Fall 2016

Assessment of Student Work from Across the Core

Scoring & Results for Communications

Report prepared by the Office of Student

Learning & Institutional Assessment

October 30, 2017

Executive Summary

The purpose of the Texas Core Curriculum (TCC) assessment is to identify the level of student attainment of the TCC core objectives and determine where to focus academic resources to improve those objective areas that are determined by the university as needing improvement.

To assess the core objectives mandated by the Texas Higher Education Coordinating Board, Stephen F. Austin State University collects student work samples in core courses. The student works samples, or artifacts, are then scored by a faculty scoring team using modified AAC&U LEAP VALUE rubrics.

This report considers the Oral and Visual Communication samples, which were collected in Spring 2016 and Fall 2016. Although VALUE rubric data is categorical in nature (if not descriptive), mean averages of each element indicated an *increase in scores* from 2014 to 2016.

This report also considers Written Communication samples and Written and Visual Communication samples, which were collected in Fall 2016. Using the same analysis, mean averages of each assessed element indicated a *decrease in scores* from Fall 2014 to Fall 2016 with significant performance declines in (a) sources and evidence; (b) organization and presentation; and (c) control of syntax and mechanics.

Some of the variation between the 2014 scores and 2016 writing scores may be explained by the differing student population characteristics that these samples represent. Nevertheless, these scores may indicate a declining rate of performance quality from SFA students over time.

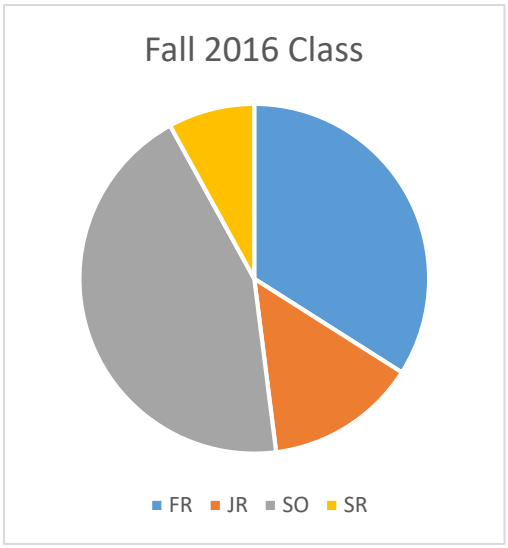
Oral & Visual Communications

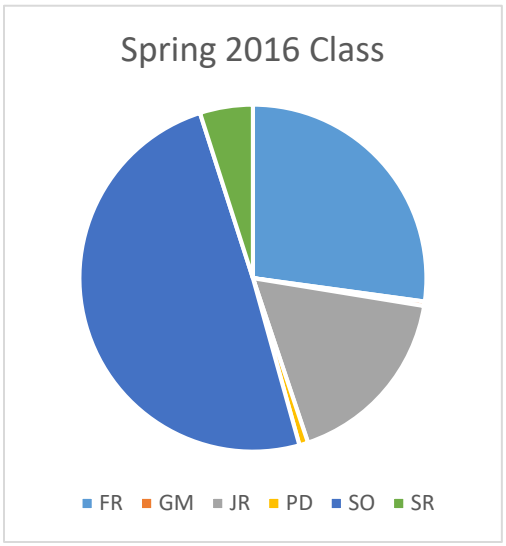
Method

Faculty members designed specific assignments for all related sections of courses designated "Core." Students then uploaded these assignments into the LiveText system online. From these collections, a Core Curriculum Scoring Team generated a random sample set for review. Each artifact was scored by a minimum of two raters (See Appendix D).

Participants

The generated sample was similar to the overall SFA student population in terms of race and gender. The plurality of participants in the Fall 2016 semester were Sophomores, while the Spring 2016 class held a plurality of Freshmen. This may infer the plurality from *both* semesters emanated from the same entering class.





Section enrollments for the participating courses were larger in Spring 2016 when compared to Fall 2016. However, submission rates increased from one semester to the next, as indicated in Table 1.

	Spring 2016	Fall 2016
Enrollment	999	549
Submission Count	660	395
Submission Rate/Percentage	66.1%	71.9%

Table 1: Course Enrollment and Submission Rates

Scoring Team and Sampling

Student work was scored by teams of faculty who were nominated by their respective departments and then selected by the Core Curriculum Assessment Committee (CCAC). The team consisted of ten members drawn from departments teaching core courses in which core objectives were assessed.

Scoring Team members were asked to report any artifacts that did not match the assignment, were plagiarized, or contained no content. These artifacts were eliminated from the scoring sample. Because of the unique nature of these artifacts (student self-made video), a higher percentage of artifacts were unusable at first. Overall, 47 samples were deemed unusable in the Spring 2016 sample. Through improvements in communication and infrastructure, the situation improved. Only three samples were unusable in the Fall 2016 semester.

Rubric

The rubrics to assess each component of the core were developed by faculty teams who modified the Association of American Colleges and Universities (AAC&U) VALUES Rubrics. The AAC&U rubrics were adapted to best fit the objectives of the SFA core. The rubric for Oral & Visual Communications can be found in Appendix A. Each rubric measures specific criteria using a 5-category continuum,
labeled 0 - 4. For purposes of this report, the data has remained consistent with the rubric's scoring system. Benchmark labels are listed in Table 2.

Score	Correlation
0	Unacceptable
1	Beginning
2	Developing
3	Accomplished
4	Capstone

Table 2: Rubric Category Scores and Corresponding Descriptions

Scoring Team Rubric Calibration

In Fall 2016, each scoring team met for two rubric calibration sessions facilitated by the Office of Student Learning and Institutional Assessment. During these sessions, the team discussed the rubric extensively and developed rules for scoring student work. The calibration sessions were used to familiarize the faculty with the rubric that they would be using for scoring, allowing them to develop shared understanding of the language used on the rubric, and to become familiar with the process of scoring using LiveText. During the session, non-sample student artifacts were scored and discussed by the team. Further scoring rules were developed if needed following the scoring of each artifact.

Scoring

The LiveText sampling tool was used to draw a random sample of student work from each objective. The Spring 2016 sample (n = 223) was drawn with the intention of having a minimum of 200 pieces of scorable student work. This was keeping with previous practice.

The Southern Association of Colleges and School Commission on Colleges (SACSCOC) recently imposed numerous sanctions on institutions based on sample size calculations. SACSCOC requires definitive reasoning behind any sampling presented to the Commission. Thus, changes were made to sampling procedures. Sample sizes were calculated with a confidence level of 80% and a margin of error of 10% using the following formula $Z^2*(p)*(1-p)/c^2$ where Z represents the Z value (in this case, 1.28), p is the population of submitted work in a specific core area, and c is the confidence interval (.1). This resulted in a sample size of 50 artifacts in the Fall of 2016.

Each artifact of student work in the sample was sent to two raters. Raters evaluated the paper in LiveText using an online copy of the rubric and following the rules developed in the calibration sessions. If the two raters had disagreement on a criterion, the artifact was then sent to a third rater to score only the criteria for which there was disagreement. A complete list of the rules for

agreement/disagreement can be found in Appendix B. Faculty on the scoring teams were given two weeks to complete their first scoring round and then an additional week to finish their second round of scoring.

Results

Inter-rater agreement (within one point in each rating) was 91.5% for the Spring 2016 semester and 96.6% for the Fall 2016 semester. For those requiring a third rater, 59.5% needed a third rater for only one of the six elements being evaluated in the Spring. The same is true for 57.9% of the Fall scores.

