General Education Assessment Report – AY 2018-19

Author: Joseph F. van Gaalen, Ph.D., Asst. VP, IR, Assessment & Effectiveness

1 INTRODUCTION

Before the beginning of AY 2014-2015, the General Education Assessment Subcommittee of the Learning Assessment Committee (LAC) adopted (see June 9, 2014 GEAS Subcommittee Meeting Minutes) the Association of American Colleges & Universities (AAC&U) Value Rubric Model (Rhodes and Finley, 2013) after an extensive review of General Education assessment models employed throughout higher education. During AY 2014-2015, the subsequent assessment during that academic year, each of the five competencies (Communication, Critical Thinking, Technology/Information Management, Global Socio-cultural Responsibility, and Scientific and Quantitative Reasoning) was assessed through assignments identified by faculty as fitting the criteria of the competency (Braselton, 2011; Rhodes and Finley, 2013) by way of a pilot study. As aligned with the AAC&U Value Rubric Model and Value Rubric Case Studies, Florida SouthWestern State College (FSW) faculty from across disciplines voluntarily submitted assignments aligned with the competencies. Assignments do not have to be uniform if outcomes, rating, and the rationale for rating (rubric interpretation) are uniform (Rhodes & Finley, 2013). Outcomes are identified by the competency definition at FSW. Calibration sessions were conducted before scoring in each competency. Inter-rater reliability studies were performed on the results (see AY 2014-2015 General Education Assessment Report). Following the completion of the pilot study, recommendations by the LAC focused on professional development opportunities in the strengthening of assignment guidelines.

Assessment continued in AY 2015-2016 using the same method to begin employing the use of the AAC&U rubrics for a comprehensive review of the Communication (COM) competency, both oral and written. Discussions pertaining to the results of the analysis led to (1) a development of the goal to strengthen dual enrollment (concurrent) participation in general education assessment and (2) professional development opportunities in supporting students' writing (see AY 2015-2016 General Education Assessment Report).

The third year in the evolution, AY 2016-2017, again using the same method, saw the use of AAC&U rubrics for another comprehensive review this time of the Critical Thinking (CT) and the Scientific and Quantitative Reasoning (QR) competencies (see AY 2016-2017 General Education Assessment Report). Note that the AAC&U Value Rubric was used for the CT competency, but an FSW developed rubric was used for QR. Discussions pertaining to the results of the analysis led to the development of FSW specific rubrics in preparation for the shift from the old competencies (Communication, Critical Thinking, Technology/Information Management, Global Socio-cultural Responsibility, and Scientific and Quantitative Reasoning) to the new competencies (Communicate, Research, Evaluate, Analyze, Think, Investigate, Visualize, and Engage {C-R-E-A-T-I-V-E}) which occurred in the Fall 2016 term.

General Education assessment continues in AY 2017-2018. The change for this year is in response to the newly adopted competencies which are based on faculty-led identification. Instead of asking faculty to volunteer assignments, once the LAC votes on which competencies to study in a given assessment, courses are randomly sampled from a list of courses which were identified by faculty as encompassing

that competency. Academic Year 2017-2018 included analyses of of 'Research' and 'Investigate' from the new C-R-E-A-T-I-V-E General Education competencies.

Academic Year 2018-2019 began the second year in which competencies were assessed based on faculty-led identification. This report details the results of FSW's General Education assessment for AY 2018-2019 which included the analysis of 'Visualize' and 'Engage' from the C-R-E-A-T-I-V-E General Education competencies.

The intent of FSW's General Education Program is to foster lifelong learning and establish academic excellence, interdisciplinary dialog, and a social responsibility among students. In that light, the purpose of the program is to: 1) measure against baseline data for the number of students receiving scores of 3 or higher on relevant dimensions of the rubric, 2) measure against baseline data for the number of students receiving scores of 3 or higher on relevant dimensions of the rubric across sites (Online, Dual Enrollment, and Traditional), 3) establish a baseline for the number of student artifacts receiving a score of 3 or higher on relevant aspects of the rubric across credit achievement level (e.g. achievement with respect to number of credits earned), 4) establish a baseline for the number of student artifacts receiving a score of 3 or higher on relevant aspects of the rubric across pre-requisite definition (e.g. achievement with respect to pre-requisite courses), and 5) develop FSW-based rubrics for assignments/assessments administered at FSW.

For additional detail on further analysis not provided in this report, please contact Dr. Joseph F. van Gaalen, Asst. VP, IR, Assessment & Effectiveness, Academic Affairs (jfvangaalen@fsw.edu; x16965).

2   VISUALIZE (V)

The outcome of the 'Visualize' competency at FSW is that by completion of the general education requirements, students will be able to visualize and engage the world from different historical, social, religious, and cultural approaches. The LAC will measure the number of artifacts scored a 3 or higher on relevant dimensions of the rubric, a mark based on the pilot results of AY 2014-2015. Figures 1 through 12 below depict achievement and inter-rater reliability for the 'Visualize' competency in college-wide, Associate of Arts (AA) cohorts, as well as value-added studies as they relate to outcome goals and objectives.

