Foreign Language Assessment Report
Fall 2020
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 modality comparisons, and comparisons by site, where applicable. This report provides achievement analysis for both fall 2020 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 Fall 2020 term, five sections of FRE 1120 were offered. An assessment first piloted in AY 2017-18 was recorded in all five section of FRE 1120 accounting for 49 artifacts, or 71% 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 several instances, 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. No achievement metric goals 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 2.0 to 2.3 (Figure 1), compared with 2.7-3.0 of spring 2020, 2.2-2.6 in fall 2019, 2.3 to 2.7 in 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.
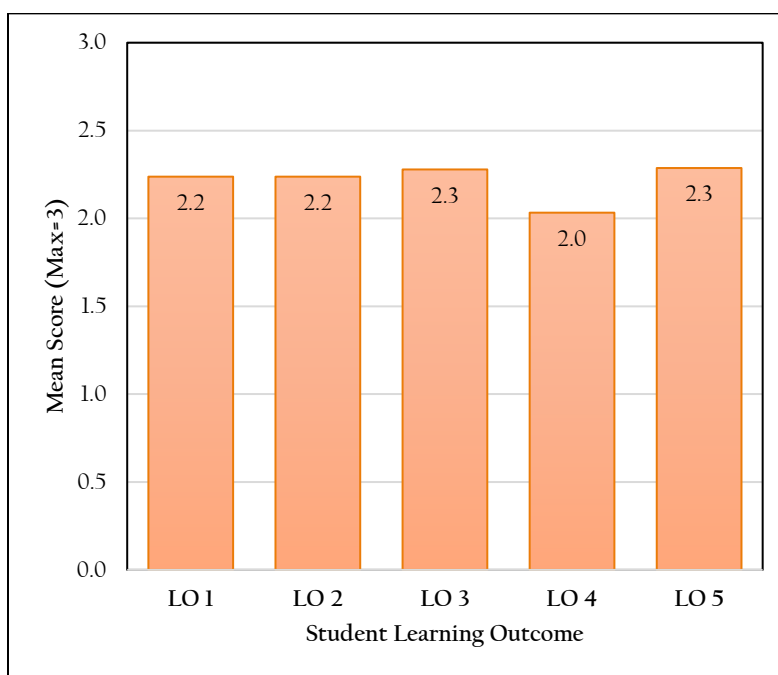

Figure 1. Mean scores by SLOs for FRE 1120.

2.1.2 Exploratory Analysis & Significance Testing

2.1.2.1 Dual Enrollment (Concurrent) to Traditional Comparison
During the Fall 2020 term, one dual enrollment (concurrent) section was offered. Overall mean scores (max = 15) are higher for concurrent dual enrollment (14.4/15) when compared with traditional course sections (10.6/15). No statistical significance study was completed due to limited sample size for dual enrollment (n=7).

2.1.2.2 Modality Comparison
During the Fall 2020 semester, 23 total asynchronous online artifacts were collected, along with 8 live online, 11 flex, 0 blended, and 0 traditional. A comparison of basic statistics is provided in Table 1. Asynchronous online mean scores are the highest of all modalities at 14.7/15. The lowest is Flex, at

12.3/15. A one-way analysis of variance was used to compare means by modality. Results of the ANOVA exhibit a statistically significant difference between sites (see Table 2). Therefore, we can reject the null hypothesis that the mean scores at each site are equal to each other and we can conclude with a 95% confidence that the differences in scores are not solely due to chance.

	Traditional	*Async Online*	*Live Online*	*Flex*	*Blended*
n	~	23	8	11	~
Max	~	21	15	15	~
Min	~	9	7	8	~
Mean	~	14.7	12.8	12.3	~
Standard deviation	~	2.25	2.87	2.18	~
Skewness	~	0.03	-1.32	-1.02	~
Kurtosis	~	4.13	1.17	0.23	~

Table 1. Comparison of mean scores for online and traditional artifacts. Positive effect sizes indicate a higher mean score for traditional artifacts.

Source of Variation	Sum of squared differences	df	Mean Squares	F_{obs}	p-value	F_{crit}
Between Sites	47.4	2	23.7	4.16	0.023	3.24
Within Sites	216.2	38	5.7			
Total	263.6	40				

Table 2. Results of one-way ANOVA of mean scores at each site for FRE 1120.

2.1.2.3   Comparison by Campus/Site

Only one site (Lee campus) had course sections run during fall 2020. All other sections were some form of online or dual enrollment so no cross-campus 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, 65% of artifacts score 14/15 or higher.
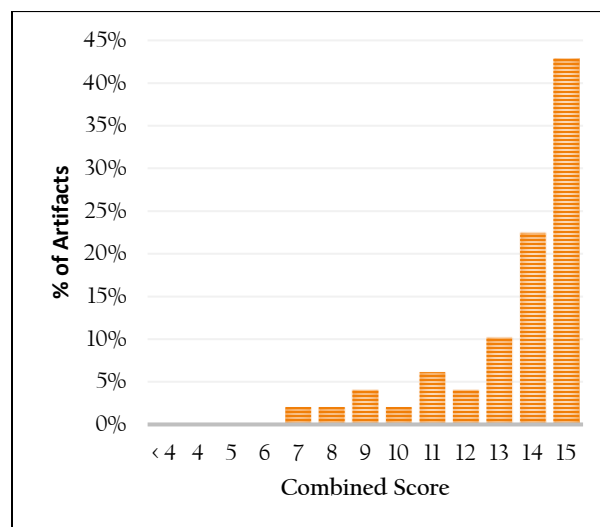

Figure 2. Distribution of combined (total) scores for assessment.

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. This gauge is reflected in the scatterplot shown in Figure 3 below. These results indicate a weak-to-moderate sense of students self-reporting a strong understanding of the topic that correlates with actual achievement. There are only a few which report a low understanding and yet score high (as seen in data point in upper left corner of graph).

Figure 3. Scatter plot of combined (total) scores based on general response to survey questions.

2.1.3.2 Longitudinal Study

Further description of achievement over time in FRE 1120 is provided in Table 3. 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 five terms, LOs 1 through 5 for the Fall 2020 rank 5th, 5th, 5th, 5th, and 5th, respectively.

	LO 1	LO 2	LO 3	LO 4	LO 5	Overall
Fall 2018 (n=26)	2.6	2.4	2.7	2.3	2.7	12.7
Spring 2019 (n=19)	2.9	3.0	3.0	2.9	2.9	14.7
Fall 2019 (n=55)	2.6	2.5	2.6	2.2	2.5	12.4
Spring 2020 (n=35)	2.9	2.9	2.9	2.7	3.0	14.4
Fall 2020 (n=49)	2.2	2.2	2.3	2.0	2.3	11.0

Table 3. Comparison of mean scores for FRE 1120 over time.

2.2 FRE 1121

During the Fall 2020 term, two sections of FRE 1121 were offered. The assessment was recorded in both, however, in one section results and scoring did not match that expected from the assessment plan, so no data could be utilized. The data collection accounts for 6 artifacts, or 25% 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 four 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 several instances, 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.

2.2.1 Descriptive Statistics, Learning Outcomes, and Objectives

In FRE 1121, a similar assessment seen in FRE 1120 is utilized. The lead French faculty has identified five areas of interest for evaluation. No achievement metric goals have yet been set. The Learning Outcomes (LOs) are:

- ➢ LO 1: Reading/Writing Introductions
- ➢ LO 2: Reading/Writing Completed Actions in the Past
- ➢ LO 3: Reading/Writing Repeated and Habitual Actions in the Past
- ➢ LO 4: Reading/Writing Hypothetical Situations
- ➢ LO 5: Future Events and Projects

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 1.8 to 2.8. 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.
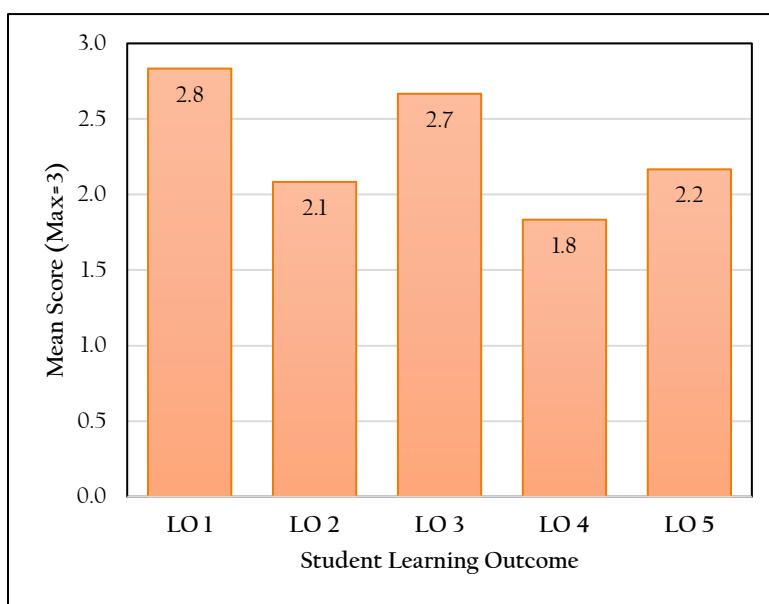

Figure 4. Mean scores by SLOs for FRE 1121.

