

# Foreign Language Assessment Report

## Spring 2021

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 spring 2021 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](mailto:jfvangaalen@fsw.edu); x16965).

## 2 FRENCH

---

### 2.1 FRE 1120

During the Spring 2021 term, two sections of FRE 1120 were offered. An assessment first piloted in AY 2017-18 was recorded in both sections of FRE 1120 accounting for 20 artifacts, or 54% 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 2.8 to 3.0 (Figure 1), compared with 2.0-2.3 in fall 2020, 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.

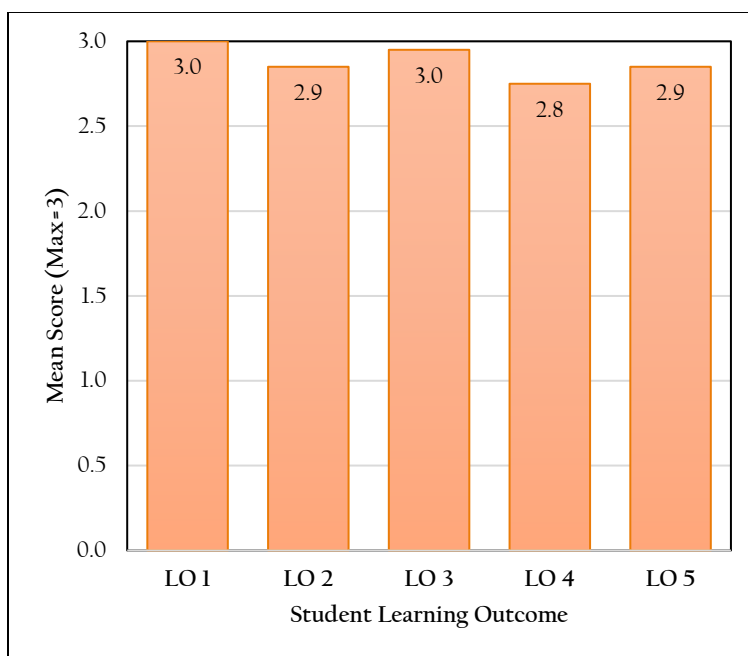


Figure 1. Mean scores by SLOs for FRE 1120.

## 2.1.2 Exploratory Analysis & Significance Testing

### 2.1.2.1 Dual Enrollment (Concurrent) to Traditional Comparison

During the Spring 2021 term, no dual enrollment (concurrent) sections were run, so no comparison could be completed.

### 2.1.2.2 Modality Comparison

During the Spring 2021 semester, 16 total asynchronous online artifacts were collected, along with 0 live online, 4 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 14.6/15. The lowest is Flex, at

13.8/15. A one-way analysis of variance is normally used to compare means by modality, however, sample size is too limited.

|                    | <i>Traditional</i> | <i>Async<br/>Online</i> | <i>Live<br/>Online</i> | <i>Flex</i> | <i>Blended</i> |
|--------------------|--------------------|-------------------------|------------------------|-------------|----------------|
| n                  | ~                  | 16                      | ~                      | 4           | ~              |
| Max                | ~                  | 15                      | ~                      | 15          | ~              |
| Min                | ~                  | 11                      | ~                      | 13          | ~              |
| Mean               | ~                  | 14.6                    | ~                      | 13.8        | ~              |
| Standard deviation | ~                  | 1.21                    | ~                      | 0.96        | ~              |
| Skewness           | ~                  | -2.62                   | ~                      | 0.85        | ~              |
| Kurtosis           | ~                  | 5.83                    | ~                      | -1.29       | ~              |

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

### 2.1.2.3 Comparison by Campus/Site

Only one site (Lee campus) had course sections run during spring 2021. The other section was online, 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, 75% of artifacts score 15/15 or higher.

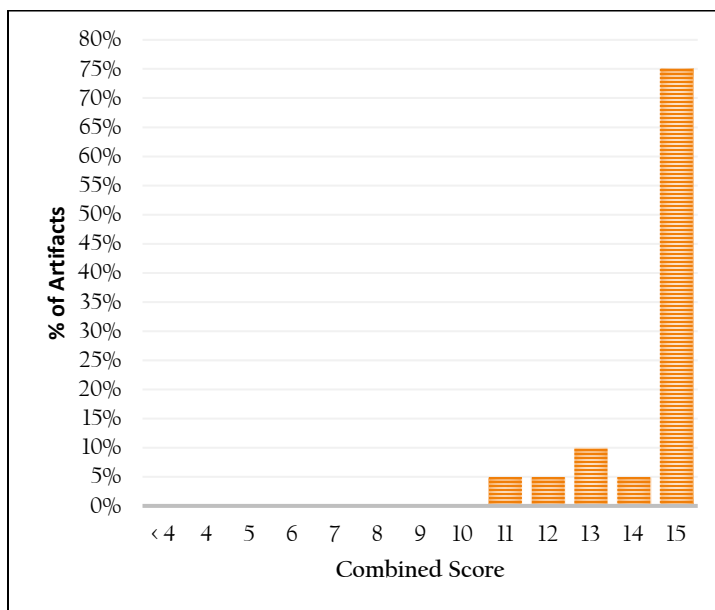


Figure 2. Distribution of combined (total) scores for assessment.

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 some which report a low understanding and yet score high (as seen in data point in upper left corner of graph).

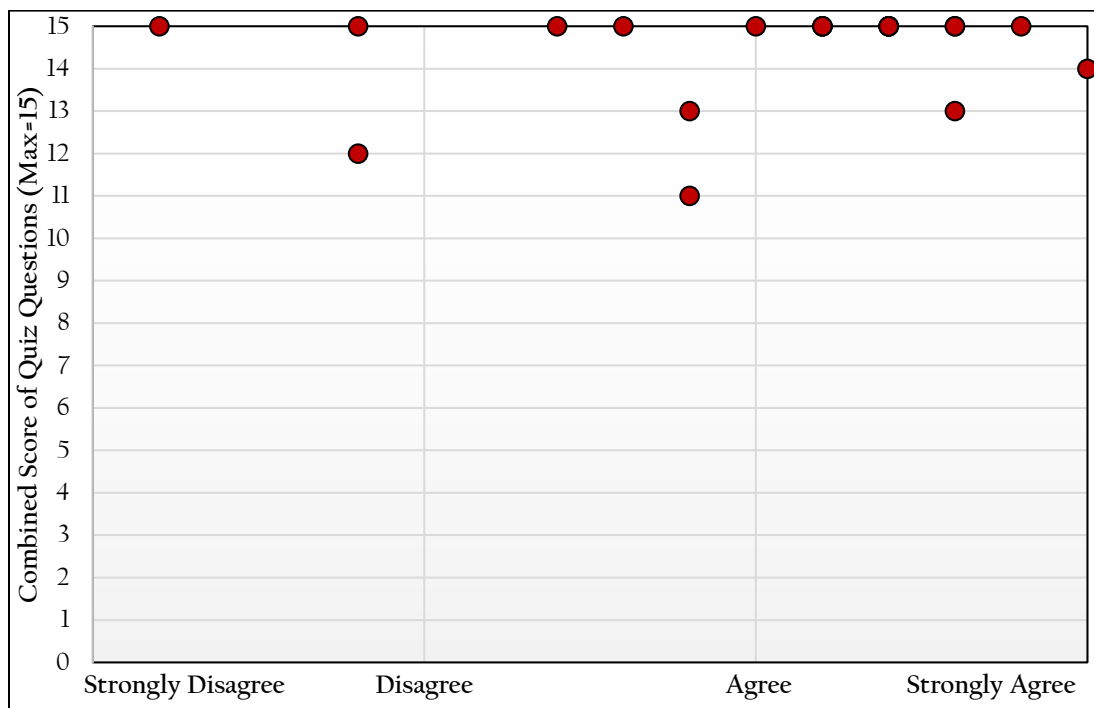


Figure 3. Scatter plot of combined (total) scores based on general response to survey questions.