Mean and mode are reported below for each rubric criterion (See Table 3 and Table 4). Frequency counts are illustrated through bar charts to assist with visualization and understanding. This is in keeping with admonishments from the Association of American Colleges & Universities:

> *Do not, to the extent possible, show means in the absence of descriptive context as that reinforces the false notion of scale. As part of scorer training on the VALUE rubrics, individuals are "forced" to select a single performance level for each dimension. They must assign a student work product to a single, albeit ordered category of performance, not assign placement on a continuum or scale. Such ordinal data may be better described by medians, frequency distributions, and bar charts. Furthermore, this also implies that some statistical procedures may be more appropriate for analyzing the data generated from VALUE rubrics (e.g., analysis of variance, etc.) than others.*
>
> *Do not average the scores assigned to each dimension on a VALUE rubric to create a total score for the rubric. The power of the VALUE rubrics rests in the ability to focus attention on the specific learning addressed within each dimension; a total score for the rubric provides little diagnostic assistance to students or faculty. Furthermore, averaging across rubric dimensions makes methodological assumptions that are inappropriate when treating the VALUE data as ordinal.*[1]

[1] On Solid Ground: VALUE Report 2017. Report. Association of American Colleges & Universities. Washington, DC, 2017. 28.

Spring 2016	Mean	Mode
Organization	2.40	3
Language	2.27	2
Delivery (oral/visual)	2.00	2
Evidence-based support	2.24	3
General purpose	2.48	3
Visual aids	1.85	2

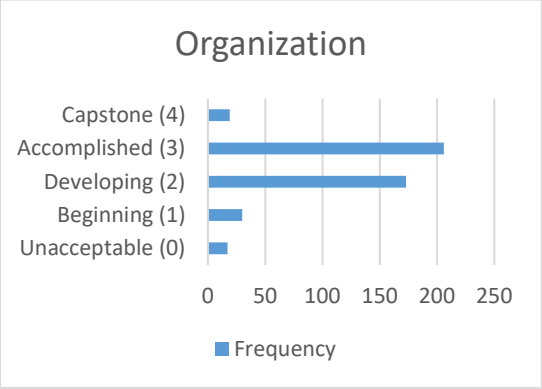
Table 3: Oral and Visual Communication Means and Modes Spring 2016

Fall 2016	Mean	Mode
Organization	2.42	3
Language	2.41	3
Delivery (oral/visual)	1.99	2
Evidence-based support	2.39	3
General purpose	2.75	3
Visual aids	2.09	3

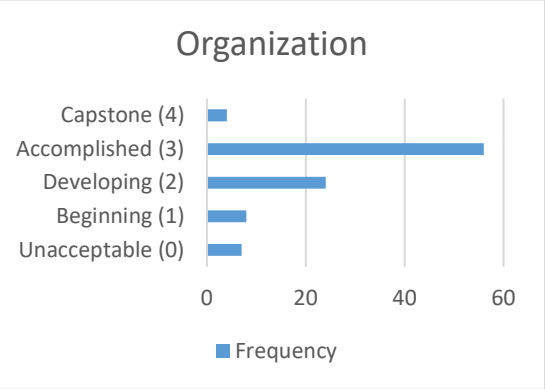
Table 4: Oral and Visual Communication Means and Modes Fall 2016

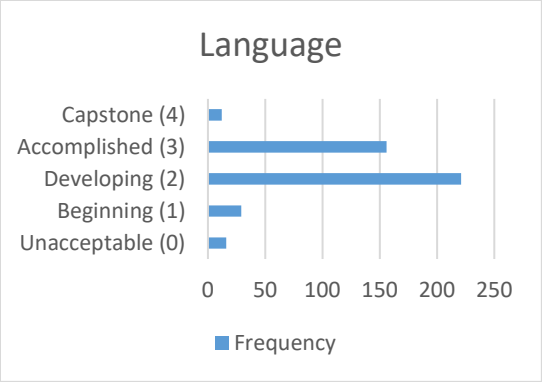
Frequency Counts: Oral and Visual Communication

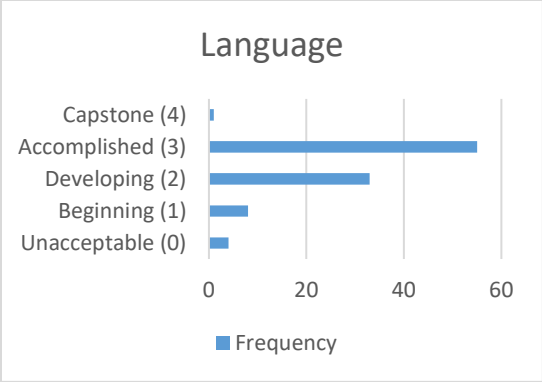
Spring 2016

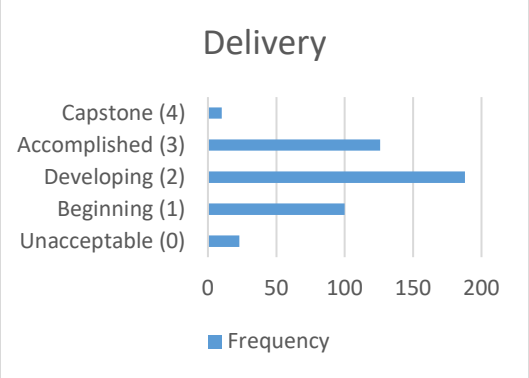


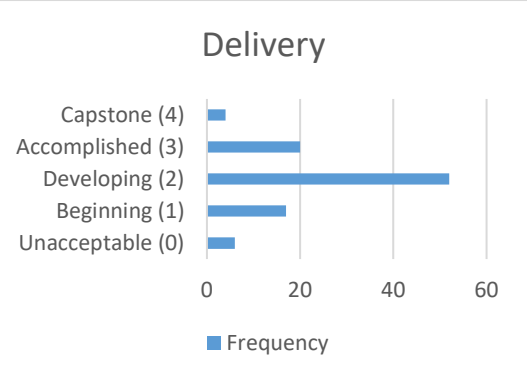
Fall 2016

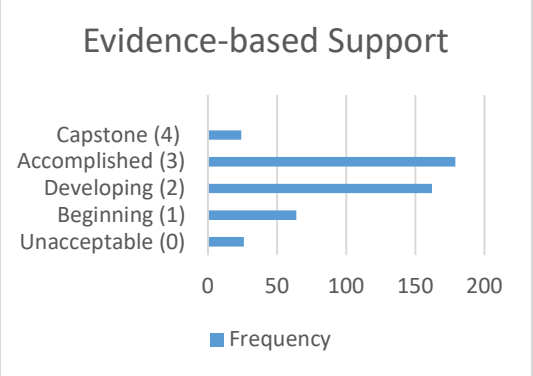


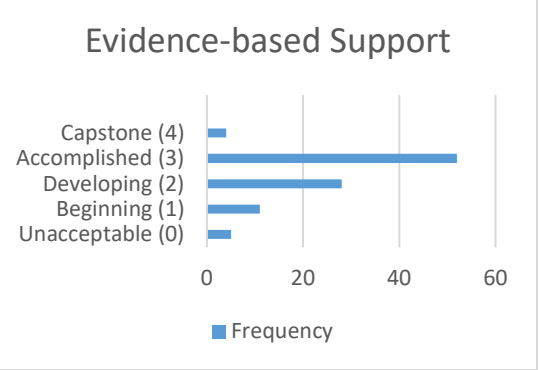

Language
Capstone (4)
Accomplished (3)
Developing (2)
Beginning (1)
Unacceptable (0)
0
50
100
150
200
250
Frequency

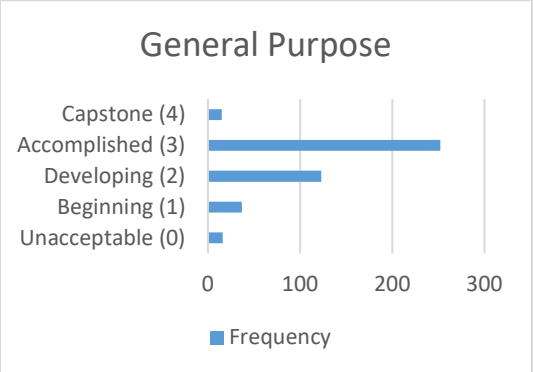

Language
Capstone (4)
Accomplished (3)
Developing (2)
Beginning (1)
Unacceptable (0)
0
20
40
60
Frequency