2.2.2   Exploratory Analysis & Significance Testing

2.2.2.1   Dual Enrollment (Concurrent) to Traditional Comparison
There was no dual enrollment (concurrent) section run during fall 2020, so no comparison study between dual enrollment and traditional courses could be completed.

2.2.2.2   Modality Comparison
While more than one modality was offered in this term, data was only able to be collected from one section, so no modality comparison could be completed.

2.2.2.3   Comparison by Campus/Site
Only one site offered sections other than online sections, so no comparison by site could be completed.

2.2.3   Data Distribution & Longitudinal Studies

2.2.3.1   Data Distribution
A distribution of combined (total) scores from the 15 achievement questions is shown below in Figure 5. Scores are centered on 12/15 with a strong negative skew meaning scores are tending strongly towards higher values (Starkweather, 2010). In total, 50% of artifacts score 12/15.
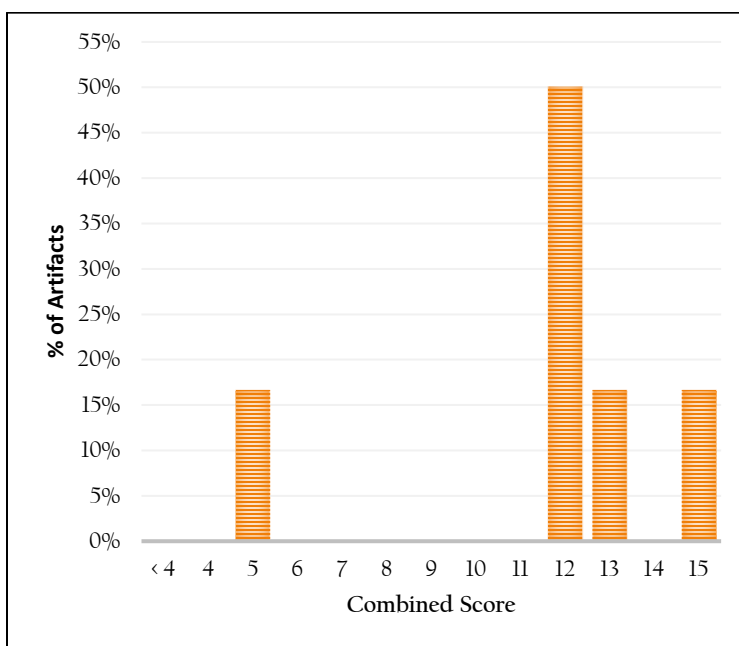

Figure 5. Distribution of combined (total) scores for assessment.

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. This gauge is reflected in the scatterplot shown in Figure 6 below. These results yield a clearer sense of just how many students self-report a strong understanding of the topic and if they appear to be capable of demonstrating it. There appears to be the beginnings of a trend of confidence with achievement, although sample size is limited (n=6).
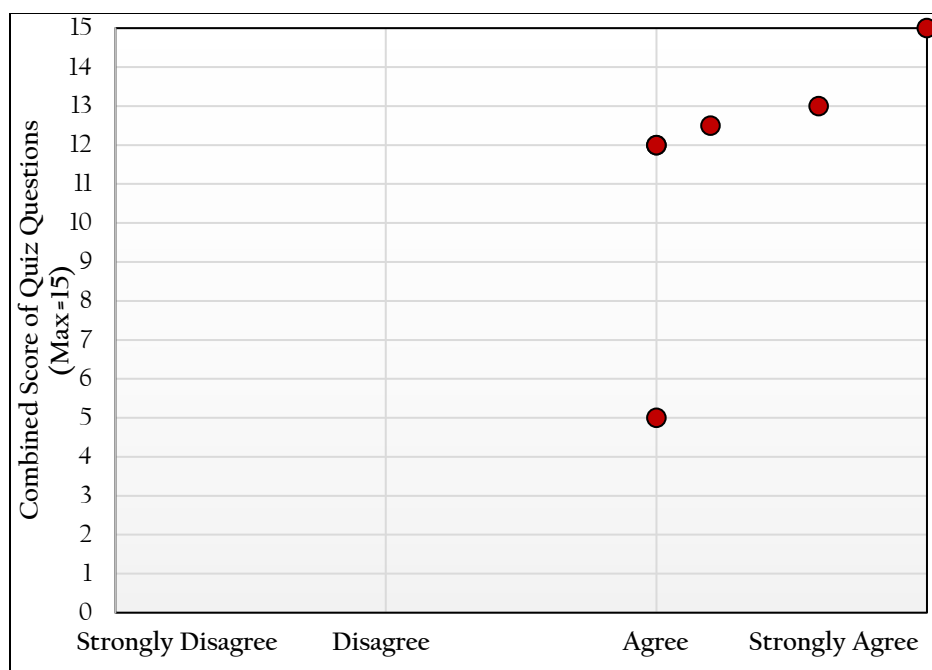
Figure 6. Scatter plot of combined (total) scores based on general response to survey questions.

2.2.3.2   Longitudinal Study

Further description of achievement over time in FRE 1120 is provided in Table 4. 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 three terms, LOs 1 through 5 for the Fall 2020 rank 2nd, 3rd, 2nd, 3rd, and 2nd, respectively.

	LO 1	LO 2	LO 3	LO 4	LO 5	Overall
Fall 2019 (n=14)	2.9	2.7	2.9	2.8	~	~
Spring 2020 (n=17)	2.4	2.5	2.1	2.2	2.4	11.6
Fall 2020 (n=6)	2.8	2.1	2.7	1.8	2.2	11.6

Table 4. Comparison of mean scores for FRE 1121 over time.

3   SPANISH

3.1   SPN 1120

3.1.1   Descriptive Statistics, Learning Outcomes, and Learning Objectives

During the Fall 2020 semester 13 sections of SPN 1120 *Beginning Spanish I* were offered. Of those, artifacts from a common final were collected from 13 sections. A total of 278 students were enrolled in SPN 1120. Of those, 197 artifacts were collected representing a sample size of 71% 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 71% of artifacts score 70% or higher in the oral competency exam section (Section I) (Table 5). 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 43% of artifacts scored 70% or higher in Section II and 81% 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 nearly met. Results exhibit 50% of artifacts scored 70% or higher in Section IV and 79% of artifacts scored 70% or higher in Section V. For a graphical representation of SLO achievement, see Figure 7.

n = 197	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%	71%	43%	81%	50%	79%	
Mean (as %)	75%	62%	83%	67%	79%	
Median (as %)	83%	63%	90%	67%	85%	
Section Score Max	15	60	15	15	20	125
Section Mean	11.2	36.9	12.4	10.1	15.7	86.4
Section Median	13	38	13.5	10	17	89.5