### 2.1.3.2 Longitudinal Study

Further description of achievement over time in FRE 1120 is provided in Table 2. 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 six terms, LOs 1 through 5 for the Spring 2021 term rank 1<sup>st</sup>, 2<sup>nd</sup>, 1<sup>st</sup>, 2<sup>nd</sup>, and 2<sup>nd</sup>, 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    |
| Spring 2021 (n=20) | 3.0  | 2.9  | 3.0  | 2.8  | 2.9  |         |

Table 2. Comparison of mean scores for FRE 1120 over time.

## 2.2 FRE 1121

During the Spring 2021 term, four sections of FRE 1121 were offered. The assessment was recorded in all four, 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 45 artifacts, or 92% 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 2.5 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.

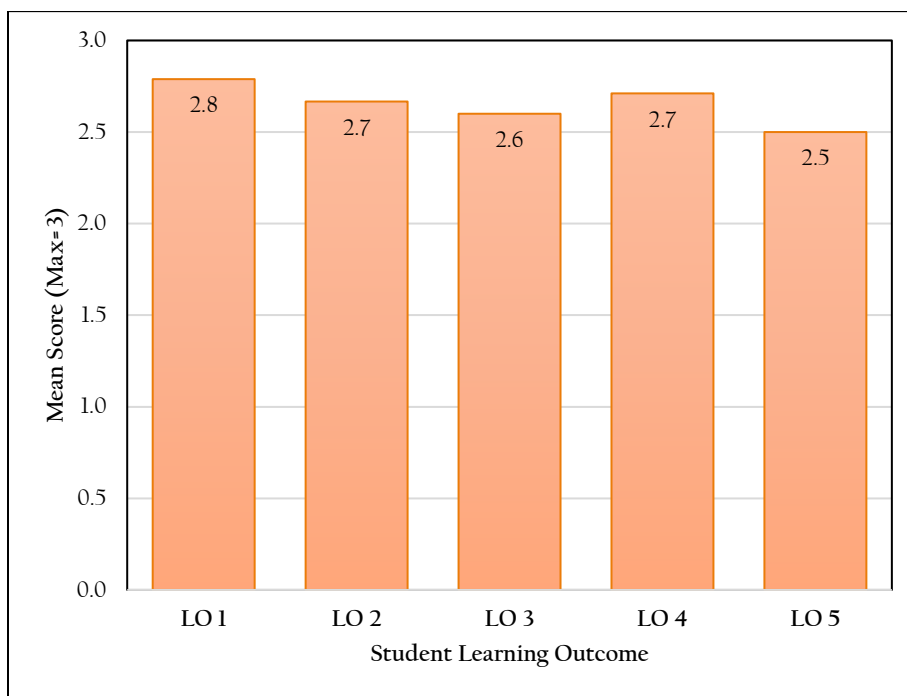


Figure 4. Mean scores by SLOs for FRE 1121.

## 2.2.2 Exploratory Analysis & Significance Testing

### 2.2.2.1 Dual Enrollment (Concurrent) to Traditional Comparison

While one dual enrollment (concurrent) section was run during spring 2021, the data could not be accurately reported, so no comparison study between dual enrollment and traditional courses could be completed.

### 2.2.2.2 Modality Comparison

During the Spring 2021 semester, 29 total asynchronous online artifacts were collected, along with 5 live online, 12 flex, 0 blended, and 0 traditional. A comparison of basic statistics is provided in Table 3. Asynchronous online mean scores are the highest of all modalities at 14.4/15. The lowest is Flex, at 11.4/15. A one-way analysis of variance found that there is a statistically significant difference between modalities ( $p=0.00001$ ).

|                    | <i>Traditional</i> | <i>Async<br/>Online</i> | <i>Live<br/>Online</i> | <i>Flex</i> | <i>Blended</i> |
|--------------------|--------------------|-------------------------|------------------------|-------------|----------------|
| n                  | ~                  | 29                      | 5                      | 12          | ~              |
| Max                | ~                  | 15                      | 14                     | 14.5        | ~              |
| Min                | ~                  | 8                       | 9                      | 7           | ~              |
| Mean               | ~                  | 14.4                    | 11.6                   | 11.4        | ~              |
| Standard deviation | ~                  | 1.52                    | 1.82                   | 2.11        | ~              |
| Skewness           | ~                  | -3.31                   | -0.27                  | -0.65       | ~              |
| Kurtosis           | ~                  | 11.93                   | 1.07                   | 0.43        | ~              |

Table 3. Comparison of mean scores for online and traditional artifacts. Positive effect sizes indicate a higher mean score for traditional artifacts.

### 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 15/15 with a strong negative skew meaning scores are tending strongly towards higher values (Starkweather, 2010). In total, over 60% of artifacts score 14/15 or higher.

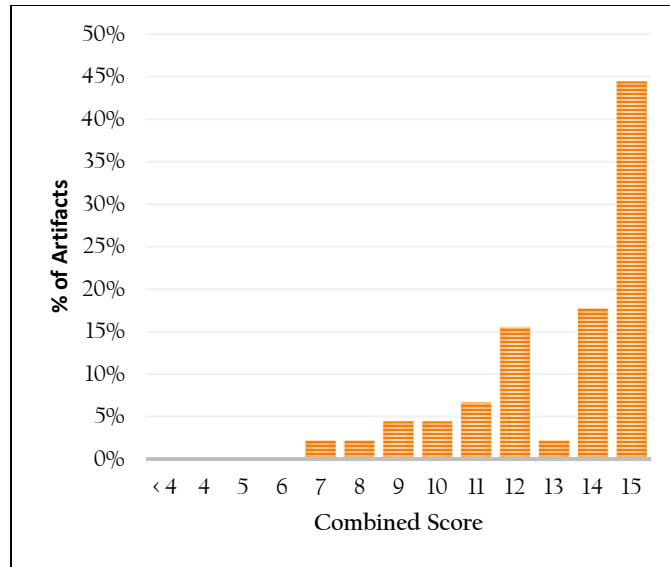


Figure 5. Distribution of combined (total) scores for assessment.

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 is a moderate correlation between self-report and achievement.

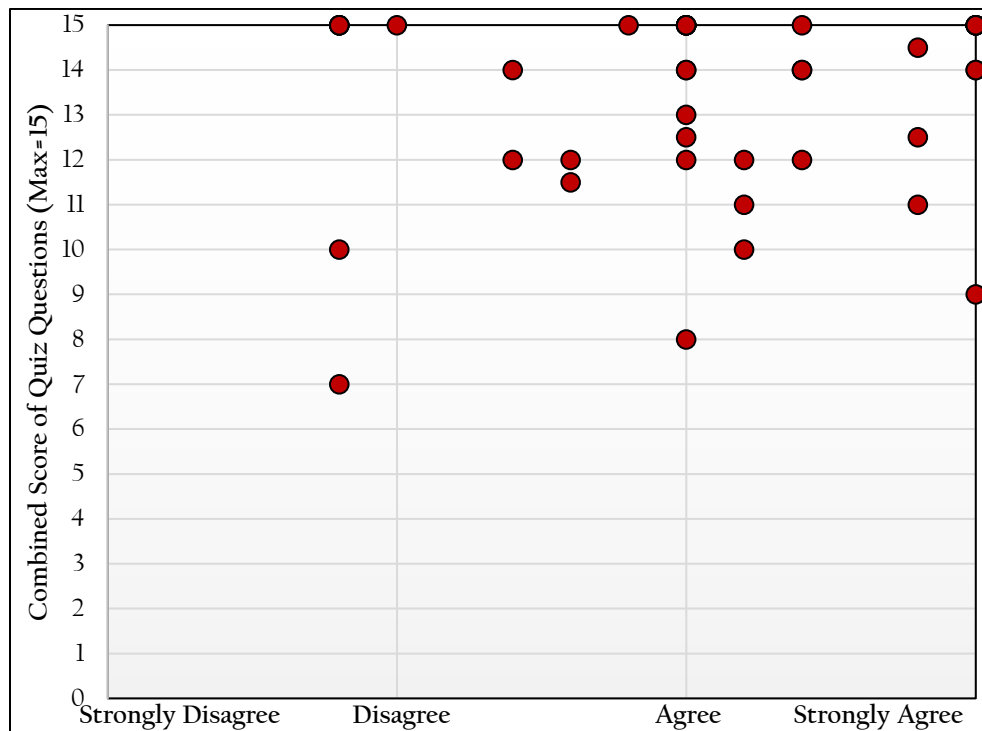


Figure 6. Scatter plot of combined (total) scores based on general response to survey questions.