For the study, the LAC selected a rubric developed as part of the AAC&U Value Rubric Model as a means of scoring achievement in this competency. Feedback from scorers regarding the AAC&U Value Rubric for the 'Visualize' competency included three main trends regarding rubric suitability. First, multiple scorers noted that the rubric addresses areas the assignment does not call for (entire dimension can't be scored). All dimensions of the rubric were at one point or another called out in varying assignments as not being able to be assessed. Second, scorers noted that there is overlap in several areas that causes problems with scoring depending on the assignment. Elements of culture, diversity, and responsibility appear in varying dimension making interpretation troublesome.   And third, scorers commented positively regarding the rubric that dimensions cover a lot of areas. So, while the rubric may overlap, it does appear to be versatile in the style of assignments. The LAC has already tended to the task of writing a new FSW 'Visualize' rubric based on the findings included in this report.
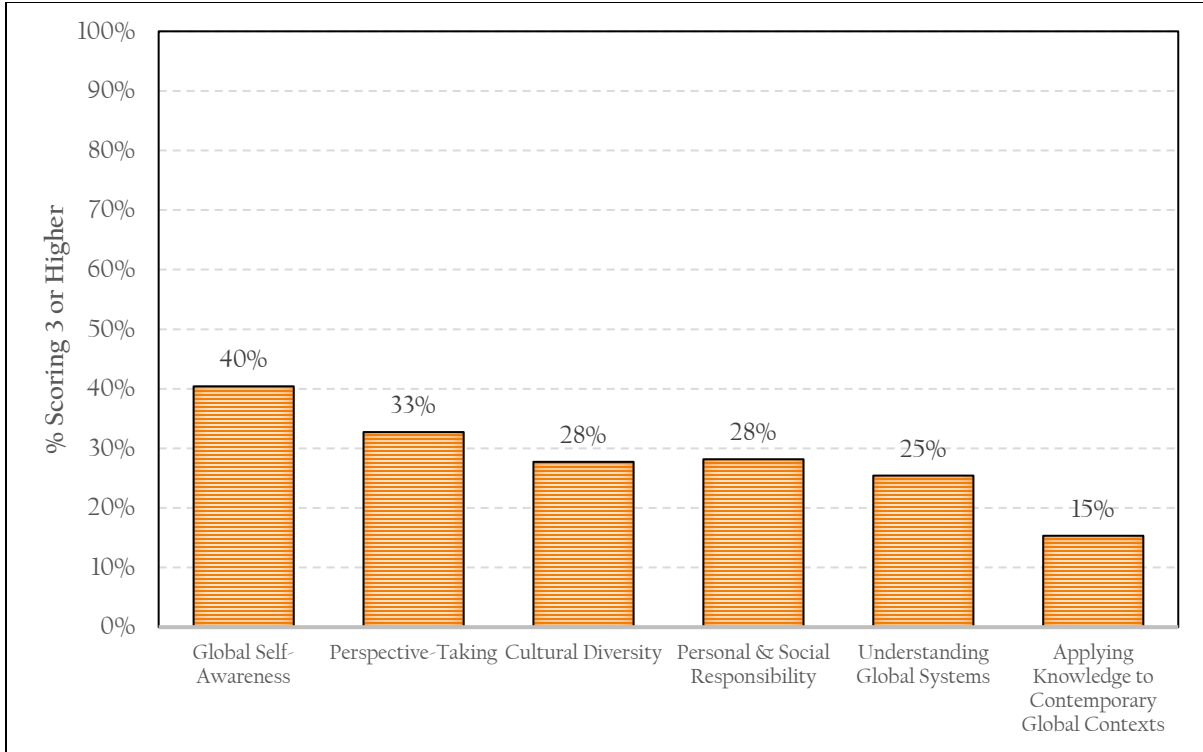
Figure 1. 'Visualize' achievement at 3 or higher across all rubric dimensions for 190 artifacts from 25 sampled course sections.
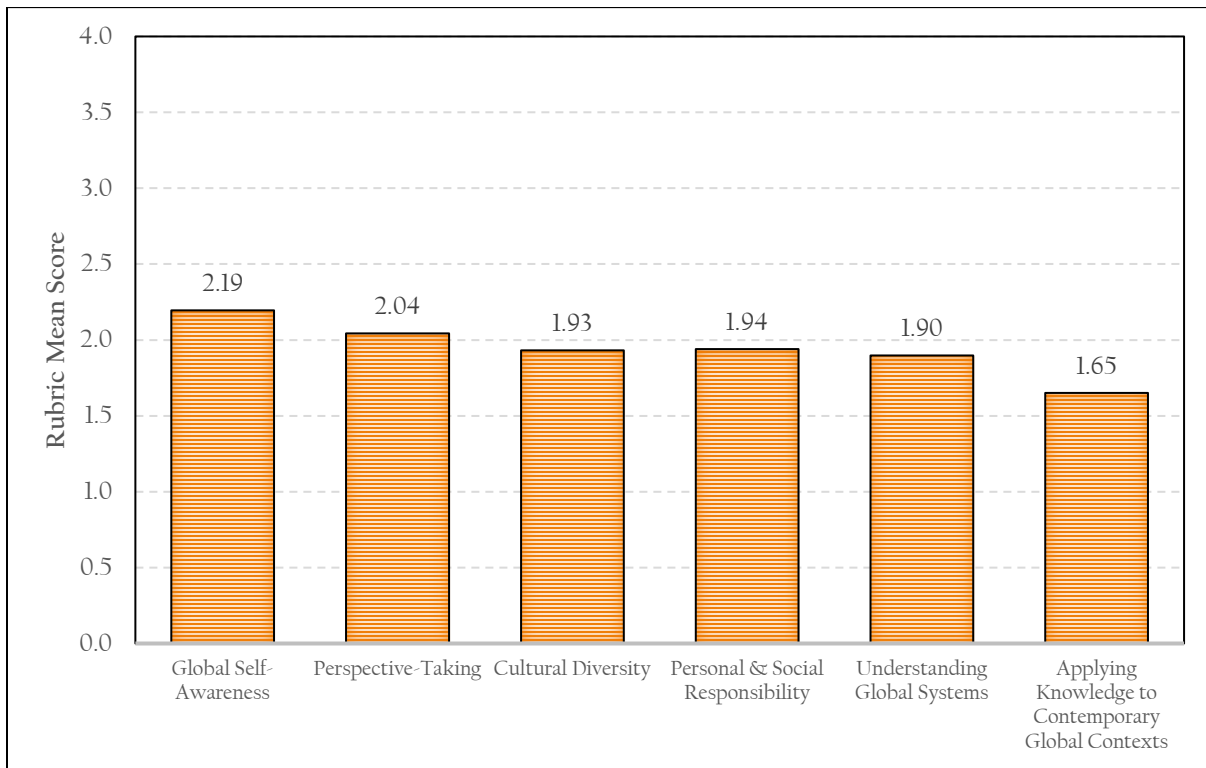