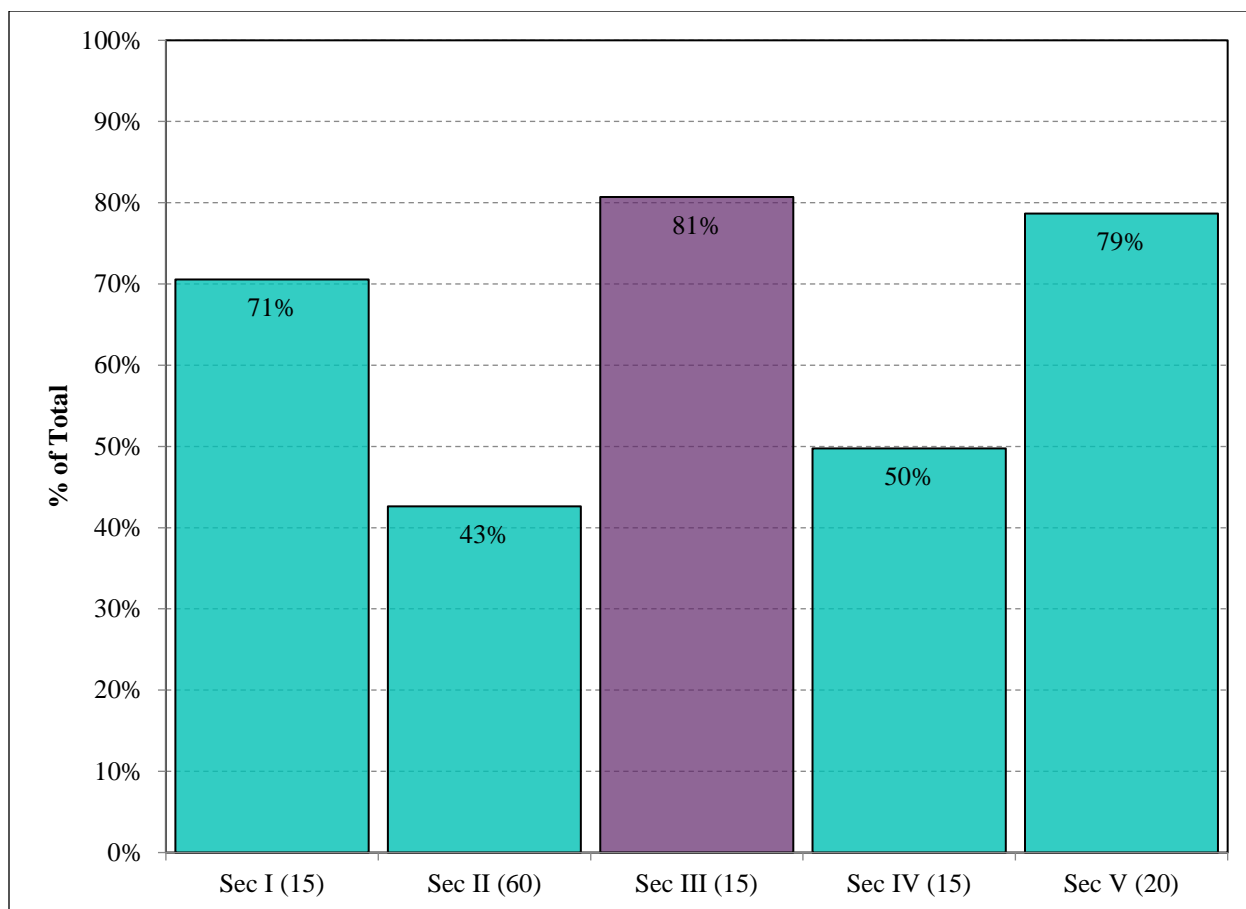Table 5. Percentage of student achievement level as per SLOs (SPN 1120).

Figure 7. 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 fall 2020 so no comparison study between dual enrollment and traditional courses could be completed.

3.1.2.2   Modality Comparison

During the Fall 2020 semester, 49 total asynchronous online artifacts were collected, along with 78 live online, 45 flex, 0 blended, and 25 traditional. A comparison of basic statistics is provided in Table 6. Asynchronous online artifacts mean scores are 4.5%-points higher than the next highest, Flex, which is 71.7%. Mean scores for traditional artifacts are similar to Flex at 71.6%. Finally, nearly 10%-points lower still, is Live Online at 62.3%. Results of the ANOVA exhibit a statistically significant difference between sites (see Table 7). Therefore, we can reject the null hypothesis that the mean scores at each site are equal to each other and we can conclude with a 95% confidence that the differences in scores are not solely due to chance. A distribution of scores by modality is shown in Figure 8.

	Traditional	*Async Online*	*Live Online*	*Flex*	*Blended*
n	25	49	78	45	0
Mean Score	89.6	95.3	77.9	89.6	~
Mean Score (as %)	71.6%	76.2%	62.3%	71.7%	~

Table 6. Comparison of descriptive statistics by modality.

Source of Variation	Sum of squared differences	df	Mean Squares	F_{obs}	p-value	F_{crit}
Between Sites	6529.1	3	2176.3	6.17	0.0005	2.65
Within Sites	68,039.7	193	352.5			
Total	74,568.8	196				

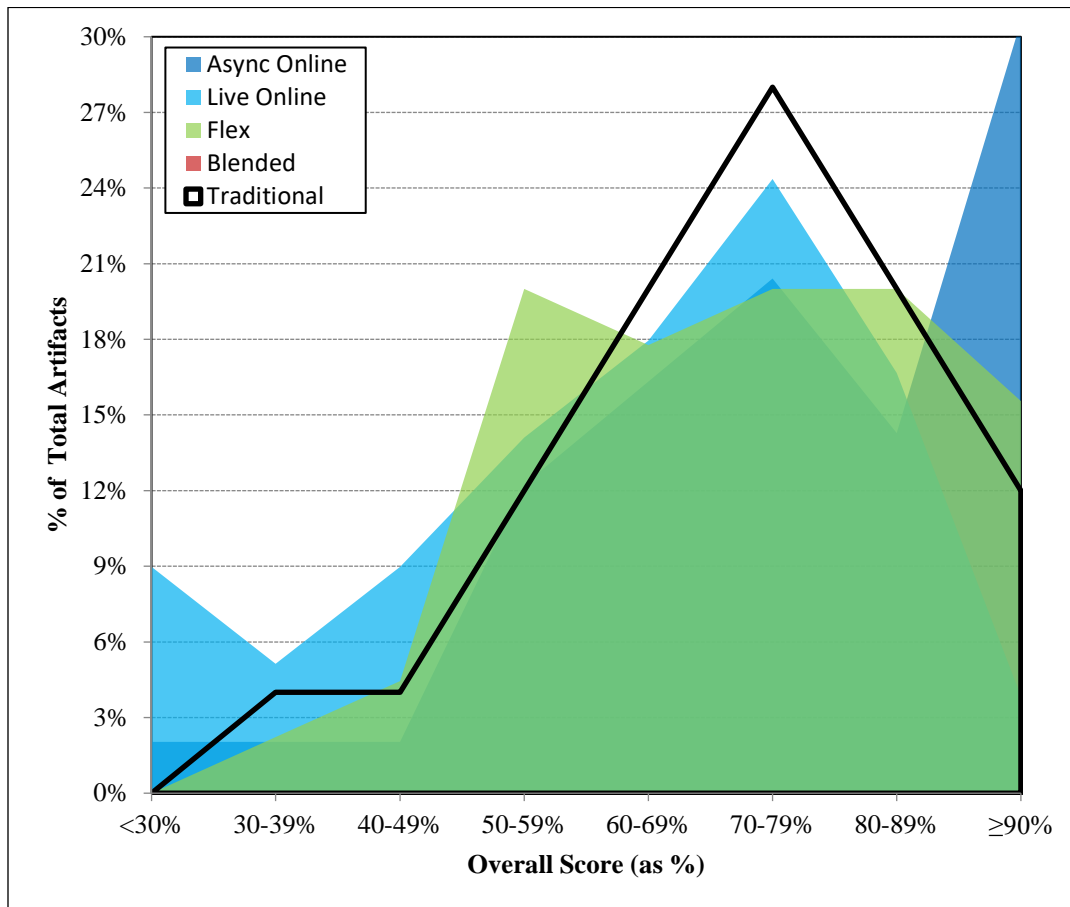*Table 7. Results of one-way ANOVA of mean scores in each modality for SPN 1120.*

Figure 8. Score distribution by modality for SPN 1120.

3.1.2.3   Comparison by Campus/Site

Of the 197 artifacts collected from SPN 1120, 0 originated from the Charlotte campus, 25 from the Collier campus, and 42 from the Thomas Edison (Lee) campus. Scores by rubric dimension varied greatly across campuses. Collier exhibits the highest mean score in 3 of 5 sections, while the Thomas Edison (Lee) campus exhibits the highest in the reaming 2/5. A comparison of mean scores by rubric dimension is provided in Table 8 and Figure 9.

	Section I	Section II	Section III	Section IV	Section V	Combined Score
Rubric Max	*15*	*60*	*15*	*15*	*20*	*125*
Charlotte	~	~	~	~	~	~
Collier	10.2	**40.1**	12.5	**11.4**	**15.4**	89.6
Hendry Glades	~	~	~	~	~	~
Thomas Edison (Lee)	**12.1**	38.5	**12.8**	10.8	15.3	89.6