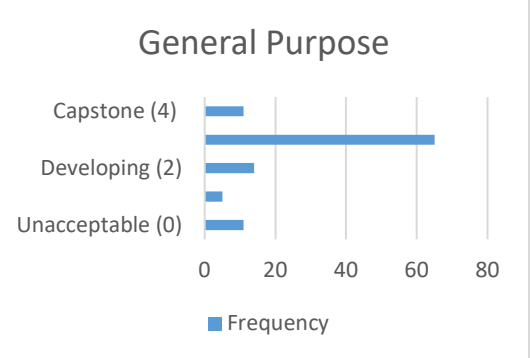

Delivery
Capstone (4)
Accomplished (3)
Developing (2)
Beginning (1)
Unacceptable (0)
0
50
100
150
200
Frequency

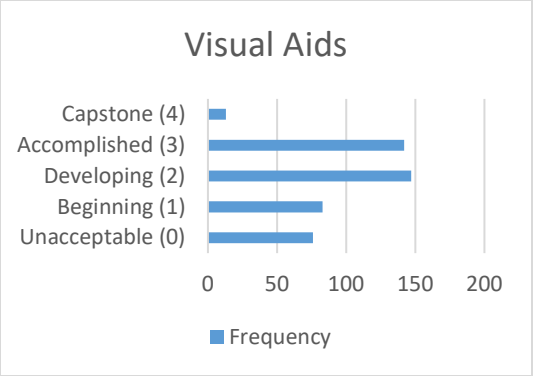

Delivery
Capstone (4)
Accomplished (3)
Developing (2)
Beginning (1)
Unacceptable (0)
0
20
40
60
Frequency

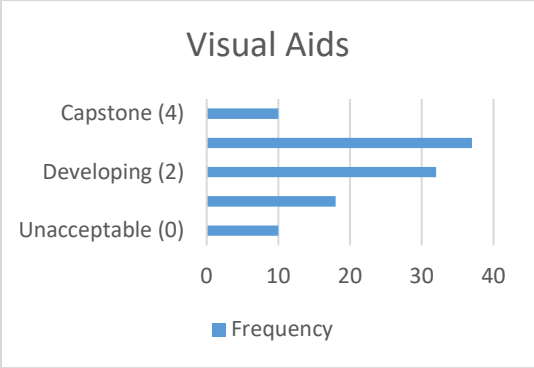

Evidence-based Support
Capstone (4)
Accomplished (3)
Developing (2)
Beginning (1)
Unacceptable (0)
0
50
100
150
200
Frequency


Evidence-based Support
Capstone (4)
Accomplished (3)
Developing (2)
Beginning (1)
Unacceptable (0)
0
20
40
60
Frequency


General Purpose
Capstone (4)
Accomplished (3)
Developing (2)
Beginning (1)
Unacceptable (0)
0
100
200
300
Frequency


General Purpose
Capstone (4)
Developing (2)
Unacceptable (0)
0
20
40
60
80
Frequency





Scoring Team ratings generally followed similar patterns from one semester to the next. It should be noted that five of the six elements considered were highly correlated with each other, while the Visual Aids category showed moderate correlations. This may be due to the fact that a large number of Visual Aids ratings were zero (0), based on the lack of any visual aid, whatsoever. The overall Cronbach's Alpha was .89. Table 5 indicates correlations between specific pairs of rubric elements.

	Organization	**Language**	**Delivery**	**Evidence**	**Gen.Purp**	**VisAids**
Organization	1.000	0.751	0.678	0.622	0.758	0.462
Language	0.751	1.000	0.695	0.626	0.702	0.399
Delivery	0.678	0.695	1.000	0.530	0.653	0.511
Evidence	0.622	0.626	0.530	1.000	0.646	0.487
General Purpose	0.758	0.702	0.653	0.646	1.000	0.488
Visual Aids	0.462	0.399	0.511	0.487	0.488	1.000

Table 5: Inter-Item Correlation Matrix

Although VALUE rubrics create ordinal and categorical data, mean averages of each element indicated an *increase in scores* from 2014 to 2016 (refer to Table 3 and Table 4). Mann-Whitney U analysis of scores is documented in Table 6. Analysis indicated statistically significant differences between semesters for two of the six elements. Language and General Purpose.

	Organization	Language	Delivery	Evidence	Gen. Purpose	Visual Aids
Mann-Whitney U	20705.500	19078.500	21842.000	20479.000	17629.000	20865.500
Wilcoxon W	119940.500	113473.500	26792.000	124219.000	115975.000	127356.500
Z	-1.020	-2.225	-.213	-1.663	-3.449	-1.849
Asymp. Sig. (2-tailed)	.308	.026	.831	.096	.001	.064

Table 6: Mann-Whitney U Comparison (Oral and Visual Communication) grouping Variable: Semester

One interesting change could be the Language element. Spring 2016 students were listed primarily as Developing (2); Fall 2016 students tended to be rated as Accomplished (3). Visual Aid usage also was rated higher in the Fall, with a smaller percentage being rated as Unacceptable (a drop from 16.5% to 9.8%). There were slightly more students rated as Accomplished in the Fall, while the Spring sample indicated more students at the Developing level. The drop in Unacceptable markings likely accounts for the difference in ratings between Fall and Spring of 2016. The Fall semester used video artifacts from only one course, while the previous scoring sample included scores from multiple courses. One potential effect could be that General Purpose may have been easier to ascertain by Scoring Team members. This singular structural change may answer most of the score increase in this element.

These three elements indicate statistically significant changes; however, the *real* change in mean scores for the three elements ranged from .14 to .27. As Hilda Bastian wrote for the *Scientific American*,

> *Statistical significance testing can easily sound as though it sorts the wheat from the chaff, telling you what's "true" and what isn't. But it can't do that on its own. What's more, "significant" doesn't mean it's important either. A sliver of an effect can reach the less-than-5% threshold.*[2]

[2] Hilda Bastian, "Statistical significance and its part in science downfalls," *Absolute Maybe*, Scientific American, November 11, 2013, https://blogs.scientificamerican.com/absolutely-maybe/statistical-significance-and-its-part-in-science-downfalls/

Moving Forward

Following each semester's artifact assessment, a debrief meeting was held with the Oral and Visual Scoring Team. At the end of the spring semester, team members noted their overall feelings on SFA students' oral and visual communication capabilities. The consensus was four words, "We're in good shape."

While these rubric data are more descriptive in nature, some general concepts can be considered:

1. Students who begin their core are typically rated as at least Developing in their level of oral communication.
2. When the Visual Aid factor is removed, Delivery seems to be the most challenging Oral Communication element for SFA students.
3. All assignments used in scoring likely need to be graded assignments. This keeps the spirit and effectiveness of the SFA VALUE rubrics.
4. SFA students may need more specific instruction on the use of Visual Aids

Written and Written & Visual Communications

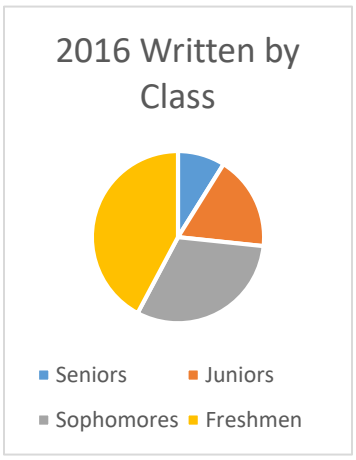
This section considers the Written Communication samples and the Written and Visual Communication samples, which were collected in Fall 2016. Faculty members designed specific assignments for all related sections of courses designated "Core." Students then uploaded these assignments into the LiveText system online. From these collections, random samples were selected for review by a Core Curriculum Scoring Team.