### 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 four terms, LOs 1 through 5 for the Spring 2020 term rank 2<sup>nd</sup>, 1<sup>st</sup>, 3<sup>rd</sup>, 2<sup>nd</sup>, and 1<sup>st</sup>, 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    |
| Spring 2021 (n=45) | 2.8  | 2.7  | 2.6  | 2.7  | 2.5  | 13.3    |

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 Spring 2021 semester 9 sections of SPN 1120 *Beginning Spanish I* were offered. Of those, artifacts from a common final were collected from 8 sections as one section did not utilize the same scoring. A total of 176 students were enrolled in SPN 1120. Of those, 109 artifacts were collected representing a sample size of 62% 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 76% 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 50% 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 nearly met. Results exhibit 52% of artifacts scored 70% or higher in Section IV and 77% of artifacts scored 70% or higher in Section V. For a graphical representation of SLO achievement, see Figure 7.



| n = 109           | Section I<br>(Oral)                                  | Section II<br>(Reading) | Section III<br>(Reading) | Section IV<br>(Written) | Section V<br>(Written) | Combined<br>(Overall) |
|-------------------|--|-------------------------|--------------------------|-------------------------|------------------------|-----------------------|
| Goal              | 80% of artifacts scored $\geq 70\%$ for all sections |                         |                          |                         |                        |                       |
| % above 70%       | 76%  | 50%                     | 85%                      | 52%                     | 77%                    |                       |
| Mean (as %)       | 77%  | 65%                     | 83%                      | 67%                     | 76%                    | 72%                   |
| Median (as %)     | 87%  | 70%                     | 87%                      | 70%                     | 83%                    | 75%                   |
| Section Score Max | 15   | 60                      | 15                       | 15                      | 20                     | 125                   |
| Section Mean      | 11.6   | 38.9                    | 12.5                     | 10.1                    | 15.2                   | 88.3                  |
| Section Median    | 13   | 42                      | 13                       | 10.5                    | 16.5                   | 92.5                  |

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

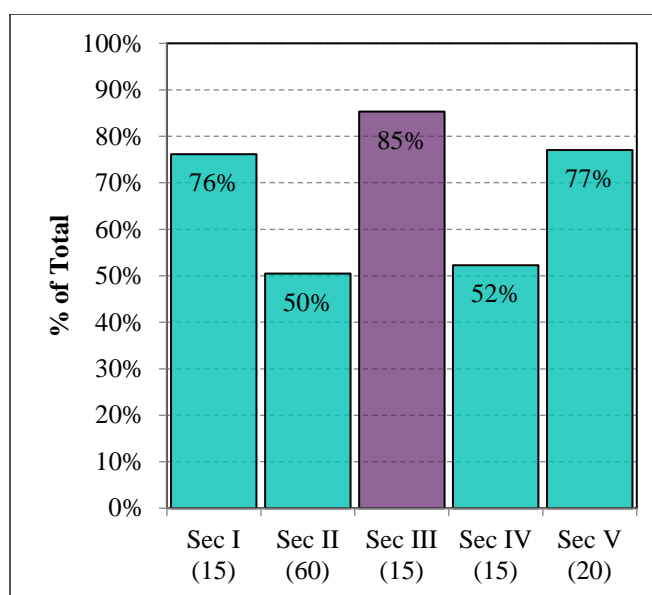


Figure 7. SLO achievement for SPN 1120 by exam section (Percentage of student achievement level as per SLOs). Purple denotes having met objective.

### 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 spring 2021 so no comparison study between dual enrollment and traditional courses could be completed.

#### 3.1.2.2 Modality Comparison

During the Spring 2021 semester, 49 total asynchronous online artifacts were collected, along with 21 live online, 31 flex, 0 blended, and 8 traditional. A comparison of basic statistics is provided in Table 6. Traditional artifacts are 6.4%-points higher than the next highest, although sample size is only n=8. Asynchronous online, live online, and flex range only 1.9%-points between the three. Results of the ANOVA do not exhibit a statistically significant difference between sites (see Table 7). Therefore, we cannot reject the null hypothesis that the mean scores in each modality 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 8 (Traditional is excluded due to limited sample size).

|                   | <i>Traditional</i> | <i>Async<br/>Online</i> | <i>Live<br/>Online</i> | <i>Flex</i> | <i>Blended</i> |
|-------------------|--------------------|-------------------------|------------------------|-------------|----------------|
| n                 | 8                  | 49                      | 21                     | 31          | 0              |
| Mean Score        | 97.6               | 86.5                    | 89.7                   | 87.7        | ~              |
| Mean Score (as %) | 78.1%              | 69.2%                   | 71.7%                  | 70.1%       | ~              |

Table 6. Comparison of descriptive statistics by modality.

| Source of Variation | Sum of squared differences | df  | Mean Squares | F <sub>obs</sub> | p-value | F <sub>crit</sub> |
|---------------------|----------------------------|-----|--------------|------------------|---------|-------------------|
| Between Sites       | 570.9                      | 3   | 190.3        | 0.49             | 0.692   | 2.69              |
| Within Sites        | 41,032.5                   | 105 | 390.8        |                  |         |                   |
| Total               | 41,603.5                   | 108 |              |                  |         |                   |

