

Foreign Language Assessment Report Fall 2015

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 the fall 2015 artifacts were collected from 5 of 6 sections of FRE 1120. Sections were taught by four different instructors, three of which were represented in the artifacts collected. A total of 103 students were enrolled in FRE 1120. Of those, 67 artifacts were collected representing a sample size of 65% of the population. Two sections of FRE 1121 *Elementary French II* were offered. Of those, artifacts were collected from two sections. All sections were taught by the same instructor. A total of 23 students were enrolled in FRE 1121. Of those, 17 artifacts were collected representing a sample size of 74%.

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 2015 exam). This tool is being piloted in spring 2016 and 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 as results exhibit 81% 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 96% 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 VIII at 58% achieving 70% or greater.

n = 67	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 ≥70% for all sections</i>									
% above Goal	96%	85%	63%	75%	73%	77%	83%	58%	81%	
Mean (as %)	91%	84%	75%	83%	75%	81%	84%	69%	80%	72%
Median (as %)	100%	83%	81%	90%	88%	100%	90%	80%	90%	81%
Section Score Max	10	6	26	20	8	10	10	14	20	124
Section Mean	9.1	5.0	19.4	16.7	6.0	8.1	8.4	9.6	16.0	89.4
Section Median	10	5	21	18	7	10	9	11.25	18	101.5

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

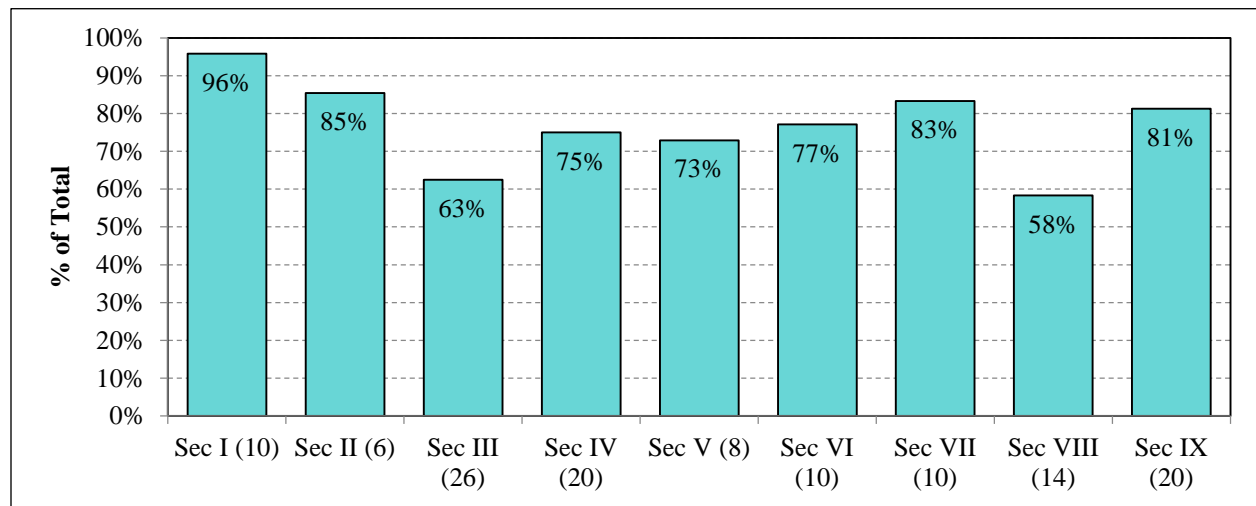


Figure 1. SLO achievement for FRE 1120 by exam section (Percentage of student achievement level as per SLOs).

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

During the fall 2015 semester, 14 artifacts were collected from dual enrollment sections in FRE 1120 and 34 traditional (non-online) artifacts were collected in FRE 1120. A comparison of mean scores is provided in Table 2. The dual enrollment mean score is 1.4 lower than traditional artifacts. The difference in the means was tested for significance using a Welch's t-test according to standard methods (Davis, 1973; McDonald, 2009; Wilkinson, 1999) and were found to not be statistically significantly different. Therefore, we cannot reject the null hypothesis that the difference in the mean scores of dual enrollment and traditional artifacts can be a result of chance 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 dual enrollment artifacts to traditional artifacts is approximately 5%. For a graphical representation of this see Figure 2.

df = 46	
Dual enrollment mean	97.3
Dual enrollment standard deviation	22.95
Traditional mean	98.7
Traditional standard deviation	20.44
Effect size	0.06
p-value	0.835

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

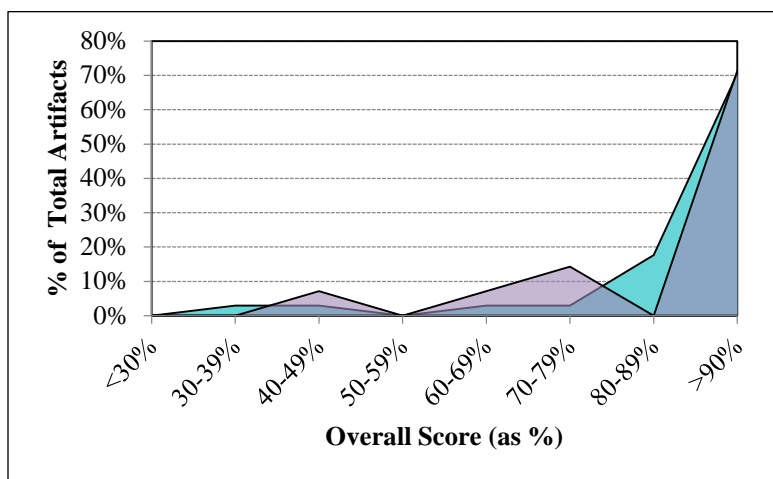


Figure 2. Score distribution for dual enrollment (purple) and traditional (teal) artifacts for FRE 1120.

2.1.2.2 Online to Traditional Comparison

During the fall 2015 semester, one online section was run. From that course section, 19 total online artifacts were collected from FRE 1120 and 34 traditional artifacts were collected from FRE 1120. A comparison of basic statistics is provided in Table 3. Online artifacts mean scores are 31.7 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 difference 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 57%. For a graphical representation of this see Figure 3.

df = 51	
Online mean	67.0
Online standard deviation	33.37
Traditional mean	98.7
Traditional standard deviation	20.44
Effect size	1.05
p-value	8.71×10^{-4}

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

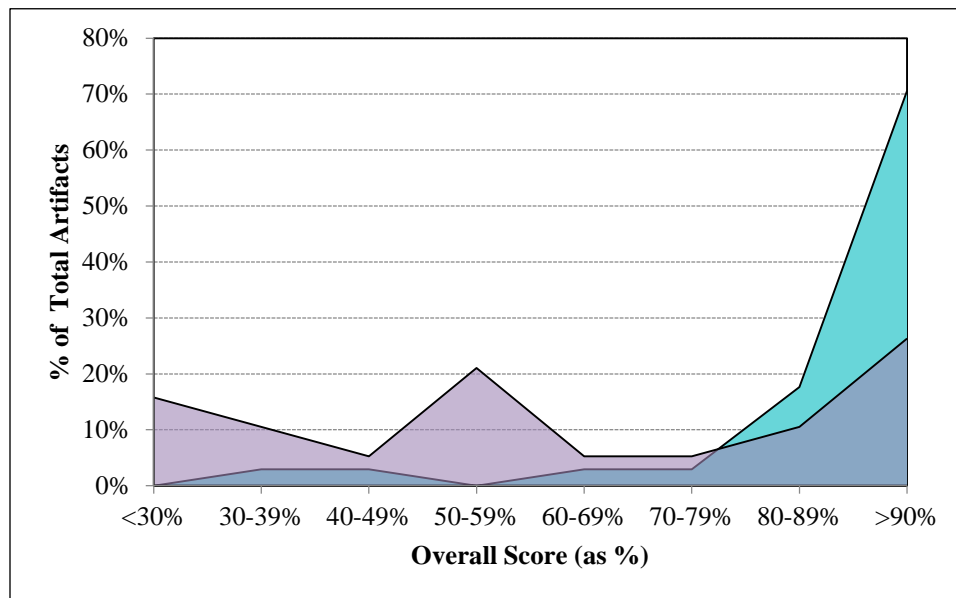


Figure 3. Score distribution for online (purple) and traditional (teal) artifacts of FRE 1120.

2.1.2.3 Comparison by Campus/Site

Artifacts were collected from only one site (Thomas Edison Campus) other than FSW Online and offsite (dual enrollment). 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 4. All sections except for Section II exhibit peaks at $\geq 90\%$. Section II exhibits a peak in the 80-89 scoring bin. Sections IV and VIII exhibit slight spikes at $< 30\%$ scores.