Table 8. Comparison of mean scores by site for SPN 1120. Bold denotes highest mean score in that dimension among all sites.
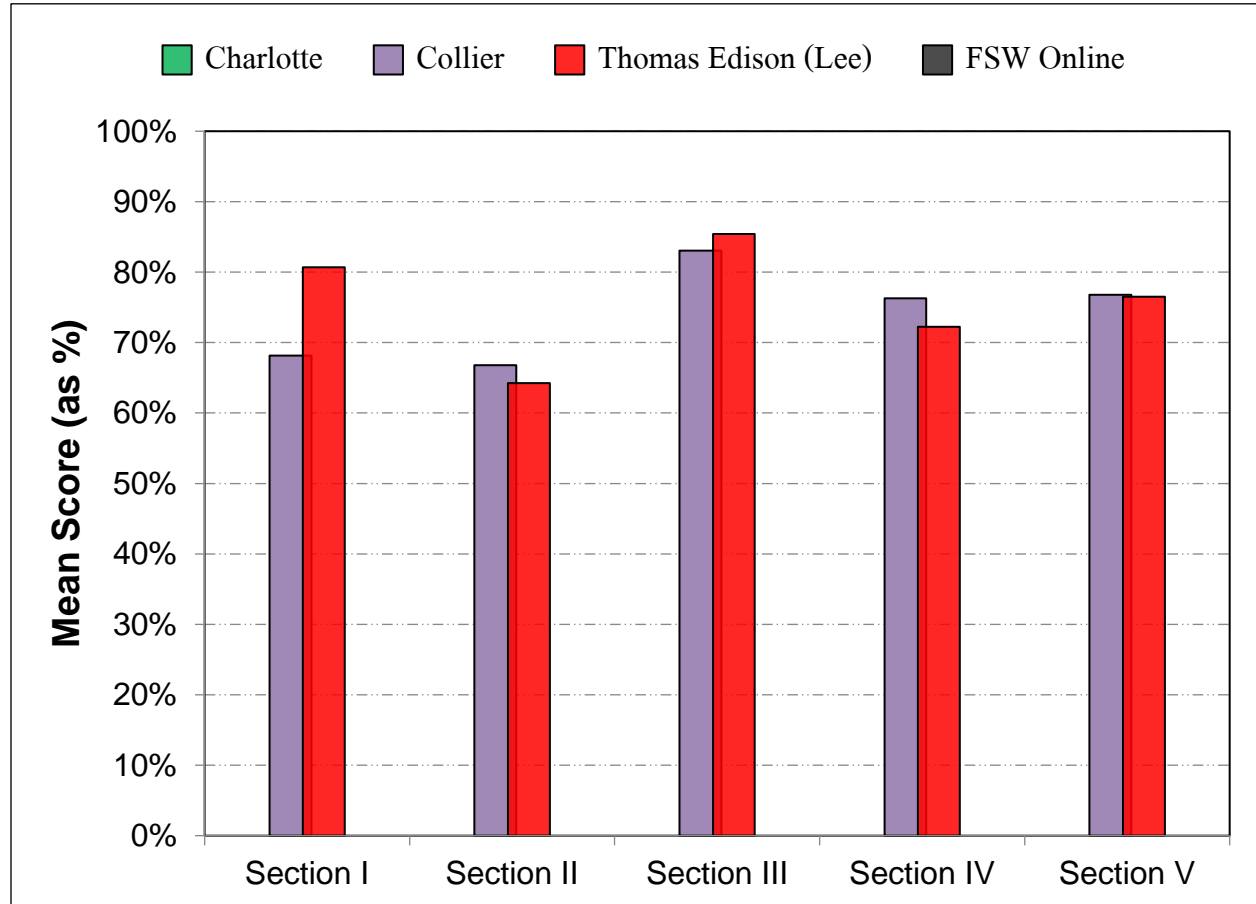

Figure 9. Comparison of mean score of exam sections by site.

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 10. All sections exhibit peaks above 90% except for Section II, which peaks at 70-79%. Section II also continues to exhibit more widely distributed scores. Section II exhibits a maximum of 12% at ≥ 90% while other sections range from 26% to 51%.
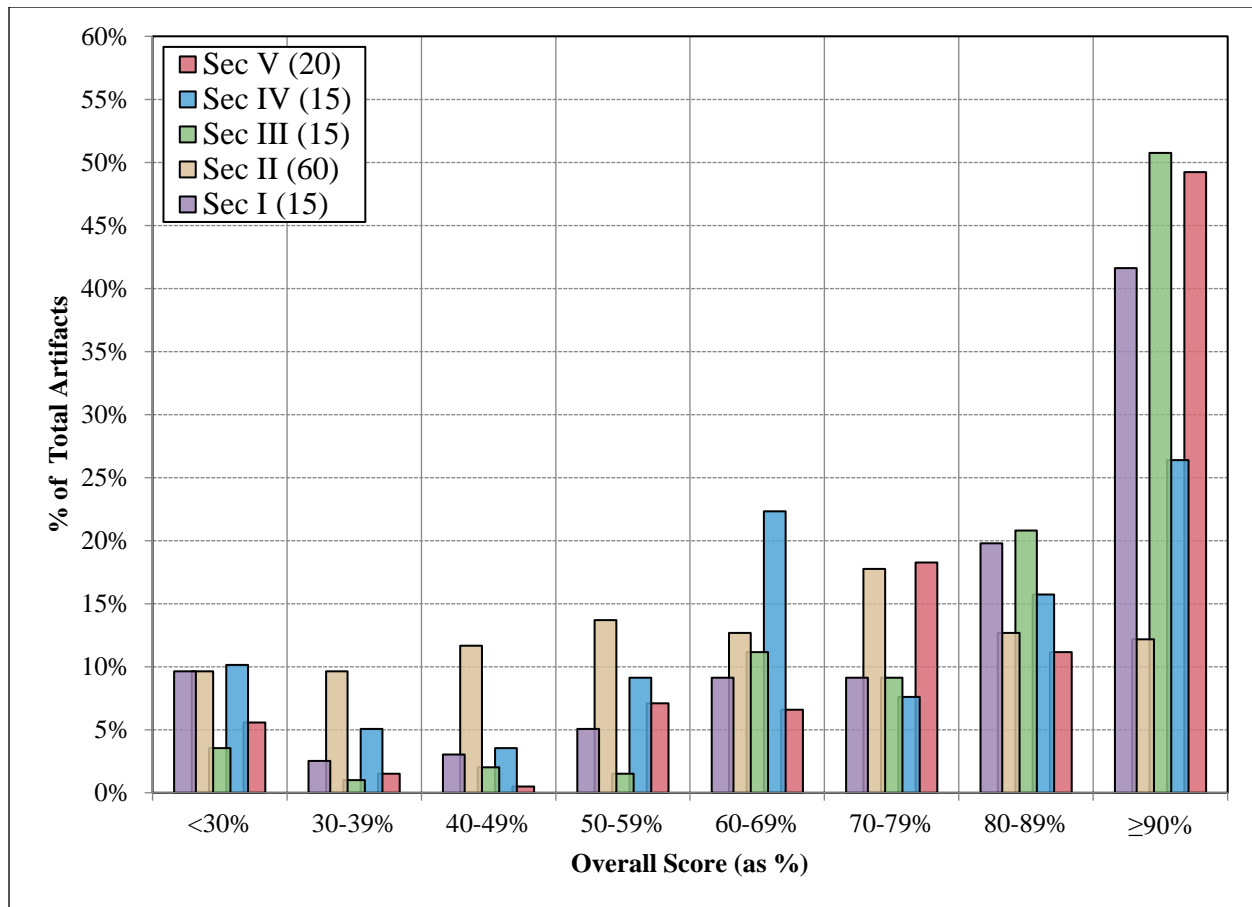
Figure 10. Histogram of SPN 1120 exam sections scores for fall 2020. Purple – Section I, Brown – Section II, Green – Section III, Blue – Section IV, and Red – Section V.

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 11). 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 11 shows that Section II remains the lowest performing compared to other sections between the ranges of 50-100%. In the overall scoring range of 80%-84%, for example, Section II exhibits a mean score of 65% while the other four sections range from 79%-91%. At the overall scoring range at 90%-94%, Section II exhibits a mean score of 85%, while the other four sections range from 90%-97%.

	Section I	Section II	Section III	Section IV	Section V		Scale
≥ 95%	99%	92%	100%	100%	100%		100%
90-94%	95%	85%	97%	90%	91%		90%
85-89%	91%	75%	94%	82%	90%		80%
80-84%	88%	65%	91%	79%	89%		70%
75-79%	79%	64%	88%	69%	86%		60%
70-74%	75%	56%	84%	66%	79%		50%
65-69%	68%	49%	81%	56%	79%		40%
60-64%	71%	51%	74%	57%	62%		30%
55-59%	52%	42%	68%	51%	70%		
50-54%	54%	48%	63%	41%	55%		
< 50%	22%	27%	43%	19%	34%		

Figure 11. 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 9 and Figure 12. Demographics and sample size vary by semester so it may be more reasonable to compare like semesters (Fall vs. Fall, Spring vs. Spring) (see http://www.fsw.edu/facultystaff/assessment/history for examples). In 15 terms, Sections I through V for the Fall 2020 term rank 13th, 14th, 12th, 15th, and 11th, respectively.