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

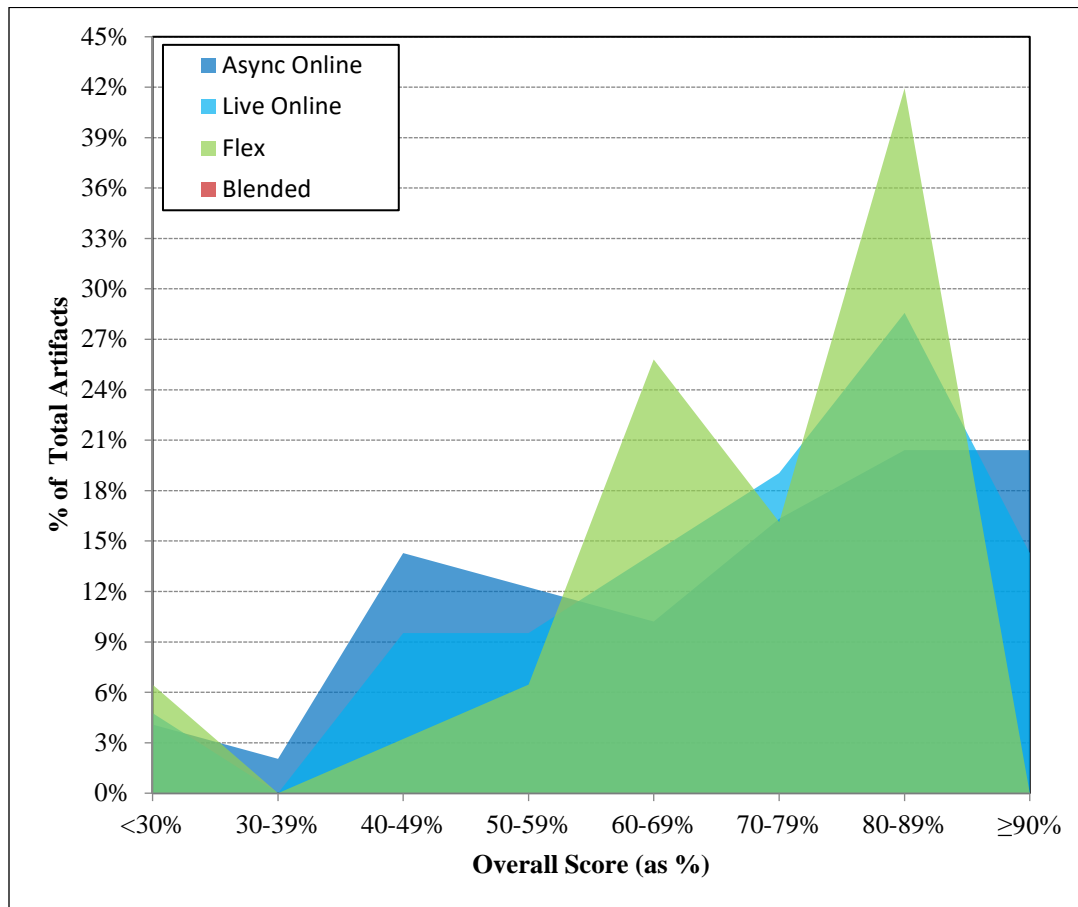


Figure 8. Score distribution by modality for SPN 1120. Traditional excluded from this figure due to limited sample size (n=8) which reduces clear interpretation of results.

### 3.1.2.3 Comparison by Campus/Site

Only one traditional site offered course sections this term, as a result, no comparison by site could be completed.

### 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 9. All sections exhibit peaks above 90%. Sections II and V exhibit more widely distributed scores. Section II is spread nearly evenly from < 30% to 60-69%. Section V exhibits a secondary peak at 70-79%.

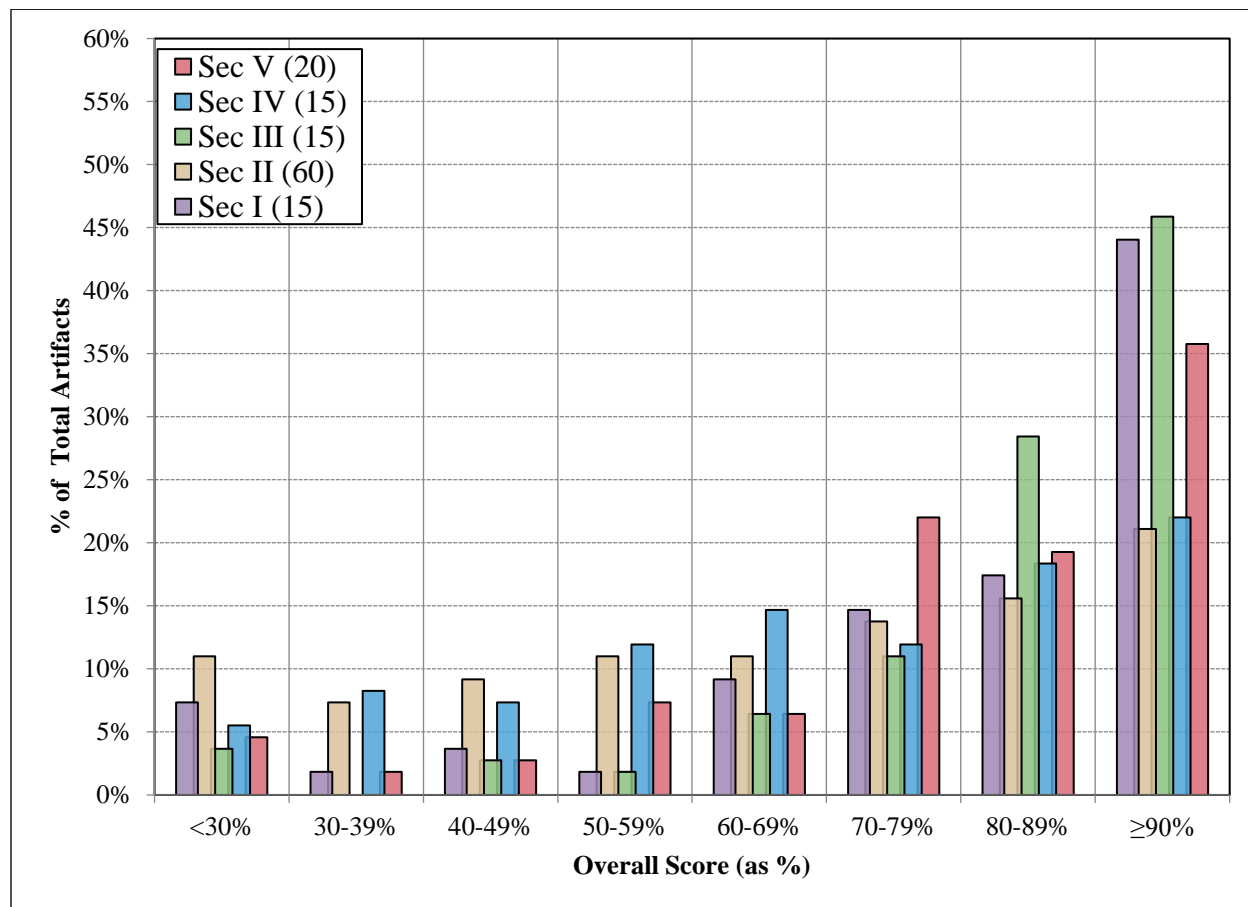


Figure 9. Histogram of SPN 1120 exam sections scores for spring 2021. Purple – Section I, Brown – Section II, Green – Section III, Blue – Section IV, and Red – Section V.

Under normal conditions, to describe the behavior of the section scores based on overall achievement, a color map, or binary raster image, is created by calculating the mean scores for each exam section as a function of combined score. Due to limited sample size, this could not be completed.

#### 3.1.3.2 Longitudinal Study

Further description of achievement over time in SPN 1120 is provided in Table 8 and Figure 11. 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 16 terms, Sections I through V for the Spring 2021 term rank 9<sup>th</sup>, 11<sup>th</sup>, 10<sup>th</sup>, 15<sup>th</sup>, and 14<sup>th</sup>, 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    |
| Spring 2021 (n=109) | 11.6      | 38.9       | 12.5        | 10.1       | 15.2      | 88.3    |

Table 8. 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 15 of 16 terms including the Spring 2021 term. Section II exhibits the lowest mean scores in 15 of 16 terms, spring 2019 being the only term in which it is not the lowest. For spring 2019, Section IV exhibits the lowest mean score.

