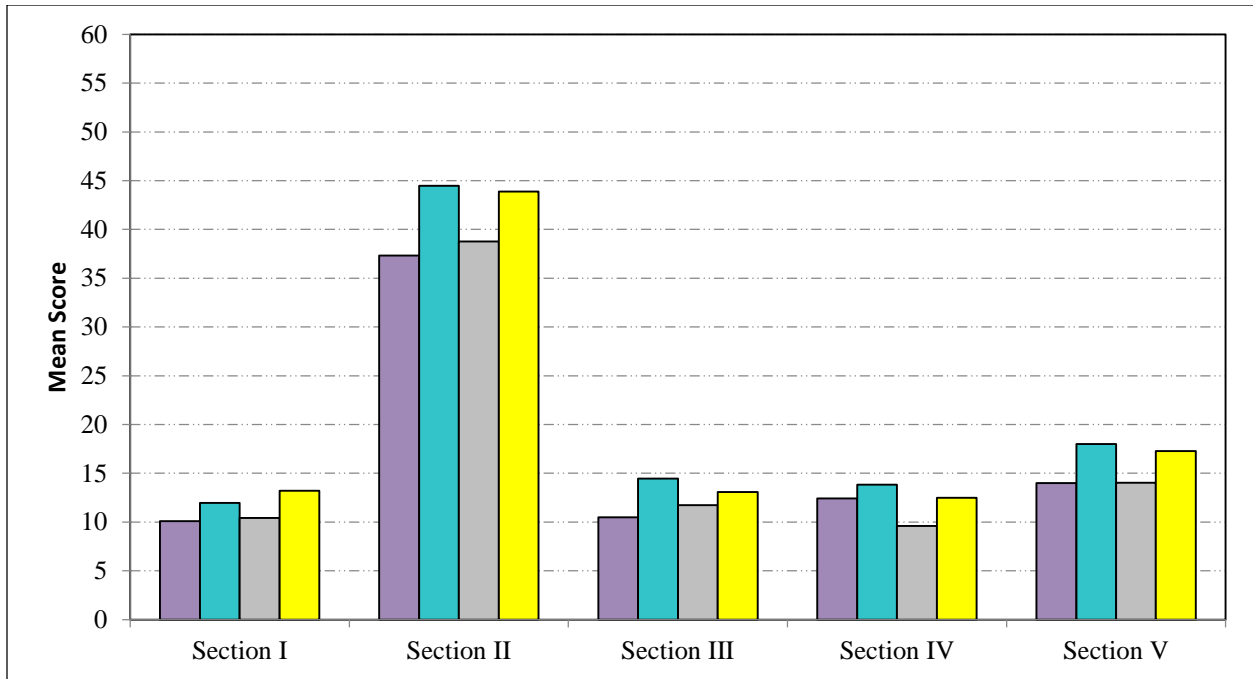
Figure 3. Comparison of mean score of exam sections by site with Charlotte in purple, Collier in aqua, Thomas Edison in gray, and FSW Online in yellow.

A plot comparing descriptive statistics of the combined (overall as a percentage) scores by site is presented in Figure 4. Sites exhibit distinct distribution variations. However, because sample size was limited at most locations, interpretation is limited.
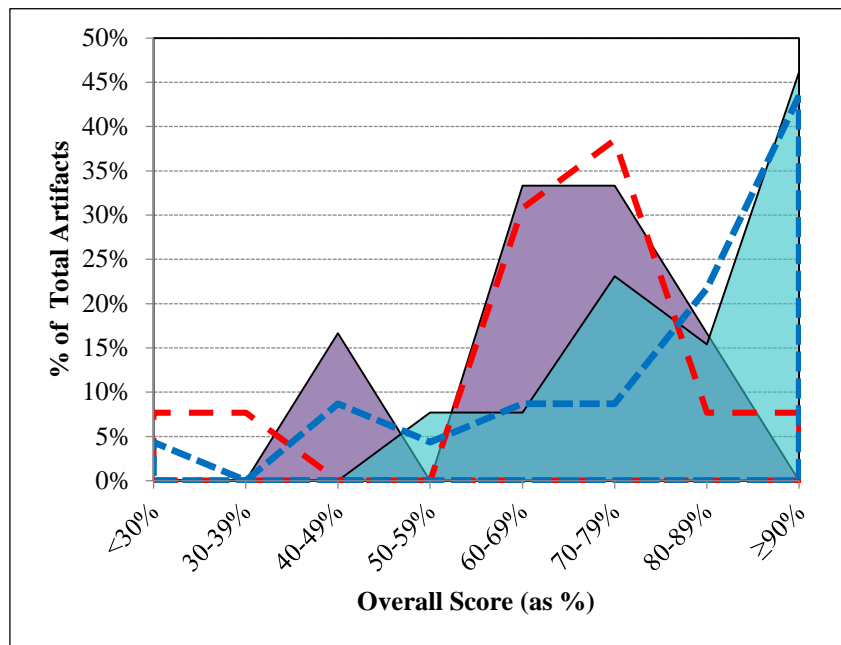

