

Foreign Language Assessment Report Spring 2016

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

1 INTRODUCTION

Florida SouthWestern's Foreign Language Department employs a common course assessment to measure student progress in course level objectives, a practice shown to be effective in establishing data driven instruction (Hall, 2010). Courses included in assessment are: FRE 1120 *Elementary French I*, FRE 1121 *Elementary French II*, SPN 1120 *Beginning Spanish I*, and SPN 1121 *Beginning Spanish II*. Through achievement of the courses students will acquire and demonstrate competency in speaking, writing, reading comprehension and listening comprehension in standard Spanish or French at the beginner's level. The assessment outcomes outlined below define the method of assessment for each course assessment as well as measure current Student Learning Objectives (SLOs) and identify areas for future SLOs to be assessed. Additionally, the plan provides information on achievement levels of Dual Enrollment artifacts compared with non-Dual Enrollment, as well as Online artifacts compared with traditional artifacts as highlighted in the course level assessment plan. This report provides achievement analysis for both fall 2015 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, Director of Academic Affairs Assessment, Academic Affairs (jfvangaalen@fsw.edu; x16965).

2 FRENCH

During spring 2016 term, artifacts were collected from 2 of 4 sections of FRE 1120. Sections were taught by two different instructors, one of which was represented in the artifacts collected. A total of 49 students were enrolled in FRE 1120. Of those, 17 artifacts were collected representing a sample size of 35% of the population. Five sections of FRE 1121 *Elementary French II* were offered. Of those, artifacts were collected from one section. All sections were taught by four different instructors. A total of 59 students were enrolled in FRE 1121. Of those, 12 artifacts were collected representing a sample size of 20%. However, results did not match the common assessment exam and therefore could not be analyzed. Assessment is expected to resume for FRE 1121, and continue with FRE 1120, in fall 2016.

2.1 FRE 1120

2.1.1 Descriptive Statistics & Learning Objectives

Using a common course assessment the FSW French faculty defined three areas of interest for evaluation that apply to FRE 1120, oral comprehension, reading, and writing. Specific objectives include acquiring and demonstrating competency in speaking, writing, reading comprehension and listening comprehension in standard French at the beginner's level. Student competency will be limited primarily to the thematic topics covered over the period of the course (i.e. work life, significant historic and personal events, media, technology, interpersonal relationships and travel). The Student Learning Outcomes (SLOs) and their measure of success are:

- SLO 1: Students will be able to understand spoken French. 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 (not yet included in data for spring 2016 exam). This tool was piloted in spring 2016 and is intended to be fully implemented during fall 2016 so no data yet exists to measure this planned goal.
- SLO 2: Students will be able to understand written French. 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 IX).
- SLO 3: Students will be able to write effectively in the French 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 I).

The faculty established measure of success for SLO 2, 80% of students scoring 70% or higher in Section IX, was met nearly met as results exhibit 71% of artifacts score 70% or higher in the reading competency exam section (Section IX) (Table 1). The faculty established measure of success for SLO 3, 80% of students scoring 70% or higher in Section I, was met. Results exhibit 94% of artifacts scored 70% or higher in Section I. For a graphical representation of SLO achievement, see Figure 1. The highest achieving SLO is Section I (96%), while the lowest is Section V at 53% achieving 70% or greater.

n = 17	Sec I	Sec II	Sec III	Sec IV	Sec V	Sec VI	Sec VII	Sec VIII	Sec IX	Combined (Overall)
Goal	<i>80% of artifacts scored \geq70% for all sections</i>									
% above Goal	94%	94%	59%	71%	53%	88%	88%	59%	71%	
Mean (as %)	94%	92%	74%	84%	76%	92%	84%	68%	79%	70%
Median (as %)	100%	100%	77%	88%	88%	100%	80%	79%	90%	71%
Section Score Max	10	6	26	20	8	10	10	14	20	124
Section Mean	9.4	5.5	19.2	16.7	6.1	9.2	8.4	9.5	15.7	87.3
Section Median	10	6	20	17.5	7	10	8	11	18	89

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

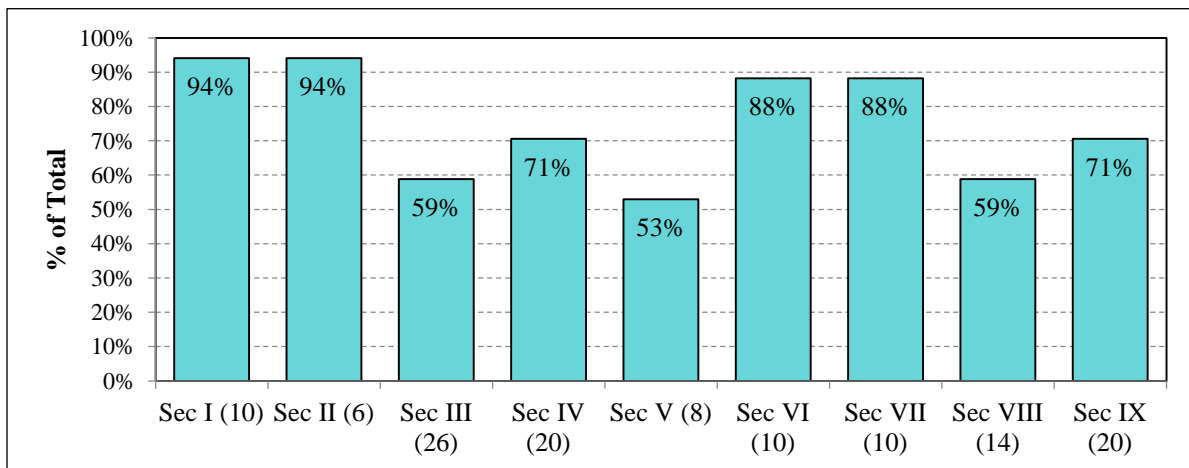


Figure 1. SLO achievement for FRE 1120 by exam section (Percentage of student achievement level as per SLOs).

2.1.2 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.

2.1.2.1 Dual Enrollment to Non-Dual Enrollment Comparison

No dual enrollment sections of FRE 1120 were offered during spring 2016 so no comparison study could be completed.

2.1.2.2 Online to Traditional Comparison

During the spring 2016 semester, two online sections were run with data collected from one. From that course section, 9 total online artifacts were collected from FRE 1120 and 17 traditional artifacts were collected from FRE 1120. A comparison of basic statistics is provided in Table 2. Online artifacts mean scores are 35.9 lower than traditional artifacts. Differences in the means were tested for significance using a Welch's t-test according to standard methods (Davis, 1973; McDonald, 2009; Wilkinson, 1999) and were found to be statistically significantly different. Therefore, we can reject the null hypothesis that the differences in the means of the online and traditional artifacts are equal to 0, and we can conclude this with a 95% confidence that the differences in scores are not solely due to chance.

Effect size was calculated using a method devised by Rosenthal and Rosnow (1991) for meta-analytical purposes in potential comparisons with other institutions (Lipsey and Wilson, 1993). The results exhibit what Cohen (1988) would consider a large effect size. In other words, non-overlap score distribution from online artifacts to traditional artifacts is approximately 72%. For a graphical representation of this see Figure 2.

df = 24		
Online mean	63.9	
Online standard deviation	25.4	
Traditional mean	99.8	
Traditional standard deviation	16.9	
Effect size	1.56	
p-value	0.002	

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

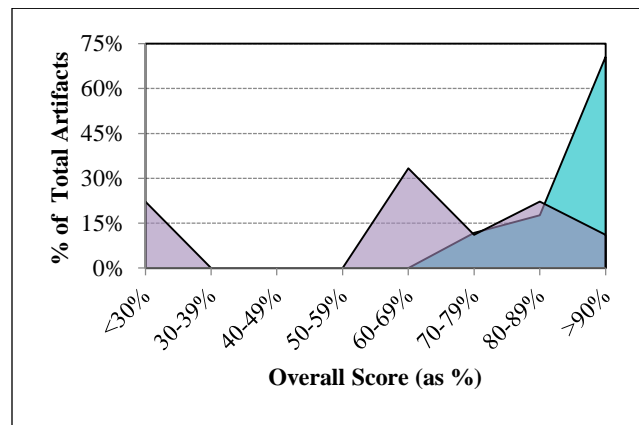


Figure 2. Score distribution for online (purple) and traditional (teal) artifacts of FRE 1120.

2.1.2.3 Comparison by Campus/Site

Artifacts were collected from only one site (Thomas Edison Campus) other than FSW Online. As a result, comparisons are included in the appropriate sections above.