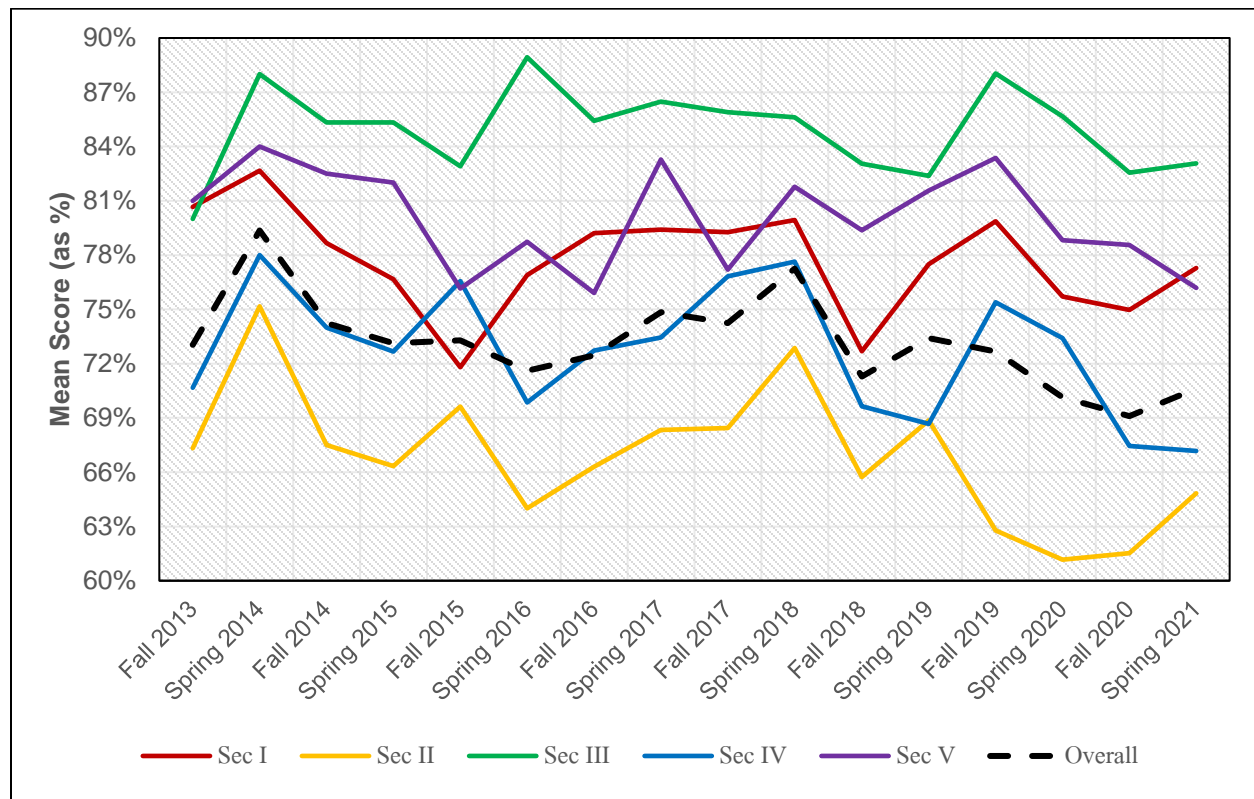


Figure 10. Comparison of mean scores (as percentage) for SPN 1120 through time from fall 2013 through the present.

### 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 12.

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 spring 2021, that gap has stabilized at 7.0.

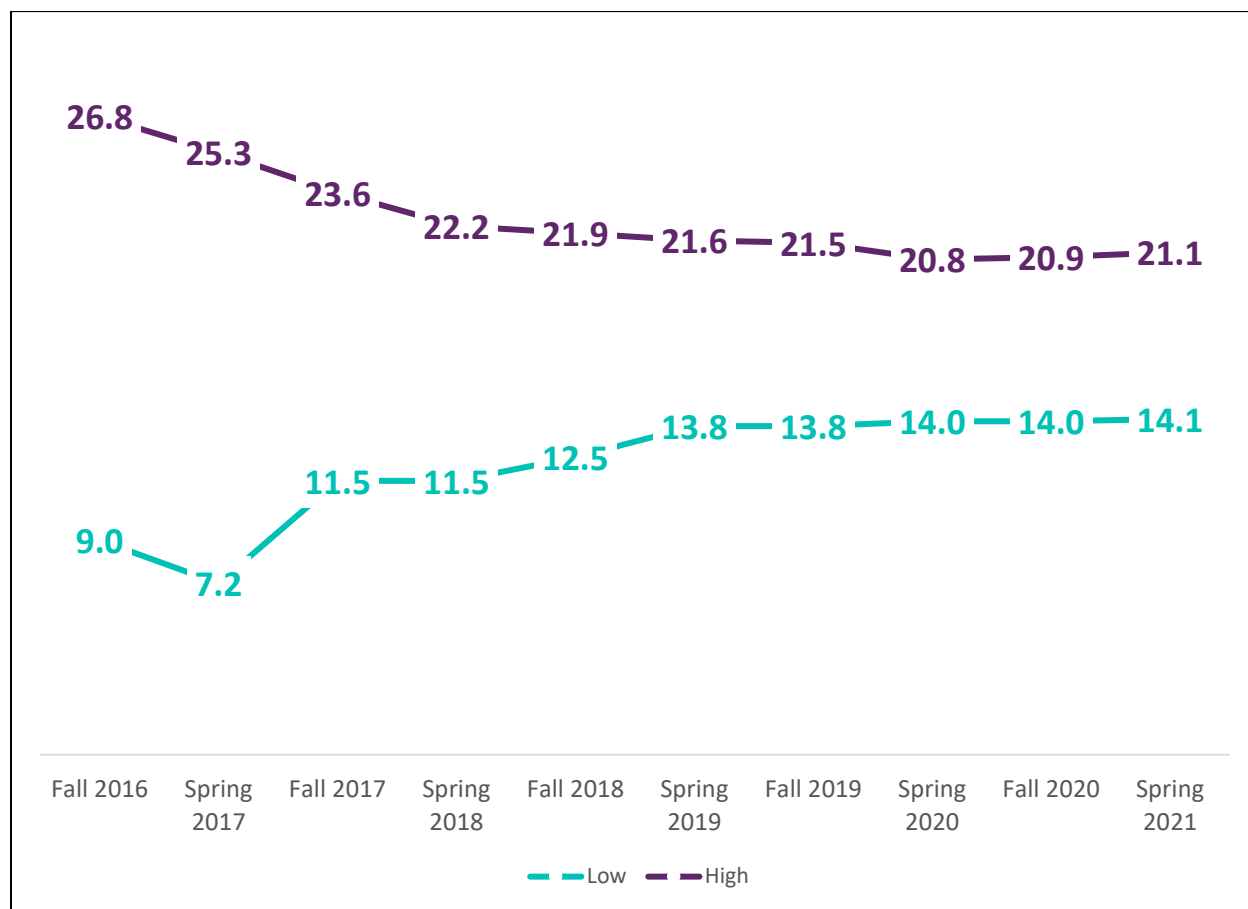


Figure 11. Comparison of achievement in entrance skills study assessment over time between highest scoring instructor (score listed is mean score) and lowest.

### 3.3 SPN 1121

#### 3.3.1 Descriptive Statistics, Learning Outcomes, and Learning Objectives

Eight sections of SPN 1121 Beginning Spanish II were offered in spring 2021. Of those, artifacts were collected from a common final from 7 of 8 sections. One section utilized a different scoring plan so scores could not be included in tally. A total of 171 students were enrolled in SPN 1121. Of those, 138 artifacts were collected representing a sample size of 81%.

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 9 below.

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 54% of artifacts score 70% or higher in the oral competency exam section (Section I) (Table 9, Figure 13). 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 41% of artifacts scored 70% or higher in Section II and 44% 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 44% of artifacts scored 70% or higher in Section V and 75% of artifacts scored 80% or higher in Section VII.

| <b>n = 138</b>                | <b>Section I<br/>(Oral)</b>   | <b>Section II<br/>(Reading)</b> | <b>Section III<br/>(Read/<br/>Write)</b> | <b>Section IV<br/>(Read/<br/>Write)</b> | <b>Section V<br/>(Writing)</b> | <b>Section VI<br/>(Reading)</b> | <b>Section<br/>VII<br/>(Writing)</b> | <b>Combined<br/>(Overall)</b> |
|-------------------------------|---|---------------------------------|--|---|--------------------------------|---------------------------------|--------------------------------------|-------------------------------|
| Goal                          | <i>80% of artifacts scored <math>\geq</math> 70% for all sections</i> |                                 |  |   |                                |                                 |                                      |                               |
| % above Goal                  | 54%   | 41%                             | 51%                                      | 65%                                     | 46%                            | 44%                             | 75%                                  |                               |
| Mean (as %)                   | 68%   | 59%                             | 62%                                      | 72%                                     | 61%                            | 60%                             | 76%                                  |                               |
| Median (as %)                 | 73%   | 62%                             | 70%                                      | 80%                                     | 67%                            | 67%                             | 89%                                  |                               |
| Section Score<br>Max Possible | 15  | 15                              | 40                                       | 15                                      | 12                             | 15                              | 20                                   | 132                           |
| Section Mean                  | 10.3  | 8.8                             | 24.7                                     | 10.8                                    | 7.3                            | 8.9                             | 15.2                                 | 86.0                          |
| Section Median                | 11  | 9.3                             | 28                                       | 12                                      | 8                              | 10                              | 17.8                                 | 93.0                          |

Table 9. Percentage of student achievement level as per SLOs (SPN 1121).