Figure 4. Comparison of artifact score distribution for Charlotte in purple, Collier in teal, Thomas Edison in red dashed line, and FSW Online in blue dashed line.

A one-way analysis of variance was used to compare means of the combined rubric scores at each site. Results of the ANOVA do not exhibit a statistically significant difference between sites (see Table 4). Therefore, we cannot reject the null hypothesis that the mean combined 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.

Source of Variation	Sum of squared differences	df	Mean Squares	F_{obs}	p-value	F_{crit}
Between Sites	3428.2	3	1142.7	2.23	0.096	2.79
Within Sites	26163.7	51	513.0			
Total	29591.9	54				

Table 4. Results of one-way ANOVA of combined rubric scores at each site for SPN 1120.

3.1.3   Data Distribution & Longitudinal Study

3.1.3.1   Data Distribution

A histogram depicting the distribution of scores across each exam section is shown in Figure 5. All sections exhibit peaks above 90%. Sections II exhibits a bimodality with a secondary peak at 60-69%. Section II also continues to exhibit more widely distributed scores. Section II exhibits a peak of 25% at ≥ 90% while other sections range from 45% to 55%.
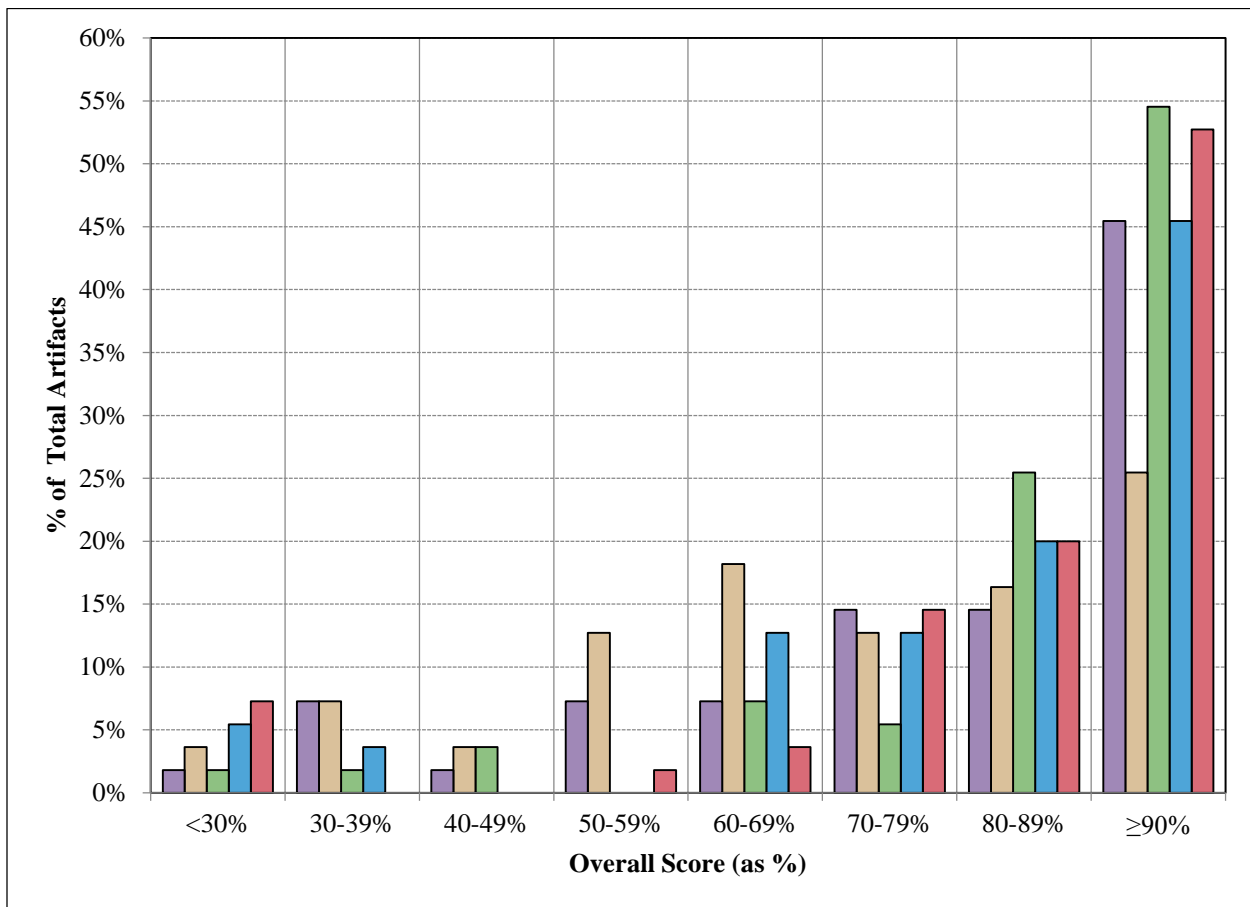

