

Foreign Language Assessment Report

Spring 2019

Author: Joseph F. van Gaalen, Ph.D., Asst. VP, IR, 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 these 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 spring 2019 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, IR, Assessment & Effectiveness (jfvangaalen@fsw.edu; x16965).

2 FRENCH

2.1 FRE 1120

During the spring 2019 term, three sections of FRE 1120 were offered. An assessment first piloted in AY 2017-18 was reported in one section of FRE 1120 accounting for 19 artifacts, or 40% of the population. The assessment utilizes a combination quiz and disposition survey. In the assessment, the students are asked for their relative levels of confidence in translating a given topic in conjunction with performing brief translations of the same topic. Each of five disposition survey questions are tethered to two questions in which students are asked to (1) translate French by answering a question written in French using an English response and (2) translate French by answering a question written in French using a French response. The resulting assessment consists of 15 questions comprised of five survey questions and 10 achievement related questions. In form, the assessment measures achievement as well as alignment with student confidence of a given topic. Note that in one instance, the scoring was altered. It is important that scoring is consistent in an assessment such as this, so it may be important to review this area.

2.1.1 Descriptive Statistics, Learning Outcomes, and Objectives

The lead French faculty has identified five areas of interest for evaluation. Given that it the assessment is currently in pilot phase, no achievement objectives have yet been set. The Learning Outcomes (LOs) are:

- LO 1: Reading/Writing Introductions
- LO 2: Reading/Writing Self, Classes & Campus
- LO 3: Reading/Writing Activity Likes & Dislikes
- LO 4: Reading/Writing Home & Rooms
- LO 5: Reading/Writing Descriptions & Family

Each LO consists of two questions. The first, worth one point, requires a translation from French using English responses. The second, worth two points, requires comprehension of French using French responses. Each question can be scored at half-point intervals. The result is a combined maximum score of 15 points. Results exhibit mean scores across the five LOs ranging from 2.9 to 3.0, substantially higher than the 2.3 to 2.7 from fall 2018. It is important to note that each LOs in this study are arranged in order of advancement for the course. LO 1 is encountered early in the term with each LO encountered sequentially to LO 5, which is studied near the term's end.

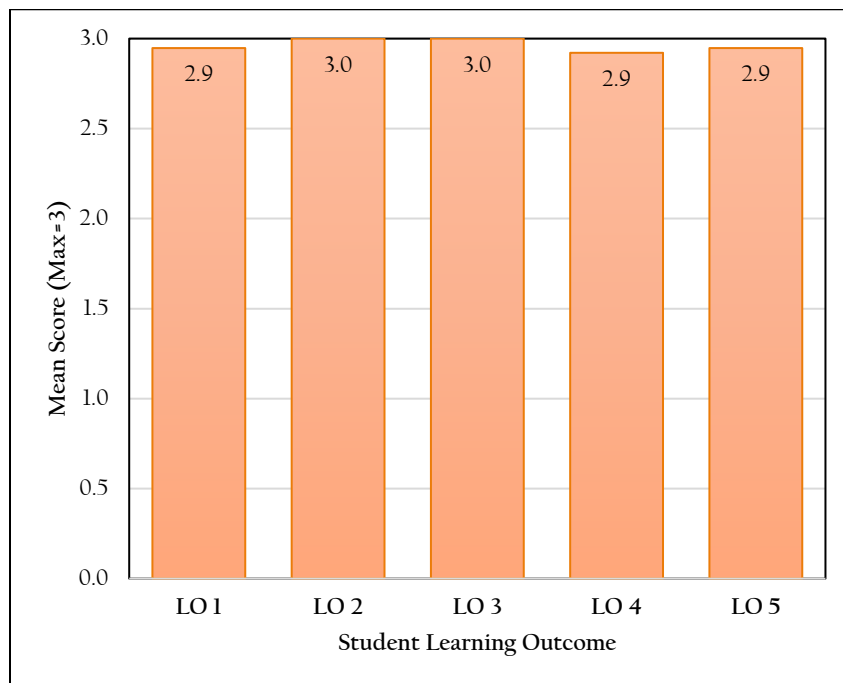


Figure 1. Mean scores by SLOs for FRE 1120.

2.1.2 Exploratory Analysis & Significance Testing

2.1.2.1 Dual Enrollment (Concurrent) to Traditional Comparison

No dual enrollment (concurrent) sections of the course were run during spring 2019 so no comparison study between dual enrollment and traditional courses could be completed.

2.1.2.2 Online to Traditional Comparison

While both online and traditional sections were offered, only data from an online section of the course was collected. As a result, no comparison could be made.

2.1.2.3 Comparison by Campus/Site

While courses were offered at more than one location in spring 2019, data was only collected from one location. As a result, no cross-site comparison could be completed.

2.1.3 Data Distribution & Longitudinal Studies

2.1.3.1 Data Distribution

A distribution of combined (total) scores from the 10 achievement questions is shown below in Figure 2. Scores are centered on 15/15 with a strong negative skew meaning scores are tending strongly towards higher values (Starkweather, 2010). In total, 95% of artifacts score 14/15 or higher.

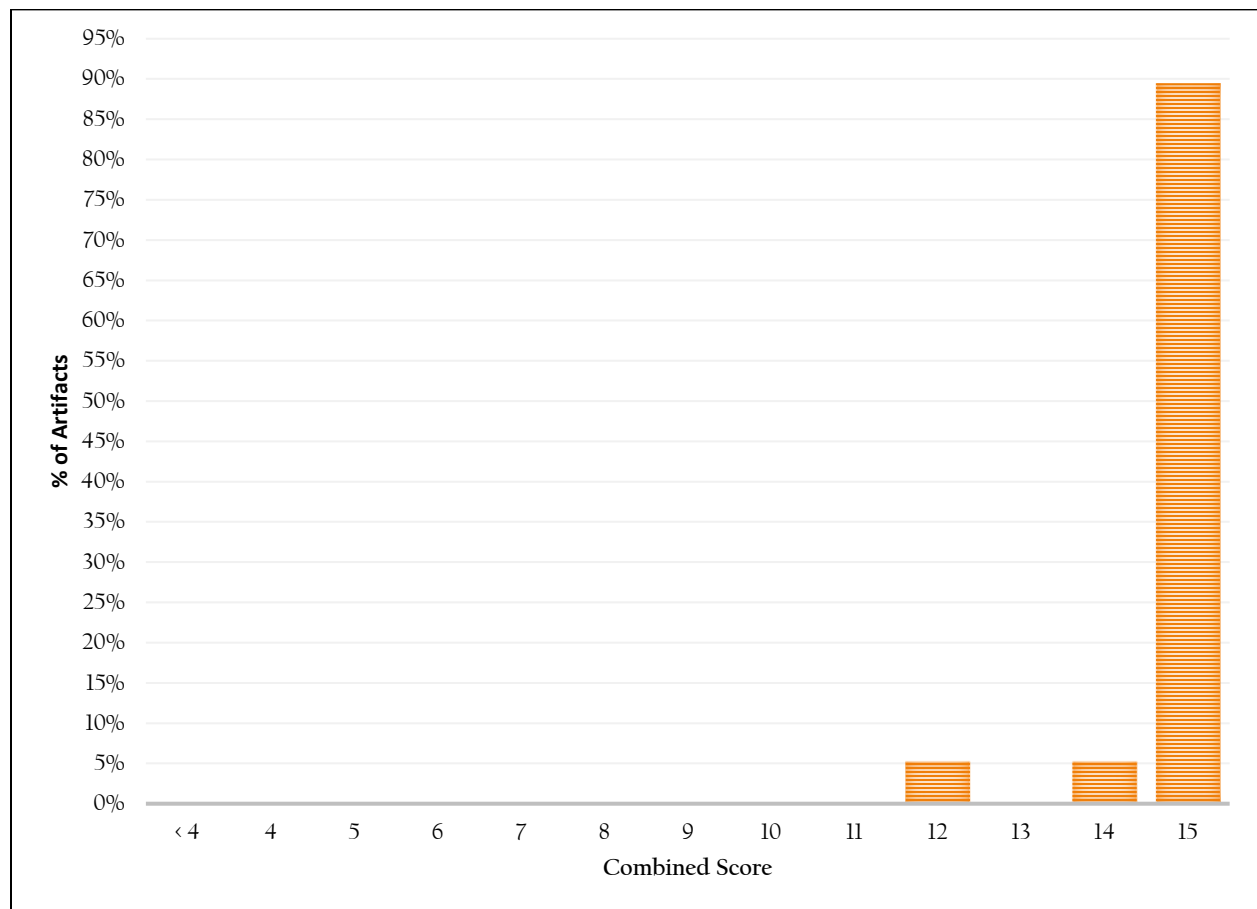


Figure 2. Distribution of combined (total) scores for assessment.