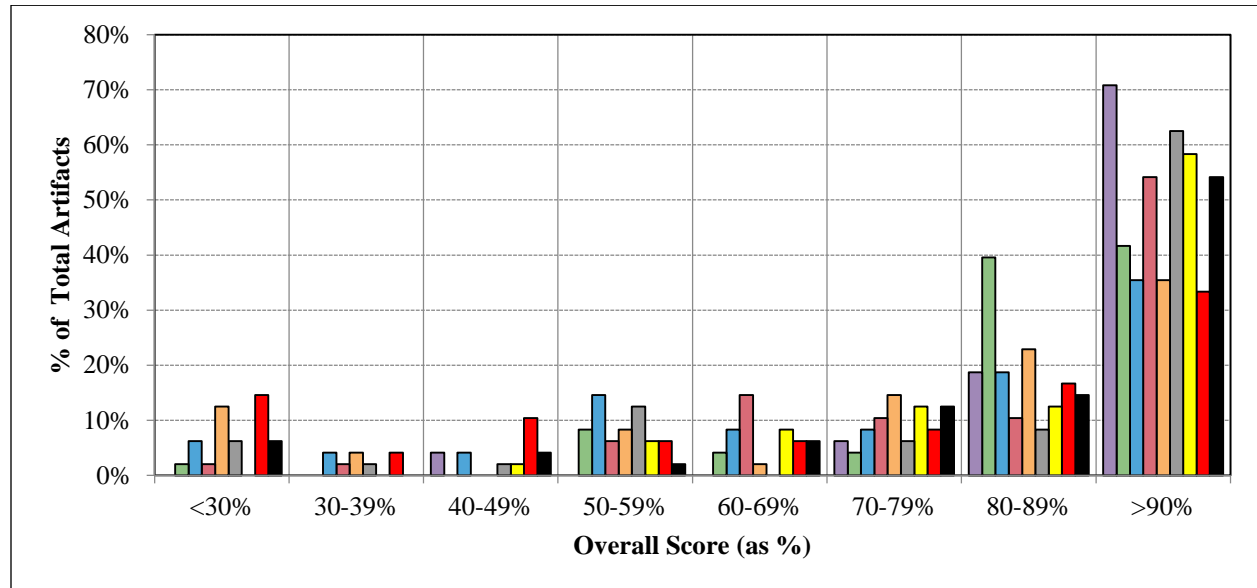


Figure 4. Histogram of FRE 1120 exam sections and combined (overall) score for fall 2015. Purple – Section I, Green – Section II, Blue – Section III, Magenta – Section IV, Peach – Section V, Gray – Section VI, Yellow – Section VII, Red – Section VIII, and Black – Section IX.

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 4 and Figure 5). Both demographics of students and student count vary by semester it may be more reasonable to compare like semesters (Fall vs. Fall, Spring vs. Spring). (see <http://www.fsw.edu/facultystaff/assessment/history> for further details). With only one term to compare, interpretation of trends is limited. From fall-to-fall, three sections exhibit slight increases while the remaining six exhibit declines (Figure 5).

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

Table 4. Comparison of mean scores for FRE 1120 for fall 2014 through fall 2015.

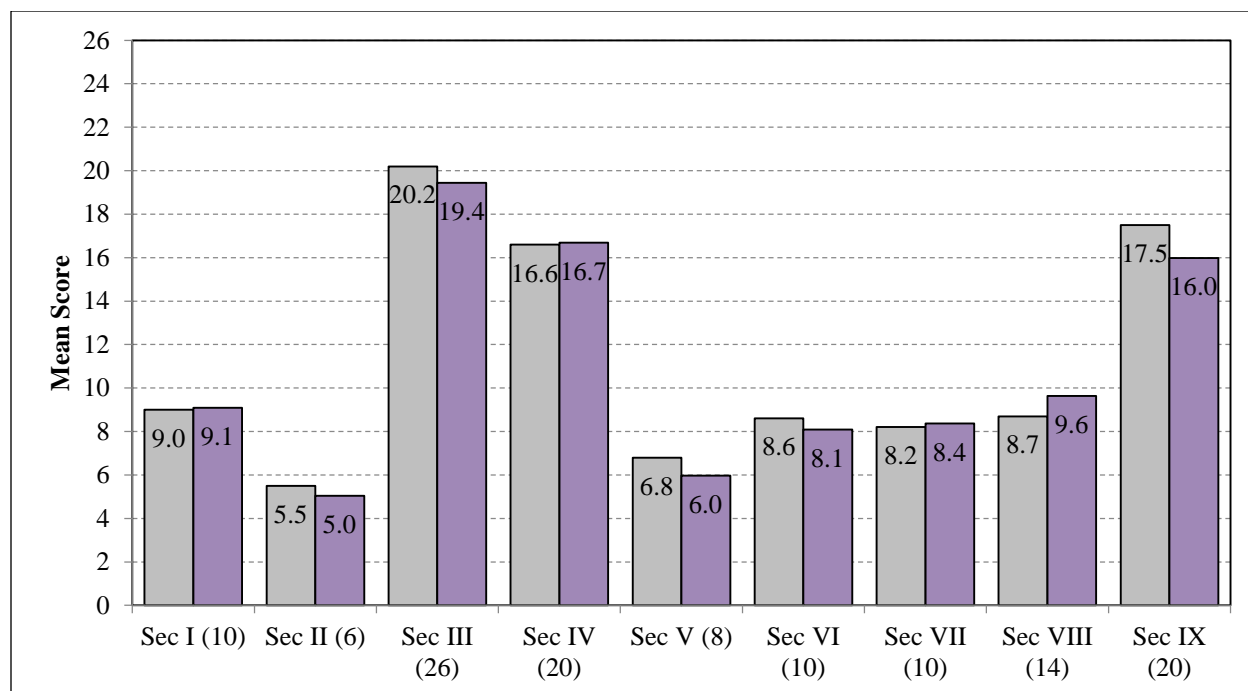


Figure 5. Comparison of mean scores for FRE 1120 through time from fall 2014 (dark gray) to fall 2015 (purple).

2.2 FRE 1121

2.2.1 Descriptive Statistics & Learning Objectives

Using a common course assessment the FSW French faculty defined the same three areas of interest for evaluation that apply to FRE 1121 as those used for FRE 1120. For details on each SLO, see 2.1.1. The only difference between FRE 1121 and FRE 1120 in terms of measuring these outcomes is that the exam sections differ slightly and are noted below.

The faculty established measure of success for SLO 2, 80% of students scoring 70% or higher in Section XI, was almost met as results exhibit 78% of artifacts score 70% or higher in the reading competency exam section (Section XI) (Table 5). The faculty established measure of success for SLO 3, 80% of students scoring 70% or higher in Section I, was met. Results exhibit 89% of artifacts scored 70% or higher in Section I. For a graphical representation of SLO achievement, see Figure 6.

n = 17	Sec I	Sec II	Sec III	Sec IV	Sec V	Sec VI	Sec VII	Sec VIII	Sec IX	Sec X	Sec XI	Combined (Overall)
Goal	<i>80% of artifacts scored $\geq 70\%$ for all sections</i>											
% above Goal	89%	78%	56%	78%	44%	67%	78%	22%	100%	78%	78%	
Mean	84%	78%	78%	88%	58%	69%	80%	58%	94%	79%	80%	64%
Median	80%	75%	81%	88%	58%	75%	100%	60%	100%	88%	90%	65%
Section Score Max	10	10	8	8	12	4	8	5	4	12	20	101
Section Mean	8.4	7.8	6.3	7.1	6.9	2.8	6.4	2.9	3.8	9.5	16.0	64.2
Section Median	8	7.5	6.5	7	7	3	8	3	4	10.5	18	65.3

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

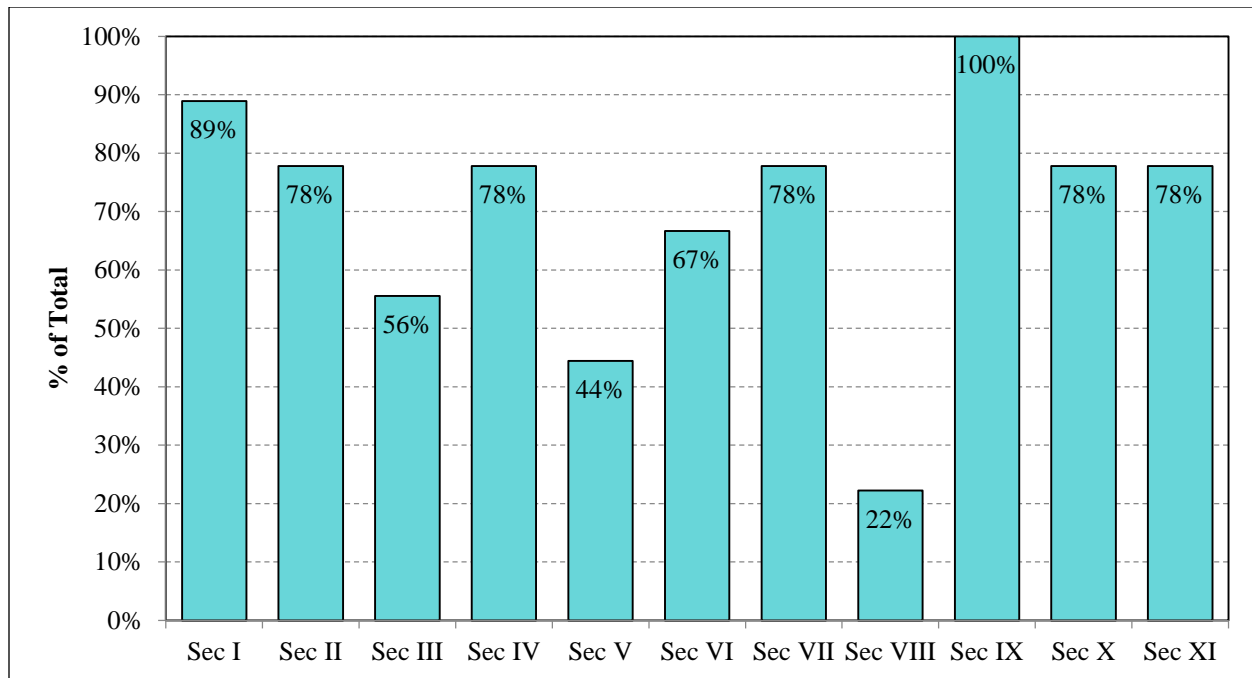


Figure 6. SLO achievement for FRE 1121 by exam section (Percentage of student achievement level as per SLOs).