2.1.3 Data Distribution & Longitudinal Study

2.1.3.1 Data Distribution

A histogram depicting the distribution of scores across each exam section is shown in Figure 3. All sections except for Section VIII exhibit peaks at $\geq 90\%$. Section VIII exhibits a peak in the 80-89 scoring bin. Section VIII also exhibits a spike at $< 30\%$ scores while sections IV, V, and 9 exhibit spikes at 50-69.

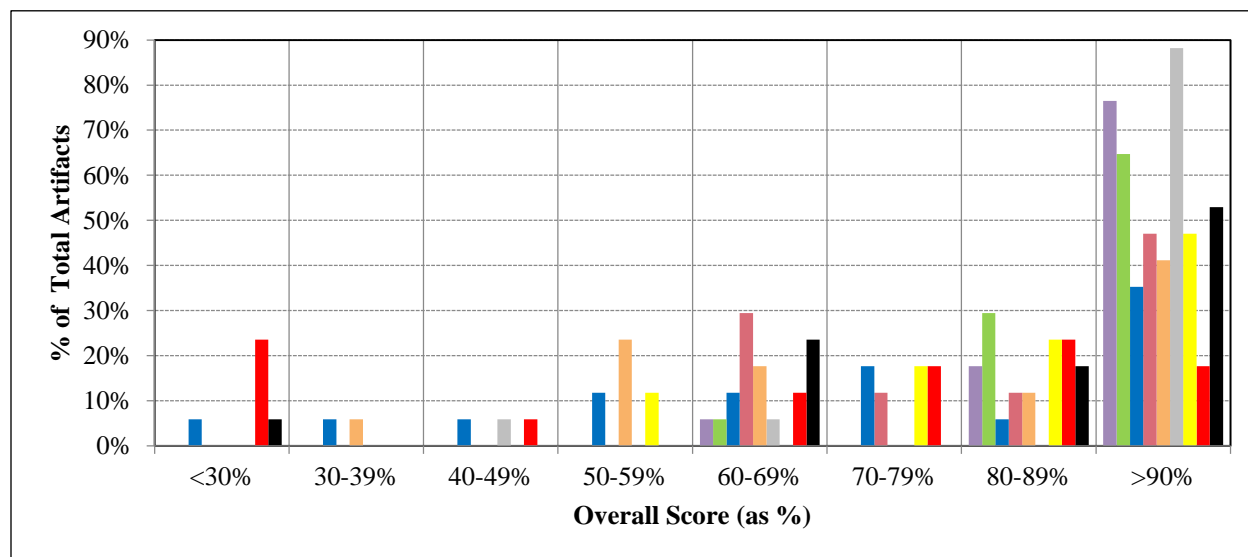


Figure 3. Histogram of FRE 1120 exam section scores for spring 2016. Purple – Section I, Green – Section II, Blue – Section III, Magenta – Section IV, Peach – Section V, Gray – Section VI, Yellow – Section VII, Red – Section VIII, and Black – Section IX.

2.1.3.2 Longitudinal Study

Description of achievement over time in FRE 1120 is provided in a comparison of fall 2014 through fall 2015 data (Table 3 and Figure 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). With only one term to compare, interpretation of trends is limited. From spring-to-spring, four sections exhibit slight increases while the remaining five exhibit declines (Figure 5). It is important to note that online data was not reported by section and is therefore only represented in overall score (see large shift in overall score from spring-to-spring in Table 4 compared with slight changes by section).

	Section Max	Fall 2014, n=10	Spring 2015, n=15	Fall 2015, n=67	Spring 2016, n=17
Section I	10	9.0	9.3	9.1	9.4
Section II	6	5.5	5.5	5.0	5.5
Section III	26	20.2	20.8	19.4	19.2
Section IV	20	16.6	17.0	16.7	16.7
Section V	8	6.8	5.2	6.0	6.1
Section VI	10	8.6	9.2	8.1	9.2
Section VII	10	8.2	7.7	8.4	8.4
Section VIII	14	8.7	9.3	9.6	9.5
Section IX	20	17.5	16.5	16.0	15.7
Overall	124	98.8	100.5	89.4	87.3

Table 3. Comparison of mean scores for FRE 1120 for fall 2014 through spring 2016.

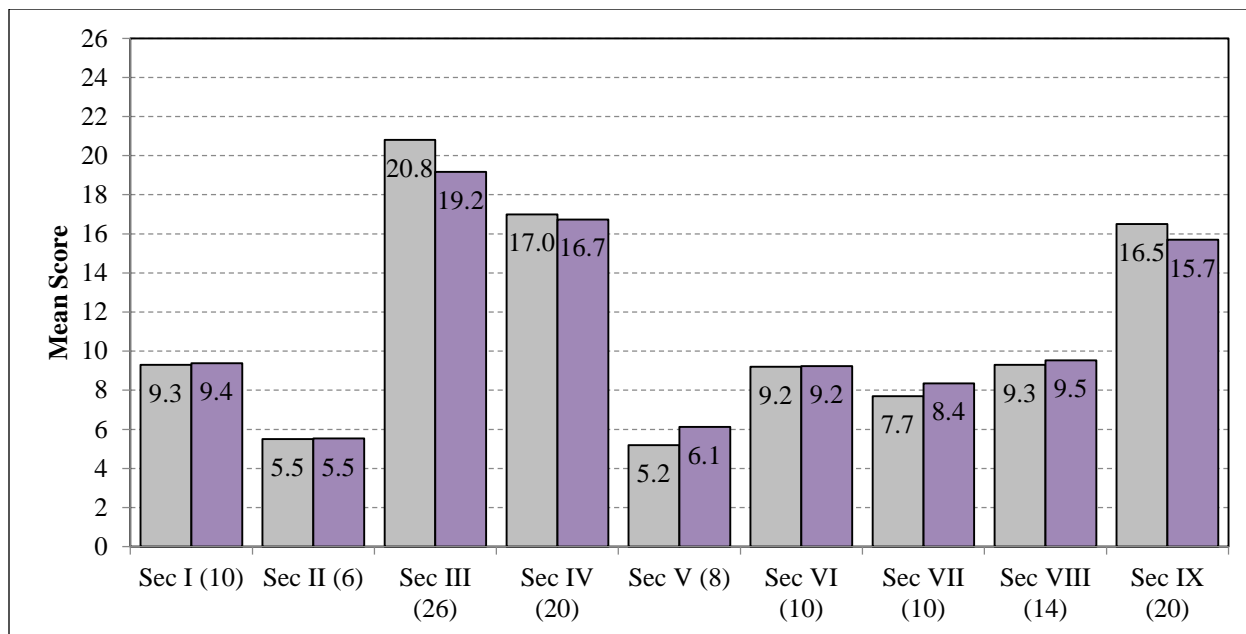


Figure 4. Comparison of mean scores for FRE 1120 through time for spring 2015 (dark gray) to spring 2016 (purple).

2.2 FRE 1121

Assessment is expected to resume for FRE 1121, and continue with FRE 1120, in fall 2016.

3 SPANISH

During the spring 2016 semester 10 sections of SPN 1120 Beginning Spanish I were offered. Of those, artifacts from a common final were collected from eight sections. Sections were taught by 10 different instructors, five of which were represented in the artifacts collected. A total of 230 students were enrolled in SPN 1120. Of those, 141 artifacts were collected representing a sample size of 61% of the population. Twelve sections of SPN 1121 Beginning Spanish II were offered. Of those, artifacts were collected from a common final from eight sections. Sections were taught by nine different instructors, five of which were represented in the artifacts collected. A total of 217 students were enrolled in SPN 1121. Of those, 109 artifacts were collected representing a sample size of 50%.

3.1 SPN 1120

3.1.1 Descriptive Statistics & Learning Objectives

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 measure 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 4). 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 45% of artifacts scored 70% or higher in Section II and 92% 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 also partially met. Results exhibit 50% of artifacts scored 70% or higher in Section IV and 80% of artifacts scored 70% or higher in Section V. For a graphical representation of SLO achievement, see Figure 5.

n = 141	Section I (Oral)	Section II (Reading)	Section III (Reading)	Section IV (Written)	Section V (Written)	Combined (Overall)
Goal	80% of artifacts scored \geq 70% for all sections					
% above 70%	71%	45%	92%	50%	80%	
Mean (as %)	77%	64%	89%	70%	79%	72%
Median (as %)	80%	65%	93%	67%	80%	72%
Section Score Max	15	60	15	15	20	125
Section Mean	11.5	38.4	13.3	10.5	15.7	89.5
Section Median	12	39	14	10	15	90.5

Table 4. Percentage of student achievement level as per SLOs (SPN 1120).