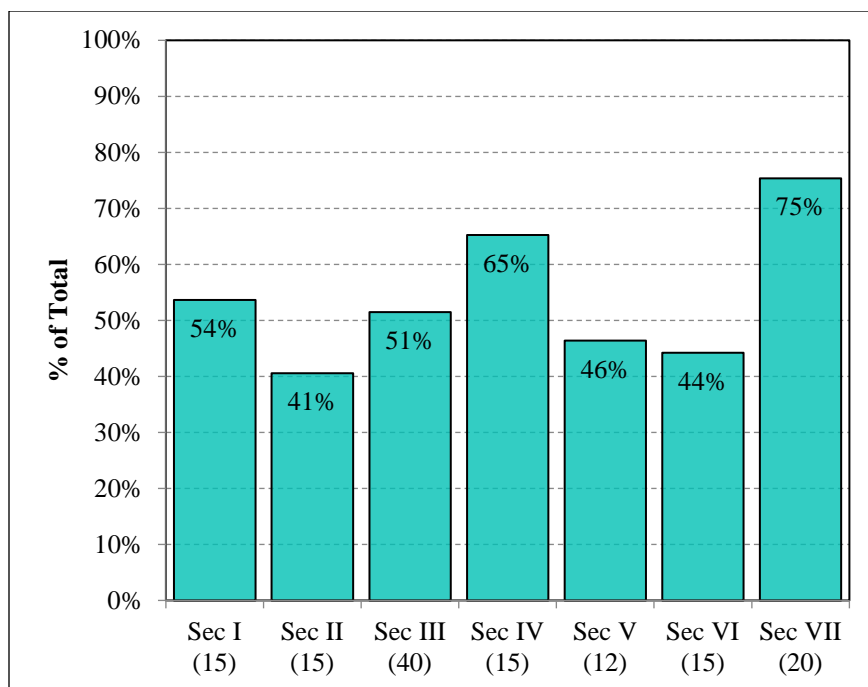


Figure 12. SLO achievement for SPN 1121 by exam section (Percentage of student achievement level as per SLOs). Purple denotes having met objective.

### 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 2021 so no comparison study between dual enrollment and traditional courses could be completed.

#### 3.3.2.2 Modality Comparison

During the Spring 2021 semester, 88 total asynchronous online artifacts were collected, along with 0 live online, 0 flex, 0 blended, and 50 traditional. A comparison of basic statistics is provided in Table 10. Traditional artifacts mean scores are 74.1% points, which is 14.0%-points higher than the next highest, asynchronous online. Results of the ANOVA exhibit a statistically significant difference between modalities (see Table 11). Therefore, we can reject the null hypothesis that the mean scores at each site are equal, 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 14.

|                   | <i>Traditional</i> | <i>Async Online</i> | <i>Live Online</i> | <i>Flex</i> | <i>Blended</i> |
|-------------------|--------------------|---------------------|--------------------|-------------|----------------|
| n                 | 50                 | 88                  | 0                  | 0           | 0              |
| Mean Score        | 97.8               | 79.3                | ~                  | ~           | ~              |
| Mean Score (as %) | 74.1%              | 60.1%               | ~                  | ~           | ~              |

Table 10. Comparison of descriptive statistics by modality.

| Source of Variation | Sum of squared differences | df  | Mean Squares | F <sub>obs</sub> | p-value | F <sub>crit</sub> |
|---------------------|----------------------------|-----|--------------|------------------|---------|-------------------|
| Between Sites       | 6262.9                     | 1   | 6262.9       | 12.29            | 0.001   | 3.91              |
| Within Sites        | 69,293.6                   | 136 | 509.5        |                  |         |                   |
| Total               | 75,556.5                   | 137 |              |                  |         |                   |

Table 11. Results of one-way ANOVA of mean scores by modality for SPN 1121.

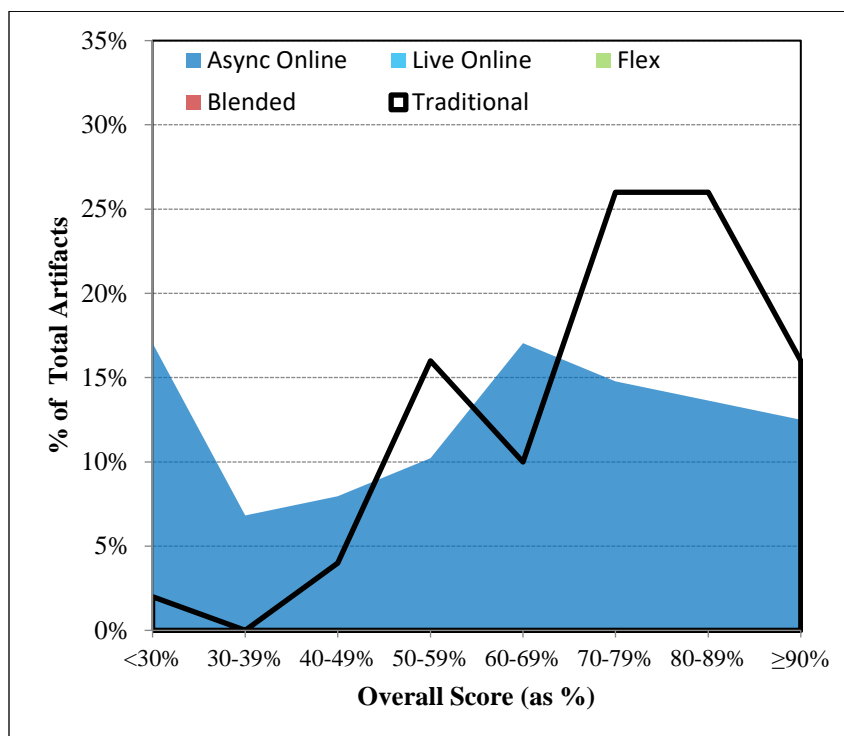


Figure 13. Score distribution by modality for SPN 1121.

### 3.3.2.3 Comparison by Campus/Site

Because only one traditional site offered sections during spring 2021 (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 15. Sections IV and VII exhibit peaks centered on ≥90%. Sections II, III, V, and VI exhibit multiple peaks at both a ≥90% and <30%. Section I exhibits peaks at both ≥90% and 60-69%.