2.2.2 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.

2.2.2.1 Dual Enrollment to Non-Dual Enrollment Comparison

No dual enrollment sections of the course were run during fall 2015 so no comparison study between dual enrollment and traditional courses could be completed.

2.2.2.2 Online to Traditional Comparison

During the fall 2015 semester, one online section was run. From that course section, 8 total online artifacts were collected from FRE 1121 and 9 traditional artifacts were collected from FRE 1121. A comparison of basic statistics is provided in Table 6. Online artifacts mean scores are 29.1 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, however, sample sizes are too low to yield meaningful results (de Winter, 2013).

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 98%. For a graphical representation of this see Figure 7.

df = 15		
Online mean		48.8
Online standard deviation		14.00
Traditional mean		77.9
Traditional standard deviation		14.36
Effect size		2.18
p-value		7.59×10^{-4}

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

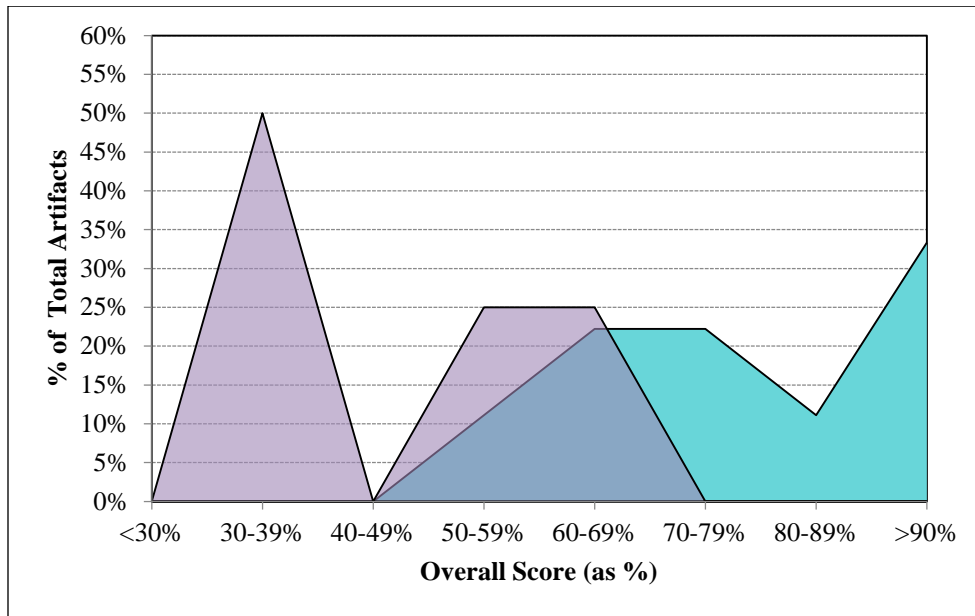


Figure 7. Score distribution for online (purple) and traditional (teal) artifacts of FRE 1121.

2.2.2.3 Comparison by Campus/Site

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

2.2.3 Data Distribution & Longitudinal Study

2.2.3.1 Data Distribution

A histogram depicting the distribution of scores across each exam section is shown in Figure 8. With sparse data (n=17) interpretation is limited. Section VIII in Red appears to be the only section substantially different than the others, with a peak centered in to 60-69% scoring bin where others are 80-89 or higher.

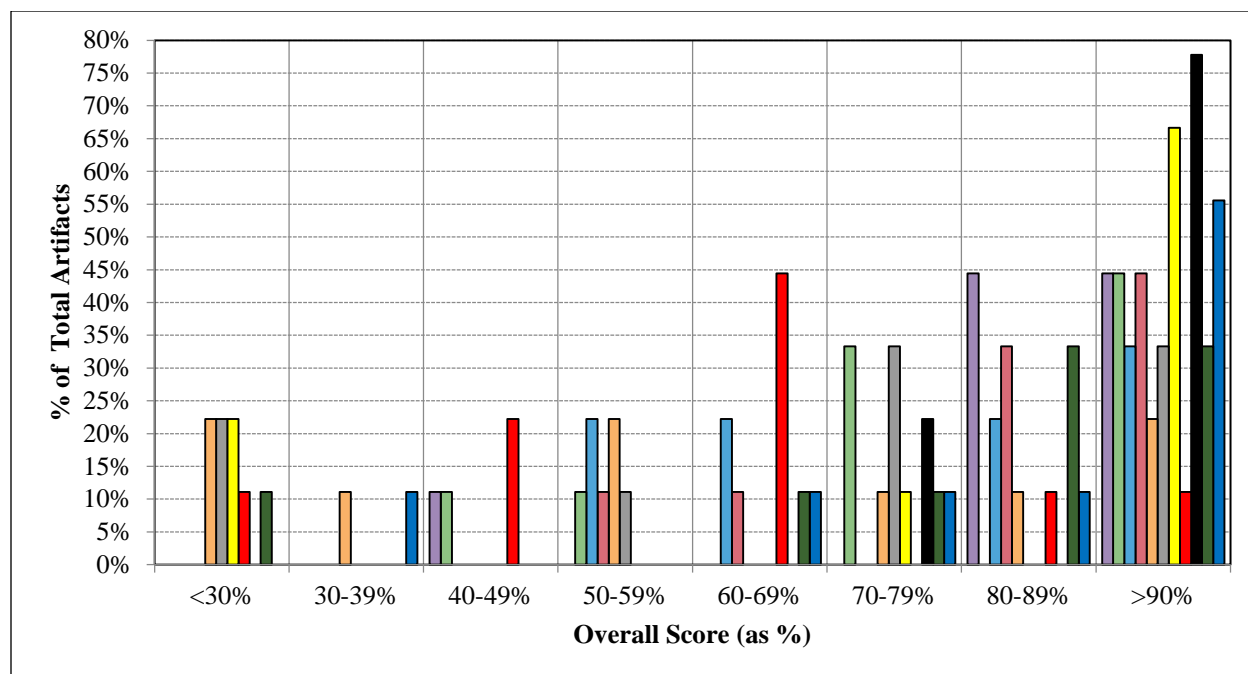


Figure 8. Histogram of FRE 1121 exam sections and combined (overall) score for fall 2015. Purple – Section I, Green – Section II, Blue – Section III, Magenta – Section IV, Peach – Section V, Gray – Section VI, Yellow – Section VII, Red – Section VIII, and Black – Section IX, Dark Green – Section X, and Dark Blue – Section XI.

2.2.3.2 Longitudinal Study

Description of achievement over time in FRE 1121 is provided in a comparison of fall 2014 through fall 2015 data (Table 7 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). With only one term to compare, interpretation of trends is limited. From fall-to-fall, seven sections exhibit slight increases while the remaining four exhibit declines (Figure 9).

	Section Max	Fall 2014 n=10	Spring 2015 n=10	Fall 2015 n=17
Section I	10	7.8	9.1	8.4
Section II	10	9.1	8.6	7.8
Section III	8	5.0	4.8	6.3
Section IV	8	5.0	5.2	7.1
Section V	12	5.3	8.4	6.9
Section VI	4	3.1	2.5	2.8
Section VII	8	6.0	7.1	6.4
Section VIII	5	4.6	2.3	2.9
Section IX	4	3.1	2.5	3.8
Section X	12	5.5	8.8	9.5
Section XI	20	16.7	15.3	16.0
Overall	101	71.0	74.5	64.2

Table 7. Comparison of mean scores for FRE 1121 for fall 2014 through fall 2015.