Figure 2. Mean score by rubric dimension for 'Visualize' for 190 artifacts from 25 sampled course sections.
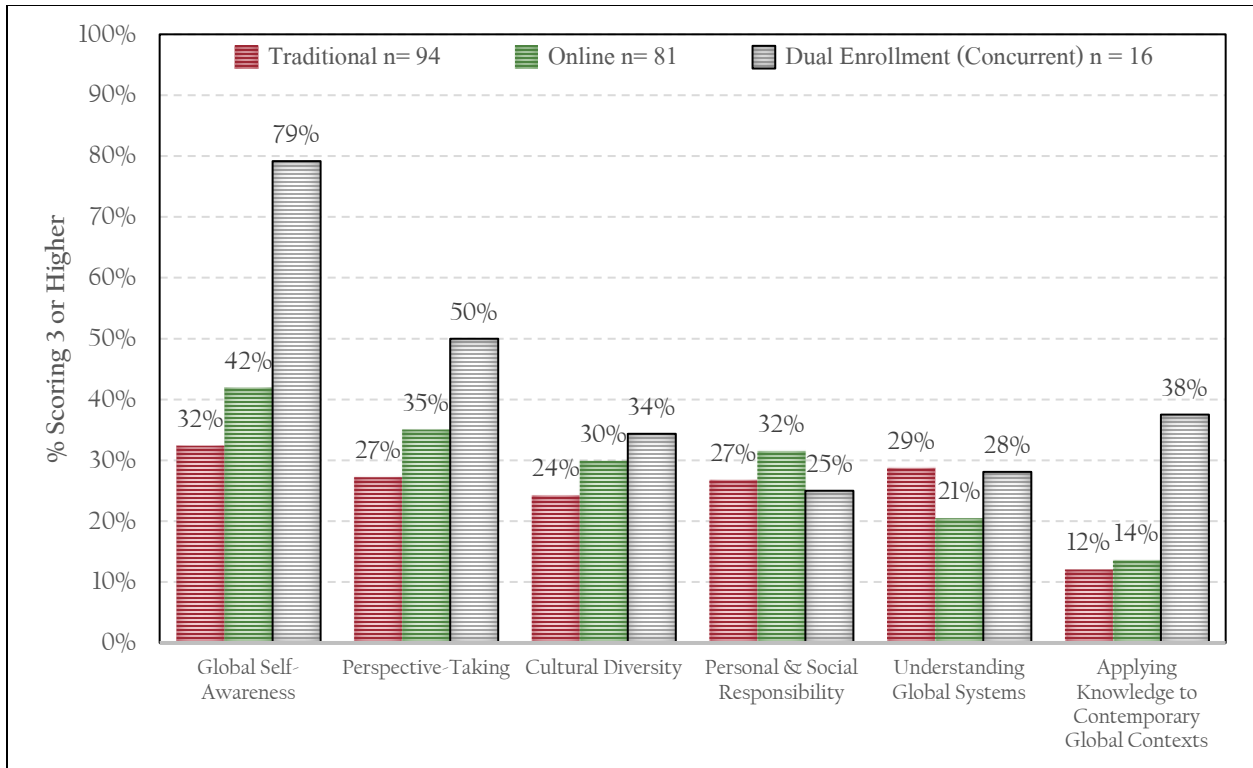
Figure 3. Comparison of 'Visualize' achievement by modality at 3 or higher across all rubric dimensions for 191 artifacts from 25 sampled course sections. Traditional (red), n=94, Online (green), n=81, Dual Enrollment (concurrent) (gray), n=16.
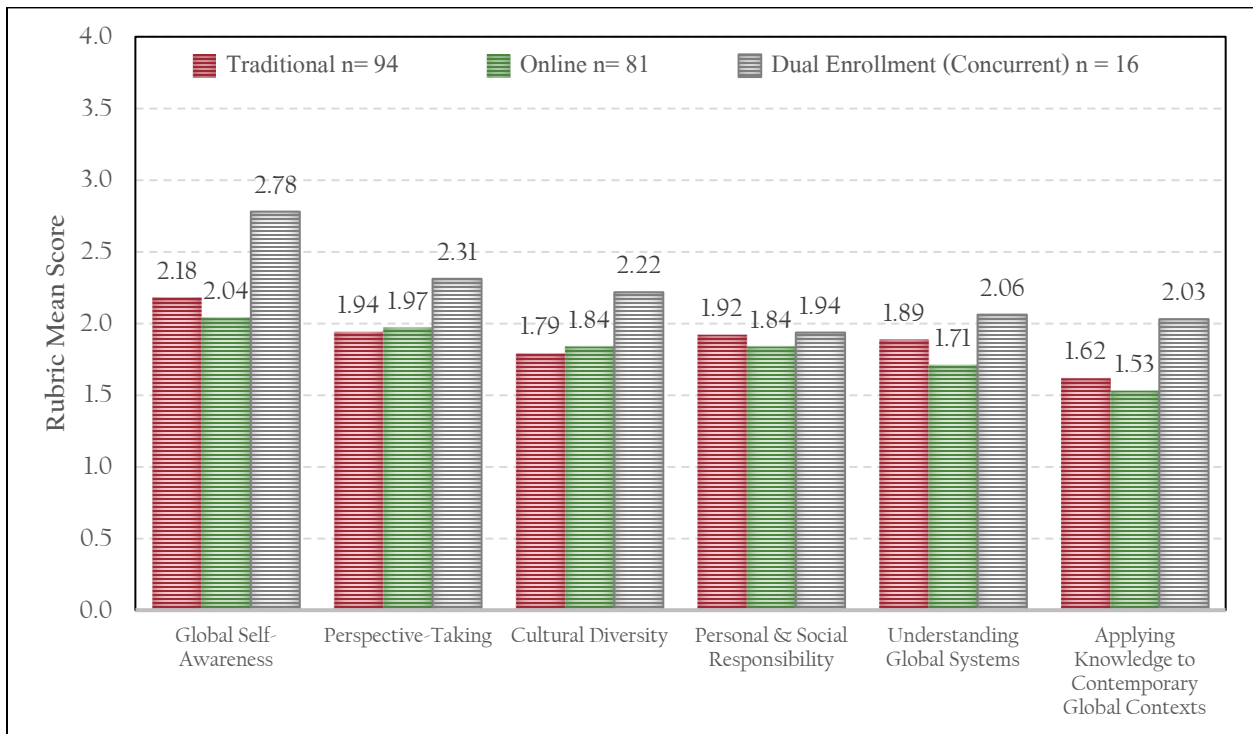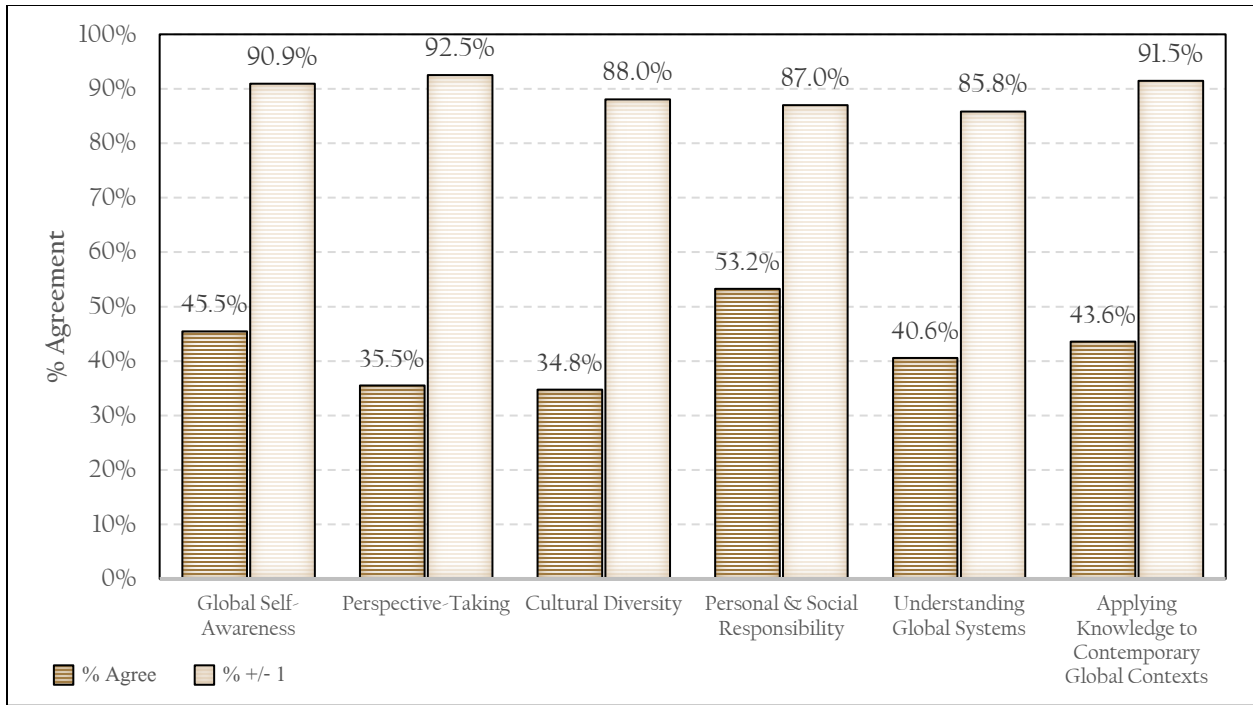