One of the strengths of a combined disposition survey and achievement assessment is the way in which the data can be analyzed. Survey questions that gauge confidence in a topic can now be compared with the results of achievement in that topic to better describe how student confidence varies with student strengths and weaknesses. Figure 3 exhibits the percent of artifacts with correct answers for each question (defined as 1a - translate French by answering a question written in French using an English

response, 1b - translate French by answering a question written in French using a French response, and so on) based on student response to the confidence of that particular topic. For example, LO 1a and 1b are coupled with a survey question regarding confidence in reading and writing introductions. Similarly, LO 2a and 2b are coupled with a survey question regarding confidence in reading and writing in daily activities.

Results for LO1b exhibit no negative responses (“Disagree” or “Strongly Disagree”). Positive responses for LOs 2, 3, 4, and 5 range from 60-70%. LO 5 exhibits the highest negative responses at 33% (Figure 3).

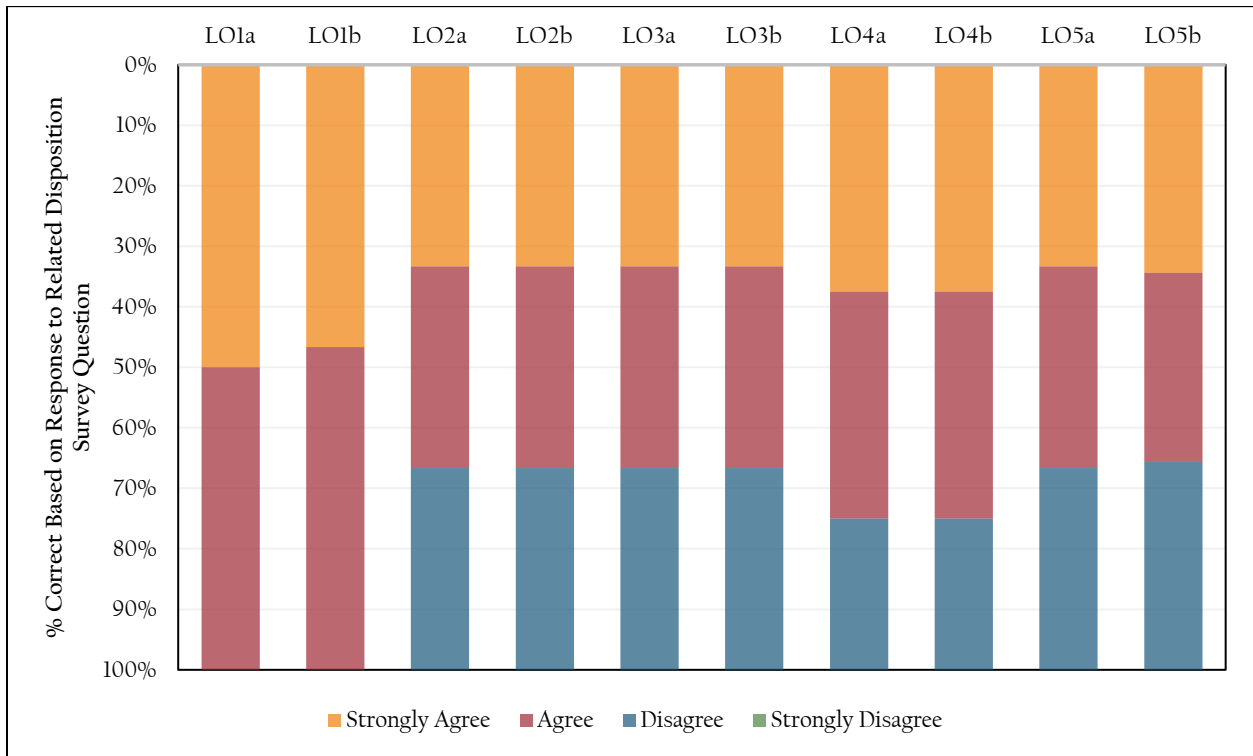


Figure 3. Comparison of % correct based on response to related disposition survey item.

Another way of gauging achievement based on disposition survey responses is through a scatter plot as shown in Figure 4 below. These results yield a clearer sense of just how many students self-report a strong understanding of the topic and yet do not appear to be capable of demonstrating it. In this study, most students scored a perfect score in the assessment, which is abnormal compared to results from previous terms. Still, the lowest score does correlate with a student exhibiting weak confidence in their skillset.

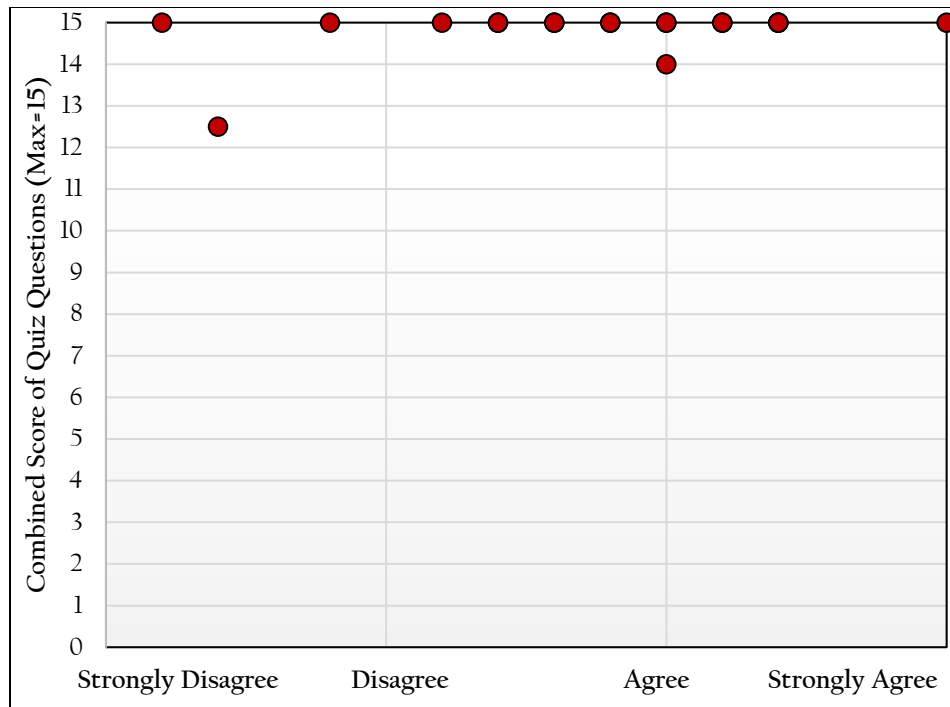


Figure 4. Scatter plot of combined (total) scores based on general response to survey questions.

2.1.3.2 Longitudinal Study

The assessment for FRE 1120 is in its first phase for fall 2018. As further data is collected in coming terms, this section will track achievement through time and highlight strengths, weaknesses and any long-term trends following data collection for fall 2019.

2.2 FRE 1121

In FRE 1121, a 'readiness' mini-test (assessed, but not for a grade in the course) was intended to be administered on the first day of classes (Spring 2019) by the Office of Academic Assessment as a means of exploring the strengths and weaknesses uncovered in the Spring 2018 Foreign Languages Assessment Report. Results would have assisted in determining the preparedness of FRE 1120 students for FRE 1121 as well as coordinating the skills learned by students across course sections, campuses, and instructors. Unfortunately, the readiness mini-test was still in development as of January, 2019, and so the administration process could not occur. The readiness mini-test is now developed and ready for first implementation in fall 2019.

3 SPANISH

3.1 SPN 1120

3.1.1 Descriptive Statistics, Learning Outcomes, and Learning Objectives

During the spring 2019 semester 10 sections of SPN 1120 Beginning Spanish I were offered. Of those, artifacts from a common final were collected from 6 sections. Two sections did not report data.