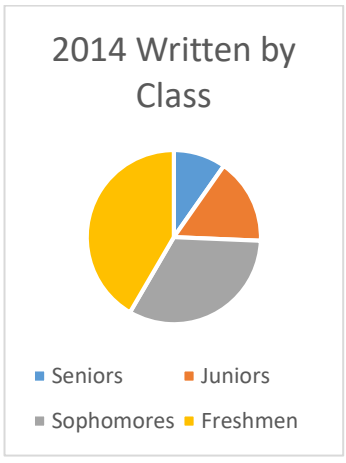
Method

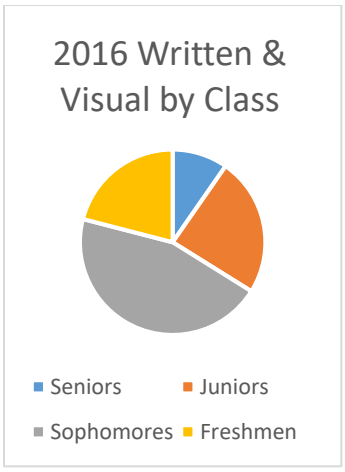
Faculty members designed specific assignments for all related sections of courses designated "Core." Students then uploaded these assignments into the LiveText system online. From these collections, a random sample set was generated for review by a Core Curriculum Scoring Team. Each artifact was scored by a minimum of two raters (See Appendix D).

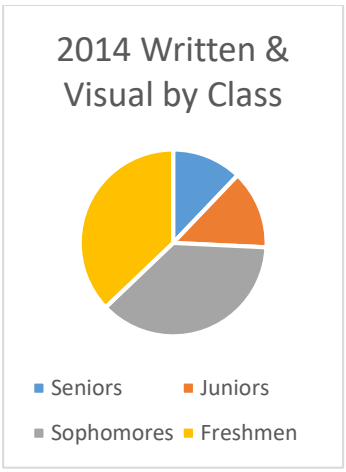
Participants

The generated sample was similar to the overall SFA student population in terms of race and gender. Both the Written Communication samples had pluralities of Freshmen; however, the Written and Visual Communications samples had at least as many Sophomores as Freshmen. In fact, the number of Freshmen *dropped* in the WV sample. Regarding class standing, the Written Communications sample from 2016 is almost a mirror of the 2014 sample. The Written & Visual Communications class data indicate an uptick in the percentage of upperclassmen (specifically, Juniors) in the sample. The jump in Sophomores may be significant, in terms of demographic effects on student outcomes.









Section enrollments for the participating courses were similar in size between years. However, submission rates fluctuated from the previous collection period, as indicated in Table 7 and Table 8.

	Fall 2016	**Fall 2014**
Enrollment	4675	4604
Submission Count	3633	3804
Submission Rate/Percentage	77.7%	82.6%

Table 7: Written Communication Submission Rates

	Fall 2016	**Fall 2014**
Enrollment	1373	6620
Submission Count	1097	4511
Submission Rate/Percentage	79.8%	68.1%

Table 8: Written & Visual Communication Submission Rates

Scoring Team and Sampling

Teams of faculty who were nominated by their respective departments and then selected by the Core Curriculum Assessment Committee (CCAC) scored student work. The team consisted of ten members drawn from departments teaching core courses.

Scoring Team members were asked to report any artifacts that did not match the assignment, were plagiarized, or contained no content. These artifacts were eliminated from the scoring sample.

Rubric

The rubrics to assess each component of the core were developed by faculty teams who modified the Association of American Colleges and Universities (AAC&U) VALUES Rubrics. The AAC&U rubrics were adapted to best fit the objectives of the SFA core. The rubric for Written Communication can be found in Appendix A. The rubric for Written & Visual Communications can be found in Appendix B. Each rubric measures specific criteria using a 5-category continuum, labeled 0 - 4. For purposes of this report, the data has remained consistent with the rubric's scoring system, with the benchmarks in Table 9

Score	Descriptor
0	Unacceptable
1	Beginning
2	Developing
3	Accomplished
4	Capstone

Table 9: Rubric Category Scores and Corresponding Descriptions

Rubric Calibration

In Spring 2017, each scoring team met for rubric calibration sessions, facilitated by the Office of Student Learning and Institutional Assessment. During these sessions, the team discussed each rubric extensively and developed rules for scoring student work. The calibration sessions were used to familiarize the faculty with the rubric that they would be using for scoring, allowing them to develop shared understanding of the language used on the rubric, and to become familiar with the process of scoring. Since most of the Scoring Team members were serving a second time, the calibration sessions functioned in more of an "update" capacity, refreshing previous knowledge. During each session, non-sample student artifacts were scored and discussed by the team.

Scoring

The LiveText sampling tool was used to draw a random sample of student work from each objective. Scoring team members were asked to report any artifacts that did not match the assignment, were plagiarized, were not scorable, or were blank documents. These artifacts were eliminated from the scoring sample. Overall, seven samples were deemed unusable in the two groups.

The Fall 2014 Written Communications sample (n = 114) and Written and Visual Communications sample (n = 135) were drawn with the intention of having a minimum of 100 pieces of scorable student work in each area. The Southern Association of Colleges and School Commission on Colleges (SACSCOC) recently imposed numerous sanctions on institutions based on sample size calculations. SACSCOC requires definitive reasoning behind any sampling presented to the Commission. Thus, changes were made to sampling procedures, effective in the Spring of 2017. Sample sizes were calculated with a confidence level of 80% and a margin of error of 10% using the following formula $Z^2*(p)*(1-p)/c^2$ where Z represents the Z value (in this case, 1.28), p is the population of submitted work in a specific core area, and c is the confidence interval (.1). The sample for Written Communication was 62, with the final sample for Written & Visual Communication at 45.

Each artifact of student work in the sample was sent to two raters. Raters evaluated the paper in LiveText using an online copy of the rubric and following the rules developed in the calibration sessions. If the two raters had disagreement on a criterion, the artifact was then sent to a third rater to score only the criteria for which there was disagreement. A complete list of the rules for agreement/disagreement can be found in Appendix C. Faculty on the scoring teams were given two weeks to complete their first scoring round and then an additional week to finish their second round of scoring.

Results

Written Communication Scoring Team agreement (within one point in each rating) was 94.9% for the Fall 2016 semester, and the Written & Visual Scoring Team agreement was 93.0% for the Fall 2016 semester. Inter-Class Correlation analysis resulted in a Cronbach's alpha of .87 for the Written scores and .89 for the Written & Visual scores. For those requiring a third rater, over half needed a third rater for only one of the elements being evaluated in Fall 2016.

Mean and mode are reported below for each rubric criterion in Table 10 and Table 11. Frequency counts are illustrated through bar charts to assist with visualization and understanding. This is in keeping with admonishments from the Association of American Colleges & Universities:

> Do not, to the extent possible, show means in the absence of descriptive context as that reinforces the false notion of scale. As part of scorer training on the VALUE rubrics, individuals are "forced" to select a single performance level for each dimension. They

must assign a student work product to a single, albeit ordered category of performance, not assign placement on a continuum or scale. Such ordinal data may be better described by medians, frequency distributions, and bar charts. Furthermore, this also implies that some statistical procedures may be more appropriate for analyzing the data generated from VALUE rubrics (e.g., analysis of variance, etc.) than others.

Do not average the scores assigned to each dimension on a VALUE rubric to create a total score for the rubric. The power of the VALUE rubrics rests in the ability to focus attention on the specific learning addressed within each dimension; a total score for the rubric provides little diagnostic assistance to students or faculty. Furthermore, averaging across rubric dimensions makes methodological assumptions that are inappropriate when treating the VALUE data as ordinal.[3]

Written Communication	2016 Mean	2016 Mode	2014 Mean	2014 Mode
Audience, Context, and Purpose	2.07	3	2.11	2
Content Development	1.70	2	1.79	2
Sources and Evidence	1.76	2	2.14	3
Organization And Presentation	1.70	1	1.89	2
Control of Syntax and Mechanics	1.88	2	2.16	2
Visual Aids	N/A	N/A	N/A	N/A

Table 10: Written Communication Means and Modes by Year

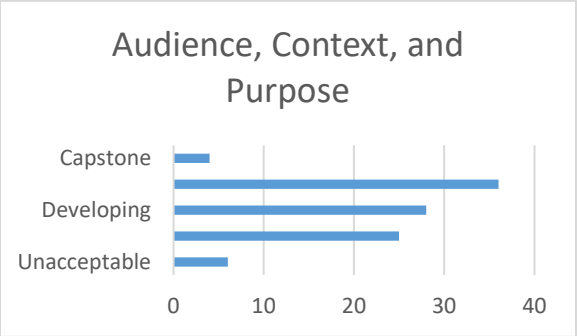
[3] [3]AAC&U, On Solid Ground (Washington: AAC&U, 2017) 28.