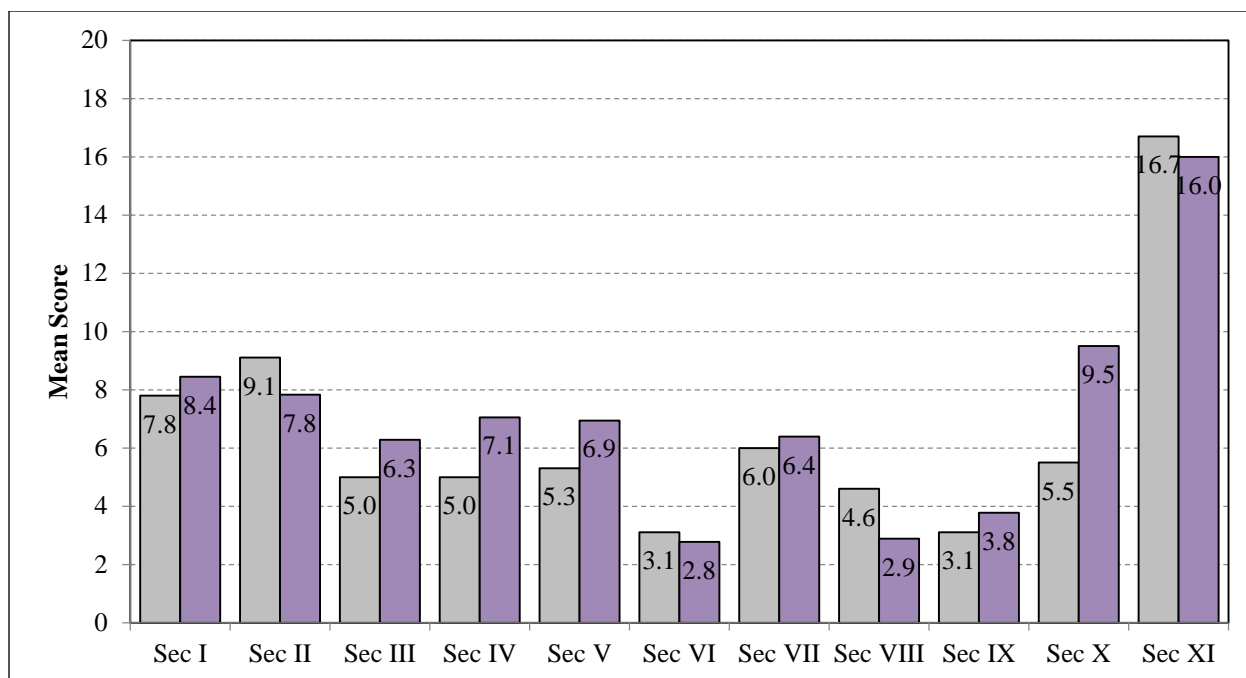


Figure 9. Comparison of mean scores for FRE 1121 through time from fall 2014 (dark gray) to fall 2015 (purple).

3 SPANISH

During fall 2015 semester 18 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 nine different instructors, four of which were represented in the artifacts collected. A total of 338 students were enrolled in SPN 1120. Of those, 122 artifacts were collected representing a sample size of 36% of the population. Six sections of SPN 1121 Beginning Spanish II were offered. Of those, artifacts were collected from a common final from one section. Sections were taught by four different instructors, one of which was represented in the artifacts collected. A total of 116 students were enrolled in SPN 1121. Of those, 17 artifacts were collected representing a sample size of 6%.

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 not met as results exhibit only 57% of artifacts score 70% or higher in the oral competency exam section (Section I) (Table 8). 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 53% of artifacts scored 70% or higher in Section II and 87% 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 not met. Results exhibit 73% of artifacts scored 70% or higher in Section IV and 71% of artifacts scored 70% or higher in Section V. For a graphical representation of SLO achievement, see Figure 10.

n = 122	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%	57%	53%	87%	73%	71%	
Mean (as %)	72%	70%	83%	77%	76%	73%
Median (as %)	75%	72%	87%	87%	85%	75%
Section Score Max	15	60	15	15	20	125
Section Mean	10.8	41.8	12.4	11.5	15.2	91.6
Section Median	11	43	13	13	17	94

Table 8. Percentage of student achievement level as per SLOs (SPN 1120).

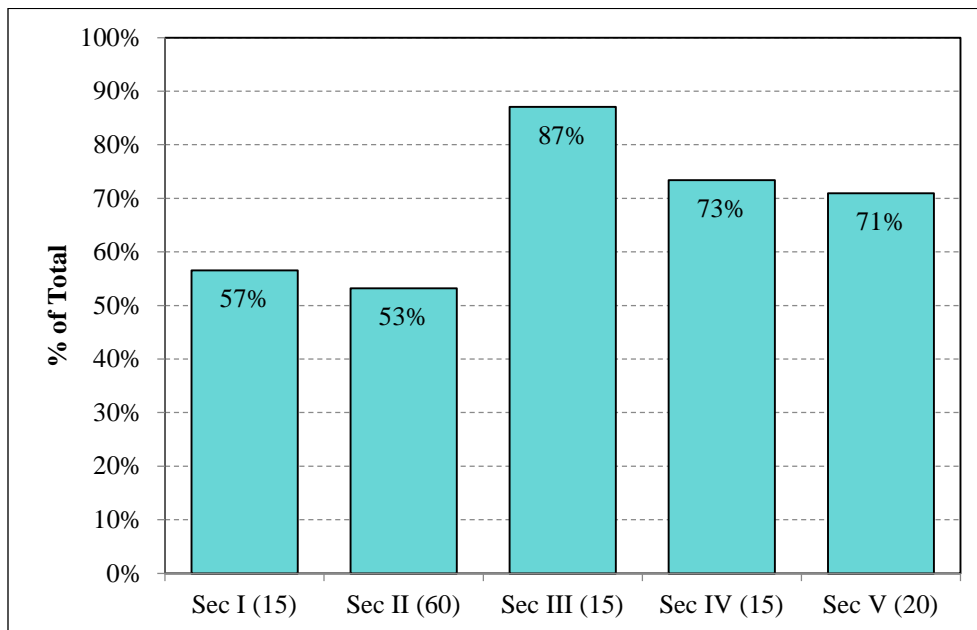


Figure 10. SLO achievement for SPN 1120 by exam section (Percentage of student achievement level as per SLOs).

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 fall 2015 so no comparison study between dual enrollment and traditional courses could be completed.

3.1.2.2 Online to Traditional Comparison

During the fall 2015 semester, one online section was run using the newly developed shell which includes the common final exam. From that course section, 5 total online artifacts were collected from SPN 1120 and 117 traditional artifacts were collected from SPN 1120. A comparison of basic statistics is provided in Table 9. Online artifacts mean scores are 24 lower than traditional artifacts. Differences in the means were tested for significance using a Welch’s t-test according to standard methods (Davis, 1973; McDonald, 2009; Wilkinson, 1999) and were found to not be statistically significantly different (online sample size was only 5). Therefore, we cannot reject the null hypothesis that the difference 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 16%. For a graphical representation of this see Figure 11.

df = 120	
Online mean	78.2
Online standard deviation	24.85
Traditional mean	92.2
Traditional standard deviation	20.98
Effect size	0.23
p-value	0.279

Table 9. Comparison of mean scores for online and traditional artifacts. Positive effect sizes indicate a higher mean score for traditional artifacts.

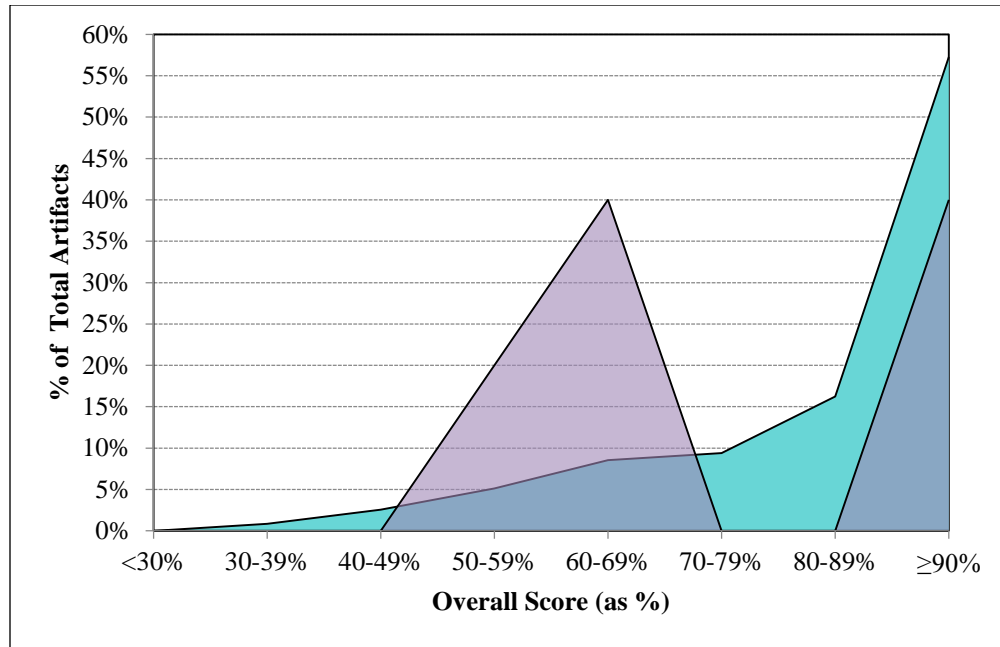


Figure 11. Score distribution for online (purple) and traditional (teal) artifacts of SPN 1120.

3.1.2.3 Comparison by Campus/Site

Of the 122 artifacts collected from SPN 1120, 4020 originated from the Charlotte Campus, 39 from the Collier Campus, 5 from FSW Online, and 38 from the Thomas Edison (Lee) Campus. A comparison of mean scores is provided in Table 10. Collier exhibits the highest mean across sites (97.4/125) while FSW Online exhibits the lowest (78.2/125).