Another two sections included an altered Section II, and excluded Section III entirely. A total of 208 students were enrolled in SPN 1120. Of those, 80 artifacts were collected representing a sample size of 38% 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 not 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 49% of artifacts scored 70% or higher in Section II and 84% 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 51% of artifacts scored 70% or higher in Section IV and 80% of artifacts scored 70% or higher in Section V. For a graphical representation of SLO achievement, see Figure 5.

n = 80	Section I (Oral)	Section II (Reading)	Section III (Reading)	Section IV (Written)	Section V (Written)	Combined (Overall)
Goal	<i>80% of artifacts scored \geq70% for all sections</i>					
% above 70%	75%	49%	84%	51%	80%	
Mean (as %)	78%	69%	82%	69%	82%	76%
Median (as %)	87%	70%	87%	73%	90%	80%
Section Score Max	15	60	15	15	20	125
Section Mean	11.6	41.3	12.4	10.3	16.3	91.8
Section Median	13	56	15	15	20	96.8

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

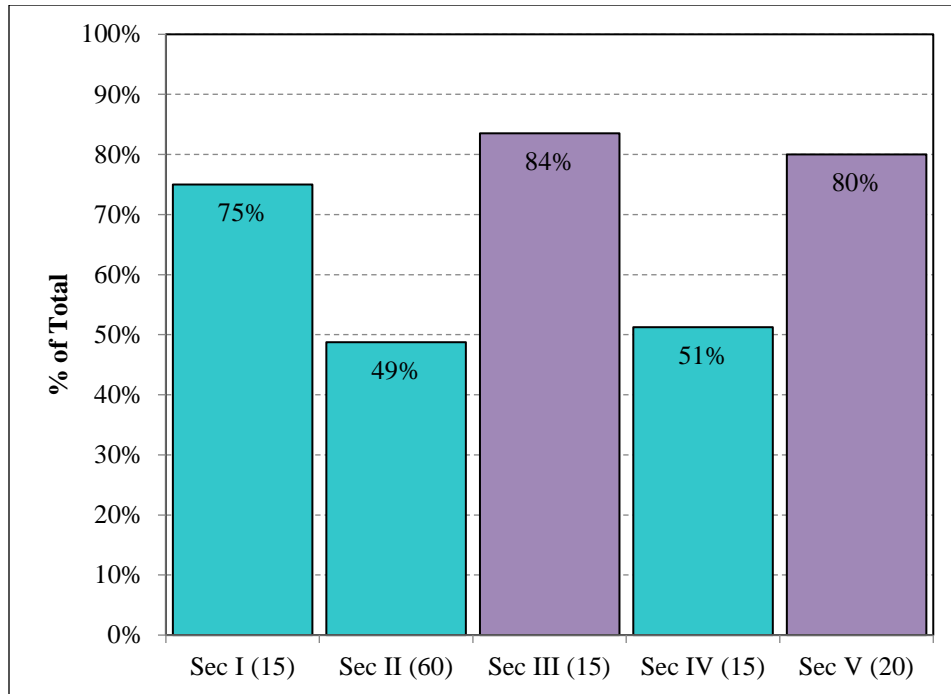


Figure 5. SLO achievement for SPN 1120 by exam section (Percentage of student achievement level as per SLOs). Purple denotes having met objective.

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 spring 2019 so no comparison study between dual enrollment and traditional courses could be completed.

3.1.2.2 Online to Traditional Comparison

During the spring 2019 semester, one online section reported data able to be used in assessment. From those course sections, 14 total online artifacts were collected from SPN 1120 and 66 traditional artifacts were collected from SPN 1120. A comparison of basic statistics is provided in Table 2. Online artifacts mean scores are 0.7 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 small effect size. In other words, non-overlap score distribution

from online artifacts to traditional artifacts is approximately 3%. For a graphical representation of this see Figure 6.

df = 78	
Online mean	74.0
Online standard deviation	23.37
Traditional mean	73.3
Traditional standard deviation	15.49
Effect size	-0.03
p-value	0.914

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

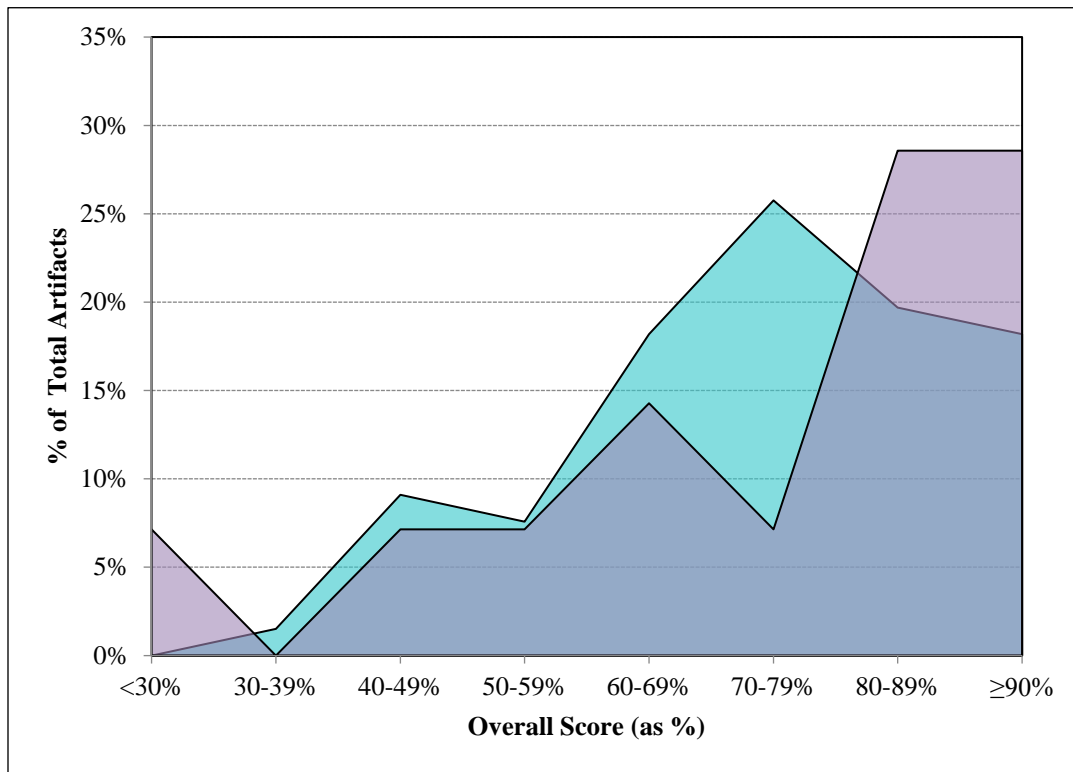


Figure 6. Score distribution for online (purple) and traditional (aqua) artifacts of SPN 1120.

3.1.2.3 Comparison by Campus/Site

Of the 80 artifacts collected from SPN 1120, 8 originated from the Charlotte campus, 0 from the Collier campus, 14 from FSW Online, and 58 from the Thomas Edison (Lee) campus. Scores by rubric dimension varied greatly across campuses. The Thomas Edison (Lee) campus exhibits the highest mean scores in 3 of 5 sections. A comparison of mean scores by rubric dimension is provided in Table 3 and Figure 7.

	Section I	Section II	Section III	Section IV	Section V	Combined Score
<i>Rubric Max</i>	15	60	15	15	20	125
Charlotte	11.9	41.4	12.0	7.8	15.5	88.6
Collier	~	~	~	~	~	~
FSW Online	11.9	42.1	12.3	10.0	16.2	92.5
Thomas Edison (Lee)	11.5	41.1	12.4	10.7	16.4	92.2

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.

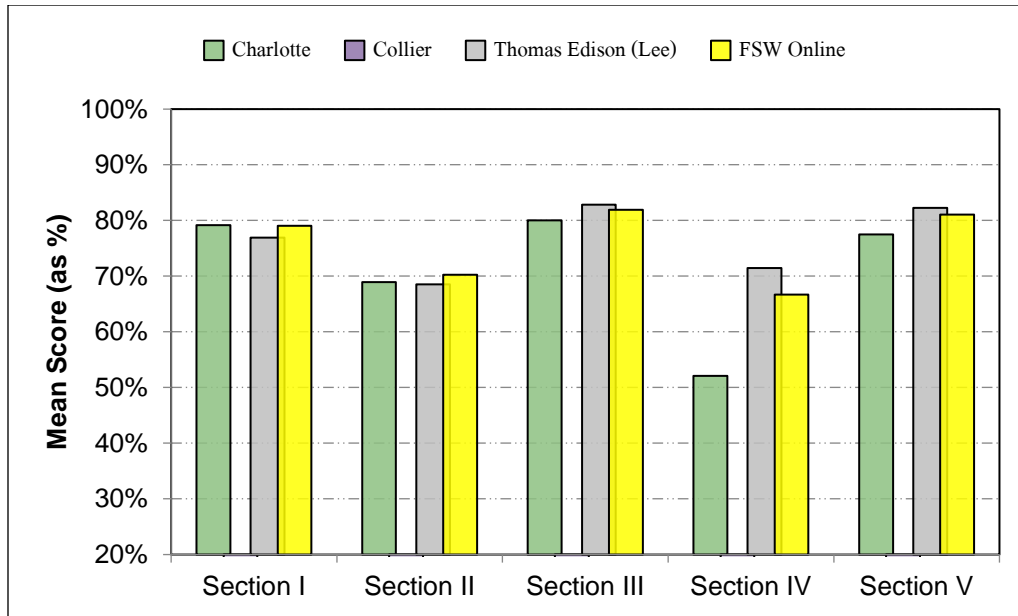


Figure 7. Comparison of mean score of exam sections by site.