	Section I	Section II	Section III	Section IV	Section V	Overall
Fall 2013 (n=58)	12.1	40.4	12.0	10.6	16.2	91.3
Spring 2014 (n=90)	12.4	45.1	13.2	11.7	16.8	99.2
Fall 2014 (n=93)	11.8	40.5	12.8	11.1	16.5	92.8
Spring 2015 (n=73)	11.5	39.8	12.8	10.9	16.4	91.4
Fall 2015 (n=122)	10.8	41.8	12.4	11.5	15.2	91.6
Spring 2016 (n=141)	11.5	38.4	13.3	10.5	15.7	89.5
Fall 2016 (n=240)	11.9	39.8	12.8	10.9	15.2	90.6
Spring 2017 (n=89)	11.9	41.0	13.0	11.0	16.7	93.6
Fall 2017 (n=266)	11.9	41.1	12.9	11.5	15.4	92.8
Spring 2018 (n=107)	12.0	43.7	12.8	11.6	16.4	96.6
Fall 2018 (n=212)	10.9	39.4	12.5	10.4	15.9	89.1
Spring 2019 (n=80)	11.6	41.3	12.4	10.3	16.3	91.8
Fall 2019 (n=240)	12.0	37.7	13.2	11.3	16.7	90.8
Spring 2020 (n=123)	11.4	36.7	12.8	11.0	15.8	87.7
Fall 2020 (n=197)	11.2	36.9	12.4	10.1	15.7	86.4

Table 9. Comparison of mean scores for SPN 1120 for fall 2013 through the present.

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 14 of 15 terms including the Fall 2020 term.

Section II exhibits the lowest mean scores in 14 of 15 terms, spring 2019 being the only term in which it is not the lowest. For spring 2019, Section IV exhibits the lowest mean score.
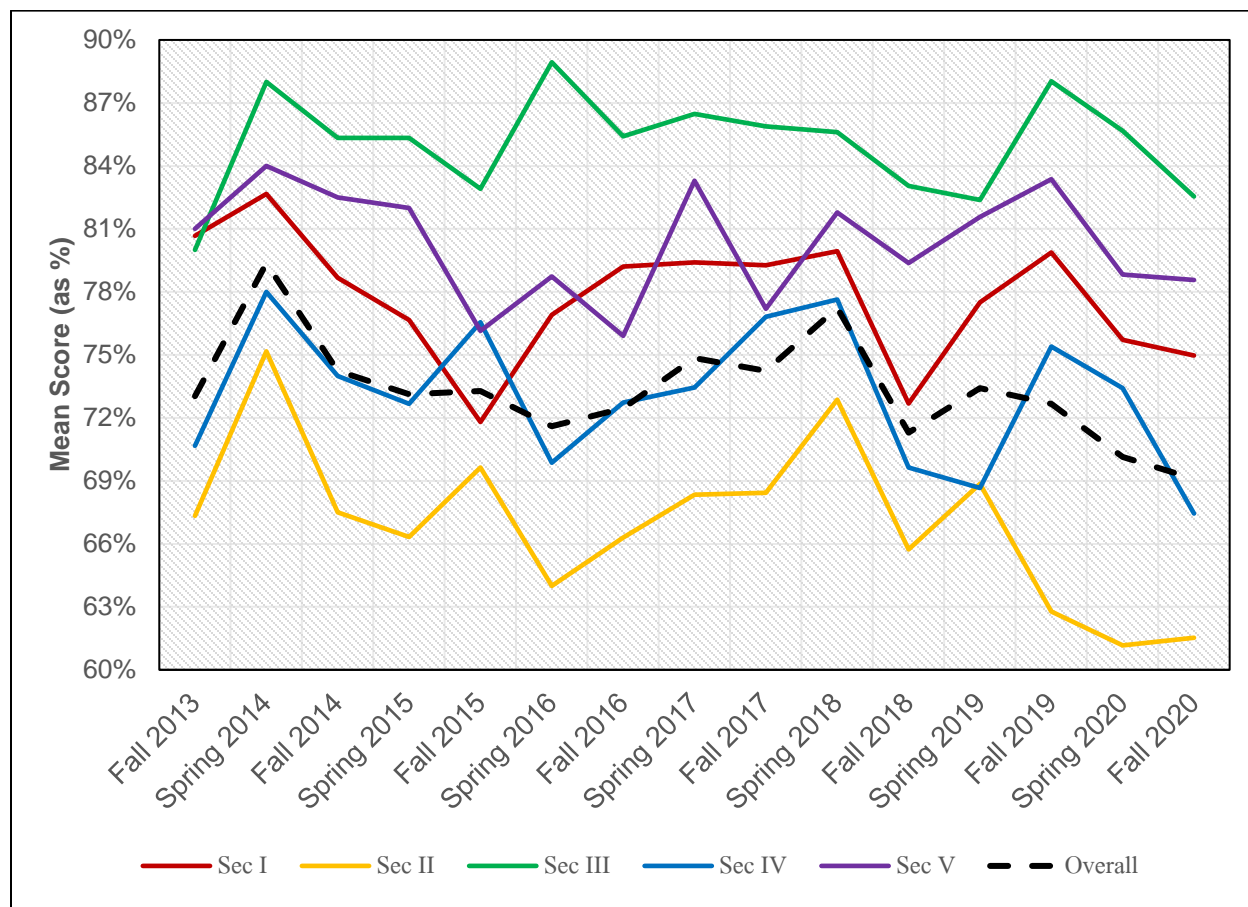

Figure 12. Comparison of mean scores (as percentage) for SPN 1120 through time from fall 2013 through the present.

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 spring 2020, 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 spring 2020 on the Charlotte, Collier, and Thomas Edison campuses. No study was conducted in summer 2020 to include in this report due to the Covid-19 pandemic. Results of the cumulative study are shown in Figure 14.

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 13). Over time, the range between highest and lowest instructors has narrowed. In fall 2016, the range was 17.8. In fall 2020, that gap has stabilized at 6.9.
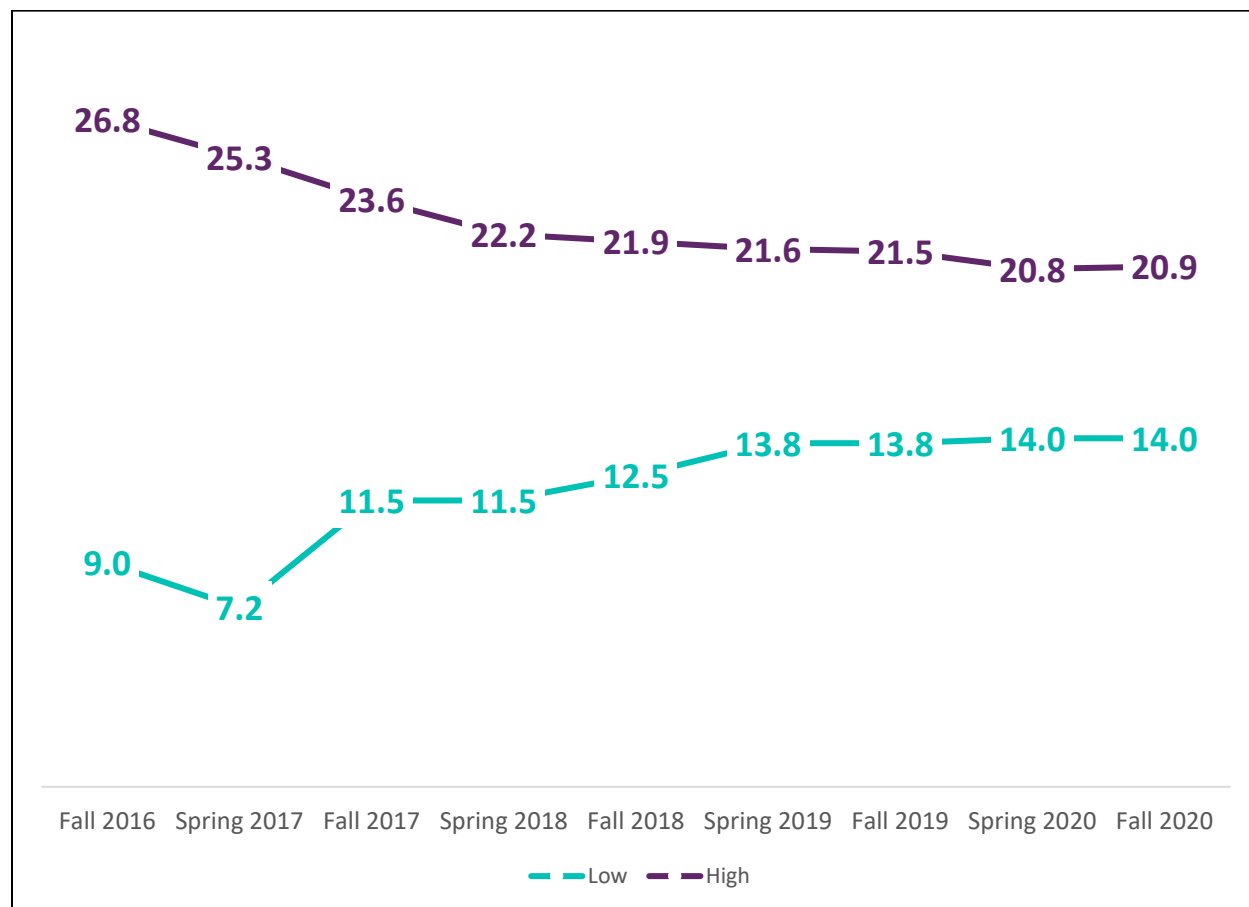

Figure 13. Comparison of achievement in entrance skills study assessment over time between highest scoring instructor (score listed is mean score) and lowest.