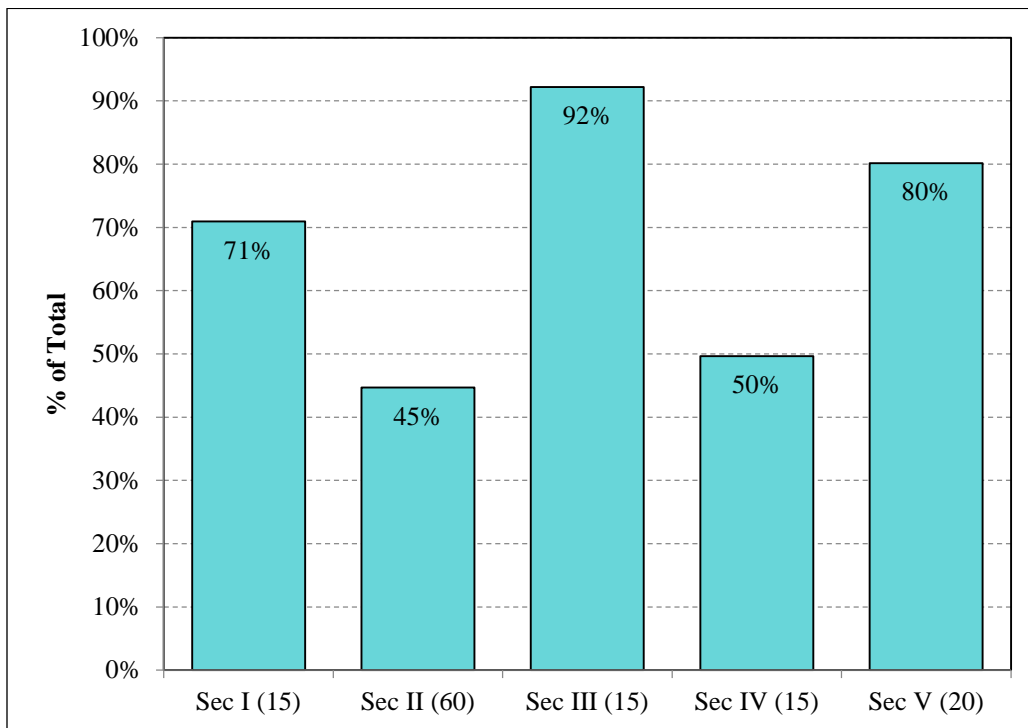


Figure 5. SLO achievement for SPN 1120 by exam section (Percentage of student achievement level as per SLOs).

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 to Non-Dual Enrollment Comparison

No dual enrollment sections of the course were run during spring 2016 so no comparison study between dual enrollment and traditional courses could be completed.

3.1.2.2 Online to Traditional Comparison

During the spring 2016 semester, two online sections reported data using the newly developed shell which includes the common final exam. From those course sections, 30 total online artifacts were collected from SPN 1120 and 111 traditional artifacts were collected from SPN 1120. A comparison of basic statistics is provided in Table 5. Online artifacts mean scores are 5.1 higher than traditional artifacts. Differences in the means were tested for significance using a Welch's t-test according to standard methods (Davis, 1973; McDonald, 2009; Wilkinson, 1999) and were found to not be statistically significantly different (online sample size was only 5). Therefore, we cannot reject the null hypothesis that the differences in the means of the online and traditional artifacts are equal to 0, and we cannot conclude this with a 95% confidence that the differences in scores are not solely due to chance.

Effect size was calculated using a method devised by Rosenthal and Rosnow (1991) for meta-analytical purposes in potential comparisons with other institutions (Lipsey and Wilson, 1993). The results exhibit what Cohen (1988) would consider a small effect size. In other words, non-overlap score distribution from online artifacts to traditional artifacts is approximately 20%. For a graphical representation of this see Figure 6.

df = 139		
Online mean		75.6
Online standard deviation		15.2
Traditional mean		70.5
Traditional standard deviation		16.9
Effect size		-0.27
p-value		0.118

Table 5. Comparison of mean scores for online and traditional artifacts. Positive effect sizes indicate a higher mean score for traditional artifacts.

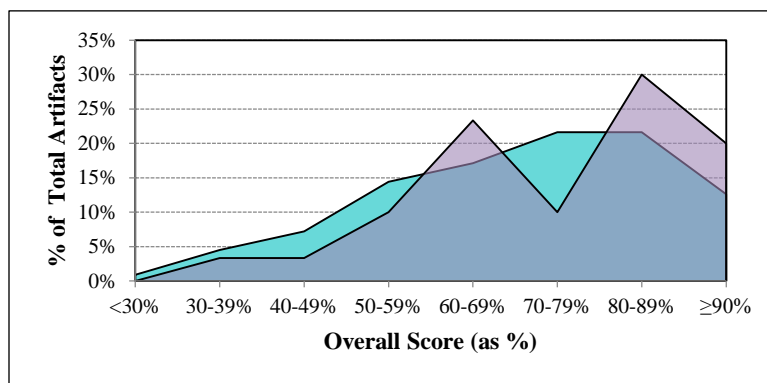


Figure 6. Score distribution for online (purple) and traditional (teal) artifacts of SPN 1120.

3.1.2.3 Comparison by Campus/Site

Because only two sites reported (online and Thomas Edison) any comparison is duplicative of a comparison of online to traditional artifacts. As such, for site comparison, please review Section 3.1.2.2.

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 7. Sections I, III, and V exhibit peaks above 90% with the distribution trailing down with decreasing score. Sections II and IV, however, exhibit more widely distributed scores. Section II exhibits between 10% and 15% of artifacts scoring in each bin from 40-49% up through $\geq 90\%$. Section IV exhibits a bimodality (two peaks) centered on $\geq 90\%$ and 60-69%.

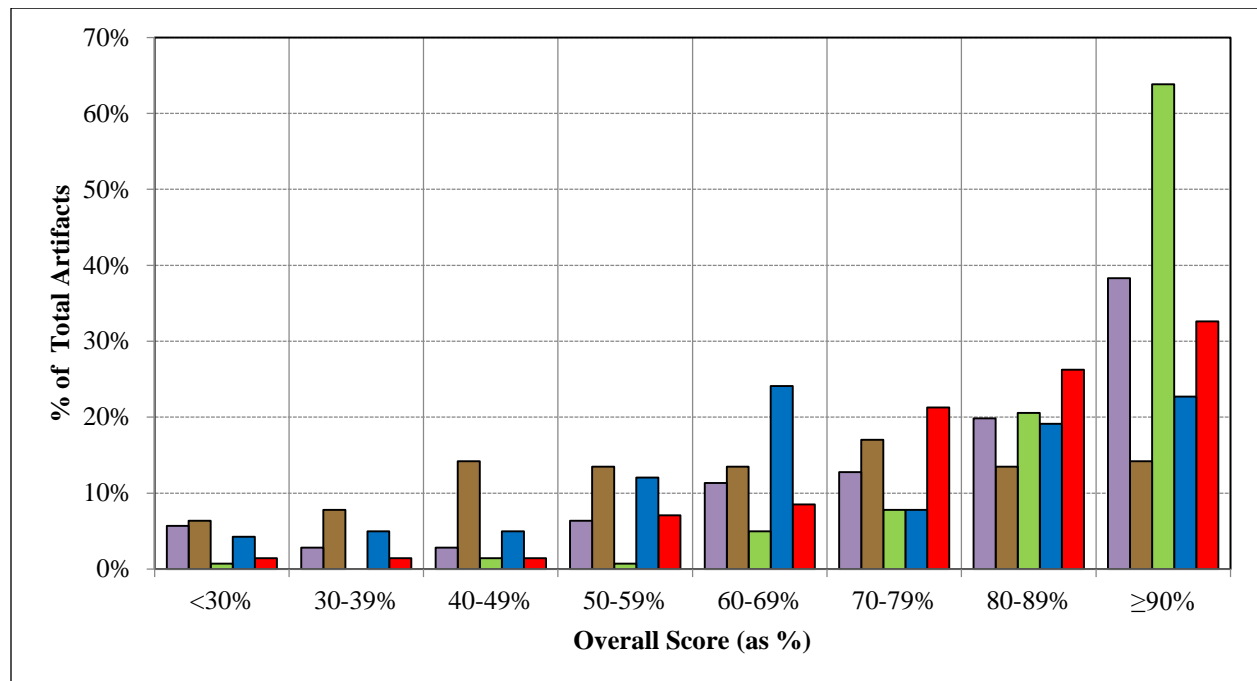


Figure 7. Histogram of SPN 1120 exam sections scores for spring 2016. Purple – Section I, Brown – Section II, Green – Section III, Blue – Section IV, and Red – Section V.