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 8. All sections exhibit peaks above 90% except for Section II, which peaks at 60-69%. Section II also continues to exhibit more widely distributed scores. Section II exhibits a maximum of 15% at $\geq 90\%$ while other sections range from 34% to 55%.

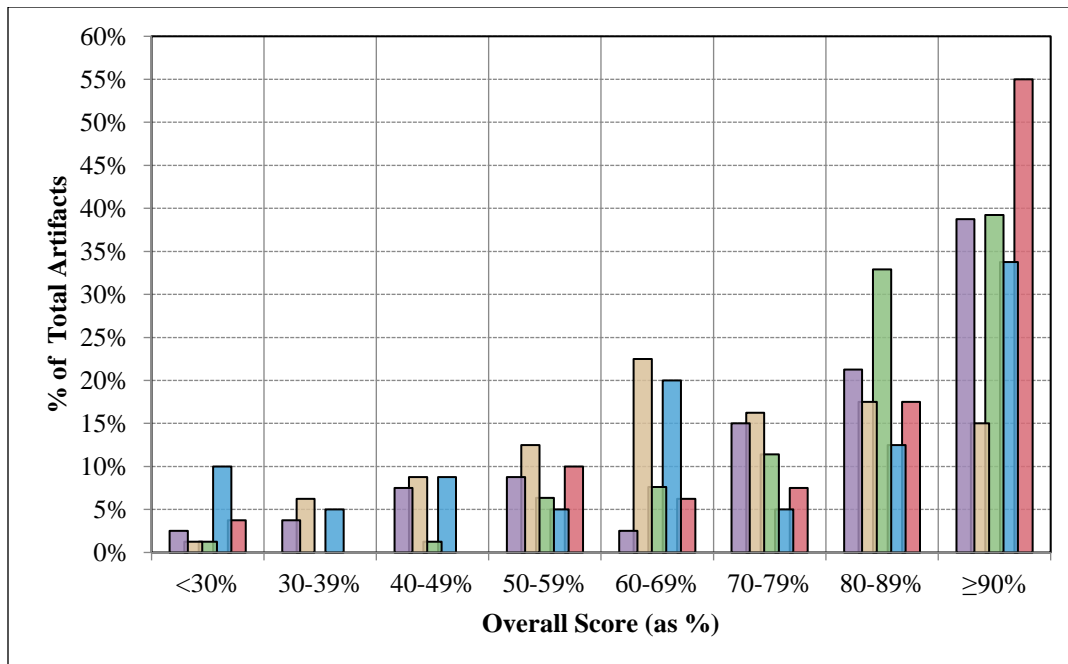


Figure 8. Histogram of SPN 1120 exam sections scores for spring 2019. Purple – Section I, Brown – Section II, Green – Section III, Blue – Section IV, and Red – Section V.

To describe the behavior of the section scores based on overall achievement, a color map, or binary raster image, was created by calculating the mean scores for each exam section as a function of combined score (Figure 9). The color represents the mean section score achieved overall score as shown in the x-axis as a percentage.

A review of the colormap in Figure 9 shows that Section II remains the lowest performing compared to other sections between the ranges of 70-100%. However, in the past this disparity has been fairly substantial. For spring 2019, for example, in the 80-84% range, the mean score for Section II is 73%, while Section IV is only slightly higher at 75%. Both Sections II and IV lag behind the other sections at 70-100% in this term.

	Section I	Section II	Section III	Section IV	Section V	Scale
≥ 95%	98%	89%	99%	98%	98%	98%
90-94%	96%	85%	95%	94%	94%	94%
85-89%	87%	77%	88%	86%	92%	90%
80-84%	83%	73%	85%	75%	92%	80%
75-79%	75%	65%	87%	70%	88%	70%
70-74%	81%	54%	81%	62%	79%	60%
65-69%	75%	63%	73%	45%	81%	50%
60-64%	63%	61%	61%	53%	69%	40%
55-59%	50%	50%	68%	59%	60%	30%
50-54%	41%	40%	73%	39%	64%	
< 50%	45%	46%	50%	14%	36%	

Figure 9. (Top) Colormap of mean scores for each exam section based on overall scoring bin for SPN 1120. An exam section with hotter colors (reds) compared with other sections means section achievement is stronger in that area than others. An exam section with colder colors (blues) compared with other sections means section achievement is weaker in that area than others.

3.1.3.2 Longitudinal Study

Further description of achievement over time in SPN 1120 is provided in Table 4 and Figure 10. 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). In 12 terms, Sections I through V for spring 2019 rank 8th, 4th, 11th, 12th, and 7th, respectively.

Sect.	Max	Fall 2013 n=58	Spring 2014 n=90	Fall 2014 n=93	Spring 2015 n=73	Fall 2015 n=122	Spring 2016 n=141	Fall 2016 n=240	Spring 2017 n=89	Fall 2017 n=266	Spring 2018 n=107	Fall 2018 n=212	Spring 2019 n=80
I	15	12.1	12.4	11.8	11.5	10.8	11.5	11.9	11.9	11.9	12.0	10.9	11.6
II	60	40.4	45.1	40.5	39.8	41.8	38.4	39.8	41.0	41.1	43.7	39.4	41.3
III	15	12.0	13.2	12.8	12.8	12.4	13.3	12.8	13.0	12.9	12.8	12.5	12.4
IV	15	10.6	11.7	11.1	10.9	11.5	10.5	10.9	11.0	11.5	11.6	10.4	10.3
V	20	16.2	16.8	16.5	16.4	15.2	15.7	15.2	16.7	15.4	16.4	15.9	16.3
Total	125	91.3	99.2	92.8	91.4	91.6	89.5	90.6	93.6	92.8	96.6	89.1	91.8

Table 4. Comparison of mean scores for SPN 1120 for fall 2013 through spring 2019.

Because some exam sections have different maximum scores (15, 20, and 60), to see which of the five sections of the exam exhibits the strongest achievement it may be best to relate them in terms of percent. As a percentage, Section III exhibit the highest mean scores in 11 of 12 terms including spring 2019. Section II exhibits the lowest mean scores in 11 of 12 terms, spring 2019 being the only term in which it is not the lowest. For spring 2019, Section IV exhibits the lowest mean score.