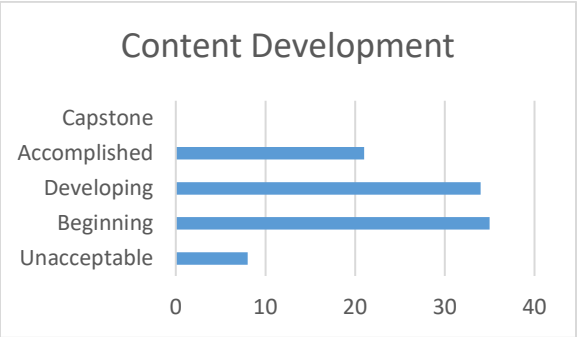
Written & Visual Communication	2016 Mean	2016 Mode	2014 Mean	2014 Mode
Audience, Context, and Purpose	1.97	2	2.06	2
Content Development	1.60	1	1.79	2
Sources and Evidence	1.19	0	1.60	2
Organization And Presentation	1.61	X	1.83	2
Control of Syntax and Mechanics	1.81	2	2.07	2
Visual Aids	1.37	2	1.59	1

Table 11: Written and Visual Communication Means and Modes by Year

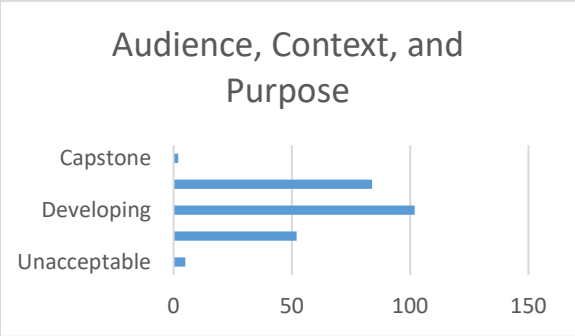
Frequency Counts: Written Communication

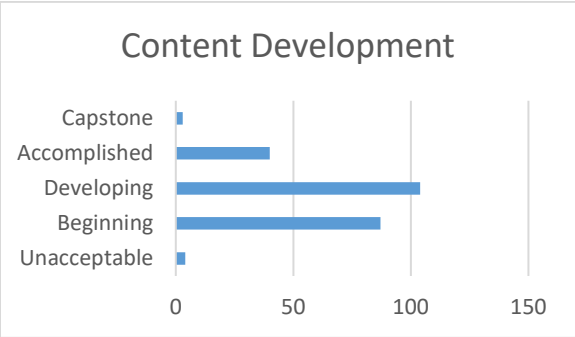
Fall 2016

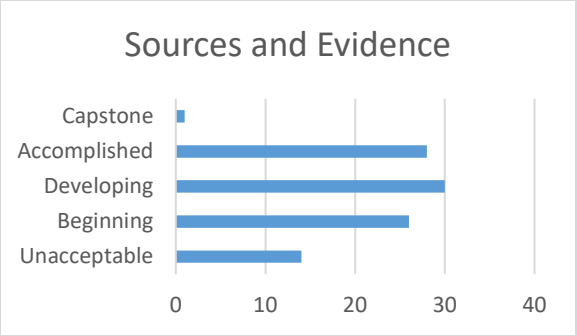


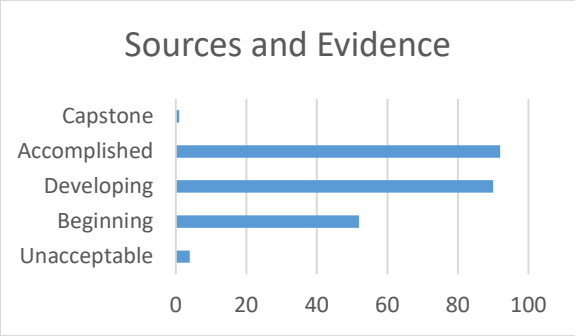


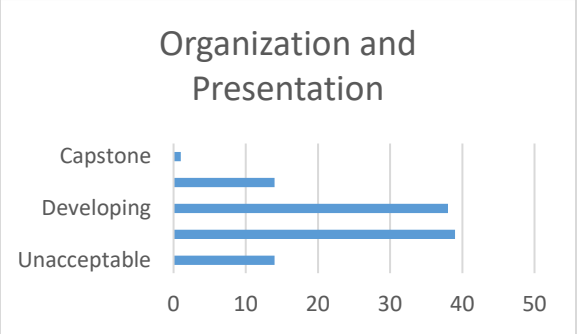
Fall 2014

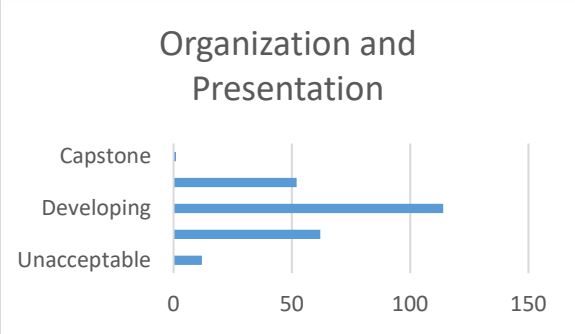


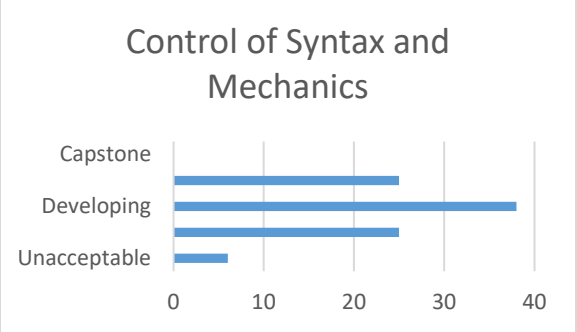


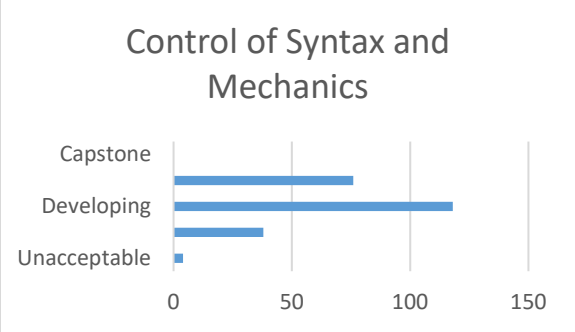






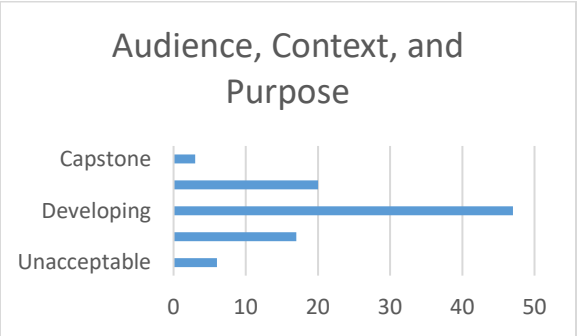




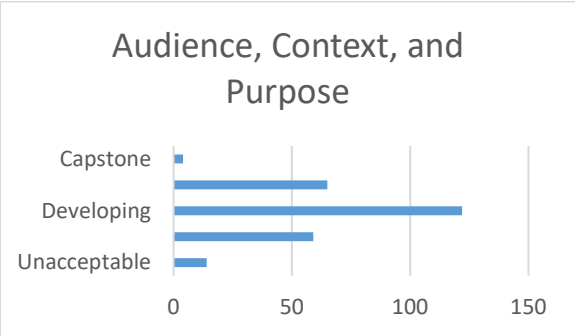


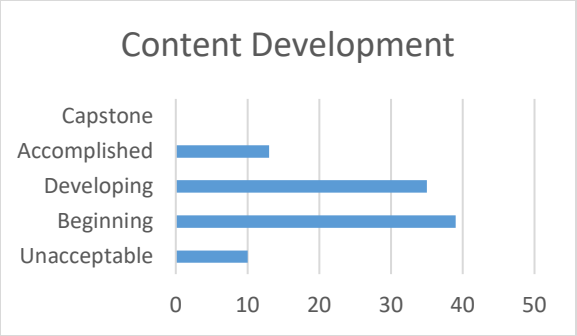
Frequency Counts: Written & Visual Communication