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 8). 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 8 shows that Section II is consistently the lowest performing compared to other sections between the ranges of 70%-84%. For example, in the 70-74% range, the mean score for Section II is 55%, while the other four sections range from 66-88%. Additionally, Section III, and to a lesser extent Section V, is over performing at the lowest overall scores. At the 55-59% range, the Section III and V mean scores are 81% and 60%, respectively, whereas other sections range from 41%-53%.

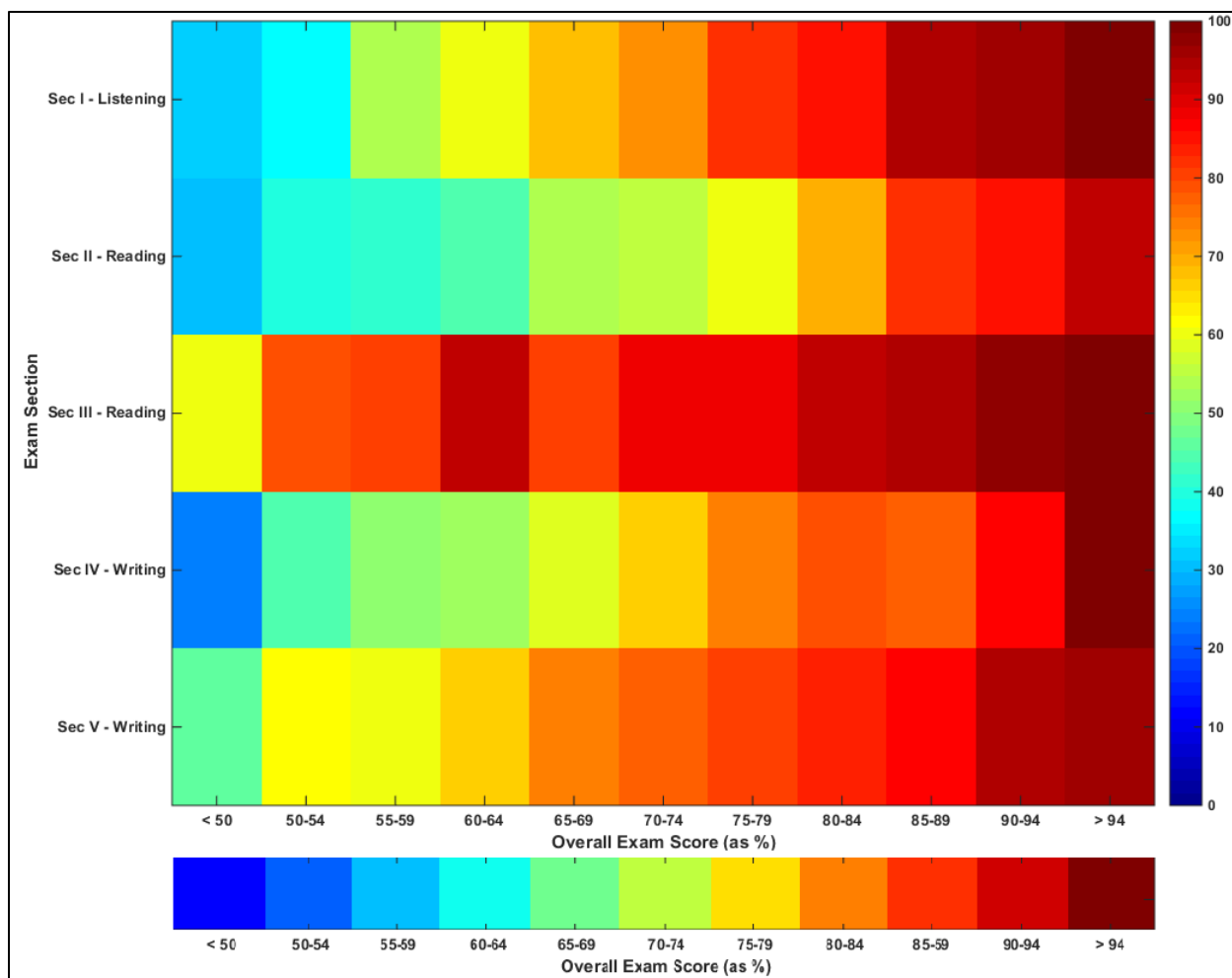


Figure 8. (Top) Colormap of mean scores for each exam section based on overall scoring bin for SPN 1120. (Bottom) Comparison exam section if section score percentage is the same as overall (i.e. artifact score is equally distributed across all sections). An exam section with hotter colors (reds/yellows) compared with the bottom bar means that section achievement exceeds the overall score and is an area of strength. An exam section with colder colors (blues/greens) compared with the bottom bar means that section achievement is lower than the overall score and is therefore an area of weakness.

3.1.3.2 Longitudinal Study

Further description of achievement over time in SPN 1120 is provided in a comparison of fall 2013 through spring 2016 data (Table 6 and Figure 9). 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). Through time, Sections II, IV, and V exhibit steady decreases. Concomitantly, Sections I and III exhibit no noticeable trend. The interpretation of the data, however, might be hindered by the samples. Unlike fall terms, where sampling has improved each year to capture an ever increasing percentage of the population, spring terms have been variable. Spring 2014 included 90 samples, which then dropped to 73 samples before jumping to 141 in spring 2016. Moreover, the smaller sample size in spring 2015 overrepresented one particular instructor, which may skew results and interpretation.

	<i>Section Max</i>	Fall 2013 n=58	Spring 2014 n=90	Fall 2014 n=93	Spring 2015 n=73	Fall 2015 n=122	Spring 2016 n=141
Section I (Oral)	15	12.1	12.4	11.8	11.5	10.8	11.5
Section II (Reading)	60	40.4	45.1	40.5	39.8	41.8	38.4
Section III (Reading)	15	12.0	13.2	12.8	12.8	12.4	13.3
Section IV (Written)	15	10.6	11.7	11.1	10.9	11.5	10.5
Section V (Written)	20	16.2	16.8	16.5	16.4	15.2	15.7
Combined (Overall)	125	91.3	99.2	92.8	91.4	91.6	89.5

Table 6. Comparison of mean scores for SPN 1120 for fall 2013 through spring 2016.

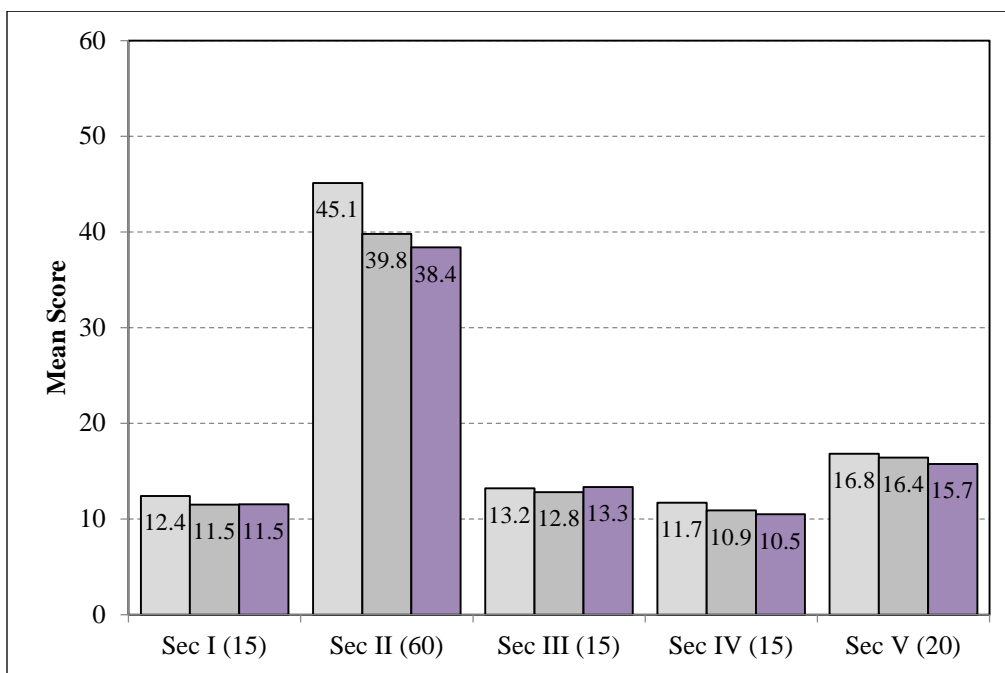


Figure 9. Comparison of mean scores for SPN 1120 through time from spring 2014 (light gray), through spring 2015 (dark gray), to spring 2016 (purple).