Figure 5. Histogram of SPN 1120 exam sections scores for summer 2018. Purple – Section I, Brown – Section II, Green – Section III, Blue – Section IV, and Red – Section V.

To describe the behavior of the section scores based on overall achievement, a color map, or binary raster image, is typically created, however due to sample size, this analysis would not be practical.

3.1.3.2   Longitudinal Study

Further description of achievement over time in for summer terms is provided in Table 5. Both demographics of students and student count vary by semester it may be more reasonable to compare like semesters (Fall vs. Fall, Spring vs. Spring) (see http://www.fsw.edu/facultystaff/assessment/history for further details). Summer 2018 exhibits substantially higher mean scores in all sections.

	Section Max	Summer 2017 n=55	Summer 2018 n=55
Section I (Oral)	15	11.5	11.9
Section II (Reading)	60	40.4	42.1
Section III (Reading)	15	12.3	12.8
Section IV (Written)	15	11.1	12.1
Section V (Written)	20	15.4	16.3
Combined (Overall)	*125*	**90.8**	**95.3**

Table 5. Comparison of mean scores for SPN 1120 for summer terms.

3.2   ENTRANCE SKILLS STUDY

In the spring 2016 term, a brief assessment was piloted in two sections of SPN 1121 on the Thomas Edison campus prior to entering into any engaged study in the course. The purpose of the assessment was to assess student skills and retention of materials from SPN 1120 prior to beginning SPN 1121. The assessment consists of a 15-minute reading comprehension quiz mirroring that which is seen in Section II of the SPN 1120 common final exam. The study compared study skill level based on the instructor they had for the previous course (SPN 1120) in an effort to align student skill level upon entry into SPN 1121. The pilot program for this study was first included in the fall 2015 assessment report. This program continues and was most recently administered at the beginning of fall 2018, before the writing of this report and is included here.

The assessment has now been administered in all traditional sections of SPN 1121 beginning with summer 2016 through fall 2018 on the Charlotte, Collier, and Thomas Edison campuses. Results of the cumulative study are shown in Figure 8.