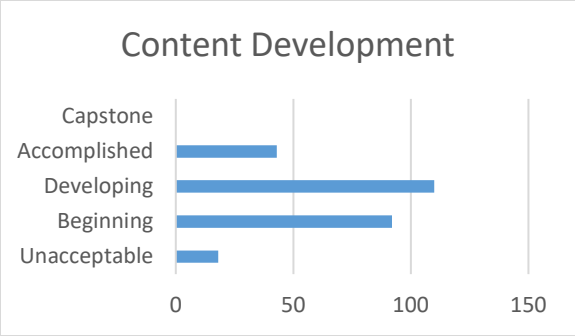
Fall 2016

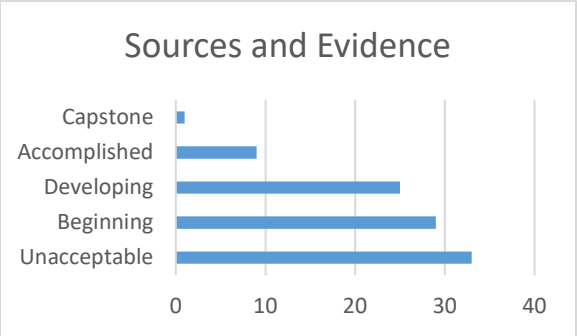


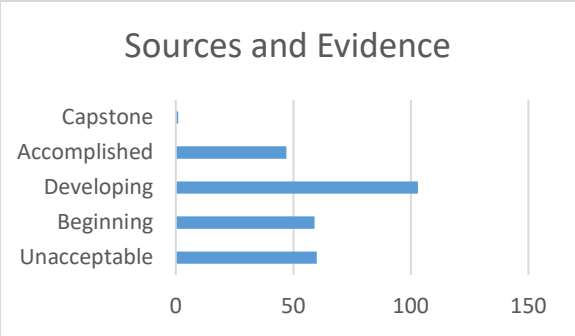
Fall 2014

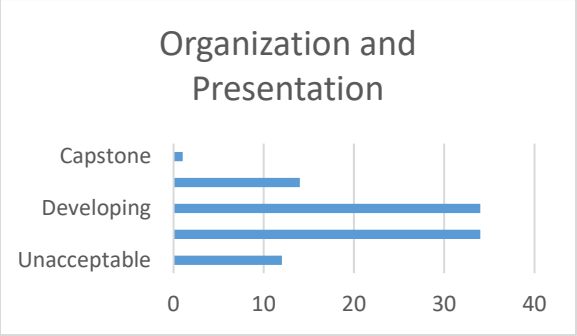


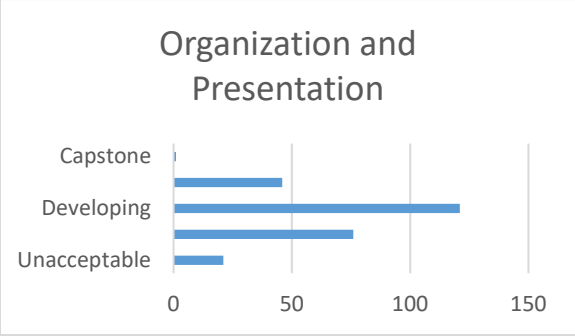

Content Development
Capstone
Accomplished
Developing
Beginning
Unacceptable
0
10
20
30
40
50

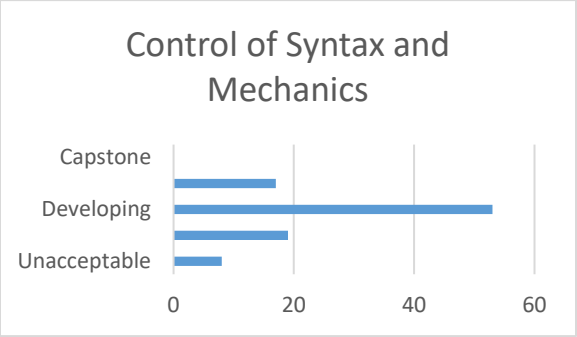

Content Development
Capstone
Accomplished
Developing
Beginning
Unacceptable
0
50
100
150

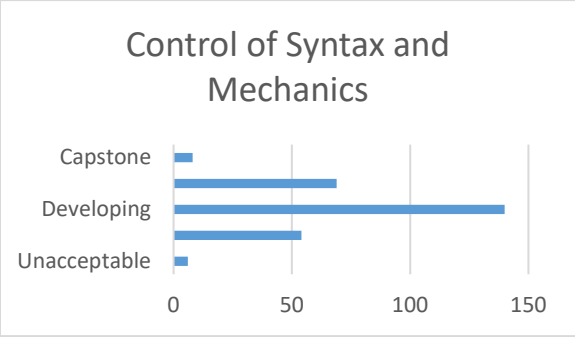

Sources and Evidence
Capstone
Accomplished
Developing
Beginning
Unacceptable
0
10
20
30
40

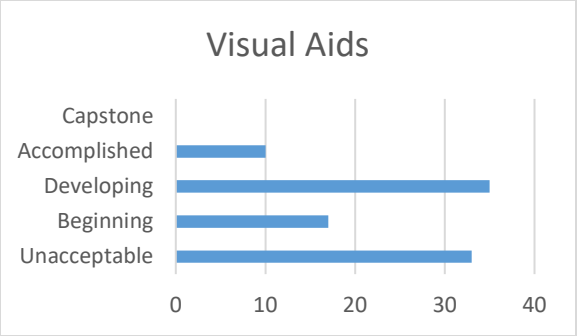

Sources and Evidence
Capstone
Accomplished
Developing
Beginning
Unacceptable
0
50
100
150

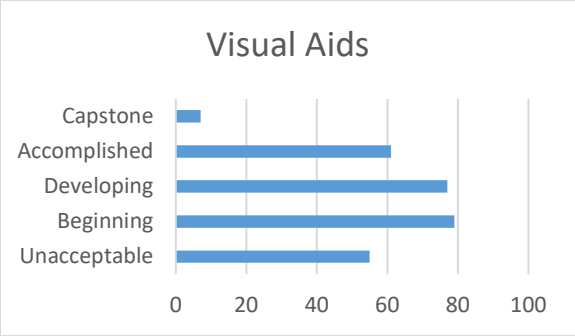

Organization and Presentation
Capstone
Developing
Unacceptable
0
10
20
30
40


Organization and Presentation
Capstone
Developing
Unacceptable
0
50
100
150


Control of Syntax and Mechanics
Capstone
Developing
Unacceptable
0
20
40
60


Control of Syntax and Mechanics
Capstone
Developing
Unacceptable
0
50
100
150





Inter-Item Correlation

Scoring Team ratings generally followed similar patterns from one semester to the next. It should be noted that both groups produced highly correlated elemental results. In Written Communications, all but one of the correlations were highly or moderately correlated with each other (Audience and Syntax were weakly correlated).

Regarding the Written and Visual Communication correlations, all areas were moderately or strongly correlated with one exception—Visual Aids and Sources were weakly correlated with each other. Specific correlations are shown in Table 12 and Table 13.