If we compare colormaps from previous years to spring 2016 the consistency of sections can be reviewed over time (Figure 10). Sections III and V exhibit increases in the lower overall scores over time. In reviewing the figures we see that both Section III and V exhibit hotter colors (higher scores) at lower overall scores (left side of the color map). From spring 2015 results exhibit a teal color for Section III (approximate Section II mean score 45% at overall score < 50%). For fall 2015, results for Section III exhibit a green color (approximate Section II mean score of 55% at overall score < 50%). And finally, for spring 2016 results exhibit a yellow color (approximate Section II mean score of 62% at overall score < 50%). This same trend is exhibited in Section V although it is less distinct.

A second trend is exhibited in Section II at higher overall scores. In spring 2015, the mean score for Section II at overall scores of 70-74% is approximately 45%. For fall 2015 that number increases to

approximately 60% before settling 55% in spring 2016. Recall that, in an exam where all sections exhibit equal difficulty, colormaps should exhibit a uniform pattern (see bottom of Figure 12 for example).

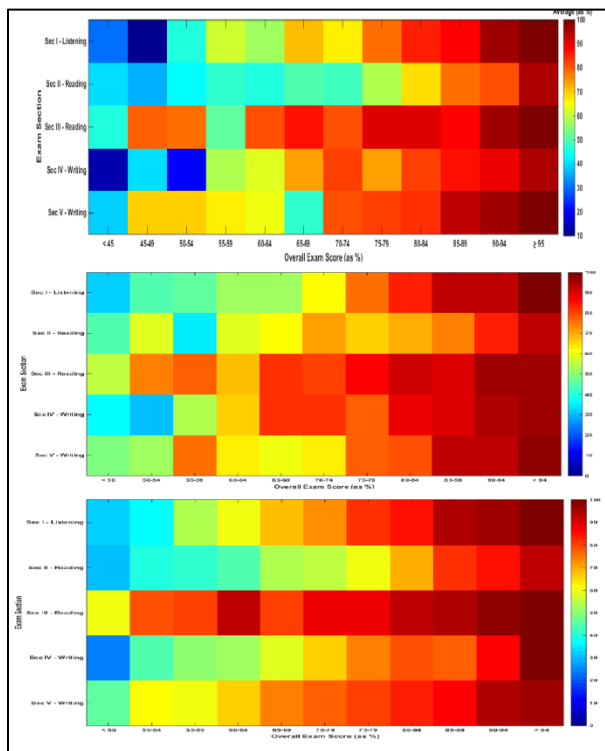


Figure 10. Comparison of spring 2015 (top), fall 2015 (middle) and spring 2016 (bottom) colormaps of SPN 1120.

3.2 ENTRANCE SKILLS STUDY

During 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 assess student skills and retention of materials from SPN 1120 prior to beginning SPN 1121. The assessment consisted 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 included in the fall 2015 assessment report. This program was continued in summer session courses to be included here.

The assessment was administered in three sections of SPN 1121 on both the Thomas Edison and Collier campuses. Results from the spring 2016 study that was originally included in the fall 2015 assessment report are shown in Figure 11 for review purposes. Results from the summer sessions are shown in Figure 12. Note that instructor numbers are not common for figures (i.e Instructor #1 is not necessarily the same as Instructor #1 across terms).

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. Much like the previous study, the study completed at the beginning of the summer sessions exhibits a wide range of mean scores across instructor (Figure 12). Results exhibit achievement levels based on previous instructor spanning as low as 9.0/30 to as high as 26.8/30. Most

glaring of these is instructor #6, which not only exhibits one of the lower scores (9.4/30) but is also the largest sample size (n=13), adding reliability to the results.

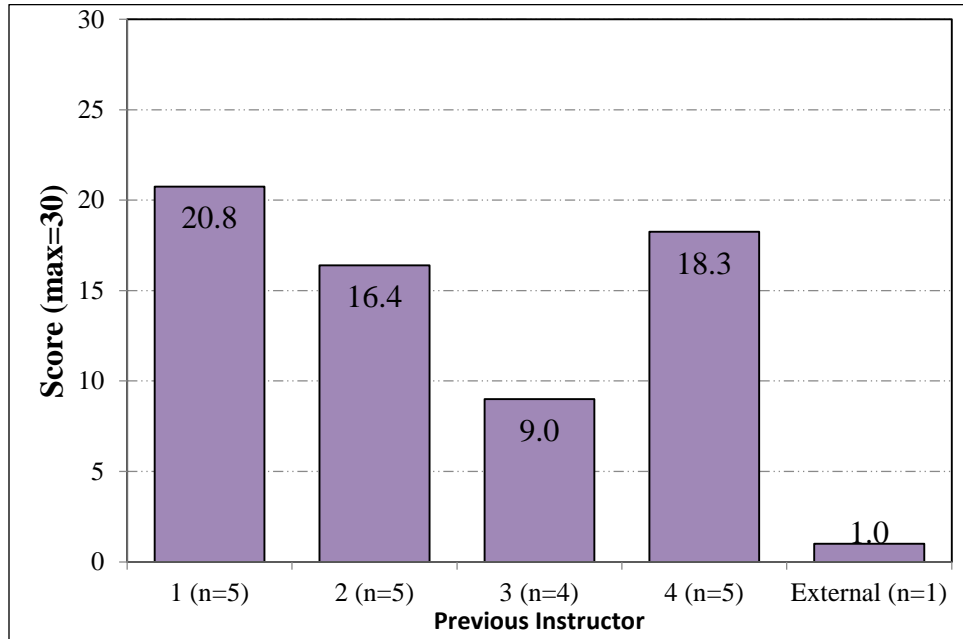


Figure 11. Comparison of achievement in entrance skills study assessment by instructor conducted at the beginning of spring 2016.

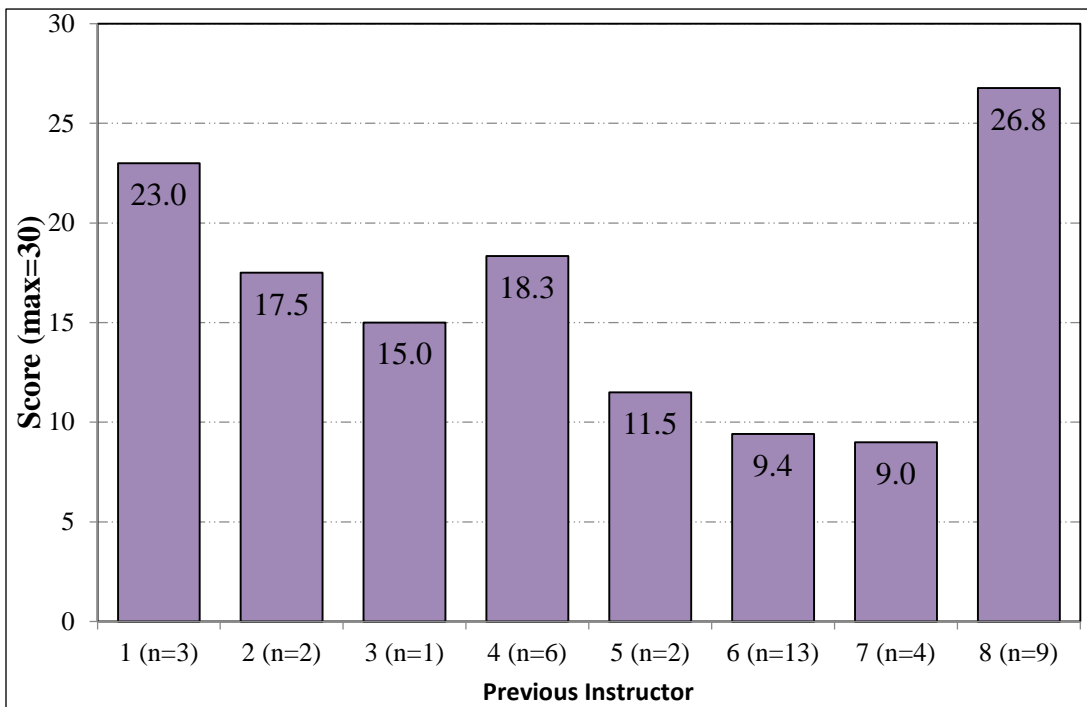


Figure 12. Comparison of achievement in entrance skills study assessment by instructor conducted at the beginning of summer 2016.