In concept, upon entry to SPN 1121, all students should have at least a passing score (or nearly so) of the final exam from SPN 1120. The entrance skills study exhibits a range of mean scores across instructor that is diminished since previous years (Figure 8). Results exhibit achievement levels based on previous instructor spanning as low as 12.5/30 (up from 11.7/30 following summer 2018) to as high as 21.9/30 (unchanged from summer 2018). This is a substantially decreased range from the earliest studies where the range was 7.2/30 to 25.3/30.
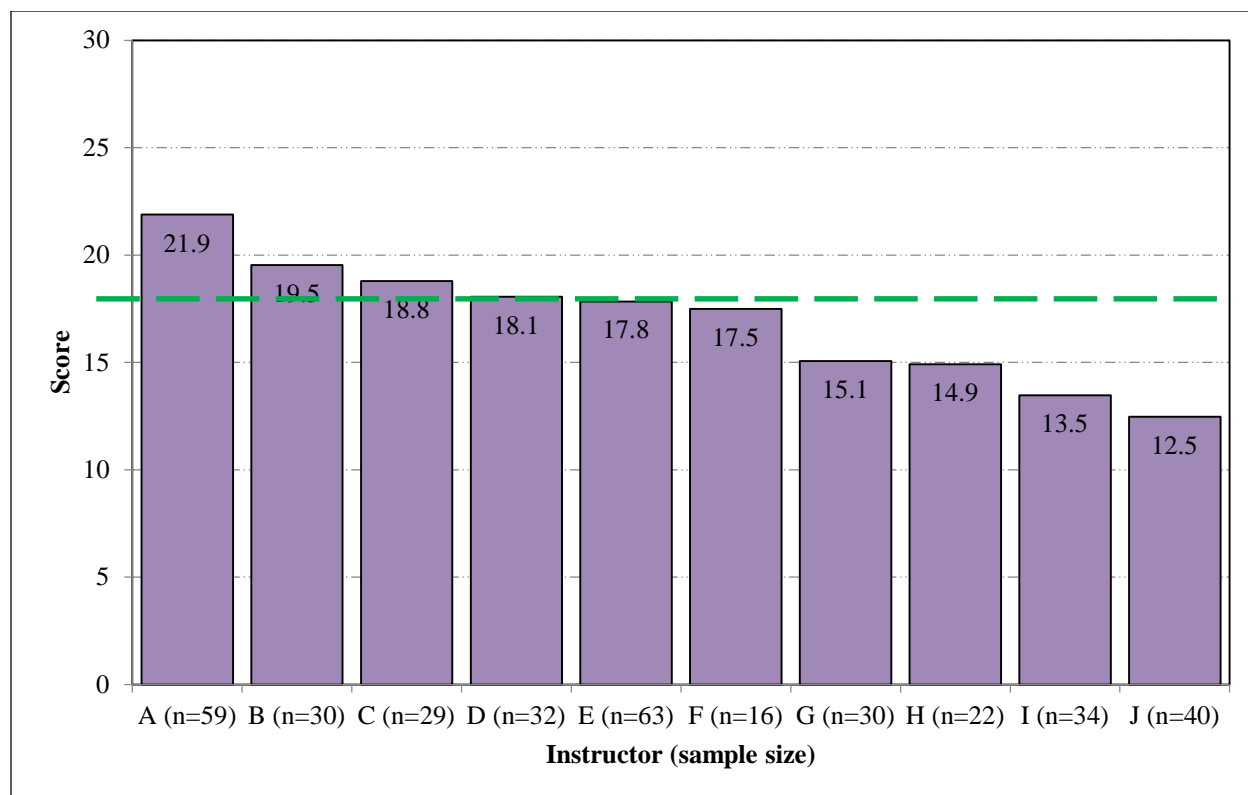
Figure 6. Comparison of achievement in entrance skills study assessment by instructor. Green dashed line denotes fall 2017 mean score for Section II of SPN 1120 (section of exam the entrance skills assessment is based).

3.3   SPN 1121

3.3.1   Descriptive Statistics, Learning Outcomes, and Learning Objectives

Six sections of SPN 1121 Beginning Spanish II were offered in summer 2018. Of those, artifacts were collected from a common final from 3 of 6 sections. A total of 79 students were enrolled in SPN 1121. Of those, 38 artifacts were collected representing a sample size of 48%.

Using a common course assessment, the FSW Spanish faculty defined the same three areas of interest for evaluation that apply to SPN 1121 as those used for SPN 1120. For details on each SLO, see 3.1.1. The only difference between SPN 1121 and SPN 1120 in terms of measuring these outcomes is that the exam sections differ slightly and are noted in Table 6 below.

The faculty established measure of success for SLO 1, 80% of students scoring 70% or higher in Section I, was nearly met as results exhibit 76% of artifacts score 70% or higher in the oral competency exam section (Section I) (Table 6, Figure 9). The faculty established measure of success for SLO 2, 80% of students scoring 70% or higher in reading only sections, Sections II, and VI, was nearly met. Results exhibit 58% of artifacts scored 70% or higher in Section II and 74% of artifacts scored 70% or higher in Section VI. The faculty established measure of success for SLO 3, 80% of students scoring 70% or higher in writing only sections, Sections V and VII, was partially met. Results exhibit 79% of artifacts scored 70% or higher in Section V and 95% of artifacts scored 80% or higher in Section VII.

n = 38	Section I (Oral)	Section II (Reading)	Section III (Read/Write)	Section IV (Read/Write)	Section V (Writing)	Section VI (Reading)	Section VII (Writing)	Combined (Overall)
Goal	80% of artifacts scored ≥70% for all sections							
% above Goal	76%	58%	79%	84%	79%	74%	95%	
Mean (as %)	85%	64%	82%	81%	82%	81%	89%	81%
Median (as %)	92%	73%	88%	85%	92%	87%	95%	87%
Section Score Max Possible	15	15	40	15	12	15	20	132
Section Mean	12.8	9.6	32.9	12.1	9.8	12.1	17.8	107.1
Section Median	13.8	11	35.3	12.8	11	13	19	114.5