	Mean	Standard Deviation
Charlotte	87.9	19.82
Collier	97.4	20.27
FSW Online	78.2	24.85
Thomas Edison (Lee)	91.4	22.25

Table 10. 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 12. Overlap across sites is somewhat inconsistent. For example, the upper 50% of Collier scores fall within the upper 25% of Charlotte scores. In the case of Collier and Thomas Edison, however, range and overlap is modestly uniform. A one-way analysis of variance was used to compare means at each site. Results of the ANOVA do not exhibit a statistically significant difference between sites (see Table 11). Therefore, we cannot reject the null hypothesis that the mean scores at each site are equal to each other and we cannot conclude with a 95% confidence that the differences in scores are not solely due to chance.

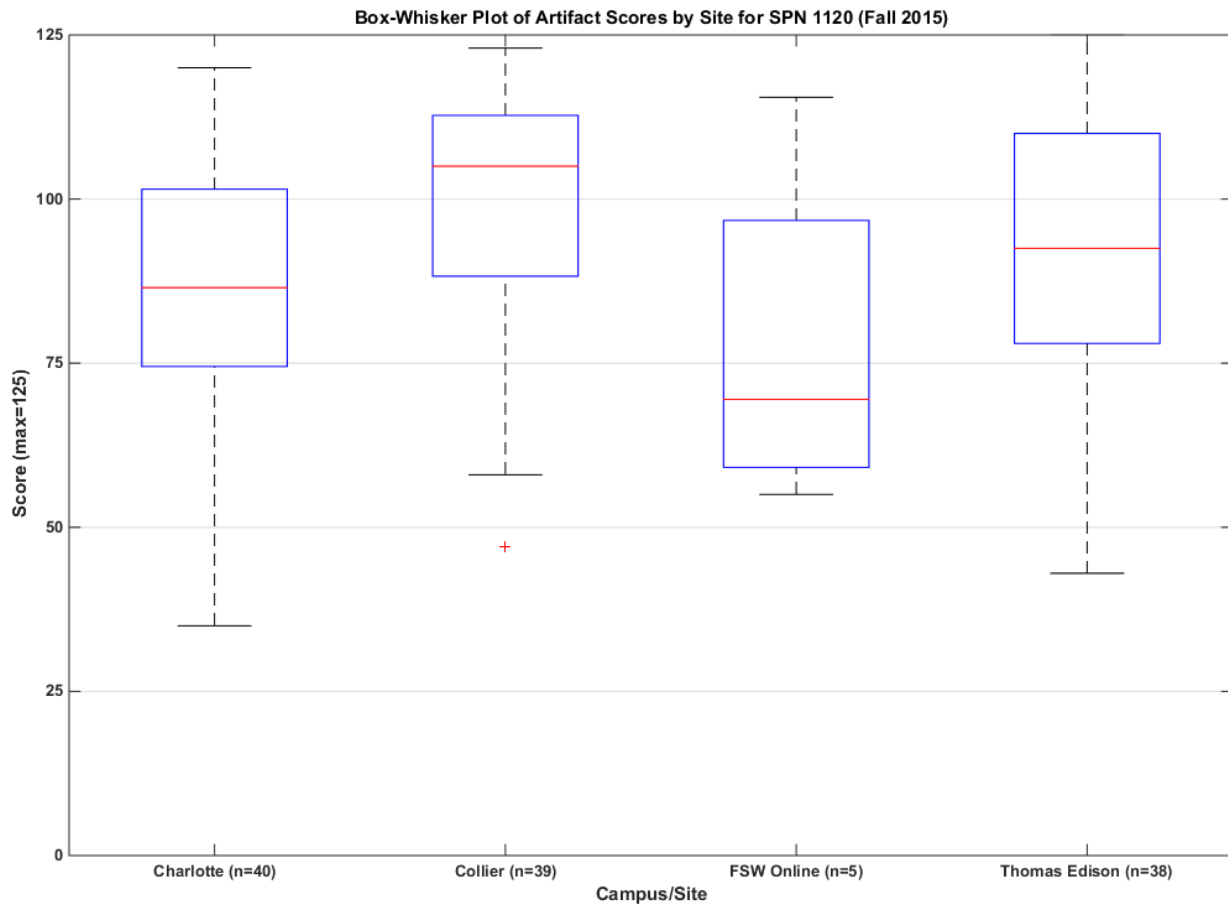


Figure 12. Box-Whisker plot of scores distributed by site for SPN 1120. Red line depicts median score. Upper and lower box boundaries indicate 75% quartile and 25% quartile (box represents central 50% of the scores). Vertical lines represent remaining scores outside central 50% that are not outliers. Red '+'s denote outliers.

Source of Variation	Sum of squared differences	df	Mean Squares	F _{obs}	p-value	F _{crit}
Between Sites	2747.9	3	916.0	2.09	0.105	2.68
Within Sites	51708.7	118	438.2			
Total	54456.6	121				

Table 11. Results of one-way ANOVA of mean scores at each site for SPN 1120.

3.1.3 Data Distribution & Longitudinal Study

3.1.3.1 Data Distribution

A histogram depicting the distribution of scores across each exam section is shown in Figure 13. Sections I and II exhibit higher percentages in lower scoring bins (40-59%) compared to other sections. Section I exhibits a bimodal distribution (scores tend to be either $\geq 90\%$ or below 70%). For Section II, the distribution exhibits far fewer artifacts in the upper scoring bins. Recall that both Section I and Section II exhibit less than 60% achievement of the SLO, while Sections III, IV, and V exhibit 70% or greater.

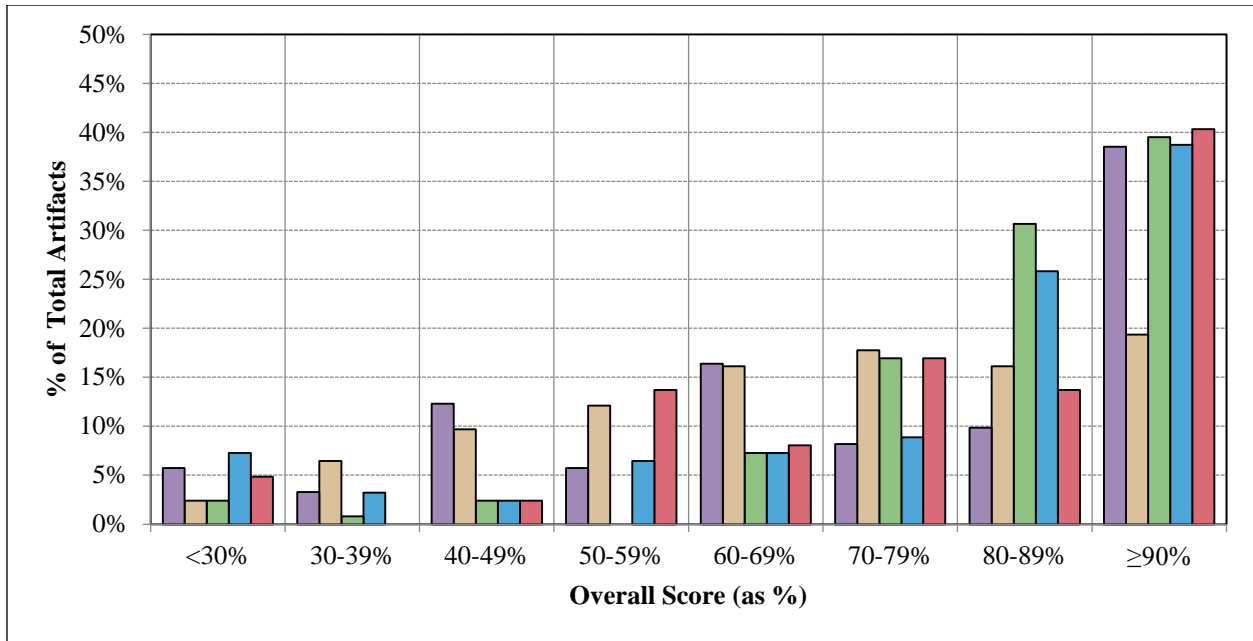


Figure 13. Histogram of SPN 1120 exam sections scores for fall 2015. Purple – Section I, Brown – Section II, Green – Section III, Blue – Section IV, and Red – Section V.

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 14). 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 14 shows that Section II is consistently the lowest performing compared to other sections between the ranges of 75%-89%. For example, in the 75-79% range, the mean score for Section II is 67%, while the other four sections range from 77-86%. Meanwhile, Section III, and to a lesser extent Section V, is over performing at the lowest overall scores. At the 55-59% range, the Section IV and V mean scores are 77% and 75%, respectively, whereas other sections range from 36%-53%.