3.3 SPN 1121

3.3.1 Descriptive Statistics & Learning Objectives

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 not met as results exhibit 63% of artifacts score 70% or higher in the oral competency exam section (Section I) (Table 7). 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 49% of artifacts scored 70% or higher in Section II and 49% 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 not met. Results exhibit 64% of artifacts scored 70% or higher in Section V and 70% of artifacts scored 70% or higher in Section VII. For a graphical representation of SLO achievement, see Figure 13.

n = 109	Section I (Oral)	Section II (Reading)	Section III (Read/ Write)	Section IV (Read/ Write)	Section V (Writing)	Section VI (Reading)	Section VII (Writing)	Combined (Overall)
Goal	80% of artifacts scored $\geq 70\%$ for all sections							
% above Goal	63%	49%	72%	58%	64%	49%	70%	
Mean (as %)	71%	63%	77%	70%	69%	60%	73%	71%
Median (as %)	80%	67%	83%	80%	75%	67%	85%	75%
Section Score Max	15	15	40	15	12	15	20	132
Section Mean	10.7	9.5	30.8	10.5	8.3	9.0	14.6	93.5
Section Median	12	10	33	12	9	10	17	99

Table 7. Percentage of student achievement level as per SLOs (SPN 1121).

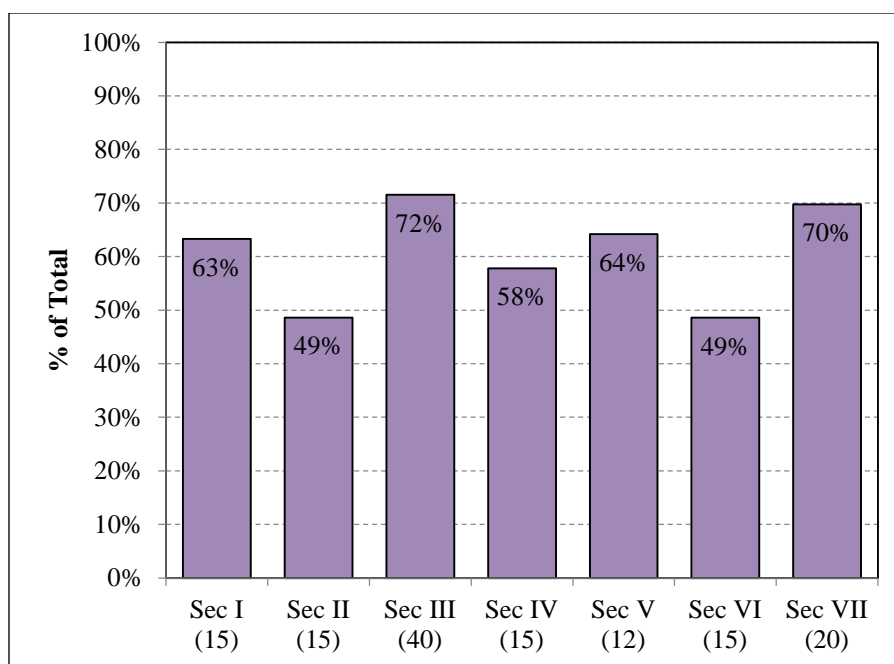


Figure 13. SLO achievement for SPN 1121 by exam section (Percentage of student achievement level as per SLOs).

3.3.2 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.

3.3.2.1 Dual Enrollment to Non-Dual Enrollment Comparison

No dual enrollment sections of the course were run during spring 2016 so no comparison study between dual enrollment and traditional courses could be completed.

3.3.2.2 Online to Traditional Comparison

While online course sections were held in spring 2016, an updated shell with the common final exam is currently under development. Once the new shell for SPN 1121 with common final exam is launched in fall 2016, assessment results will be collected.

3.3.2.3 Comparison by Campus/Site

Of the 109 artifacts collected from SPN 1121, 29 from the Charlotte campus, 27 from the Collier campus, and 53 from the Thomas Edison (Lee) campus. A comparison of mean scores is provided in Table 8. Collier exhibits the highest mean score (77.9%) when compared with other sites while Charlotte exhibits the lowest (57.5%).

	Mean	Standard Deviation
Charlotte	57.5	21.0
Collier	77.9	12.4
Thomas Edison (Lee)	74.5	17.2

Table 8. Comparison of mean scores by site. Bold denotes highest among all sites.

A plot comparing descriptive statistics of scores by site is presented in Figure 14. Collier and Thomas Edison exhibit similar distributions while Charlotte's central 50% is substantially lower. A one-way analysis of variance was used to compare means at each site. Results of the ANOVA exhibit a statistically significant difference between sites (see Table 9). Therefore, we 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.

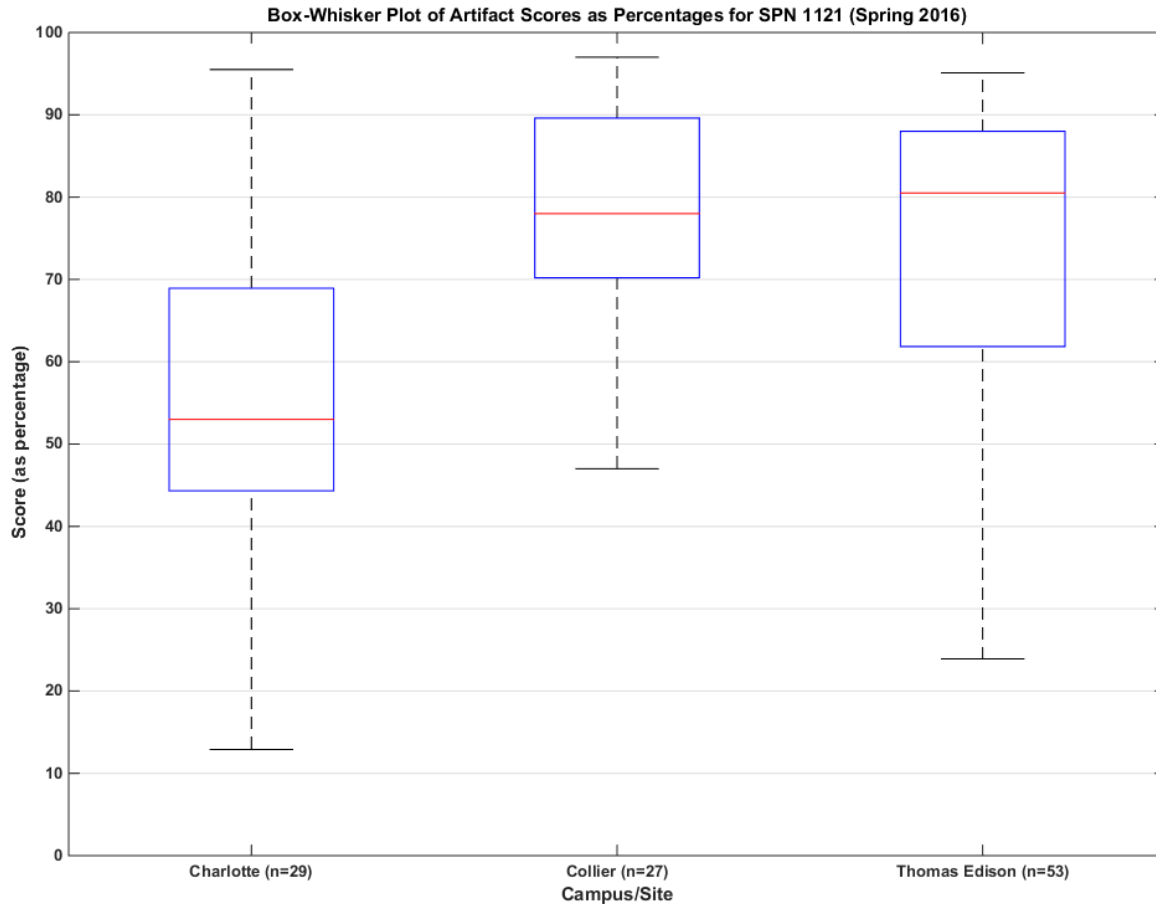


Figure 14. Box-Whisker plot of scores distributed by site for SPN 1121. Red line depicts median score. Upper and lower box boundaries indicate 75% quartile and 25% quartile (box represents central 50% of the scores). Vertical lines represent remaining scores outside central 50% that are not outliers. Red '+'s denote outliers.

Source of Variation	Sum of squared differences	df	Mean Squares	F _{obs}	p-value	F _{crit}
Between Sites	7175.1	2	3587.5	12.00	2.00x10 ⁻⁵	3.08
Within Sites	31,680.7	106	298.9			
Total	38,855.8	108				

Table 9. Results of one-way ANOVA of mean scores at each site for SPN 1121.