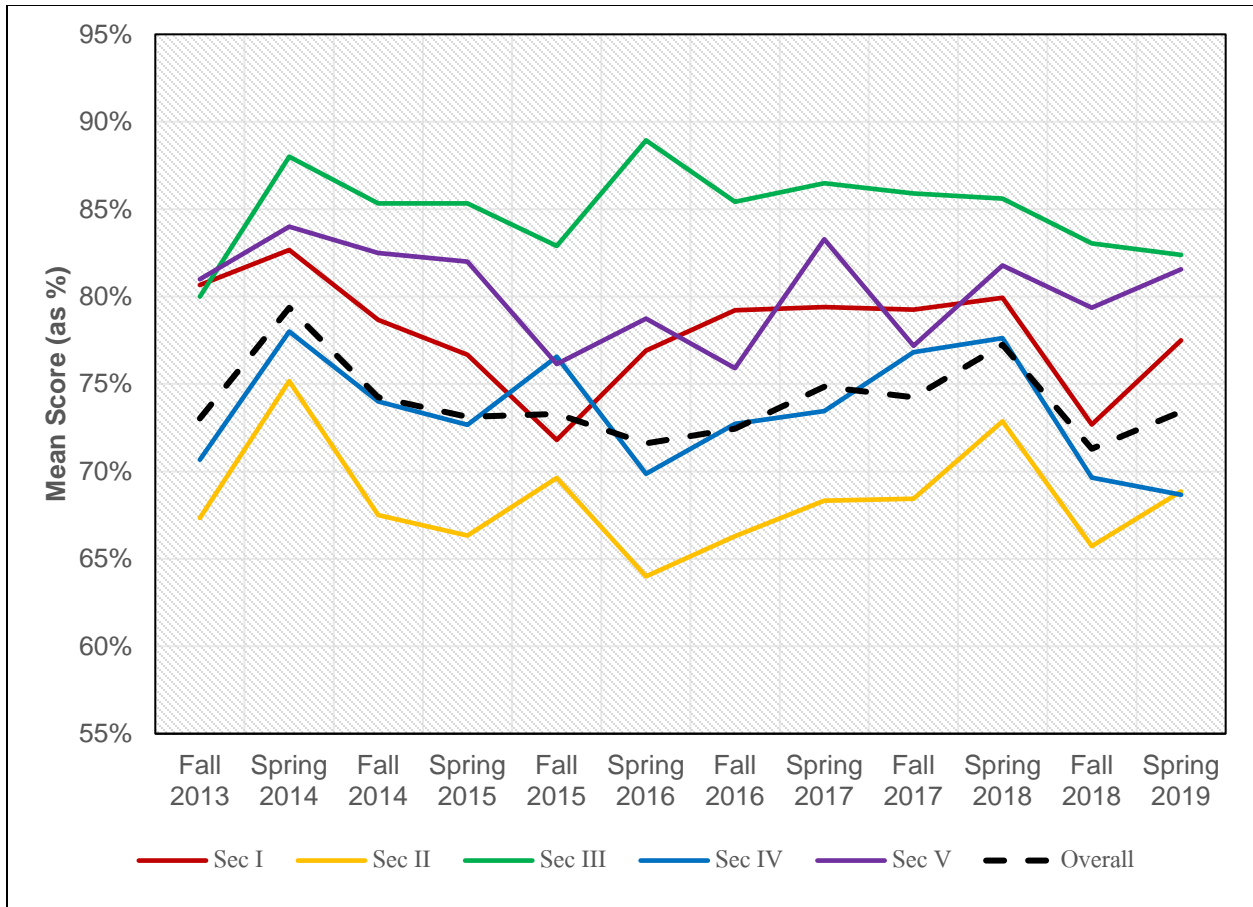


Figure 10. Comparison of mean scores (as percentage) for SPN 1120 through time from fall 2013 through spring 2019.

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 summer 2019, 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 summer 2019 on the Charlotte, Collier, and Thomas Edison campuses. Results of the cumulative study are shown in Figure 11.

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 11). Results exhibit achievement levels based on previous instructor spanning as low as 13.4/30, down from 13.8/30 in spring 2019 but up from 12.5/30 following

fall 2018, to as high as 20.8/30, down from 21.6/30 following spring 2019. This is a substantially decreased range from the earliest studies where the range was 7.2/30 to 25.3/30.

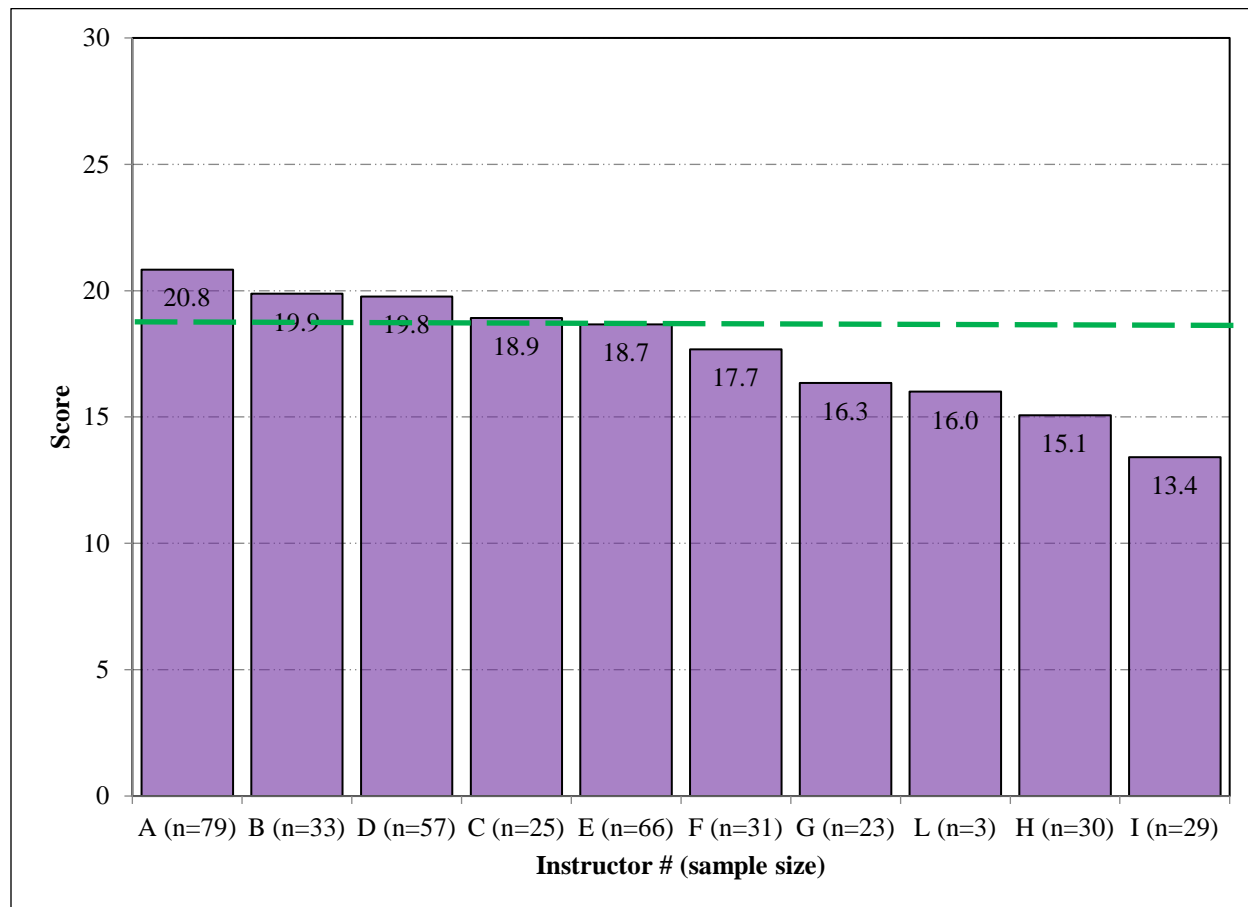


Figure 11. Comparison of achievement in entrance skills study assessment by instructor. Green dashed line denotes spring 2019 mean score for Section II of SPN 1120 (section of exam the entrance skills assessment is based).

3.3 SPN 1121

3.3.1 Descriptive Statistics, Learning Outcomes, and Learning Objectives

Eleven sections of SPN 1121 Beginning Spanish II were offered in spring 2019. Of those, artifacts were collected from a common final from 11 of 11 sections. A total of 216 students were enrolled in SPN 1121. Of those, 185 artifacts were collected representing a sample size of 86%.

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 5 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 72% of artifacts score 70% or higher in the oral competency exam section (Section I) (Table 5, Figure 12). 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 not met. Results exhibit

43% of artifacts scored 70% or higher in Section II and 50% 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 nearly met. Results exhibit 61% of artifacts scored 70% or higher in Section V and 91% of artifacts scored 80% or higher in Section VII.

n = 185	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 \geq 70% for all sections							
% above Goal	72%	43%	74%	66%	61%	50%	91%	
Mean (as %)	78%	61%	77%	74%	74%	65%	86%	75%
Median (as %)	80%	63%	80%	80%	75%	67%	90%	76%
Section Score Max Possible	15	15	40	15	12	15	20	132
Section Mean	11.7	9.2	30.8	11.1	8.9	9.8	17.3	98.7
Section Median	12	9.5	32	12	9	10	18	99.8

Table 5. Percentage of student achievement level as per SLOs (SPN 1121).