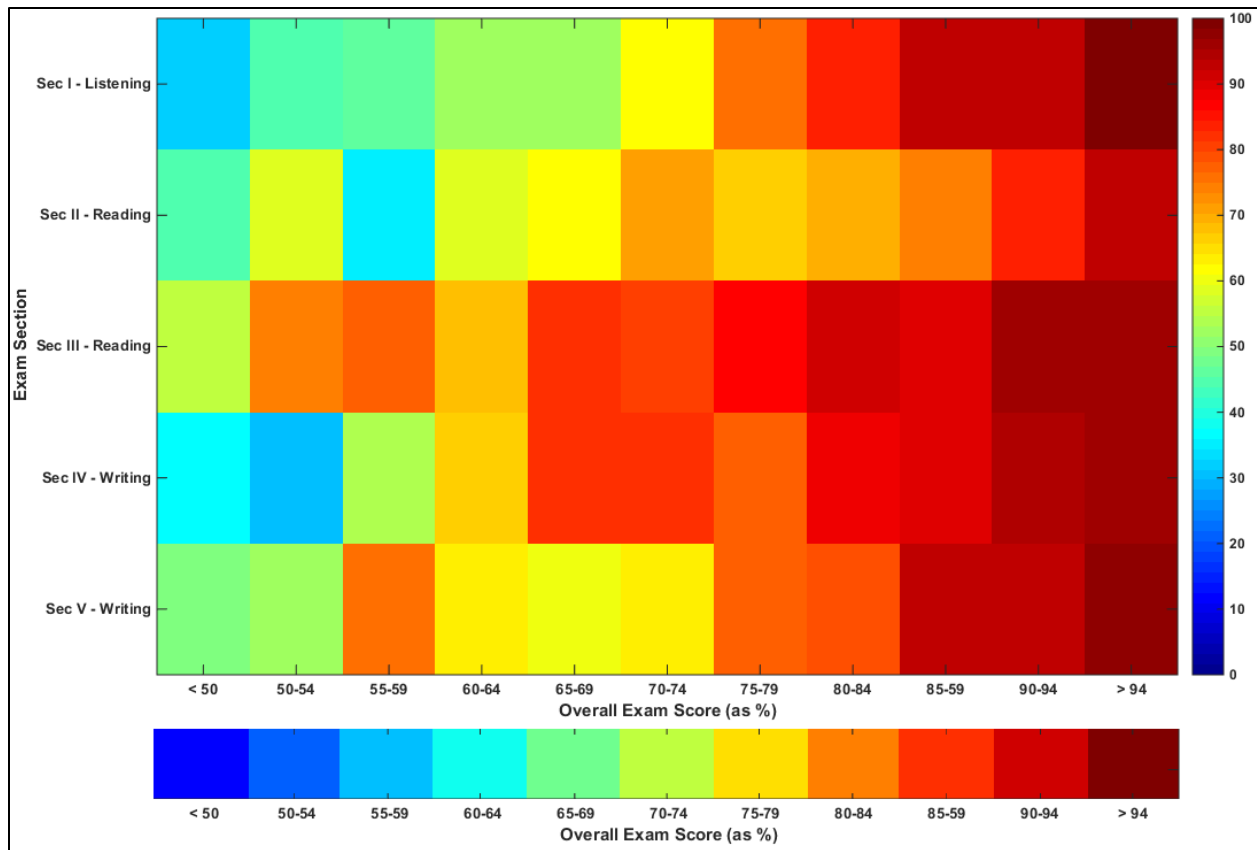


Figure 14. (Top) Colormap of mean scores for each exam section based on overall scoring bin for SPN 1120. (Bottom) Comparison exam section if section score percentage is the same as overall (i.e. artifact score is equally distributed across all sections). An exam section with hotter colors (reds/yellows) 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 fall 2015 data (Table 12 and Figure 15). 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 and IV exhibit steady increases. Section II increased from 40.4/60 in fall 2013 to 41.8/60 in fall 2015, a gradual improvement of 2%. Meanwhile, Section IV increased from 10.6/15 in fall 2013 to 11.5/15 in fall 2015, a gradual improvement of 6%. Section I exhibits a steady decline from 12.1/15 in fall 2013 to 10.8/15 in fall 2015, a steady decline of nearly 9%.

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

Table 12. Comparison of mean scores for SPN 1120 for fall 2013 through fall 2015.

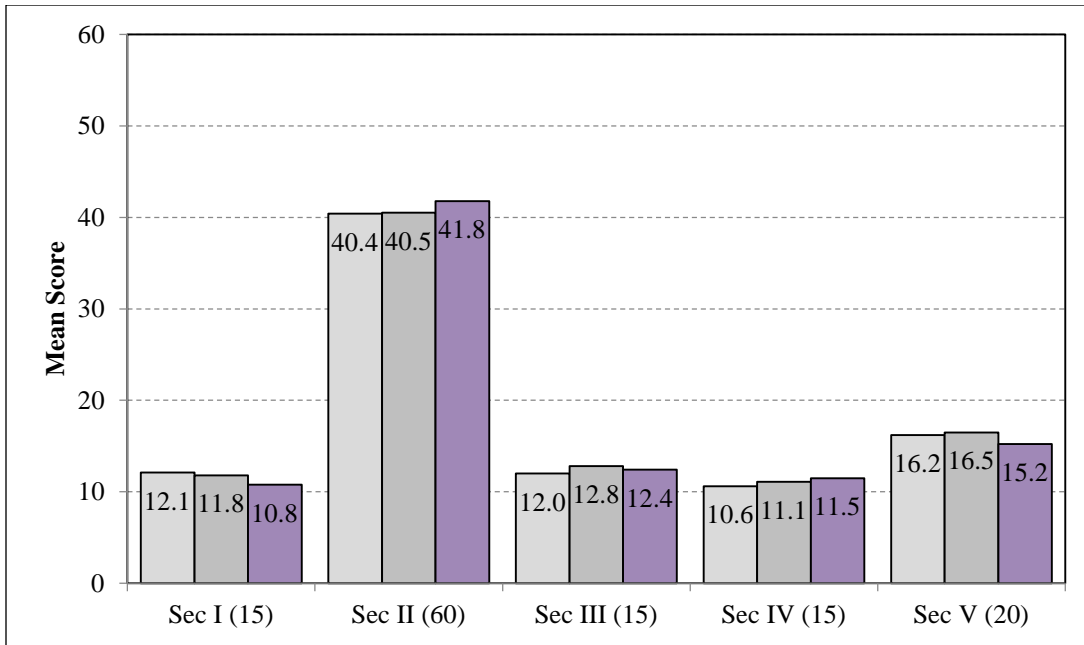


Figure 15. Comparison of mean scores for SPN 1120 through time from fall 2013 (light gray), through fall 2014 (dark gray), to fall 2015 (purple).

If we compare colormaps from previous years to fall 2015 the gradual increases and decreases by sections can be reviewed based on overall achievement (Figure 16). In other words, what achievement demographic is improving/declining? Are strong students improving further, or are weak students improving and thereby improving the averages? The nearly 9% decline of Section I is clearly visible in the upper color rows on each colormap (left-fall 2014 vs. right-fall 2015). While high performing students remain almost unchanged, in fall 2014 underperforming students achieved higher scores in Section I than II compared with fall 2015, where that relationship has reversed.

The improvements of Sections II and IV are also clearly evident. In the case of Section II, while this section continues to be underperforming compared to others at higher overall scores, the lag at the lower overall scores has decreased and is no longer the lowest at lower overall scores. Section IV shows ample improvement at the lowest overall achievement level as well.

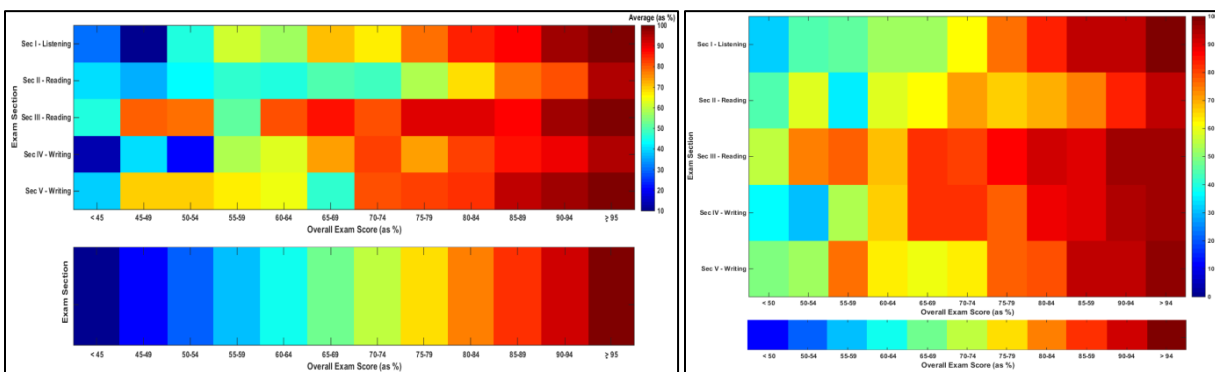


Figure 16. Comparison of fall 2014 colormap of SPN 1120 (left) with fall 2015 colormap (right).

3.2 PILOT STUDY

During the fall 2015 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 evaluate 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. Results exhibit a wide range of achievement levels based on previous instructor spanning from as low as 9.0/30 (for FSW instructors) to as high as 20.8/30 (Figure 17).