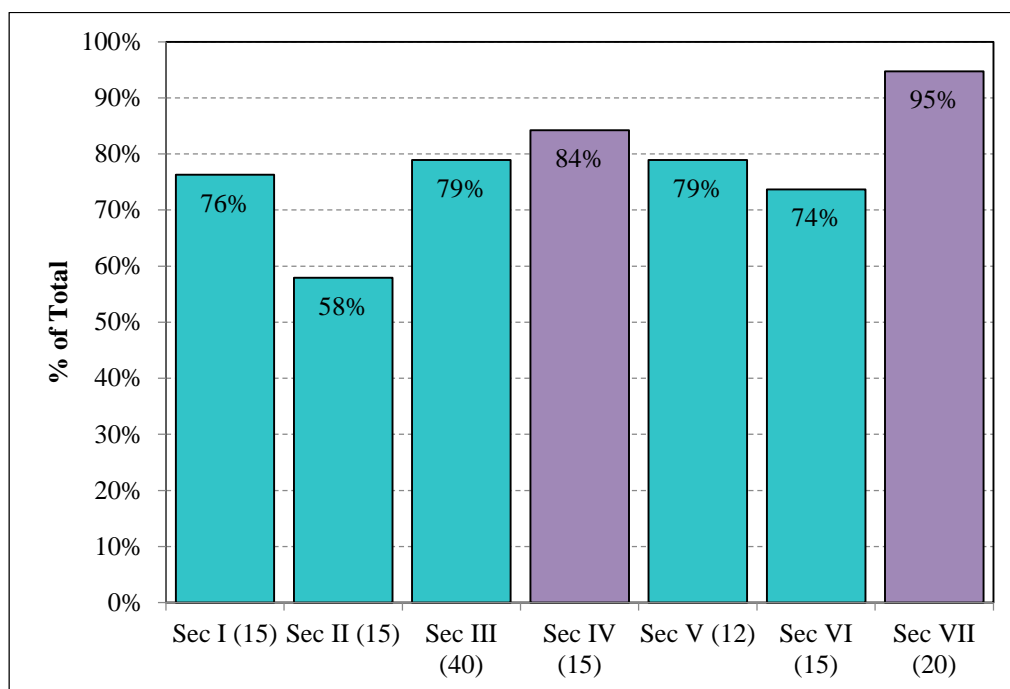Table 6. Percentage of student achievement level as per SLOs (SPN 1121).

Figure 7. SLO achievement for SPN 1121 by exam section (Percentage of student achievement level as per SLOs). Purple denotes having met objective.

3.3.2 Exploratory Analysis & Significance Testing

Multiple comparisons of artifact scores across varying formats, campuses, and student types were made, where possible, in order to add depth to the causes of the distribution of the artifacts. Each course was divided into the appropriate subgroups to perform the analysis. In cases where a subgroup is not represented in the course comparisons were not conducted and are noted for comprehensiveness.

3.3.2.1 Dual Enrollment (Concurrent) to Non-Dual Enrollment Comparison

No dual enrollment (concurrent) sections of the course were run during summer 2018 so no comparison study between dual enrollment and traditional courses could be completed.

3.3.2.2 Online to Traditional Comparison

During the summer 2018 semester, one online sections reported data. From those course sections, 12 total online artifacts were collected from SPN 1121 and 26 traditional artifacts were collected from SPN 1121. A comparison of basic statistics is provided in Table 7. Online artifacts mean scores are 13.2 lower than traditional artifacts. Differences in the means were tested for significance using a Welch's t-test according to standard methods (Davis, 1973; McDonald, 2009; Wilkinson, 1999) and were found to

not be statistically significantly different. Therefore, we cannot reject the null hypothesis that the differences in the means of the online and traditional artifacts are equal to 0, and we cannot conclude this with a 95% confidence that the differences in scores are not solely due to chance.