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 fall 2020. Of those, artifacts were collected from a common final from 6 of 6 sections. A total of 123 students were enrolled in SPN 1121. Of those, 83 artifacts were collected representing a sample size of 67%.

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 10 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 75% of artifacts score 70% or higher in the oral competency exam section (Section I) (Table 10, Figure 14). 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 60% 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 53% of artifacts scored 70% or higher in Section V and 82% of artifacts scored 80% or higher in Section VII.

n = 83	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	75%	43%	43%	72%	53%	60%	82%	
Mean (as %)	75%	59%	56%	76%	62%	71%	81%	
Median (as %)	83%	60%	58%	87%	75%	80%	90%	
Section Score Max Possible	*15*	*15*	*40*	*15*	*12*	*15*	*20*	*132*
Section Mean	11.2	8.8	22.5	11.3	7.4	10.7	16.1	87.9
Section Median	12.5	9	23	13	9	12	18	92.5

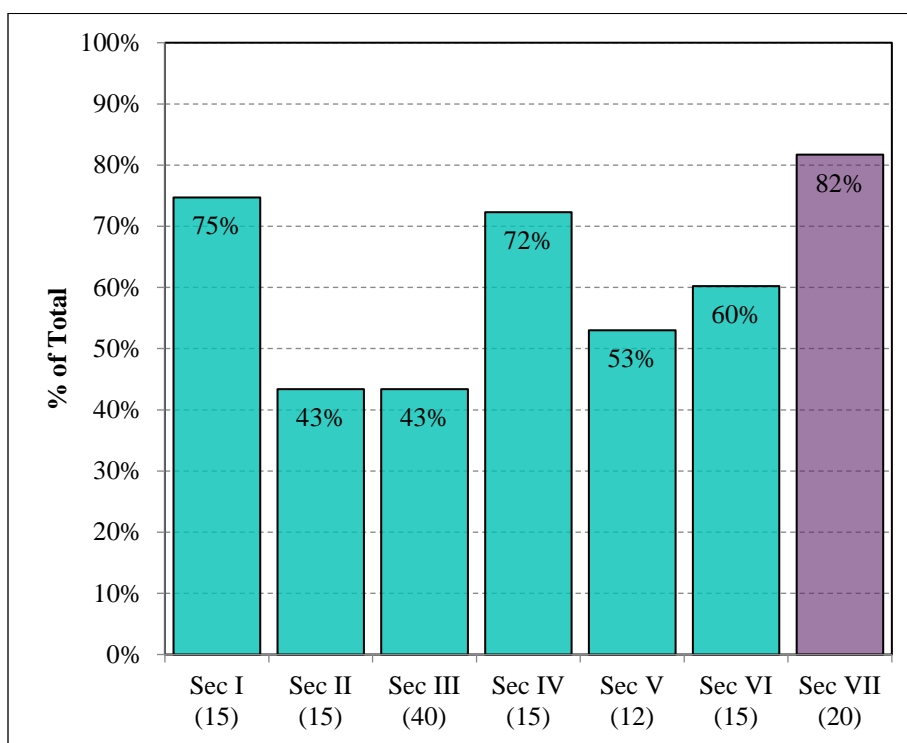*Table 10. Percentage of student achievement level as per SLOs (SPN 1121).*

Figure 14. 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 fall 2020 so no comparison study between dual enrollment and traditional courses could be completed.

3.3.2.2 Modality Comparison

During the Fall 2020 semester, 51 total asynchronous online artifacts were collected, along with 16 live online, 16 flex, 0 blended, and 0 traditional. A comparison of basic statistics is provided in Table 11. Flex artifacts mean scores are 73.0% points, which is 5.6%-points higher than the next highest, Asynchronous online. Live online is a further 9.9%-points lower, at 57.5%. Results of the ANOVA do not exhibit a statistically significant difference between sites (see Table 12). Therefore, we cannot reject the null hypothesis that the mean 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. A distribution of scores by modality is shown in Figure 15.

	Traditional	*Async Online*	*Live Online*	*Flex*	*Blended*
n	0	51	16	16	0
Mean Score	~	89.0	75.9	96.4	~
Mean Score (as %)	~	67.4%	57.5%	73.0%	~

Table 11. Comparison of descriptive statistics by modality.

Source of Variation	Sum of squared differences	df	Mean Squares	F_{obs}	p-value	F_{crit}
Between Sites	2015.4	2	1007.7	2.37	0.100	3.11
Within Sites	34,045.2	80	425.6			
Total	36,060.6	82				

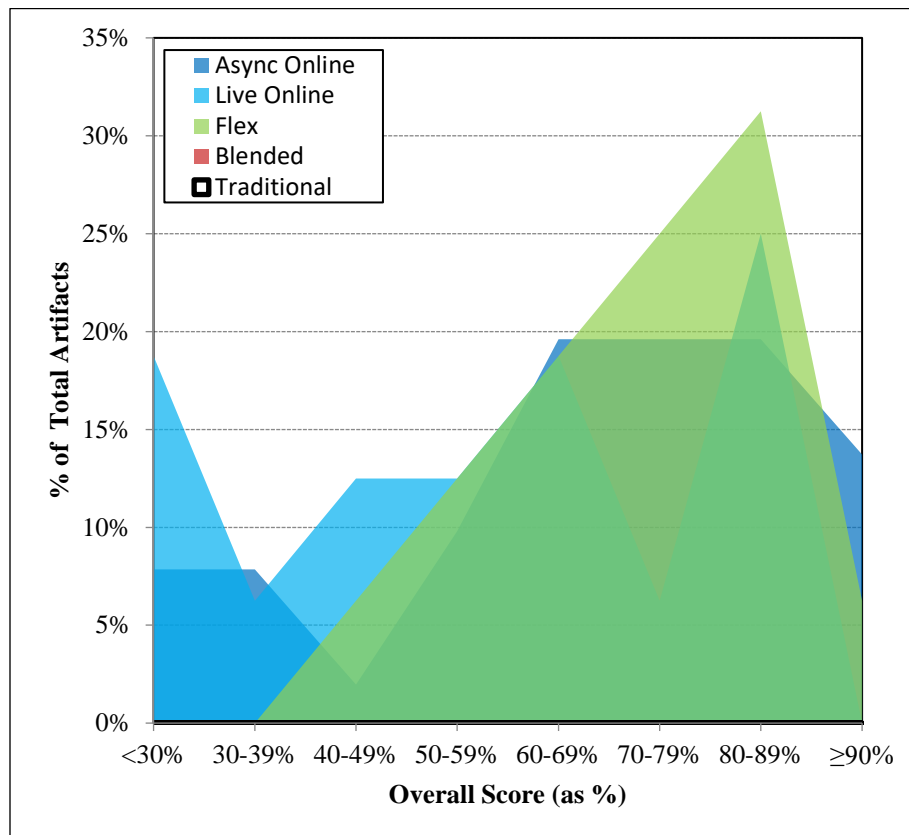*Table 12. Results of one-way ANOVA of mean scores by modality for SPN 1121.*

Figure 15. Score distribution by modality for SPN 1121.