Although low, the sample size is on the lower end of significance tests potentially providing accurate results (de Winter, 2013). 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 ($df=18$, $p=0.018$). Therefore, we can reject the null hypothesis that the mean scores at each site are equal to each other and we can conclude with a 95% confidence that the differences in scores are not solely due to chance. Based on the work of Johnson (2013), there is a 17-25% chance that the marginally significant result may be false positives (i.e. Type I errors).

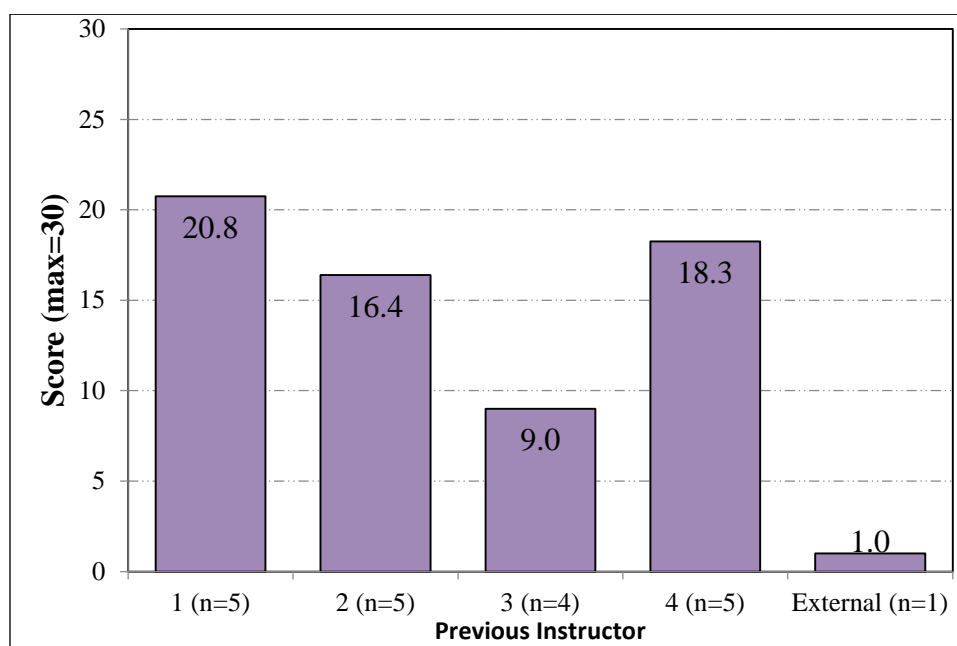


Figure 17. Comparison of achievement in pilot study assessment by instructor.

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 6 below.

The faculty established measure of success for SLO 1, 80% of students scoring 70% or higher in Section I, was nearly met as results exhibit 76% of artifacts score 70% or higher in the oral competency exam

section (Section I) (Table 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 65% of artifacts scored 70% or higher in Section II and 53% of artifacts scored 70% or higher in Section VI. The faculty established measure of success for SLO 3, 80% of students scoring 70% or higher in writing only sections, Sections V and VII, was partially met. Results exhibit 65% of artifacts scored 70% or higher in Section V and 94% of artifacts scored 70% or higher in Section VII. For a graphical representation of SLO achievement, see Figure 18.

n = 17	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	76%	65%	82%	71%	65%	53%	94%	
Mean (as %)	76%	71%	75%	75%	67%	66%	82%	74%
Median (as %)	83%	77%	78%	73%	75%	70%	88%	78%
Section Score Max	15	15	40	15	12	15	20	132
Section Mean	11.5	10.6	29.9	11.3	8.1	9.9	16.4	97.7
Section Median	12.5	11.5	31	11	9	10.5	17.5	103.5

Table 13. Percentage of student achievement level as per SLOs (SPN 1121).

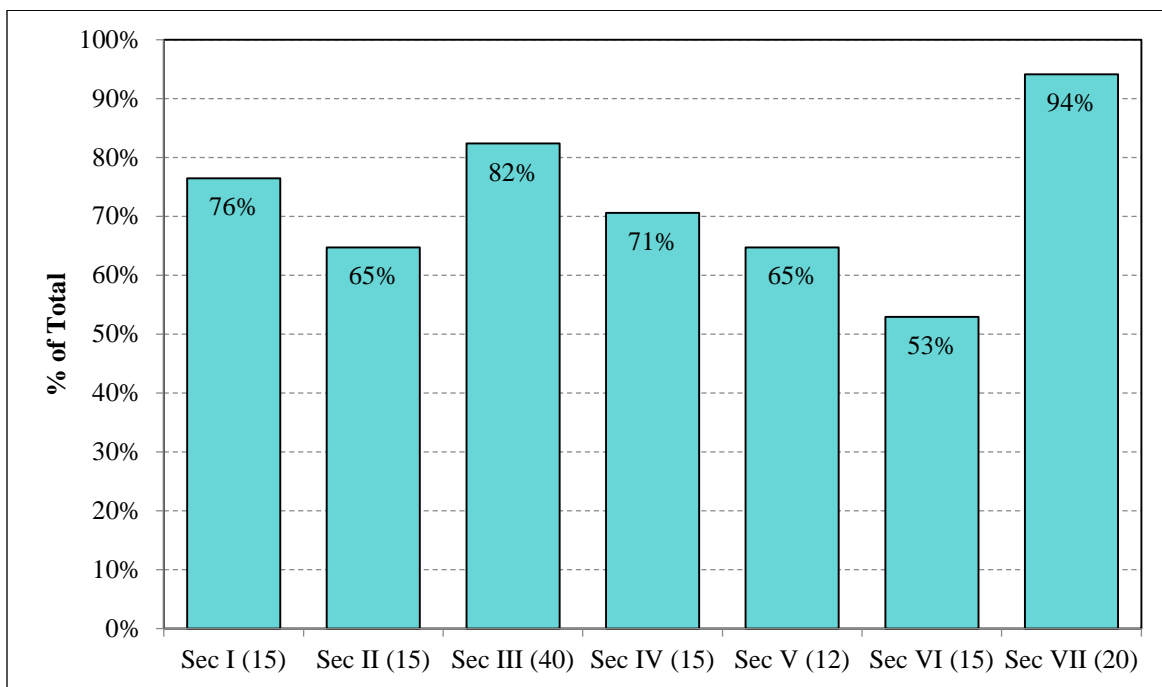


Figure 18. SLO achievement for SPN 1121 by exam section (Percentage of student achievement level as per SLOs).

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 fall 2015 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 fall 2015, 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

Since only one section of the course reported results of the common final exam, no cross-campus comparison study 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 19. Sections I, II, III, IV, V, and VII exhibit scores centered on 80-89%. Section VI exhibits a bimodality with a second peak at 50-59%.

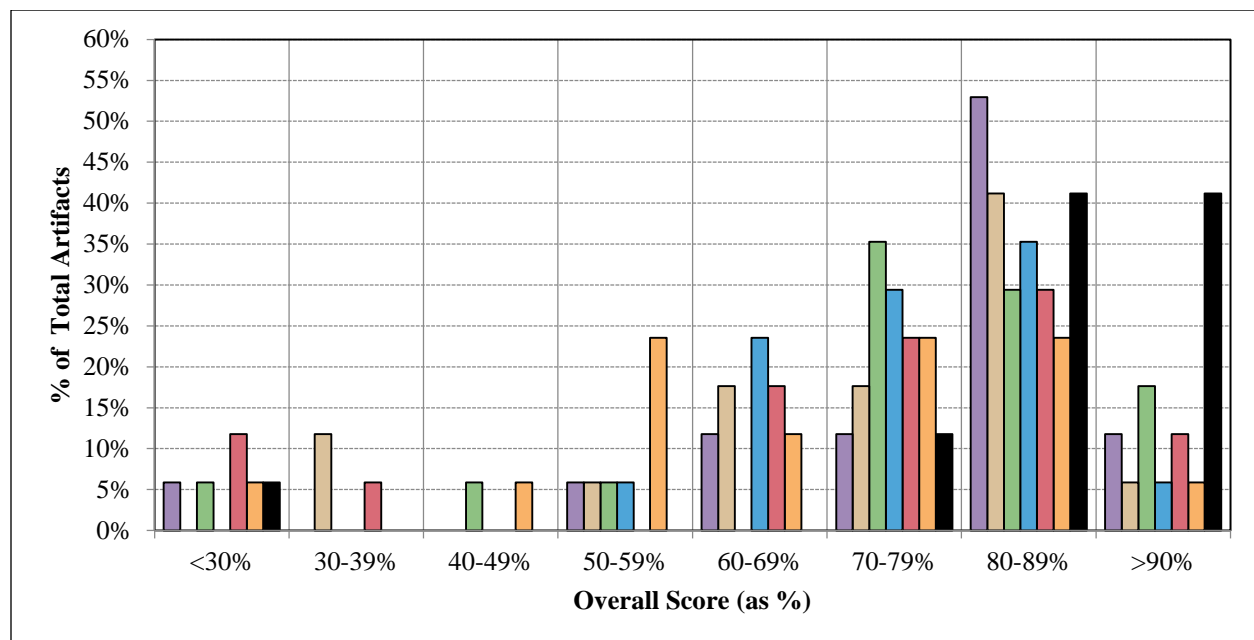


Figure 19. Histogram of SPN 1121 exam sections scores for fall 2015. Purple – Section I, Brown – Section II, Green – Section III, Blue – Section IV, Red – Section V, Orange – Section VI, and Black – Section VII.