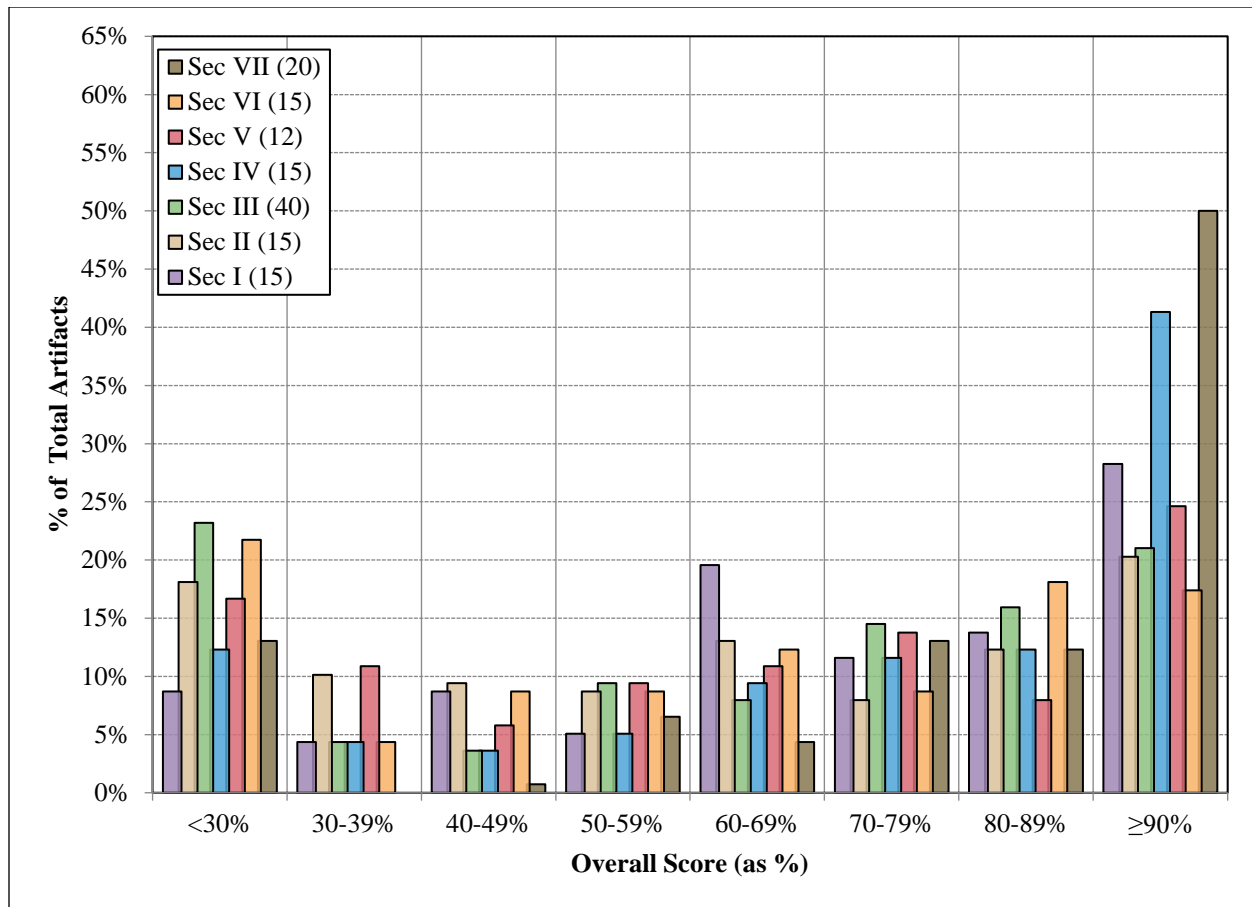


Figure 14. Histogram of SPN 1121 exam sections scores for spring 2021. Purple – Section I, Brown – Section II, Green – Section III, Blue – Section IV, Red – Section V, Orange – Section VI, and Black – Section VII.

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 is shown in Figure 16. Here, we can see that Sections IV and VII over perform at low overall scores compared to the other sections. For example, for artifacts with an overall mean score of 60-64%, the two sections exhibit mean scores of 80% and 79%, respectively, while the other sections range from 45% to 69%.

|        | Section I | Section II | Section III | Section IV | Section V | Section VI | Section VII |       |
|--------|-----------|------------|-------------|------------|-----------|------------|-------------|-------|
| ≥ 95%  | 100%      | 95%        | 89%         | 98%        | 95%       | 95%        | 99%         | Scale |
| 90-94% | 96%       | 84%        | 96%         | 97%        | 91%       | 84%        | 97%         | 100%  |
| 85-89% | 88%       | 84%        | 74%         | 95%        | 86%       | 84%        | 96%         | 90%   |
| 80-84% | 79%       | 73%        | 86%         | 90%        | 80%       | 72%        | 94%         | 80%   |
| 75-79% | 82%       | 64%        | 68%         | 83%        | 76%       | 73%        | 89%         | 70%   |
| 70-74% | 71%       | 69%        | 75%         | 77%        | 66%       | 58%        | 86%         | 60%   |
| 65-69% | 66%       | 61%        | 68%         | 74%        | 72%       | 52%        | 79%         | 50%   |
| 60-64% | 69%       | 58%        | 56%         | 80%        | 45%       | 52%        | 79%         | 40%   |
| 55-59% | 68%       | 54%        | 49%         | 61%        | 53%       | 42%        | 73%         | 30%   |
| 50-54% | 48%       | 35%        | 60%         | 61%        | 38%       | 53%        | 78%         |       |
| < 50%  | 38%       | 24%        | 24%         | 36%        | 24%       | 32%        | 37%         |       |

Figure 15. (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 12 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 16 terms, Sections I through VII for spring 2021 ranks 16<sup>th</sup>, 13<sup>th</sup>, 14<sup>th</sup>, 10<sup>th</sup>, 14<sup>th</sup>, 14<sup>th</sup>, and 13<sup>th</sup> highest, respectively.

|                 | Section I<br>Oral | Section II<br>Reading | Section III<br>Read/Write | Section IV<br>Read/Write | Section V<br>Writing | Section VI<br>Reading | Section VII<br>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    |
| Sp 2021 (n=138) | 10.3              | 8.8                   | 24.7                      | 10.8                     | 7.3                  | 8.9                   | 15.2                   | 86.0    |

Table 12. Comparison of mean scores for SPN 1121 for fall 2013 through the present.

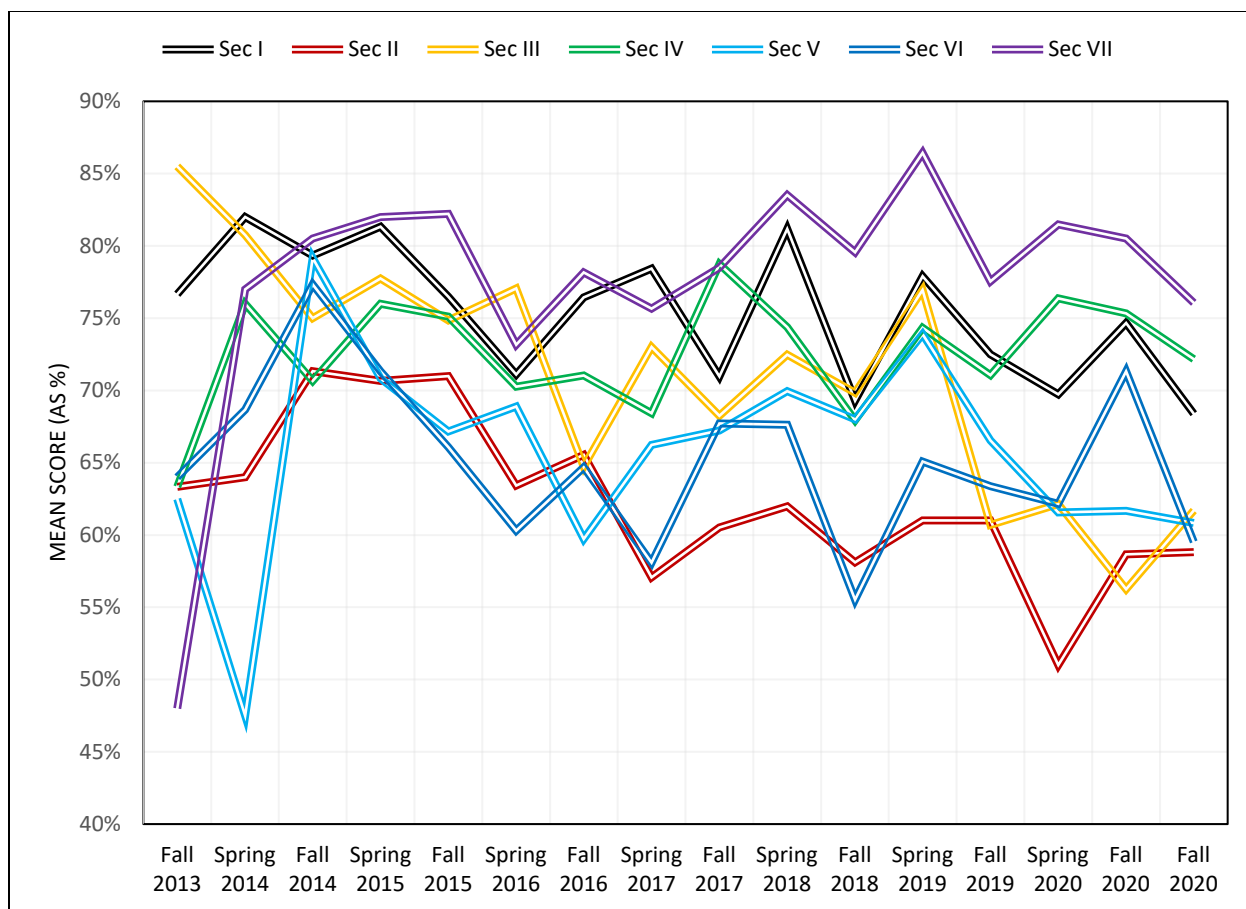


Figure 16. Comparison of mean scores (as percentage) for SPN 1121 through time.