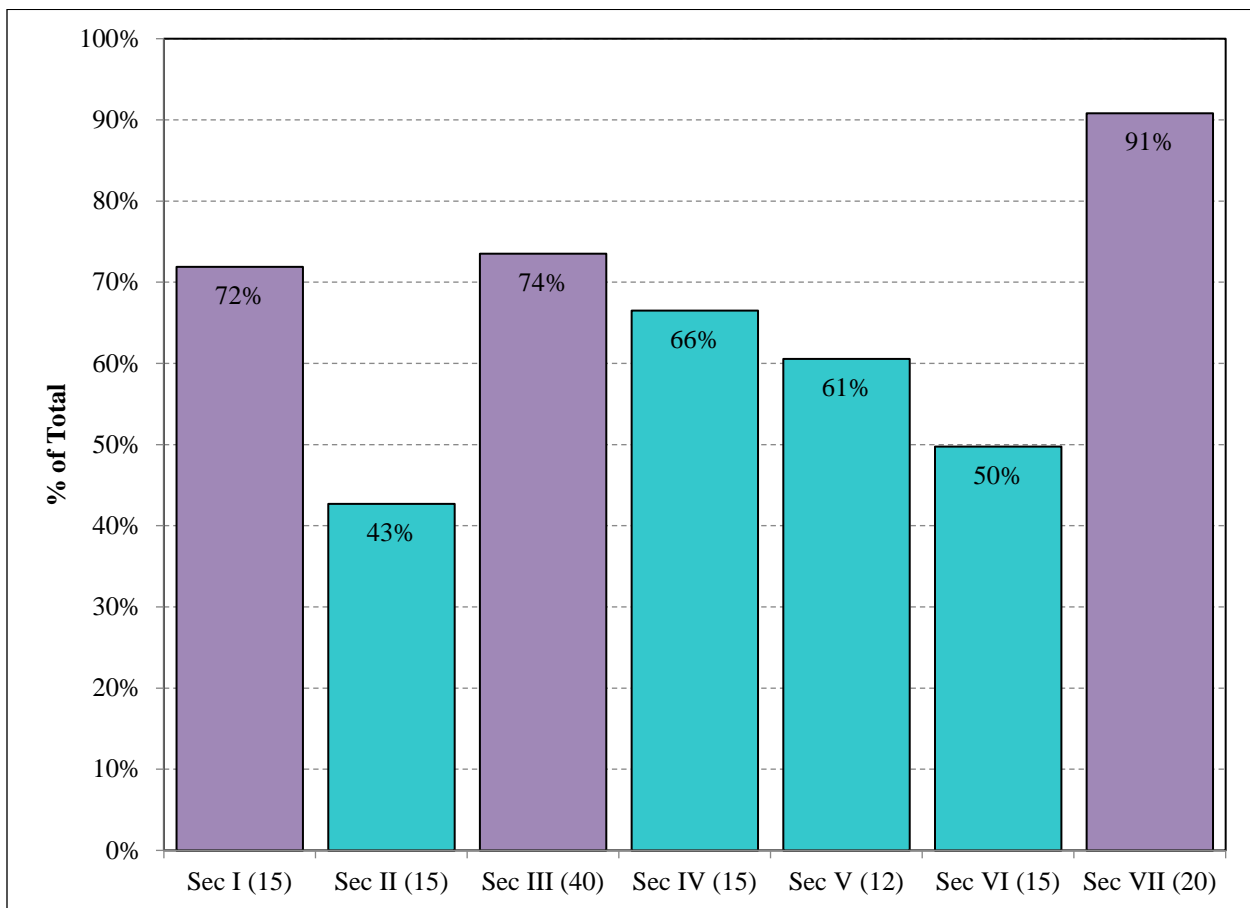


Figure 12. SLO achievement for SPN 1121 by exam section (Percentage of student achievement level as per SLOs). Purple denotes having met objective.

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 spring 2019 so no comparison study between dual enrollment and traditional courses could be completed.

3.3.2.2 Online to Traditional Comparison

During the spring 2019 semester, two online sections reported data. From those course sections, 32 total online artifacts were collected from SPN 1121 and 152 traditional artifacts were collected from SPN 1121. A comparison of basic statistics is provided in Table 6. Online artifacts mean scores are 7.0 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 small effect size. In other words, non-overlap score distribution from online artifacts to traditional artifacts is approximately 21%. For a graphical representation of this see Figure 13.

df = 182		
Online mean		69.0
Online standard deviation		18.95
Traditional mean		76.0
Traditional standard deviation		14.37
Effect size		-0.29
p-value		0.057

Table 6. Comparison of mean scores (as %) for online and traditional artifacts. Positive effect sizes indicate a higher mean score for traditional artifacts.

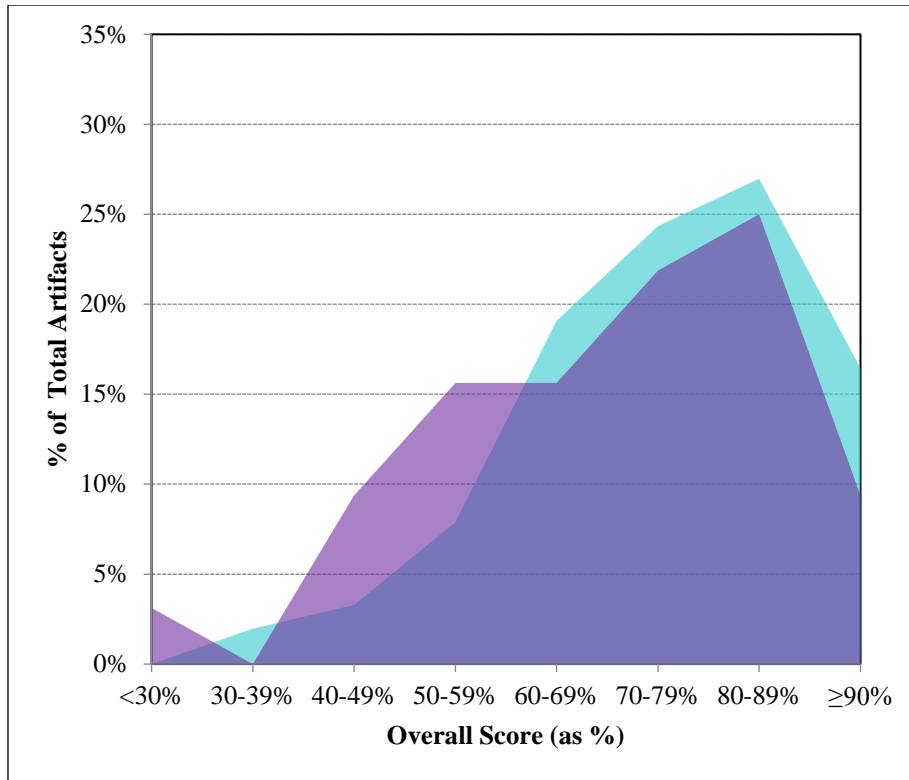


Figure 13. Score distribution for online (purple) and traditional (aqua) artifacts of SPN 1121.

3.3.2.3 Comparison by Campus/Site

Of the 185 artifacts collected from SPN 1120, 22 originated from the Charlotte campus, 30 from the Collier campus, 32 from FSW Online, and 100 from the Thomas Edison (Lee) campus. Score distribution varies greatly by site. The Charlotte campus exhibits scores centered on 70-79% with little spread in the data (relatively speaking). Collier and Thomas Edison exhibit score distributions shifted towards higher values with tails down to lower scores. FSW Online exhibits a similar distribution, but not as sharply peaked at higher scores and more scores in lower scoring bins.

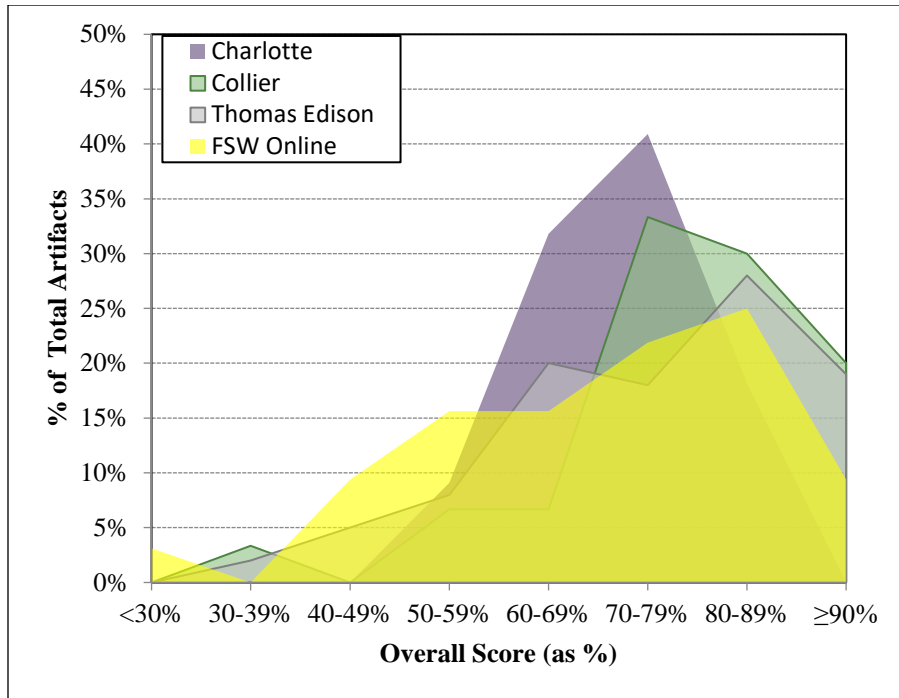


Figure 14. Comparison of score distribution by site.

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 15. Sections I, III, IV, V, VI, and VII exhibit peaks centered on $\geq 90\%$. Section II exhibits two peaks centered at 80-89% and 60-69%.