Normally, 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. However, since SPN 1121 has a sample size of just 11, such a study could not be conducted.

3.3.3.2 Longitudinal Study

Further description of achievement over time in SPN 1121 is provided in a comparison of fall 2013 through fall 2015 data (Table 14 and Figure 20). 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 IV and VII exhibit steady increases. Section IV increased from 9.5/15 in fall 2013 to 11.3/15 in fall 2015, a

gradual improvement of 12%. Meanwhile, Section VII increased from 9.6/20 in fall 2013 to 16.4/20 in fall 2015, a gradual improvement of 34%. Section III exhibits a steady decline from 34.2/40 in fall 2013 to 29.9/40 in fall 2015, a steady decline of nearly 11%.

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

Table 14. Comparison of mean scores for SPN 1121 for fall 2013 through fall 2015.

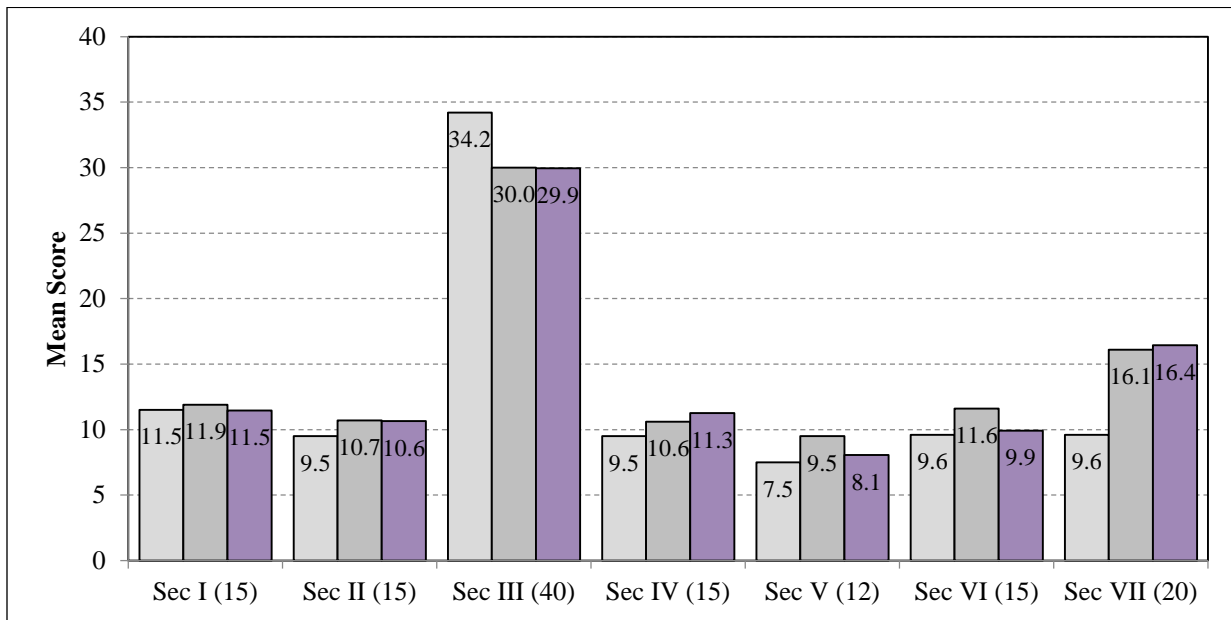


Figure 20. Comparison of mean scores for SPN 1121 through time from fall 2013 (light gray), through fall 2014 (dark gray), to fall 2015 (purple).

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 fall 2015 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 semester.
2. Achievement of 80% of artifacts scoring 70% or better (SLO 2): Achievement goal was met as results exhibit 81% 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 96% 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, Section III and VIII, at 63% and 58% achieving 70% or greater, are the lowest scoring sections.
5. In a comparison of dual enrollment to traditional artifacts, dual enrollment artifacts score 1.4 points lower than traditional. However, results were not statistically significantly different.
6. In a comparison of online to traditional artifacts there is a statistically significantly lower mean score for online artifacts (67.0/124 compared with 98.7/124).
7. No cross-campus comparison was completed because artifacts reported originated from only one site other than online or dual enrollment.
8. In a study of score distribution by Section, section II exhibits a peak in the 80-89 scoring bin. Sections IV and VIII exhibit slight spikes at < 30% scores.
9. In a longitudinal study of data distribution, from fall 2014 to fall 2015 three sections exhibit slight increases while the remaining six exhibit declines.

A drill-down of FRE 1121 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 semester.
2. Achievement of 80% of artifacts scoring 70% or better (SLO 2): Achievement goal was almost met as results exhibit 78% of artifacts score 70% or higher in the reading competency exam section.
3. Achievement of 80% of artifacts scoring 70% or better (SLO 3): Achievement goal was met as results exhibit 89% 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, Section VIII, at 22% achieving 70% or greater, was the lowest scoring sections.
5. No comparison of dual enrollment to traditional artifacts was completed because no dual enrollment sections were offered during fall 2015.
6. In a comparison of online to traditional artifacts there is a statistically significantly lower mean score for online artifacts (48.8/101 compared with 77.9/101).
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, Section VIII appears to be the only section substantially different than the others, with a peak centered in to 60-69% scoring bin where others are 80-89 or higher.
9. In a longitudinal study of data distribution, from fall 2014 to fall 2015 seven sections exhibit slight increases while the remaining four exhibit declines.

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 not met as results exhibit only 57% 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 53% of artifacts scored 70% or higher in the first reading section, Section II and 87% of artifacts scored 70% or higher in the second reading section, Section III.
3. Achievement of 80% of artifacts scoring 70% or better (SLO 3): Achievement was not met as results exhibit 73% of artifacts scored 70% or higher in the first writing section, Section IV, and 71% of artifacts scored 70% or higher in the second writing section, Section V.
4. No comparison of dual enrollment to traditional artifacts was completed because no dual enrollment sections were offered during fall 2015.
5. In a comparison of online to traditional artifacts, online artifacts score 24 points lower than traditional. However, results were not statistically significantly different.
6. In a cross-campus comparison, Collier exhibits the highest mean across sites (97.4/125) while FSW Online exhibits the lowest (78.2/125) out of four sites in which artifacts were collected (Charlotte, Collier, FSW Online, and Thomas Edison (Lee)). Results were not statistically significantly different.
7. In a study of score distribution by section, Section II exhibits far fewer artifacts in the upper scoring bins than other sections. In both Section I and Section II exhibit less than 60% achievement of the SLO, while Sections III, IV, and V exhibit 70% or greater.
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 75%-89%. 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, from fall 2013 to fall 2015 Sections II and IV exhibit steady increases. Section II increased from 40.4/60 in fall 2013 to 41.8/60 in fall 2015, a gradual improvement of 2%. Section IV increased from 10.6/15 in fall 2013 to 11.5/15 in fall 2015, a gradual improvement of 6%. Section I exhibits a steady decline from 12.1/15 in fall 2013 to 10.8/15 in fall 2015, a steady decline of nearly 9%.
10. In a pilot study assessing student skills and retention of materials from SPN 1120 prior to beginning SPN 1121, results exhibit a wide range of achievement levels based on previous instructor spanning from as low as 9.0/30 (for FSW instructors) to as high as 20.8/30 that were statistically significant.

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

1. Achievement of 80% of artifacts scoring 70% or better (SLO 1): Achievement was met as results exhibit 76% of artifacts score 70% or higher in the oral competency exam section (Section I).
2. Achievement of 80% of artifacts scoring 70% or better (SLO 2): Achievement was not met as results exhibit 65% of artifacts scored 70% or higher in the first reading section, Section II and 53% of artifacts scored 70% or higher in the second reading section, Section VI.
3. Achievement of 80% of artifacts scoring 70% or better (SLO 3): Achievement was partially met as results exhibit 65% of artifacts scored 70% or higher in the first writing section, Section V, and 94% of artifacts scored 70% or higher in the second writing section, Section VII.
4. No comparison of dual enrollment to traditional artifacts was completed because no dual enrollment sections were offered during fall 2015.

5. While online course sections were held in fall 2015, 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. No cross-campus comparison was completed because only one section of the course reported results of the common final exam.
7. In a study of score distribution by section, Sections I, II, III, IV, V, and VII exhibit scores centered on 80-89%. Section VI exhibits a bimodality with a second peak at 50-59%.
8. In a longitudinal study of data distribution, from fall 2013 to fall 2015 Sections IV and VII exhibit steady increases. Section IV increased from 9.5/15 in fall 2013 to 11.3/15 in fall 2015, a gradual improvement of 12%. Section VII increased from 9.6/20 in fall 2013 to 16.4/20 in fall 2015, a gradual improvement of 34%. Section III exhibits a steady decline from 34.2/40 in fall 2013 to 29.9/40 in fall 2015, a steady decline of nearly 11%.

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.
- de Winter, J.C.F. 2013. Using the Student's T-Test with Extremely Small Sample Sizes. *Practical Assessment, Research, and Evaluation*, 18(10), 1-12.
- 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.