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 results exhibit what Cohen (1988) would consider a medium effect size. In other words, non-overlap score distribution from online artifacts to traditional artifacts is approximately 42%. For a graphical representation of this see Figure 10.

df = 36	
Online mean	72.1
Online standard deviation	21.37
Traditional mean	85.3
Traditional standard deviation	12.30
Effect size	-0.67
p-value	0.065

Table 7. Comparison of mean scores (as %) for online and traditional artifacts. Positive effect sizes indicate a higher mean score for traditional artifacts.
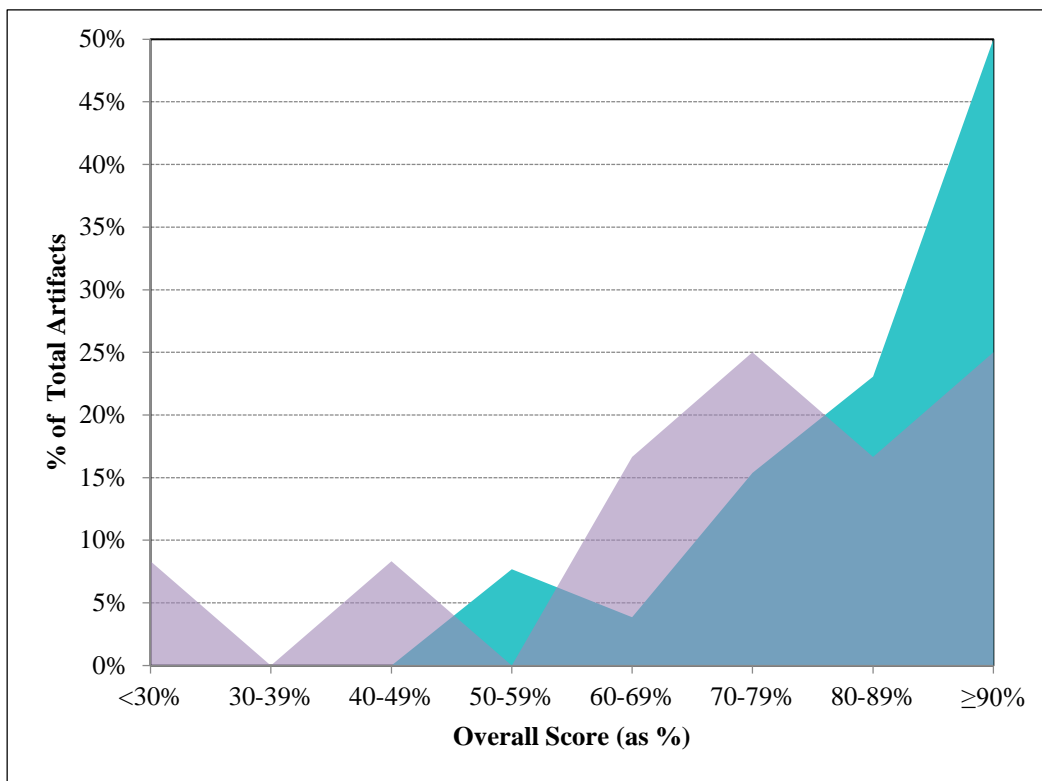

Figure 8. Score distribution for online (purple) and traditional (aqua) artifacts of SPN 1121.

3.3.2.3   Comparison by Campus/Site
While three sites offered the course, sample size was limited for one of the sites. As a result, no cross-campus comparison is completed.