Figure 4. Mean score of 'Visualize' for each rubric dimension by modality at 3 or higher across all rubric dimensions for 191 artifacts from 25 sampled course sections. Traditional (red), n=94, Online (green), n=81, Dual Enrollment (concurrent) (gray), n=16.

Figure 5. Inter-rater reliability (as %) for the 'Visualize' competency. Each artifact was scored by two scorers. Percentage (%) of agreement (dark beige) is defined as cases where scores by each scorer were identical. Percentage (%) +/- 1 agreement (light beige) is defined as cases where scores by each scorer were within 1 of each other. κ-statistic for the study exhibits similar results. Results are herein presented as percentages for reader convenience.
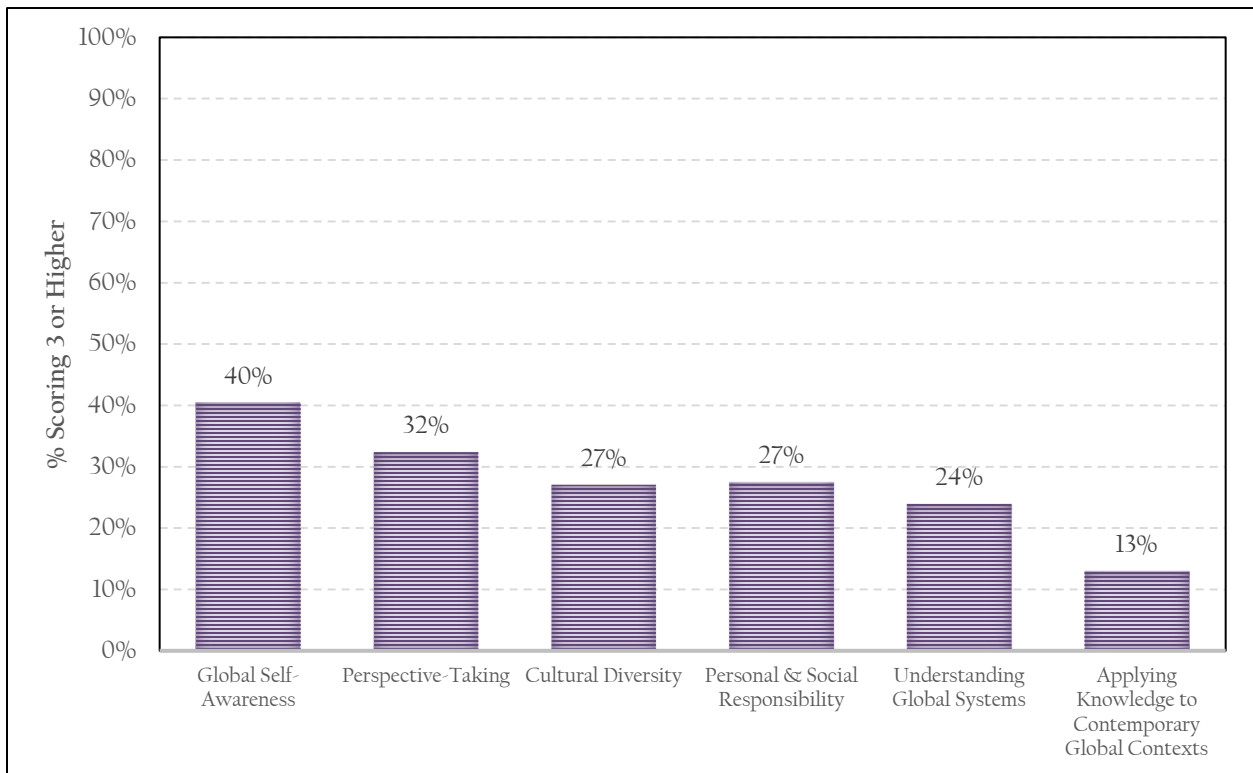