	Audience	**Content**	**Sources**	**Organization**	**Syntax**
Audience	1				
Content	0.695335	1			
Sources	0.596216	0.6387	1		
Organization	0.701443	0.712408	0.621084	1	
Syntax	0.298557	0.47004	0.54061	0.484303	1

Table 12: Inter-Item Correlation Matrix (Written Communications)

	Audience	Content	Sources	Organization	Syntax	Vis. Aids
Audience	1					
Content	0.748952	1				
Sources	0.656448	0.672758	1			
Org	0.84353	0.778054	0.633115	1		
Syntax	0.676257	0.759458	0.52771	0.696103	1	
Visual Aids	0.536904	0.457217	0.303929	0.463311	0.490093	1

Table 13: Inter-Item Correlation Matrix (Written and Visual Communications)

Although VALUE rubric data is ordinal in nature (if not descriptive), mean averages of each element indicated a decrease in scores from 2014 to 2016 (See Table 4 and Table 5). This decrease merited further investigation. Results of Mann-Whitney U analysis results are shown in Table 14 and Table 15.

	Audience	Content	Sources	Organization	Syntax	Visual Aids
Mann-Whitney U	12134	11484	9644	10696	9398	NA
Wilcoxon W	17184	16635	14694	15847	13958	NA
Z	-0.146	-0.692	-2.957	-2.057	-2.695	NA
Asymp. Sig. (2-tailed)	0.884	0.489	0.003	0.04	0.007	NA

Table 14: Mann-Whitney U Comparison (Written Communication) Grouping Variable: Year

	Audience	Context	Sources	Organization	Syntax	Visual Aids
Mann-Whitney U	11023	10742	9481	10243.5	10085.5	11101
Wilcoxon W	15394	15792	14431	14996.5	14838.5	16151
Z	-1.067	-2.333	-3.813	-2.527	-2.821	-1.955
Asymp. Sig. (2-tailed)	0.286	0.02	0	0.012	0.005	0.051

Table 15: Mann-Whitney U Comparison (Written and Visual Communication) Grouping Variable: Year

Written scores indicated significant differences between years for Sources and Evidence; Organization and Presentation; and Control of Syntax and Mechanics. It should be noticed that Elements 3, 4, and 5 were statistically different, as per Mann-Whitney U analysis. In all three instances, a score from the 2014 sample was higher than the corresponding score in the 2016 sample.

Mann-Whitney analysis of Written & Visual scores indicated significant differences between years for four elements: Content Development; Sources and Evidence; Organization and Presentation; and Control of Syntax and Mechanics. Again, the scores in the 2014 sample were higher than corresponding scores in the 2016 sample.

Visual investigation of the related bar charts (above) substantiates somewhat large discrepancies between years. Most obvious are the differences in the Organization and Presentation element and the Control of Syntax and Mechanics. Generally, the Scoring Team members in 2014 were more likely to rate an artifact as Accomplished (3), while the 2016 raters seemed to lean more toward Beginning (1) and Unacceptable (0) ratings. In fact, the plurality of the Written & Visual sample scored at a zero (0) in the 2016 sample on the Sources and Evidence element.

Discussion

The combined plurality of students in the assessed classes were Sophomores. However, the Written & Visual sample from 2016 contained a larger number of females and a lower number of Freshmen. There was a drop in the percentage of transfer students in the samples, as well. Some of the variation may be explained by the differing student population characteristics that these samples represent. Nevertheless, these scores may indicate a declining rate of performance quality from SFA students over time.

Moving Forward

As Linda Suskie recently posted, "Decisions are made with some level of uncertainty. Assessment results should reduce uncertainty but won't eliminate it."[4] These are only two samples, so new policies may be premature. While these rubric data are descriptive in nature, some general concepts can be found throughout:

1. Further investigation into student writing abilities may be warranted, as scores seem to be lowering over time.
2. Organization and Presentation is an element that may require specific attention by SFA students, faculty, and staff.
3. Use of Sources and Evidence seems to be the most challenging written and visual element for SFA students. This element showed the largest drop from both scoring teams.
4. Control of Syntax and Mechanics, which has a focus on choice of language and wording, seems to show a decline. Neither rubric addresses skills such as grammar and punctuation. This may be a topic for further discussion.
5. All assignments used in scoring should be graded assignments. This keeps the spirit and effectiveness of the SFA VALUE rubrics.
6. SFA students may need more specific instruction on the use of Visual Aids.

Appendix A: Oral and Visual Communication Rubric

2 Linda Suskie, *How to Assess Anything without Killing Yourself…Really*, online, Linda Suskie Blog, Internet, 30 May, 2017. Available: http://www.lindasuskie.com/apps/blog/show/44560748-how-to-assess-anything-without-killing-yourself-really-

Appendix B: Stephen F. Austin State University Written Communications Rubric

ORAL AND VISUAL COMMUNICATION RUBRIC

Definitions: Oral communication is, most generally, a prepared, purposeful presentation designed to increase knowledge, to foster understanding, or to promote change in the listeners' attitudes, values, beliefs, or behaviors. In some disciplines, oral communication is a direct interaction between a "professional" speaker and specific audience (such as nurse to patient, in clinical or therapeutic settings), where the effect of the presentation of ideas is to foster understanding or learning.

Visual communication is a prepared, purposeful presentation and delivery of supporting visual aids, typically relative to the oral communication and intended to benefit or amplify an audience's understanding of a central message or purpose. Thus, visual communication is facilitated through such media as poster presentations, power point presentations, video presentations, among others available and appropriate to specific disciplines and classroom experiences.

	Capstone 4	Accomplished 3	Developing 2	Beginning 1	Unacceptable 0
Organization	Organizational development is clearly and consistently observable; skillfully makes content and expression of ideas in the presentation cohesive.	Organizational development and expression of ideas are clearly and consistently observable within the presentation; content is expressed reasonably well as a result.	Organizational development and expression of ideas are observable within the presentation.	Organizational development and expression of ideas are occasionally observable.	Organizational development and/or expression of ideas are not observable within the presentation; lack of coherence and unity exist.
Language	Language choices are imaginative, memorable, and compelling; choices enhance presentation effectiveness. Language is appropriate to audience and aids the clear expression of ideas.	Language choices are thoughtful and generally support the effectiveness of the presentation. Language is appropriate to audience and is useful to the expression of ideas.	Language choices are mundane and commonplace and partially support the effectiveness of the presentation and the expression of ideas.	Language choices are sometimes unclear and minimally support the effectiveness of the presentation. Language appropriateness is inconsistent. Expression of ideas is hindered.	Language choices are unclear and fail to support the effectiveness of the presentation. Language is not appropriate to audience; ideas are not expressed clearly.
Delivery (oral/visual)	Delivery techniques make the presentation compelling; speaker appears polished and confident; speaker energy and emphases foster interpretation of ideas expressed. Dependency upon notes, if applicable, is not evident or intrusive. Non-verbal cues aid significantly.	Delivery techniques make the presentation interesting, and speaker appears comfortable; speaker tends toward conversational tone, and dependency upon notes is minimally noticeable. Non-verbal cues are appropriate and useful.	Delivery techniques make the presentation understandable; speaker appears tentative; speaker tends to be a bit casual, as evidenced in word choices; non-verbal cues do not particularly elevate audience's level of understanding or interpretation.	Delivery techniques sometimes detract from audience comprehension; speaker appears uncomfortable; speaker seems unenthusiastic, monotonic, or hesitancies suggest unpreparedness. Verbal cues include unnecessary gestures and purposeless body language.	Delivery techniques are either distracting from understandability of the presentation or fail to be effective; the speaker is clearly uncomfortable or unprepared.

WRITTEN COMMUNICATION RUBRIC

Definition: Written communication is the development and expression of ideas in writing. Written communication involves learning to work in many genres and styles. It can involve working with many different writing technologies, and mixing texts, data, and images. Written communication abilities develop through iterative experiences across the curriculum.