3.3.3   Data Distribution & Longitudinal Study

3.3.3.1   Data Distribution

A histogram depicting the distribution of scores across each exam section is shown in Figure 11. All sections exhibit peaks centered on ≥90%. Sections II, III, and IV exhibit a bimodality. Section II exhibits a secondary peak at < 30%. Section III exhibits a secondary peak at 50-59%. And Section IV exhibits a secondary peak centered on 70-79%.
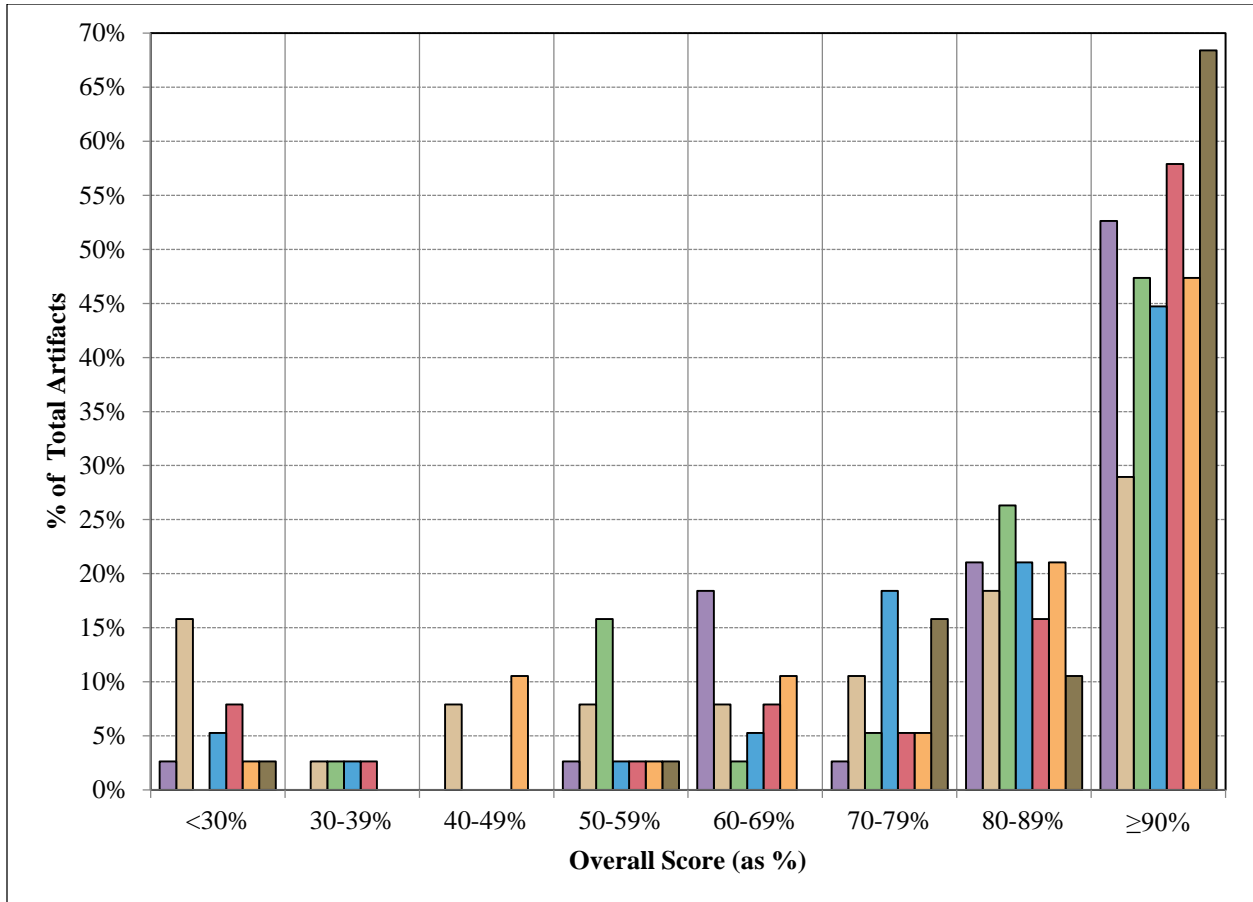

Figure 9. Histogram of SPN 1121 exam sections scores for summer 2018. Purple – Section I, Brown – Section II, Green – Section III, Blue – Section IV, Red – Section V, Orange – Section VI, and Black – Section VII.

To describe the behavior of the section scores based on overall achievement, a color map, or binary raster image, is typically created, however due to sample size, this analysis would not be practical.

3.3.3.2   Longitudinal Study

Further description of achievement over time in SPN 1121 is provided in Table 8). Both demographics of students and student count vary by semester. It may be more reasonable to compare like semesters (Fall vs. Fall, Spring vs. Spring). (see http://www.fsw.edu/facultystaff/assessment/history for further details). Summer 2018 exhibits higher mean scores in all sections except Section IV.

	Section Max	Summer 2017 n=57	Summer 2018 n=38
Section I (Oral)	15	12.6	12.8
Section II (Reading)	15	9.4	9.6
Section III (Read/Write)	40	28.5	32.9
Section IV (Read/Write)	15	12.2	12.1
Section V (Writing)	12	8.8	9.8
Section VI (Reading)	15	11.3	12.1
Section VII (Writing)	20	16.8	17.8
Combined (Overall)	**132**	**99.5**	**107.1**

Table 8. Comparison of mean scores for SPN 1121 for summer terms.

4 CONCLUSIONS

Florida SouthWestern's Foreign Language Department employs a common course assessment in both French and Spanish courses to measure student progress in course level objectives in an effort to improve instruction. What follows is a drilldown of findings for both disciplines (French and Spanish) for the summer 2018 assessment.