Figure 6. 'Visualize' achievement at 3 or higher across all rubric dimensions for AA courses only for 175 artifacts from 24 sampled course sections.
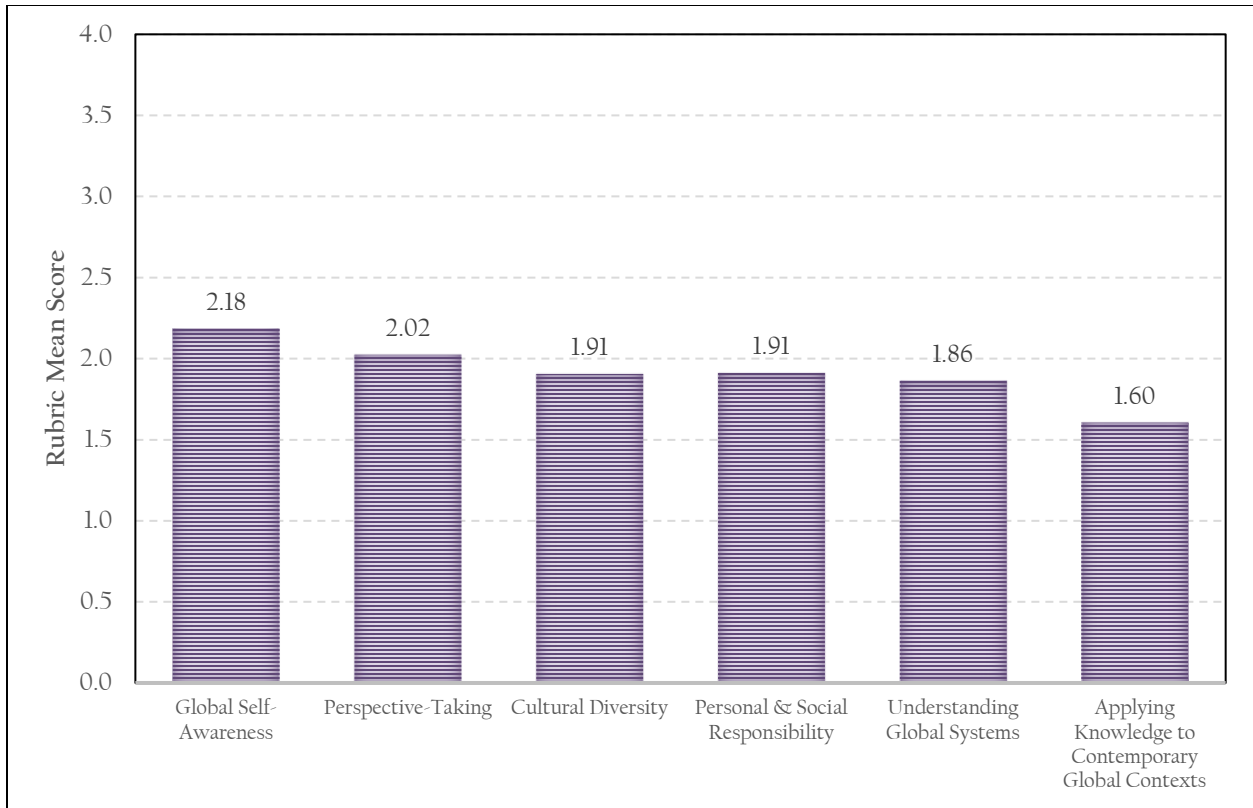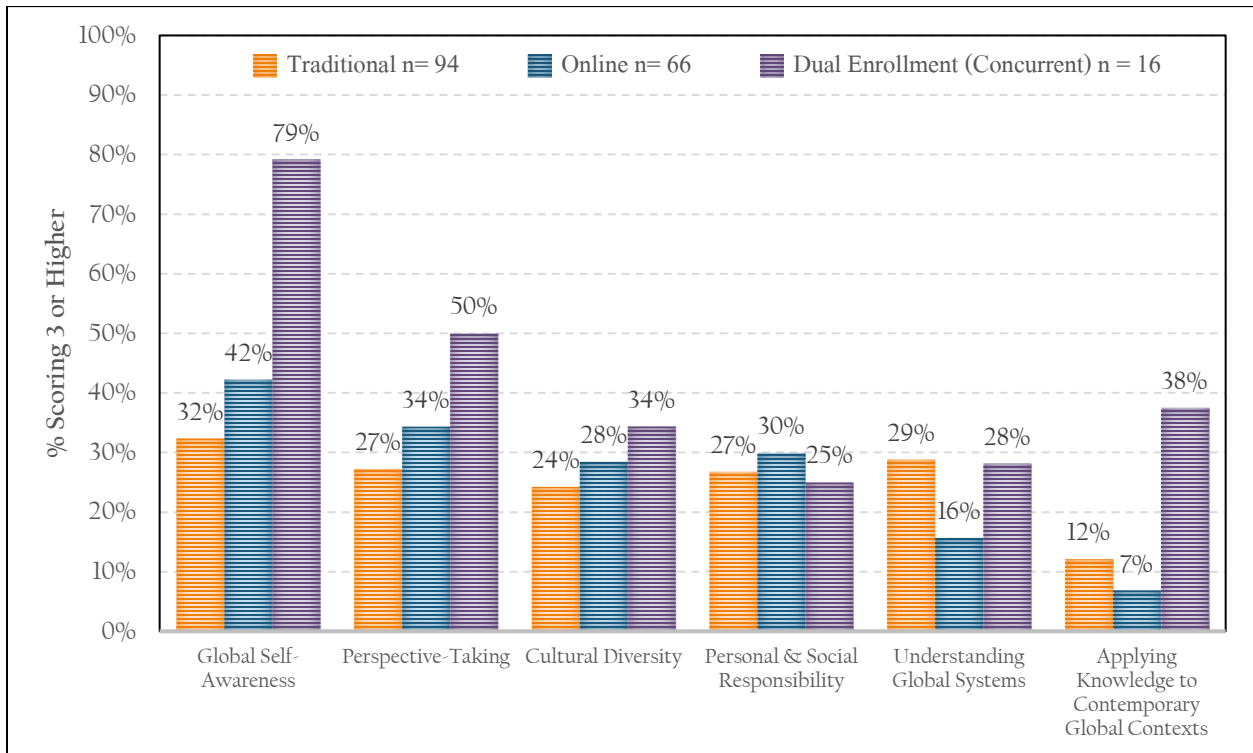
Figure 7. Mean score by rubric dimension for 'Visualize' for AA courses only for 175 artifacts from 24 sampled course sections.