3.3.3 Data Distribution & Longitudinal Study

3.3.3.1 Data Distribution

A histogram depicting the distribution of scores across each exam section is shown in Figure 15. Sections I, III, IV, and VII exhibit scores centered on 90% or higher. Section V exhibits a bimodality with a second peak at 70-79%. All sections except Section III exhibit minor peaks at < 30%.

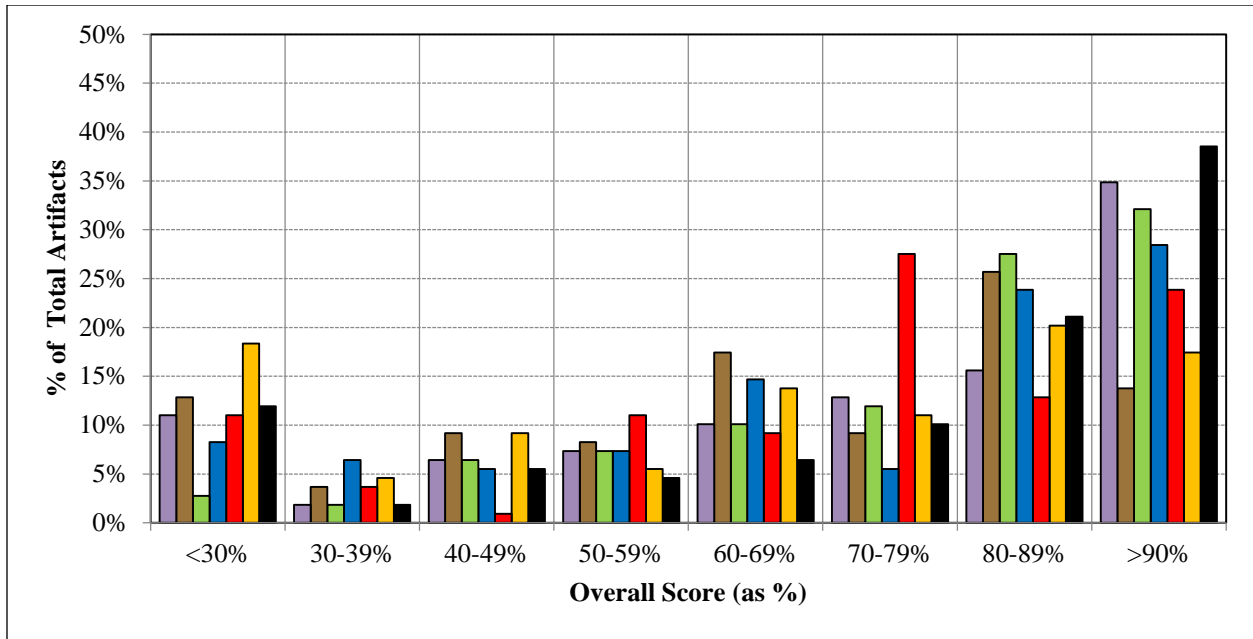


Figure 15. Histogram of SPN 1121 exam sections scores for spring 2016. Purple – Section I, Brown – Section II, Green – Section III, Blue – Section IV, Red – Section V, Orange – Section VI, and Black – Section VII.

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

A review of the colormap in Figure 16 shows that Sections II and V are consistently the lowest performing compared to other sections between the ranges of 60%-79%. Additionally, Section VI is the lowest scoring area at high overall achievement (90 and higher). In other words, average-to-strong performing students tend to struggle with Sections II and V. At the same time, exceptional students struggle most with Section VI.

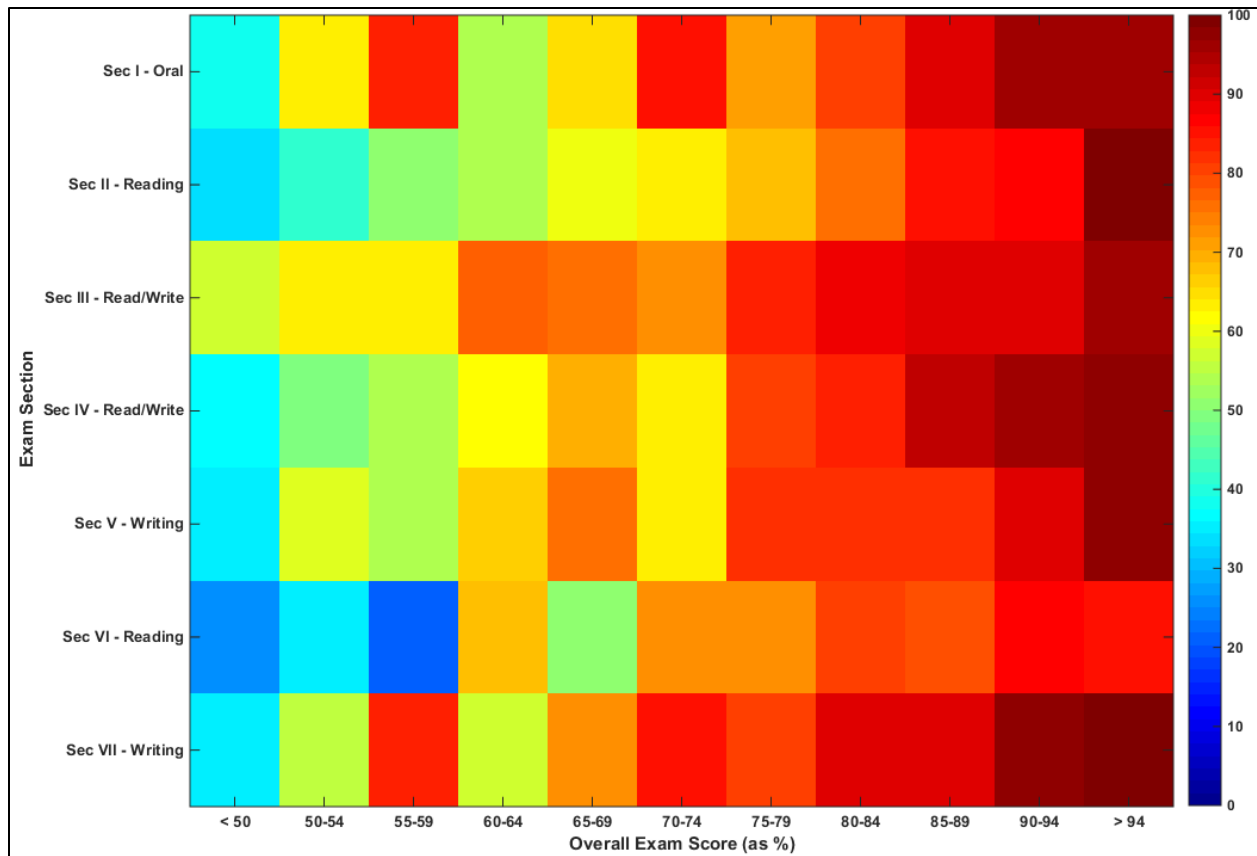


Figure 16. (Top) Colormap of mean scores for each exam section based on overall scoring bin for SPN 1121. (Bottom) Comparison exam section if section score percentage is the same as overall (i.e. artifact score is equally distributed across all sections). An exam section with hotter colors (reds/yellows) compared with the bottom bar means that section achievement exceeds the overall score and is an area of strength. An exam section with colder colors (blues/greens) compared with the bottom bar means that section achievement is lower than the overall score and is therefore an area of weakness.

3.3.3.2 Longitudinal Study

Further description of achievement over time in SPN 1121 is provided in a comparison of fall 2013 through spring 2016 data (Table 10 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). Through time, Sections I, III, and IV exhibit steady decreases. The other sections exhibit no discernable trend. Spring 2014 included 115 samples, which then dropped to 58 samples before jumping to 109 in spring 2016, which may skew results and interpretation.

	Section Max	Fall 2013 n=10	Spring 2014 n=115	Fall 2014 n=25	Spring 2015 n=58	Fall 2015 n=17	Spring 2016 n=109
Section I (Oral)	15	11.5	12.3	11.9	12.2	11.5	10.7
Section II (Reading)	15	9.5	9.6	10.7	10.6	10.6	9.5
Section III (Read/Write)	40	34.2	32.3	30.0	31.1	29.9	30.8
Section IV (Read/Write)	15	9.5	11.4	10.6	11.4	11.3	10.5
Section V (Writing)	12	7.5	5.7	9.5	8.5	8.1	8.3
Section VI (Reading)	15	9.6	10.3	11.6	10.7	9.9	9.0
Section VII (Writing)	20	14.2	15.4	16.1	16.4	16.4	14.6
Combined (Overall)	128	96.0	97.0	100.5	100.9	97.7	93.5

Table 10. Comparison of mean scores for SPN 1121 for fall 2013 through spring 2016.