4.1 FRENCH

A drill-down of FRE 1120 results are as follows:
1. Full implementation of the new assessment and/or modification for FRE 1120 will continue in the fall 2018 term.

A drill-down of FRE 1121 results are as follows:
1. Full implementation of the new assessment and/or modification for FRE 1121 will continue in the fall 2018 term.

4.2 SPANISH

A drill-down of SPN 1120 results are as follows:
1. Achievement of 80% of artifacts scoring 70% or better (SLO 1): Achievement was nearly met as results exhibit 75% of artifacts score 70% or higher in the oral competency exam section (Section I).
2. Achievement of 80% of artifacts scoring 70% or better (SLO 2): Achievement was partially met. Results exhibit 55% of artifacts scored 70% or higher in Section II and 85% of artifacts scored 70% or higher in Section III.
3. Achievement of 80% of artifacts scoring 70% or better (SLO 3): Achievement was partially met. Results exhibit 78% of artifacts scored 70% or higher in Section IV and 87% of artifacts scored 70% or higher in Section V.
4. No dual enrollment (concurrent) sections of the course were run during summer 2018 so no comparison study between dual enrollment and traditional courses could be completed.
5. In a comparison of online to traditional artifacts, online artifact mean scores are 6.4 higher than traditional artifacts. Results were not statistically significantly different.
6. In a cross-campus comparison, the Collier campus exhibits the highest mean score in four of five sections (II, III, IV, and V) and in the overall. FSW Online exhibits the highest mean score in Section I.
7. In a study of score distribution by section, sections exhibit peaks above 90%. Sections II exhibits a bimodality with a secondary peak at 60-69%. Section II also continues to exhibit more widely

distributed scores. Section II exhibits a peak of 25% at ≥ 90% while other sections range from 45% to 55%.

8. In a longitudinal study of data distribution through summer terms, Summer 2018 exhibits substantially higher mean scores in all sections.

9. In a study assessing student skills and retention of materials from SPN 1120 prior to beginning SPN 1121, results exhibit achievement levels based on previous instructor spanning as low as 12.5/30 (up from 11.7/30 following summer 2018) to as high as 21.9/30 (unchanged from summer 2018). This is a substantially decreased range from the earliest studies where the range was 7.2/30 to 25.3/30.

A drill-down of SPN 1121 results are as follows:

1. Achievement of 80% of artifacts scoring 70% or better (SLO 1): Achievement was nearly met as results exhibit 76% of artifacts score 70% or higher in the oral competency exam section (Section I).

2. Achievement of 80% of artifacts scoring 70% or better (SLO 2): Achievement was nearly met. Results exhibit 58% of artifacts scored 70% or higher in Section II and 74% of artifacts scored 70% or higher in Section VI.

3. Achievement of 80% of artifacts scoring 70% or better (SLO 3): Achievement was partially met. Results exhibit 79% of artifacts scored 70% or higher in Section V and 95% of artifacts scored 80% or higher in Section VII.

4. No dual enrollment (concurrent) sections of the course were run during summer 2018 so no comparison study between dual enrollment and traditional courses could be completed.

5. In a comparison of online to traditional artifacts, online artifacts score 13.2 lower than traditional. Results were not statistically significantly different.

6. No cross-campus comparison was completed because while three sites offered the course, sample size was limited for one of the sites. As a result, no cross-campus comparison is completed.

7. In a study of score distribution by section, all sections exhibit peaks centered on ≥90%. Sections II, III, and IV exhibit a bimodality. Section II exhibits a secondary peak at < 30%. Section III exhibits a secondary peak at 50-59%. And Section IV exhibits a secondary peak centered on 70-79%.

8. In a longitudinal study of data distribution through summer terms, Summer 2018 exhibits higher mean scores in all sections except Section IV.

5 REFERENCES

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

Davis, J.C. 1973. Statistics and Data Analysis in Geology. John Wiley & Sons, New York, New York, 564 pp.

Hall, J.C. 2010. Doing More with Less: Making Colleges Work Better. Springer, New York, New York, 250 pp.

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.

Starkweather, J. D. 2010. Introduction to Statistics for the Social Sciences. In: Research and Statistical Support. Retrieved from http://www.unt.edu/rss/class/Jon/ISSS_SC/.

Wilkinson, L. 1999. APA Task Force on Statistical Inference. Statistical Methods in Psychology Journals: Guidelines and Explanations. American Psychologist 54 (8), 594–604.