Figure 8. 'Visualize' achievement at 3 or higher across all rubric dimensions for AA courses only for 175 artifacts from 24 sampled course sections. Traditional (orange), n=94, Online (blue), n=66, Dual Enrollment (concurrent) (purple), n=16.
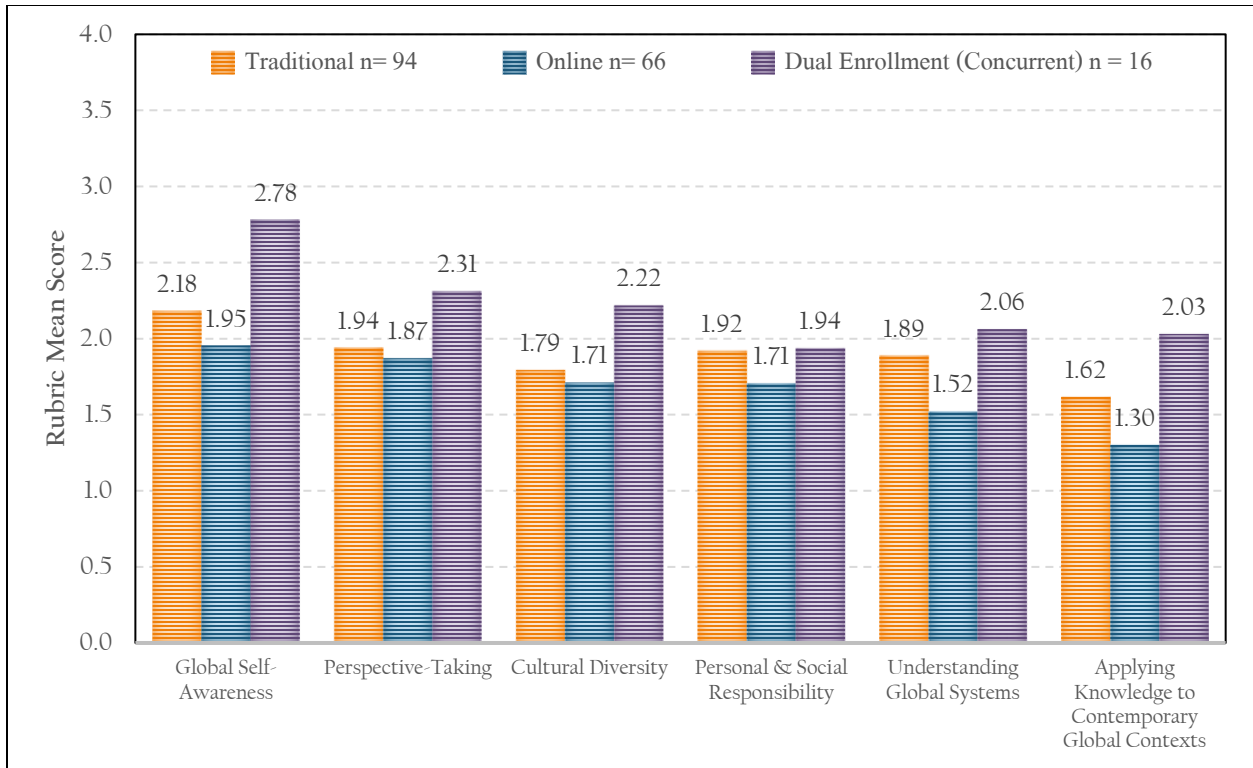
Figure 9. Mean score of 'Visualize' for each rubric dimension by modality at 3 or higher across all rubric dimensions for AA courses only for 175 artifacts from 24 sampled course sections. Traditional (orange), n=94, Online (blue), n=66, Dual Enrollment (concurrent) (purple), n=16.
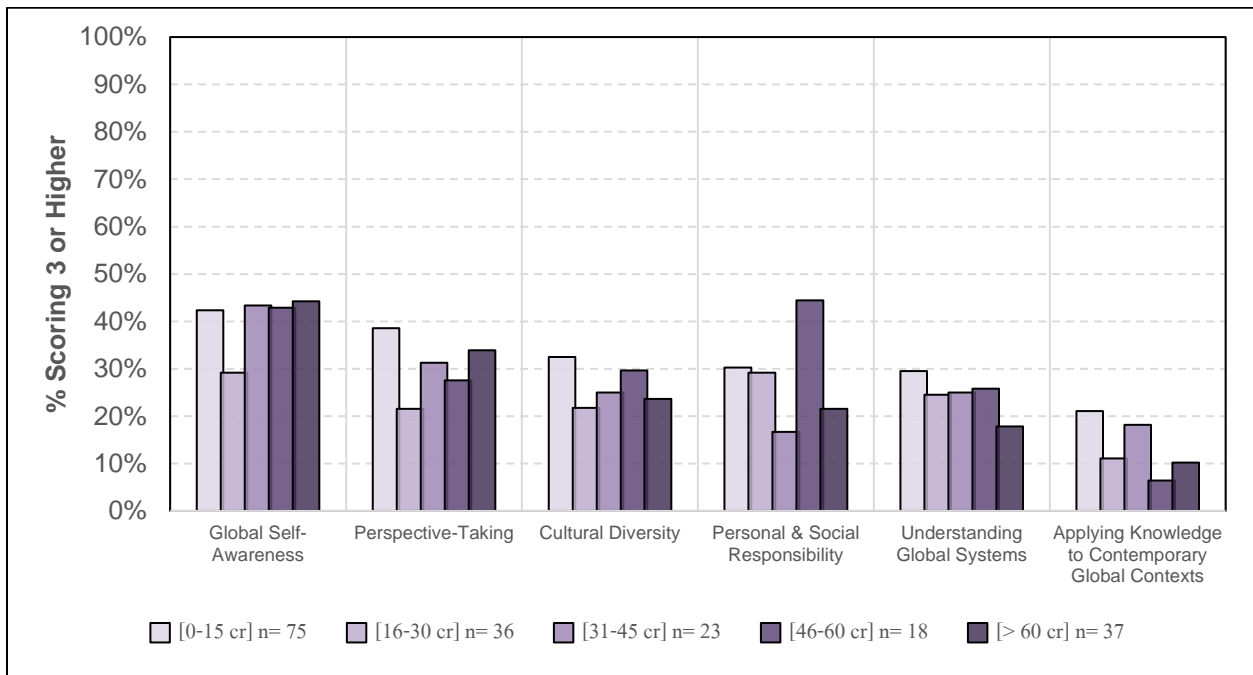