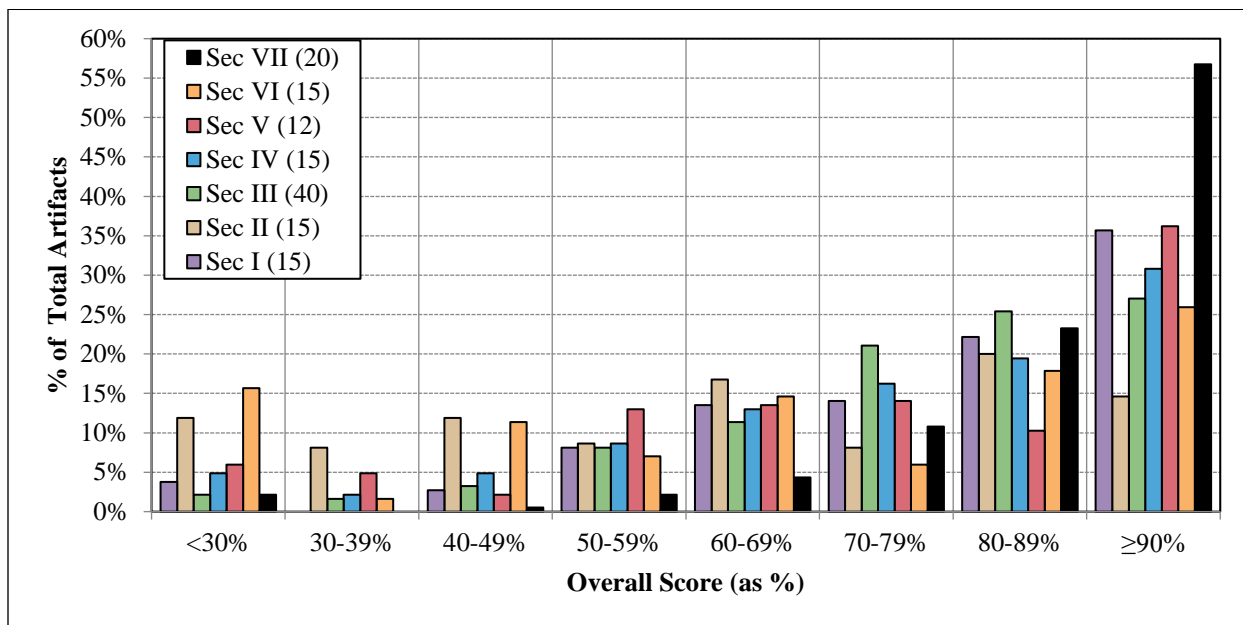


Figure 15. Histogram of SPN 1121 exam sections scores for spring 2019. Purple – Section I, Brown – Section II, Green – Section III, Blue – Section IV, Red – Section V, Orange – Section VI, and Black – Section VII.

To describe the behavior of the section scores based on overall achievement, a color map, or binary raster image, was created by calculating the mean scores for each exam section as a function of combined score (Figure 16). The color represents the mean section score achieved overall score as shown in the x-axis as a percentage.

A review of the colormap in Figure 16 shows several trends. First, Section VII exhibits a strong performance at lower overall scores. For example, in the 55-59% range, Section VII exhibits a mean score of 85%, while other sections range from 26%-66%. Second, Sections II and VI exhibit the weakest performance at mid-range scores. For example, in the range of 75-79%, Sections II and VI exhibit scores of 65% and 60%, while other sections range from 75%-91%. And third, at high scores (85 or higher) all sections score fairly evenly. In the 90-94% range, section mean scores range from 85% to 96%, a range of 11% points. By comparison, in the 75-79% range, section mean scores range from 60-91%, or a range of 31% points.

	Section I	Section II	Section III	Section IV	Section V	Section VI	Section VII	
≥ 95%	100%	95%	98%	98%	97%	99%	99%	Scale
90-94%	93%	85%	90%	92%	94%	93%	96%	100%
85-89%	91%	76%	87%	87%	85%	85%	91%	90%
80-84%	89%	71%	88%	77%	85%	71%	94%	80%
75-79%	83%	65%	81%	79%	75%	60%	91%	70%
70-74%	74%	61%	73%	74%	75%	65%	82%	60%
65-69%	71%	48%	75%	66%	70%	60%	81%	50%
60-64%	68%	47%	60%	65%	63%	43%	88%	40%
55-59%	63%	26%	66%	62%	52%	46%	85%	30%
50-54%	54%	38%	63%	55%	47%	35%	73%	
< 50%	43%	24%	47%	41%	35%	25%	60%	

Figure 16. (Top) Colormap of mean scores for each exam section based on overall scoring bin for SPN 1121. An exam section with hotter colors (reds) compared with other sections means section achievement is stronger in that area than others. An exam section with colder colors (blues) compared with other sections means section achievement is weaker in that area than others.

3.3.3.2 Longitudinal Study

Further description of achievement over time in SPN 1121 is provided in Table 7 and Figure 17). 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). In 12 terms, Sections I through VII for spring 2019 rank 4th, 9th, 4th, 6th, 2nd, 7th, and 1st highest, respectively.

	<i>Section Max</i>	Fall 2013 n=10	Sp 2014 n=115	Fall 2014 n=25	Sp 2015 n=58	Fall 2015 n=17	Sp 2016 n=109	Fall 2016 n=42	Sp 2017 n=140	Fall 2017 n=45	Sp 2018 n=136	Fall 2018 n=67	Sp 2019 n=185
Section I (Oral)	15	11.5	12.3	11.9	12.2	11.5	10.7	11.5	11.8	10.6	12.2	10.4	11.7
Section II (Reading)	15	9.5	9.6	10.7	10.6	10.6	9.5	9.8	8.6	9.1	9.3	8.7	9.2
Section III (Read/Write)	40	34.2	32.3	30.0	31.1	29.9	30.8	25.9	29.2	27.3	29.0	27.9	30.8
Section IV (Read/Write)	15	9.5	11.4	10.6	11.4	11.3	10.5	10.7	10.3	11.8	11.2	10.2	11.1
Section V (Writing)	12	7.5	5.7	9.5	8.5	8.1	8.3	7.2	7.9	8.1	8.4	8.2	8.9
Section VI (Reading)	15	9.6	10.3	11.6	10.7	9.9	9.0	9.7	8.7	10.2	10.1	8.3	9.8
Section VII (Writing)	20	14.2	15.4	16.1	16.4	16.4	14.6	15.6	15.1	15.7	16.7	15.9	17.3
Combined (Overall)	132	96.0	97.0	100.5	100.9	97.7	93.5	90.3	91.6	92.8	96.9	89.6	98.7

Table 7. Comparison of mean scores for SPN 1121 for fall 2013 through spring 2019.

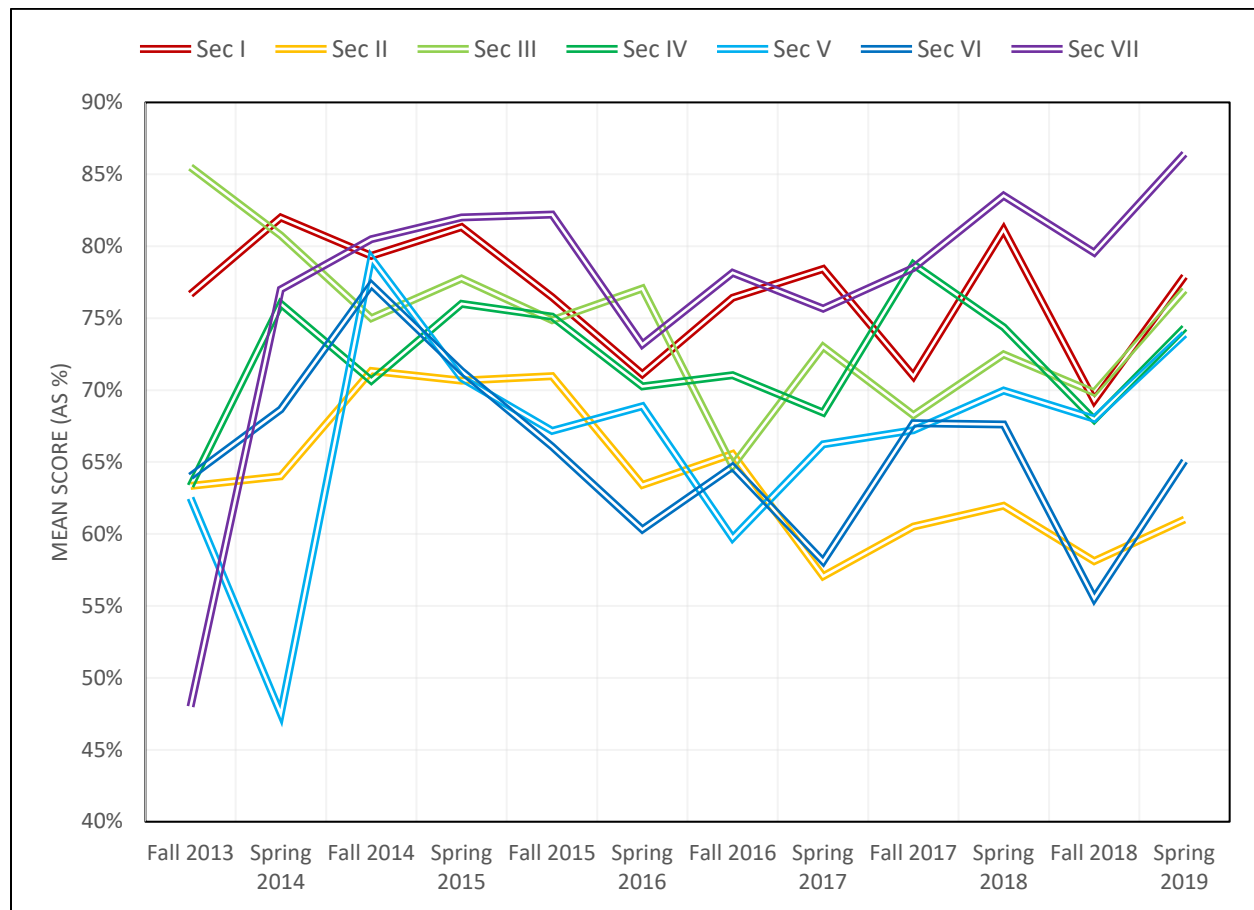


Figure 17. Comparison of mean scores (as percentage) for SPN 1121 through time.