3.3.2.3   Comparison by Campus/Site

Because only one traditional site offered sections during fall 2020 (Thomas Edison campus) no comparison by site could be 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 16. Sections I, IV, V, VI, and VII exhibit peaks centered on ≥90%. Section II exhibits a peak centered at 80-89%. Section III exhibits a peak centered at 50-59%.
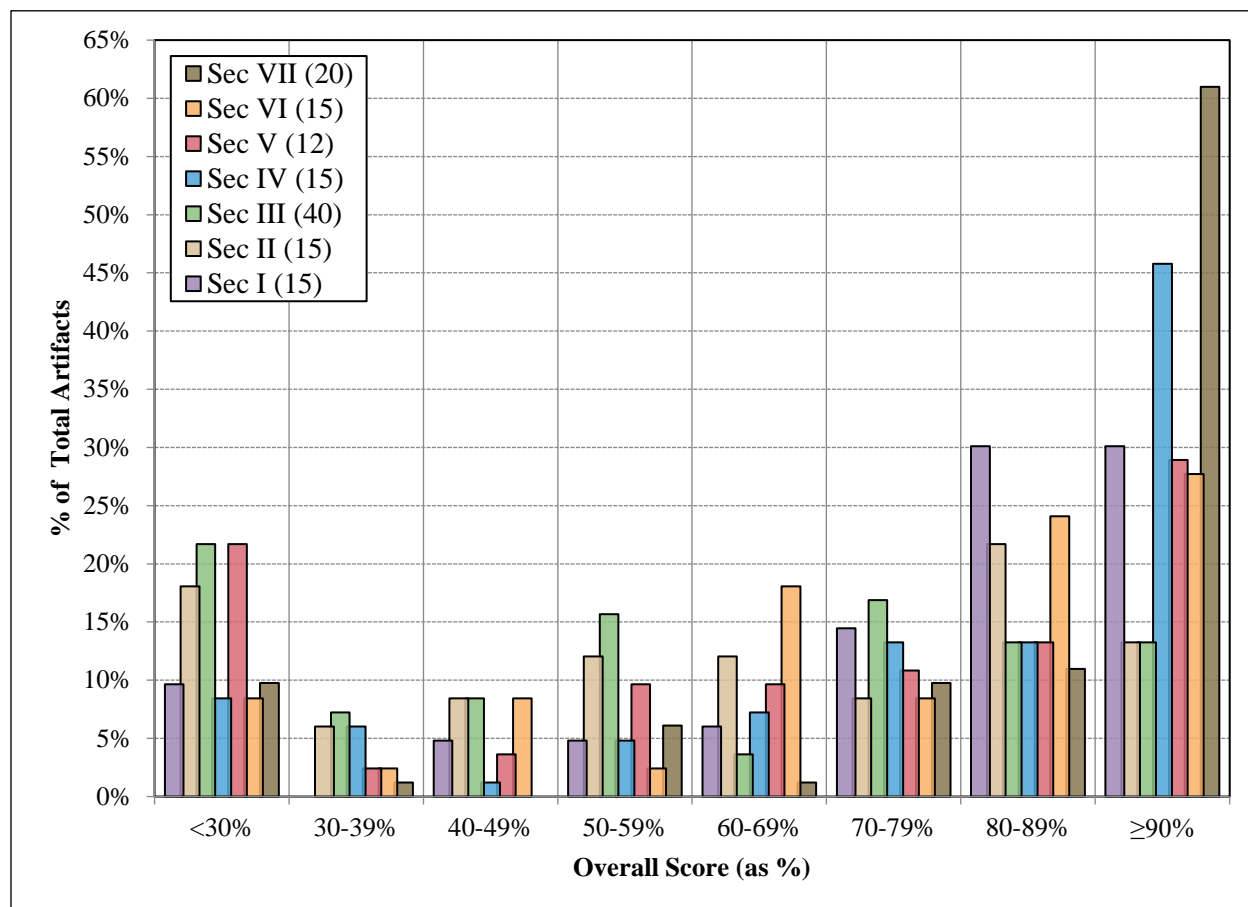

Figure 16. Histogram of SPN 1121 exam sections scores for fall 2020. Purple – Section I, Brown – Section II, Green – Section III, Blue – Section IV, Red – Section V, Orange – Section VI, and Black – Section VII.

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 17). 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 17 shows two main trends. First, Section VII exhibits a strong performance at lower overall scores. For example, in the 70-44% range, Section VII exhibits a mean score of 94%, while other sections range from 48%-85%. And second, Sections I exhibits the lowest scores in

most ranges. For example, in the 85-89% range, Section III scores 73%, while all others range from 84-98%.

	Section I	Section II	Section III	Section IV	Section V	Section VI	Section VII		Scale
≥ 95%	100%	97%	86%	100%	100%	100%	100%		100%
90-94%	95%	85%	76%	98%	93%	96%	98%		90%
85-89%	85%	89%	73%	93%	87%	84%	94%		80%
80-84%	86%	69%	70%	89%	82%	79%	93%		70%
75-79%	87%	61%	71%	81%	72%	78%	93%		60%
70-74%	72%	72%	48%	85%	70%	68%	94%		50%
65-69%	82%	57%	69%	82%	51%	66%	63%		40%
60-64%	67%	46%	50%	73%	58%	68%	74%		30%
55-59%	77%	49%	42%	69%	42%	52%	68%		
50-54%	53%	50%	39%	49%	47%	40%	83%		
< 50%	46%	13%	19%	30%	7%	47%	47%		

Figure 17. (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 13 and Figure 18). 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 15 terms, Sections I through VII for fall 2020 ranks 10th, 12th, 15th, 5th, 12th, 2nd, and 5th highest, respectively.

	Section I Oral	Section II Reading	Section III Read/Write	Section IV Read/Write	Section V Writing	Section VI Reading	Section VII Writing	Overall
F 2013 (n=10)	11.5	9.5	34.2	9.5	7.5	9.6	14.2	96.0
Sp 2014 (n=115)	12.3	9.6	32.3	11.4	5.7	10.3	15.4	97.0
F 2014 (n=25)	11.9	10.7	30.0	10.6	9.5	11.6	16.1	100.5
Sp 2015 (n=58)	12.2	10.6	31.1	11.4	8.5	10.7	16.4	100.9
F 2015 (n=17)	11.5	10.6	29.9	11.3	8.1	9.9	16.4	97.7
Sp 2016 (n=109)	10.7	9.5	30.8	10.5	8.3	9.0	14.6	93.5
F 2016 (n=42)	11.5	9.8	25.9	10.7	7.2	9.7	15.6	90.3
Sp 2017 (n=140)	11.8	8.6	29.2	10.3	7.9	8.7	15.1	91.6
F 2017 (n=45)	10.6	9.1	27.3	11.8	8.1	10.2	15.7	92.8
Sp 2018 (n=136)	12.2	9.3	29.0	11.2	8.4	10.1	16.7	96.9
F 2018 (n=67)	10.4	8.7	27.9	10.2	8.2	8.3	15.9	89.6
Sp 2019 (n=185)	11.7	9.2	30.8	11.1	8.9	9.8	17.3	98.7
F 2019 (n=63)	10.9	9.2	24.3	10.7	8.0	9.5	15.5	88.0
Sp 2020 (n=168)	10.5	7.7	24.9	11.5	7.4	9.3	16.3	87.4
F 2020 (n=83)	11.2	8.8	22.5	11.3	7.4	10.7	16.1	87.9

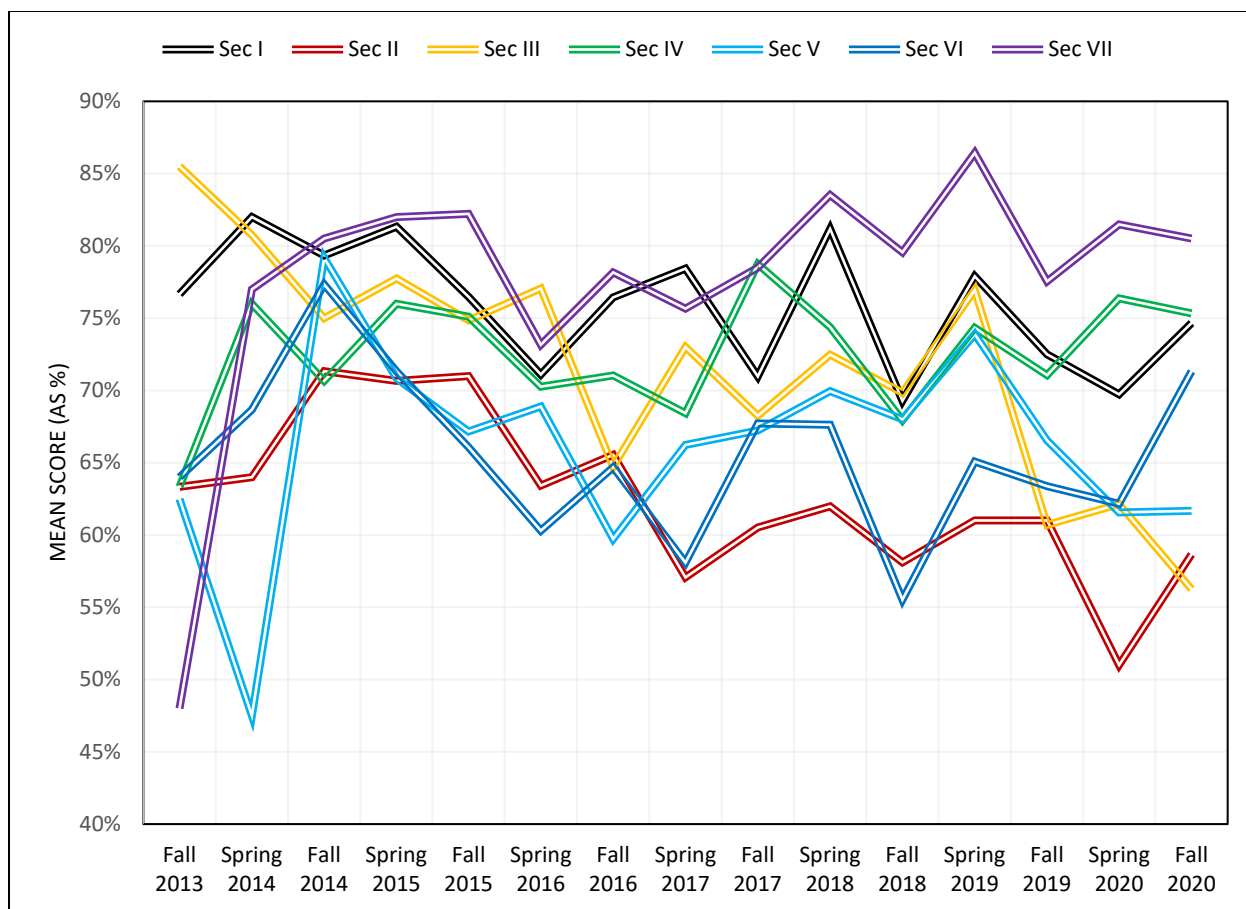*Table 13. Comparison of mean scores for SPN 1121 for fall 2013 through the present.*