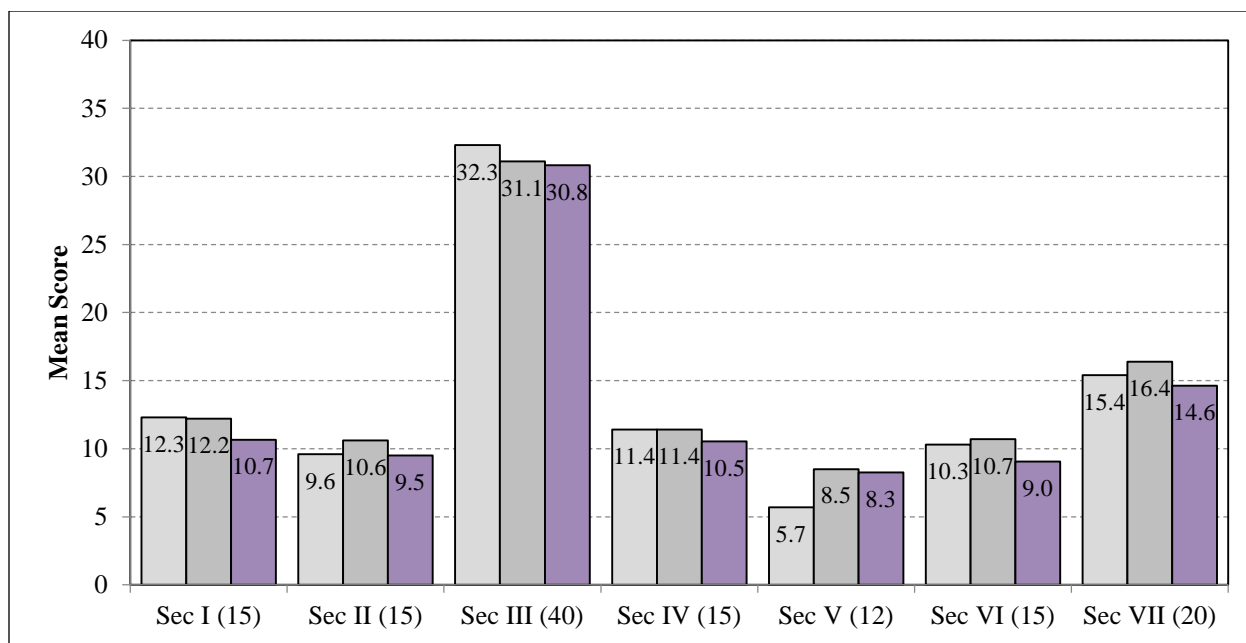


Figure 17. Comparison of mean scores for SPN 1121 through time from spring 2014 (light gray), through spring 2015 (dark gray), to spring 2016 (purple).

4 CONCLUSIONS

Florida SouthWestern’s Foreign Language Department employs a common course assessment in both French and Spanish courses to measure student progress in course level objectives in an effort to improve instruction. What follows is a drilldown of findings for both disciplines (French and Spanish) for the spring 2016 assessment.

4.1 FRENCH

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

1. Achievement of 80% of artifacts scoring 70% or better (SLO 1): Achievement was not analyzed in this report as a new section to be included in assessment was in development during the fall 2015 and spring 2016 semester.
2. Achievement of 80% of artifacts scoring 70% or better (SLO 2): Achievement goal was met as results exhibit 71% of artifacts score 70% or higher in the reading competency section (Section X) of the exam.
3. Achievement of 80% of artifacts scoring 70% or better (SLO 3): Achievement goal was met as results exhibit 94% of artifacts score 70% or higher in the writing effectiveness section (Section I) of the exam.
4. In a complete study of all sections of the assessed exam, the highest achieving SLO is Section I (96%), while the lowest is Section V at 53% achieving 70% or greater.
5. No comparison of dual enrollment to traditional artifacts was made because no FRE 1120 dual enrollment sections were offered in spring 2016.
6. In a comparison of online to traditional artifacts there is a statistically significantly lower mean score for online artifacts (63.9/124 compared with 99.8/124).

7. No cross-campus comparison was completed because artifacts reported originated from only one site other than online.
8. In a study of score distribution by Section, all sections except for Section VIII exhibit peaks at $\geq 90\%$. Section VIII exhibits a peak in the 80-89 scoring bin. Section VIII also exhibits a spike at $< 30\%$ scores while sections IV, V, and 9 exhibit spikes at 50-69.
9. In a longitudinal study of data distribution, from spring-to-spring, four sections exhibit slight increases while the remaining five exhibit declines.

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

1. Assessment is expected to resume for FRE 1121 following changes made to the assessment process.

4.2 SPANISH

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

1. Achievement of 80% of artifacts scoring 70% or better (SLO 1): Achievement was nearly met as results exhibit 71% of artifacts score 70% or higher in the oral competency exam section (Section I).
2. Achievement of 80% of artifacts scoring 70% or better (SLO 2): Achievement was partially met as results exhibit 45% of artifacts scored 70% or higher in Section II and 92% of artifacts scored 70% or higher in Section III.
3. Achievement of 80% of artifacts scoring 70% or better (SLO 3): Achievement was partially met as results exhibit 50% of artifacts scored 70% or higher in Section IV and 80% of artifacts scored 70% or higher in Section V.
4. No comparison of dual enrollment to traditional artifacts was completed because no dual enrollment sections were offered during spring 2016.
5. In a comparison of online to traditional artifacts, online artifacts score 5.1 points higher than traditional. However, results were not statistically significantly different.
6. No cross-campus comparison was made because only two sites reported (online and Thomas Edison) and so comparisons between these are encompasses in #5 above.
7. In a study of score distribution by section, Sections I, III, and V exhibit peaks above 90% with the distribution trailing down with decreasing score. Sections II and IV, however, exhibit more widely distributed scores.
8. In a study of section score distribution based on overall score, Section II is consistently the lowest performing compared to other sections between the ranges of 70%-84%. Also, Section III, and to a lesser extent Section V, is over performing at the lowest overall scores.
9. In a longitudinal study of data distribution, through time, Sections II, IV, and V exhibit steady decreases. Sections I and III exhibit no noticeable trend. Spring 2014 included 90 samples, which then dropped to 73 samples before jumping to 141 in spring 2016. Moreover, the smaller sample size in spring 2015 overrepresented one particular instructor, which may skew results and interpretation.
10. In a study assessing student skills and retention of materials from SPN 1120 prior to beginning SPN 1121, results exhibit a wide range of mean scores across instructor. Results exhibit achievement levels based on previous instructor spanning as low as 9.0/30 to as high as 26.8/30.

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

1. Achievement of 80% of artifacts scoring 70% or better (SLO 1): Achievement was not met as results exhibit 63% of artifacts score 70% or higher in the oral competency exam section (Section I).
2. Achievement of 80% of artifacts scoring 70% or better (SLO 2): Achievement was not met as results exhibit 49% of artifacts scored 70% or higher in Section II and 49% of artifacts scored 70% or higher in Section VI.
3. Achievement of 80% of artifacts scoring 70% or better (SLO 3): Achievement was not met as results exhibit 64% of artifacts scored 70% or higher in Section V and 70% of artifacts scored 70% or higher in Section VII.
4. No comparison of dual enrollment to traditional artifacts was completed because no dual enrollment sections were offered during spring 2016.
5. While online course sections were held in spring 2016, an updated shell with the common final exam is currently under development. Once the new shell for SPN 1121 with common final exam is launched in fall 2016, assessment results will be collected.
6. In a cross-campus comparison, Collier exhibits the highest mean score (77.9%) when compared with other sites while Charlotte exhibits the lowest (57.5%).
7. In a study of score distribution by section, Sections I, III, IV, and VII exhibit scores centered on 90% or higher. Section V exhibits a bimodality with a second peak at 70-79%. All sections except Section III exhibit minor peaks at < 30%.
8. In a study of section score distribution based on overall score, average-to-strong performing students tend to struggle with Sections II and V. At the same time, exceptional students struggle most with Section VI.
9. In a longitudinal study of data distribution, through time, Sections I, III, and IV exhibit steady decreases. The other sections exhibit no discernable trend. Spring 2014 included 115 samples, which then dropped to 58 samples before jumping to 109 in spring 2016, which may skew results and interpretation.

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