## 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 to improve instruction. What follows is a drilldown of findings for both disciplines (French and Spanish) for the spring 2021 assessment.

### 4.1 FRENCH

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

1. During the Spring 2021 term, two sections of FRE 1120 were offered. An assessment first piloted in AY 2017-18 was recorded in both sections of FRE 1120 accounting for 20 artifacts, or 54% of the population.
2. In a study of Learning Outcome (LO) achievement, results exhibit mean scores across the five LOs ranging from 2.8 to 3.0, compared with 2.0-2.3 in fall 2020, 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 Spring 2021 term, no dual enrollment (concurrent) sections were run, so no comparison could be completed.
4. During the Spring 2021 semester, 16 total asynchronous online artifacts were collected, along with 0 live online, 4 flex, 0 blended, and 0 traditional. Asynchronous online mean scores are the highest of all modalities at 14.6/15. The lowest is Flex, at 13.8/15. A one-way analysis of variance is normally used to compare means by modality, however, sample size is too limited.
5. Only one site (Lee campus) had course sections run during spring 2021. The other section was online, 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. There are some which report a low understanding and yet score high (as seen in data point in upper left corner of graph).
7. In a longitudinal study, in six terms, LOs 1 through 5 for the Spring 2021 term rank 1<sup>st</sup>, 2<sup>nd</sup>, 1<sup>st</sup>, 2<sup>nd</sup>, and 2<sup>nd</sup>, respectively.

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

1. During the Spring 2021 term, four sections of FRE 1121 were offered. The assessment was recorded in all four, 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 45 artifacts, or 92% of the population.
2. In a study of Learning Outcome (LO) achievement, results exhibit mean scores across the five LOs ranging from 2.5 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. While one dual enrollment (concurrent) section was run during spring 2021, the data could not be accurately reported, so no comparison study between dual enrollment and traditional courses could be completed.
4. During the Spring 2021 semester, 29 total asynchronous online artifacts were collected, along with 5 live online, 12 flex, 0 blended, and 0 traditional. Asynchronous online mean scores are the highest of all modalities at 14.4/15. The lowest is Flex, at 11.4/15. A one-way analysis of variance found that there is a statistically significant difference between modalities ( $p=0.00001$ ).
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 is a moderate correlation between self-report and achievement.
7. In a longitudinal study, in four terms, LOs 1 through 5 for the Spring 2020 term rank 2<sup>nd</sup>, 1<sup>st</sup>, 3<sup>rd</sup>, 2<sup>nd</sup>, and 1<sup>st</sup>, respectively.

## 4.2 SPANISH

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

1. During the Spring 2021 semester 9 sections of SPN 1120 *Beginning Spanish I* were offered. Of those, artifacts from a common final were collected from 8 sections as one section did not utilize the same scoring. A total of 176 students were enrolled in SPN 1120. Of those, 109 artifacts were collected representing a sample size of 62% of the population.

2. 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.
3. Achievement of 80% of artifacts scoring 70% or better (SLO 2): Achievement was partially met. Results exhibit 50% of artifacts scored 70% or higher in Section II and 85% 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 52% of artifacts scored 70% or higher in Section IV and 77% of artifacts scored 70% or higher in Section V.
5. No dual enrollment (concurrent) sections of the course were run during spring 2021 so no comparison study between dual enrollment and traditional courses could be completed.
6. During the Spring 2021 semester, 49 total asynchronous online artifacts were collected, along with 21 live online, 31 flex, 0 blended, and 8 traditional. Traditional artifacts are 6.4%-points higher than the next highest, although sample size is only n=8. Asynchronous online, live online, and flex range only 1.9%-points between the three. Results of the ANOVA do not exhibit a statistically significant difference between sites.
7. Only one traditional site offered course sections this term, as a result, no comparison by site could be completed.
8. In a study of score distribution by section, all sections exhibit peaks above 90%. Sections II and V exhibit more widely distributed scores. Section II is spread nearly evenly from < 30% to 60-69%. Section V exhibits a secondary peak at 70-79%.
9. Under normal conditions, to describe the behavior of the section scores based on overall achievement, a color map, or binary raster image, is created by calculating the mean scores for each exam section as a function of combined score. Due to limited sample size, this could not be completed.
10. In a longitudinal study of data distribution through time, in 16 terms, Sections I through V for the Spring 2021 term rank 9<sup>th</sup>, 11<sup>th</sup>, 10<sup>th</sup>, 15<sup>th</sup>, and 14<sup>th</sup>, respectively.
11. In a study assessing student skills and retention of materials from SPN 1120 prior to beginning SPN 1121, over time, Section III exhibit the highest mean scores in 15 of 16 terms including the Spring 2021 term. Section II exhibits the lowest mean scores in 15 of 16 terms, spring 2019 being the only term in which it is not the lowest. For spring 2019, Section IV exhibits the lowest mean score.

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

1. Eight sections of SPN 1121 Beginning Spanish II were offered in spring 2021. Of those, artifacts were collected from a common final from 7 of 8 sections. One section utilized a different scoring plan so scores could not be included in tally. A total of 171 students were enrolled in SPN 1121. Of those, 138 artifacts were collected representing a sample size of 81%.
2. Achievement of 80% of artifacts scoring 70% or better (SLO 1): Achievement was not met as results exhibit 54% 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 not met. Results exhibit 41% of artifacts scored 70% or higher in Section II and 44% 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 44% of artifacts scored 70% or higher in Section V and 75% of artifacts scored 80% or higher in Section VII.
5. No dual enrollment (concurrent) sections of the course were run during spring 2021 so no comparison study between dual enrollment and traditional courses could be completed.

6. During the Spring 2021 semester, 88 total asynchronous online artifacts were collected, along with 0 live online, 0 flex, 0 blended, and 50 traditional. Traditional artifacts mean scores are 74.1% points, which is 14.0%-points higher than the next highest, asynchronous online. Results of the ANOVA exhibit a statistically significant difference between modalities.
7. Because only one traditional site offered sections during spring 2021 (Thomas Edison campus) no comparison by site could be completed.
8. In a study of score distribution by section, Sections IV and VII exhibit peaks centered on  $\geq 90\%$ . Sections II, III, V, and VI exhibit multiple peaks at both a  $\geq 90\%$  and  $< 30\%$ . Section I exhibits peaks at both  $\geq 90\%$  and 60-69%.
9. In a study of section score distribution based on overall score, we can see that Sections IV and VII over perform at low overall scores compared to the other sections. For example, for artifacts with an overall mean score of 60-64%, the two sections exhibit mean scores of 80% and 79%, respectively, while the other sections range from 45% to 69%.
10. In a longitudinal study of data distribution through time, in 16 terms, Sections I through VII for spring 2021 ranks 16<sup>th</sup>, 13<sup>th</sup>, 14<sup>th</sup>, 10<sup>th</sup>, 14<sup>th</sup>, 14<sup>th</sup>, and 13<sup>th</sup> highest, respectively.

## 5 REFERENCES

---

- Cohen, J. 1988. Statistical power analysis for the behavioral sciences (2<sup>nd</sup> 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 (2<sup>nd</sup> 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/](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.