Figure 18. 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 fall 2020 assessment.

4.1 FRENCH

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

1. During the Fall 2020 term, five sections of FRE 1120 were offered. An assessment first piloted in AY 2017-18 was recorded in all five section of FRE 1120 accounting for 49 artifacts, or 71% of the population.

2. In a study of Learning Outcome (LO) achievement, results exhibit mean scores across the five LOs ranging from 2.0 to 2.3 (Figure 1), compared with 2.7-3.0 of spring 2020, 2.2-2.6 in fall 2019, 2.3 to 2.7 in 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. During the Fall 2020 term, one dual enrollment (concurrent) section was offered. Overall mean scores (max = 15) are higher for concurrent dual enrollment (14.4/15) when compared with traditional course sections (10.6/15). No statistical significance study was completed due to limited sample size for dual enrollment (n=7).
4. During the Fall 2020 semester, 23 total asynchronous online artifacts were collected, along with 8 live online, 11 flex, 0 blended, and 0 traditional. Asynchronous online mean scores are the highest of all modalities at 14.7/15. The lowest is Flex, at 12.3/15. Results of the ANOVA exhibit a statistically significant difference between sites.
5. Only one site (Lee campus) had course sections run during fall 2020. All other sections were some form of online or dual enrollment so no cross-campus comparison could be completed.
6. In a study of achievement based on disposition survey question response, results indicate a weak-to-moderate sense of students self-reporting a strong understanding of the topic that correlates with actual achievement.
7. In a longitudinal study, in five terms, LOs 1 through 5 for the Fall 2020 rank 5th, 5th, 5th, 5th, and 5th, respectively.

A drill-down of FRE 1121 results are as follows:
1. During the Fall 2020 term, two sections of FRE 1121 were offered. The assessment was recorded in both, however, in one section results and scoring did not match that expected from the assessment plan, so no data could be utilized. The data collection accounts for 6 artifacts, or 25% of the population.
2. In a study of Learning Outcome (LO) achievement, results exhibit mean scores across the five LOs ranging from 1.8 to 2.8. 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. There was no dual enrollment (concurrent) section run during fall 2020, so no comparison study between dual enrollment and traditional courses could be completed.
4. While more than one modality was offered in this term, data was only able to be collected from one section, so no modality comparison could be completed.
5. Only one site offered sections other than online sections, so no comparison by site could be completed.
6. In a 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 if they appear to be capable of demonstrating it. There appears to be the beginnings of a trend of confidence with achievement, although sample size is limited (n=6).
7. In a longitudinal study, in three terms, LOs 1 through 5 for the Fall 2020 rank 2nd, 3rd, 2nd, 3rd, and 2nd, respectively.

4.2 SPANISH
A drill-down of SPN 1120 results are as follows:
1. During the Fall 2020 semester 13 sections of SPN 1120 *Beginning Spanish I* were offered. Of those, artifacts from a common final were collected from 13 sections. A total of 278 students were enrolled in SPN 1120. Of those, 197 artifacts were collected representing a sample size of 71% of the population.
2. Achievement of 80% of artifacts scoring 70% or better (SLO 1): Achievement was nearly met as results exhibit 71% of artifacts score 70% or higher in the oral competency exam section.

3. Achievement of 80% of artifacts scoring 70% or better (SLO 2): Achievement was partially met. Results exhibit 43% of artifacts scored 70% or higher in Section II and 81% of artifacts scored 70% or higher in Section III.

4. Achievement of 80% of artifacts scoring 70% or better (SLO 3): Achievement was nearly met. Results exhibit 50% of artifacts scored 70% or higher in Section IV and 79% of artifacts scored 70% or higher in Section V.

5. No dual enrollment (concurrent) sections of the course were run during fall 2020 so no comparison study between dual enrollment and traditional courses could be completed.

6. During the Fall 2020 semester, 49 total asynchronous online artifacts were collected, along with 78 live online, 45 flex, 0 blended, and 25 traditional. Asynchronous online artifacts mean scores are 4.5%-points higher than the next highest, Flex, which is 71.7%. Mean scores for traditional artifacts are similar to Flex at 71.6%. Finally, nearly 10%-points lower still, is Live Online at 62.3%. Results of the ANOVA exhibit a statistically significant difference between sites.

7. Of the 197 artifacts collected from SPN 1120, 0 originated from the Charlotte campus, 25 from the Collier campus, and 42 from the Thomas Edison (Lee) campus. Scores by rubric dimension varied greatly across campuses. Collier exhibits the highest mean score in 3 of 5 sections, while the Thomas Edison (Lee) campus exhibits the highest in the reaming 2/5.

8. In a study of score distribution by section, all sections exhibit peaks above 90% except for Section II, which peaks at 70-79%. Section II also continues to exhibit more widely distributed scores. Section II exhibits a maximum of 12% at ≥ 90% while other sections range from 26% to 51%.

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 50-100%. In the overall scoring range of 80%-84%, for example, Section II exhibits a mean score of 65% while the other four sections range from 79%-91%. At the overall scoring range at 90%-94%, Section II exhibits a mean score of 85%, while the other four sections range from 90%-97%.

10. In a longitudinal study of data distribution through time, in 15 terms, Sections I through V for the Fall 2020 term rank 13th, 14th, 12th, 15th, and 11th, respectively.

11. In a study assessing student skills and retention of materials from SPN 1120 prior to beginning SPN 1121, over time, the range between highest and lowest instructors has narrowed. In fall 2016, the range was 17.8. In fall 2020, that gap has stabilized at 6.9.

A drill-down of SPN 1121 results are as follows:
1. Six sections of SPN 1121 Beginning Spanish II were offered in fall 2020. Of those, artifacts were collected from a common final from 6 of 6 sections. A total of 123 students were enrolled in SPN 1121. Of those, 83 artifacts were collected representing a sample size of 67%.

2. 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).

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 60% 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 53% of artifacts scored 70% or higher in Section V and 82% of artifacts scored 80% or higher in Section VII.

5. No dual enrollment (concurrent) sections of the course were run during fall 2020 so no comparison study between dual enrollment and traditional courses could be completed.

6. During the Fall 2020 semester, 51 total asynchronous online artifacts were collected, along with 16 live online, 16 flex, 0 blended, and 0 traditional. Flex artifacts mean scores are 73.0% points,

which is 5.6%-points higher than the next highest, Asynchronous online. Live online is a further 9.9%-points lower, at 57.5%. Results of the ANOVA do not exhibit a statistically significant difference between sites.

7. Because only one traditional site offered sections during fall 2020 (Thomas Edison campus) no comparison by site could be completed.

8. In a study of score distribution by section, Sections I, IV, V, VI, and VII exhibit peaks centered on ≥90%. Section II exhibits a peak centered at 80-89%. Section III exhibits a peak centered at 50-59%.

9. In a study of section score distribution based on overall score, two main trends. First, Section VII exhibits a strong performance at lower overall scores. For example, in the 70-44% range, Section VII exhibits a mean score of 94%, while other sections range from 48%-85%. And second, Sections I exhibits the lowest scores in most ranges. For example, in the 85-89% range, Section III scores 73%, while all others range from 84-98%.

10. In a longitudinal study of data distribution through time, in 15 terms, Sections I through VII for fall 2020 ranks 10th, 12th, 15th, 5th, 12th, 2nd, and 5th 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.