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 spring 2019 assessment.

4.1 FRENCH

A drill-down of FRE 1120 results are as follows:

1. A newly developed assessment first piloted in spring 2019 was used in 1 of 3 sections of FRE 1120 accounting for 19 artifacts, or 40% of the population. The assessment utilizes a combination quiz and disposition survey.
2. In a study of Learning Outcome (LO) achievement, results exhibit mean scores across the five LOs ranging from 2.9 to 3.0, substantially higher than the 2.3 to 2.7 from fall 2018. It is important to note that each LOs in this study are arranged in order of advancement for the course. LO 1 is encountered early in the term with each LO encountered sequentially to LO 5, which is studied near the term's end.
3. No comparison of dual enrollment (concurrent) to traditional artifacts was completed because no dual enrollment sections were offered during spring 2019.
4. While both online and traditional sections were offered, only data from an online section of the course was collected. As a result, no comparison could be made.
5. While courses were offered at more than one location in spring 2019, data was only collected from one location. As a result, no cross-site comparison could be completed.
6. In a study of score distribution by section, scores are centered on 15/15 with a strong negative skew meaning scores are tending strongly towards higher values. In total, 95% of artifacts score 14/15 or higher.
7. In a study of achievement based on disposition survey question response, results for LO1b exhibit no negative responses ("Disagree" or "Strongly Disagree"). Positive responses for LOs 2, 3, 4, and 5 range from 60-70%. LO 5 exhibits the highest negative responses at 33%.
8. In an additional study of achievement based on disposition survey question response, results yield a clearer sense of just how many students self-report a strong understanding of the topic and yet do not appear to be capable of demonstrating it. In this study, most students scored a perfect score in the assessment, which is abnormal compared to results from previous terms. Still, the lowest score does correlate with a student exhibiting weak confidence in their skillset.
9. The assessment for FRE 1120 is in its first phase for fall 2018. As further data is collected in coming terms, this section will track achievement through time and highlight strengths, weaknesses and any long-term trends following data collection for fall 2019.

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

1. In FRE 1121, a 'readiness' mini-test (assessed, but not for a grade in the course) was intended to be administered on the first day of classes (Spring 2019) by the Office of Academic Assessment as a means of exploring the strengths and weaknesses uncovered in the Spring 2018 Foreign Languages Assessment Report. Results would have assisted in determining the preparedness of FRE 1120 students for FRE 1121 as well as coordinating the skills learned by students across course sections, campuses, and instructors. Unfortunately, the readiness mini-test was still in development as of January, 2019, and so the administration process could not occur. The readiness mini-test is now developed and ready for first implementation in fall 2019.

4.2 SPANISH

A drill-down of SPN 1120 results are as follows:

1. During the fall 2018 semester 10 sections of SPN 1120 Beginning Spanish I were offered. Of those, artifacts from a common final were collected from 6 sections. Two sections did not report data. Another two sections included an altered Section II, and excluded Section III entirely. A total of 208 students were enrolled in SPN 1120. Of those, 80 artifacts were collected representing a sample size of 38% of the population.
2. Achievement of 80% of artifacts scoring 70% or better (SLO 1): Achievement was not met as results exhibit 75% of artifacts score 70% or higher in the oral competency exam section (Section I).
3. Achievement of 80% of artifacts scoring 70% or better (SLO 2): Achievement was partially met. Results exhibit 49% of artifacts scored 70% or higher in Section II and 84% of artifacts scored 70% or higher in Section III.
4. Achievement of 80% of artifacts scoring 70% or better (SLO 3): Achievement was partially met. Results exhibit 51% of artifacts scored 70% or higher in Section IV and 80% of artifacts scored 70% or higher in Section V.
5. No dual enrollment (concurrent) sections of the course were run during spring 2019 so no comparison study between dual enrollment and traditional courses could be completed.
6. In a comparison of online to traditional artifacts, online artifacts scores are 0.7 higher than traditional. Results were not statistically significantly different.
7. In a cross-campus comparison study, the Thomas Edison (Lee) campus exhibits the highest mean scores in 3 of 5 sections.
8. In a study of score distribution by section, all sections exhibit peaks above 90% except for Section II, which peaks at 60-69%. Section II also continues to exhibit more widely distributed scores. Section II exhibits a maximum of 15% at $\geq 90\%$ while other sections range from 34% to 55%.
9. In a study of section score distribution based on overall score, Section II remains the lowest performing compared to other sections between the ranges of 70-100%. However, in the past this disparity has been fairly substantial. For spring 2019, for example, in the 80-84% range, the mean score for Section II is 73%, while Section IV is only slightly higher at 75%. Both Sections II and IV lag behind the other sections at 70-100% in this term.
10. In a longitudinal study of data distribution through time, in 12 terms, Sections I through V for spring 2019 rank 8th, 4th, 11th, 12th, and 7th, respectively.
11. 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 13.4/30, down from 13.8/30 in spring 2019 but up from 12.5/30 following fall 2018, to as high as 20.8/30, down from 21.6/30 following spring 2019. 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. Eleven sections of SPN 1121 Beginning Spanish II were offered in spring 2019. Of those, artifacts were collected from a common final from 11 of 11 sections. A total of 216 students were enrolled in SPN 1121. Of those, 185 artifacts were collected representing a sample size of 86%.
2. Achievement of 80% of artifacts scoring 70% or better (SLO 1): Achievement not nearly met as results exhibit 72% of artifacts score 70% or higher in the oral competency exam section (Section I).

3. Achievement of 80% of artifacts scoring 70% or better (SLO 2): Achievement was not met. Results exhibit 43% of artifacts scored 70% or higher in Section II and 50% of artifacts scored 70% or higher in Section VI.
4. Achievement of 80% of artifacts scoring 70% or better (SLO 3): Achievement was nearly met. Results exhibit 61% of artifacts scored 70% or higher in Section V and 91% of artifacts scored 80% or higher in Section VII.
5. No dual enrollment (concurrent) sections of the course were run during spring 2019 so no comparison study between dual enrollment and traditional courses could be completed.
6. In a comparison of online to traditional artifacts, online artifacts score 7.0 lower than traditional. Results were not statistically significantly different.
7. In a cross-campus comparison, score distribution varies greatly by site. The Charlotte campus exhibits scores centered on 70-79% with little spread in the data (relatively speaking). Collier and Thomas Edison exhibit score distributions shifted towards higher values with tails down to lower scores. FSW Online exhibits a similar distribution, but not as sharply peaked at higher scores and more scores in lower scoring bins.
8. In a study of score distribution by section, Sections I, III, IV, V, VI, and VII exhibit peaks centered on $\geq 90\%$. Section II exhibits two peaks centered at 80-89% and 60-69%.
9. In a study of section score distribution based on overall score, several trends are present. First, Section VII exhibits a strong performance at lower overall scores. For example, in the 55-59% range, Section VII exhibits a mean score of 85%, while other sections range from 26%-66%. Second, Sections II and VI exhibit the weakest performance at mid-range scores. For example, in the range of 75-79%, Sections II and VI exhibit scores of 65% and 60%, while other sections range from 75%-91%. And third, at high scores (85 or higher) all sections score fairly evenly. In the 90-94% range, section mean scores range from 85% to 96%, a range of 11% points. By comparison, in the 75-79% range, section mean scores range from 60-91%, or a range of 31% points.
10. In a longitudinal study of data distribution through time, in 12 terms, Sections I through VII for spring 2019 rank 4th, 9th, 4th, 6th, 2nd, 7th, and 1st highest, respectively.

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.
- 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.