	Capstone 4	**Accomplished** 3	**Developing** 2	**Beginning** 1	**Unacceptable** 0
Audience, Context, and Purpose	Demonstrates a thorough understanding of context, audience, and purpose that is wholly responsive to the assigned task(s) and applied consistently through all elements of the work.	Demonstrates adequate consideration of context, audience, and purpose and a clear focus on the assigned task(s).	Demonstrates some attention to context, audience, purpose, and to the assigned task(s).	Demonstrates minimal attention to context, audience, purpose, and to the assigned task(s).	Fails to meet minimum criteria in addressing the audience, context, and purpose for writing.
Content Development	Uses appropriate, relevant, and compelling content and ideas that illustrate the writer's command and deep understanding of the subject, skillfully shaping the whole work.	Uses appropriate, relevant, and compelling content to accurately explore ideas within the subject and shape the whole work.	Uses appropriate and relevant content to develop and accurately explore ideas through most of the work.	Uses appropriate and relevant content to accurately develop simple ideas in some parts of the work.	Fails to meet minimum criteria in addressing content development.
Sources and Evidence	Demonstrates skillful use of high-quality, credible, relevant sources to develop ideas that are appropriate for the assignment.	Demonstrates consistent use of credible, relevant sources to support ideas that are appropriate for the assignment.	Demonstrates an attempt to use credible and relevant sources to support ideas that are appropriate for the assignment.	Demonstrates an attempt to use sources to support ideas in the assignment.	Fails to meet minimum criteria in demonstrating the use of sources to support ideas in the assignment.
Organization And Presentation	Demonstrates consistent, skillful, and thoroughly detailed attention to organization, presentation, and stylistic choices as appropriate to the assignment.	Demonstrates consistent and skillful organization and presentation as appropriate to the assignment.	Follows expectations for a consistent system of basic organization and presentation as appropriate to the assignment.	Attempts to use a consistent system for basic organization and presentation as appropriate to the assignment.	Fails to meet minimum criteria in organization and presentation.
Control of Syntax and Mechanics	Uses graceful language that skillfully communicates meaning to readers with clarity and fluency, and is nearly error-free.	Uses straightforward language that conveys meaning to readers with clarity. The language in the work has few errors.	Uses language that generally conveys meaning to readers, although writing may include some errors.	Uses language that sometimes impedes meaning because of errors in usage.	Fails to use language that demonstrates control of syntax and mechanics.

Appendix C: Stephen F. Austin State University Written and Visual Communications Rubric

WRITTEN AND VISUAL COMMUNICATION RUBRIC

Definition: Written communication is the development and expression of ideas in writing. Written communication involves learning to work in many genres and styles. It can involve working with many different writing technologies, and mixing texts, data, and images. Written communication abilities develop through iterative experiences across the curriculum.

	Capstone 4	Accomplished 3	Developing 2	Beginning 1	Unacceptable 0
Audience, Context, and Purpose	Demonstrates a thorough understanding of context, audience, and purpose that is wholly responsive to the assigned task(s) and applied consistently through all elements of the work.	Demonstrates adequate consideration of context, audience, and purpose and a clear focus on the assigned task(s).	Demonstrates some attention to context, audience, purpose, and to the assigned task(s).	Demonstrates minimal attention to context, audience, purpose, and to the assigned task(s).	Fails to meet minimum criteria in addressing the audience, context, and purpose for writing.
Content Development	Uses appropriate, relevant, and compelling content and ideas that illustrate the writer's command and deep understanding of the subject, skillfully shaping the whole work.	Uses appropriate, relevant, and compelling content to accurately explore ideas within the subject and shape the whole work.	Uses appropriate and relevant content to develop and accurately explore ideas through most of the work.	Uses appropriate and relevant content to accurately develop simple ideas in some parts of the work.	Fails to meet minimum criteria in addressing content development.
Sources and Evidence	Demonstrates skillful use of high-quality, credible, relevant sources to develop ideas that are appropriate for the assignment.	Demonstrates consistent use of credible, relevant sources to support ideas that are appropriate for the assignment.	Demonstrates an attempt to use credible and relevant sources to support ideas that are appropriate for the assignment.	Demonstrates an attempt to use sources to support ideas in the assignment.	Fails to meet minimum criteria in demonstrating the use of sources to support ideas in the assignment.
Organization And Presentation	Demonstrates consistent, skillful, and thoroughly detailed attention to organization, presentation, and stylistic choices as appropriate to the assignment.	Demonstrates consistent and skillful organization and presentation as appropriate to the assignment.	Follows expectations for a consistent system of basic organization and presentation as appropriate to the assignment.	Attempts to use a consistent system for basic organization and presentation as appropriate to the assignment.	Fails to meet minimum criteria in organization and presentation.
Control of Syntax and Mechanics	Uses graceful language that skillfully communicates meaning to readers with clarity and fluency, and is nearly error-free.	Uses straightforward language that conveys meaning to readers with clarity. The language in the work has few errors.	Uses language that generally conveys meaning to readers, although writing may include some errors.	Uses language that sometimes impedes meaning because of errors in usage.	Fails to use language that demonstrates control of syntax and mechanics.

Visual aids	Visual aids effectively support the communication of purposes and ideas; aids are integrated into the presentation seamlessly, thus fostering a full understanding of the message's content.	Visual aids generally support the communication of the student's ideas and purposes; the aids effectively amplify or resonate the presentation of ideas and foster a good understanding of the message's content.	Visual aids support the communication of the student's ideas and purposes but are only partially useful or informative.	Visual aids do not particularly support the communication of the student's ideas and purpose; they are insufficient to be of much use as they do little to elevate understanding.	Visual aids are virtually non-existent, serve no purpose, or are not credible

Appendix D: Rules for Scoring Student Work

Procedures for assessment of student work:

1. Two raters will initially assess each piece of student work.

2. If the two raters agree on their rating on any element/criterion of a rubric then there is no need for a third rater on that element/criterion.

3. If the first two raters are no more than one integer apart on their ratings on an element/criterion of a rubric, then there is no need for a third rater on that element/criterion.

   For example, if Rater A gives a piece of student work a 2 on element/criterion of Audience, Context, and Purpose, and Rater B gives the piece of student work a 3 on Audience, Context, and Purpose, then the two ratings are averaged together to give a 2.5 on the Audience, Context, and Purpose element/criterion. If the two raters are more than one integer apart on their ratings on any element/criterion of a rubric, a third rater is asked to rate only the element(s)/criteria where there was disagreement.

For example, if Rater A gives a piece of student work a 1 on the element/criterion Audience, Context, and Purpose, and Rater B gives the piece of student work a 3 on Audience, Context, and Purpose. In addition, rater A also gives the same piece of student work a 4 on Sources and Evidence, and Rater B gives that same piece of student work a 2. Then a third rater (Rater C) is asked to rate the student work only on the elements/criteria of Audience, Context, and Purpose and Sources and Evidence.

4. If Rater C's rating agrees with one of the other two ratings, then that rating is used and the rating that does not agree is discarded.

   For example, if Rater C and Rater A each rate a piece of student work a 2 on Content Development, but Rater B rates the work a 4, then Rater B's rating is discarded and the student work received a rating of 2 on Content Development.

5. If Rater C's rating does not agree with one of the other two ratings, and is no more than one integer from only one of the other ratings, then the rating that is more than one integer from the other ratings is discarded, and the two ratings that are no more than one integer apart are averaged.

   For example, if Rater C rates a piece of student work 2, Rater A rated the work a 1, and Rater B rated the work 4 on Content Development. Rater B's rating of 4 is discarded and the ratings of Rater C and Rater A are averaged to get a rating of 1.5.

6. If Rater C's rating is no more than one integer from the other two ratings, then all of the ratings are averaged.

For example, if Rater C rates a piece of student work 3, Rater A rated the work a 2, and Rater B rated the work 4 on Content Development. All of the ratings are averaged for a rating of 3.

7. If Rater C's rating does not agree with one of the other two ratings and is more than one integer apart from the other two ratings, then Rater C's rating is discarded, and the other two ratings are averaged.

   For example, if Rater C rates a piece of student work 4, Rater A rated the work a 0, and Rater B rated the work a 2 on Content Development. Rater C's rating of 4 is discarded, and the other two ratings are averaged to get a rating of 1.