*Figure 10. Comparison of 'Visualize' achievement at 3 or higher across all rubric dimensions for 189 artifacts in which credit information could be matched to artifact score. From light purple to dark, 0-15 credits earned n=75, 16-30 credits earned n=36, 31-45 credits earned n=23, 46-60 credits earned n=57, and > 60 credits earned n=37. *Credits earned based on number of credits earned entering fall 2018 term.*
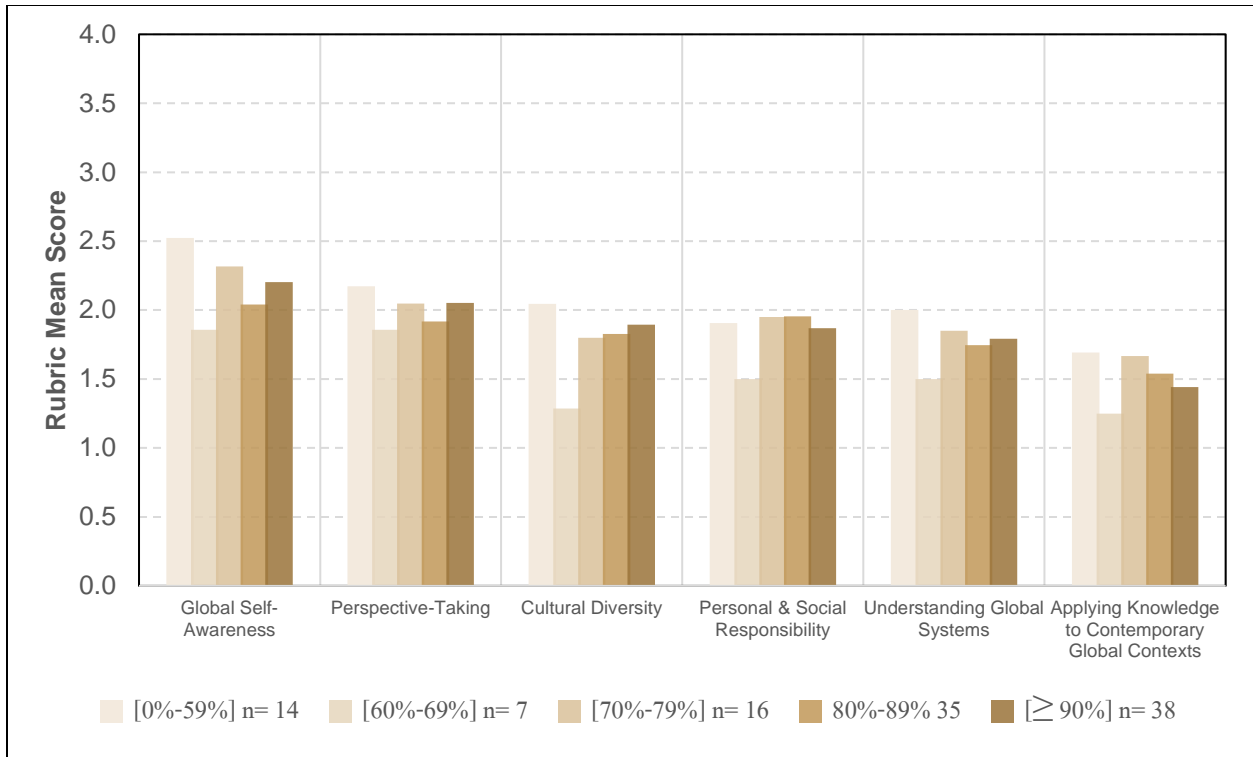
Figure 11. Comparison of 'Visualize' achievement at 3 or higher across all rubric dimensions based on course success rates of students. From light beige to dark, students with 0-59% n=14, 60-69% n=7, 70-79% n=16, 80-89% n=35, and 90% or above n=38.
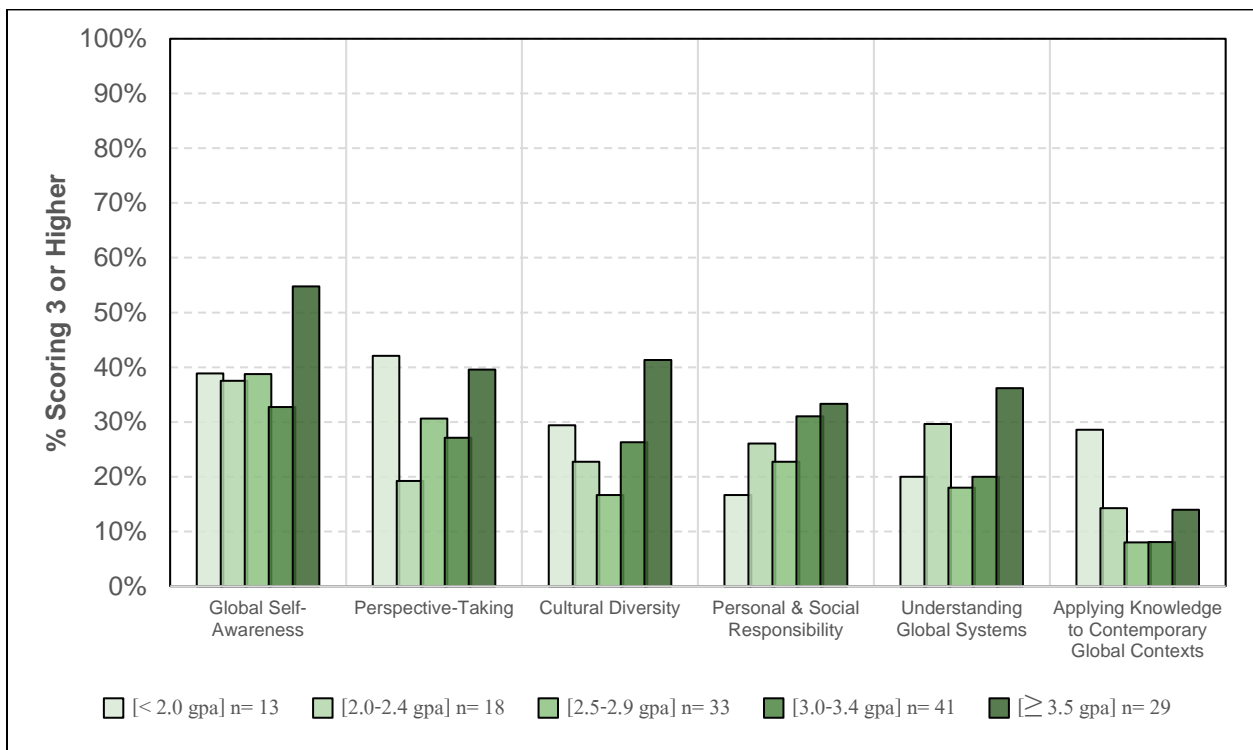

*Figure 12. Comparison of 'Visualize' achievement at 3 or higher across all rubric dimensions based on GPA. From light green to dark, GPA < 2.0 n=13, GPA 2.0-2.4 n=18, GPA 2.5-2.9 n=33, GPA 3.0-3.4 n=41, GPA ≥ 3.5 n=29. *Credits earned based on number of credits earned entering fall 2018 term.*

3 ENGAGE (E)

The outcome of the 'Engage' competency at FSW is that by completion of the general education requirements, students will be able to engage meanings of active citizenship in one's community, nation, and the world. The FSW Learning Assessment Committee will measure the number of artifacts scored a 3 or higher on relevant dimensions of the rubric against the pilot results (AY 2014-2015). Figures 13 through 24 below depict achievement and inter-rater reliability for the 'Engage' competency in college-wide, Associate of Arts (AA) cohorts, as well as value-added studies.

For the study, the LAC selected a rubric developed as part of the AAC&U Value Rubric Model as a means of scoring achievement in this competency. Feedback from scorers regarding the AAC&U Value rubric for the 'Engage' competency included three main trends regarding rubric suitability. First, scorers noted some overlap between the Action/Reflection and Context/Structures dimension. Second, scorers noted that the scoring levels '1' and '2' in the 'Analysis of Knowledge' dimension are too similar. And third, very few assignments appear to capture the rubric entirely. This is not necessarily a problem as the assignment should be tied to the competency and not the rubric. The LAC has already tended to the task of writing a new FSW 'Engage' rubric based on the findings included in this report.
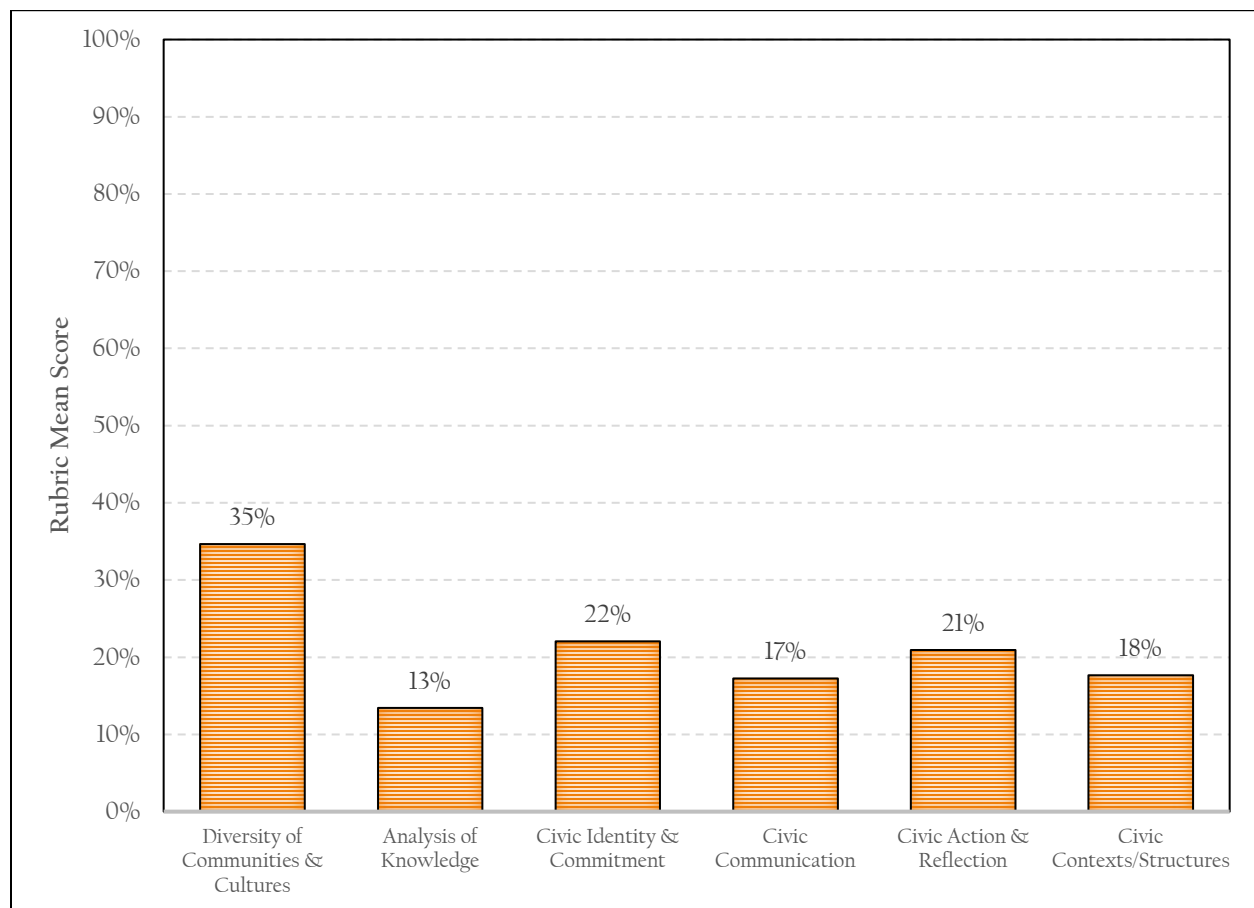

Figure 13. 'Engage' achievement at 3 or higher across all rubric dimensions for 177 artifacts from 28 sampled course sections.
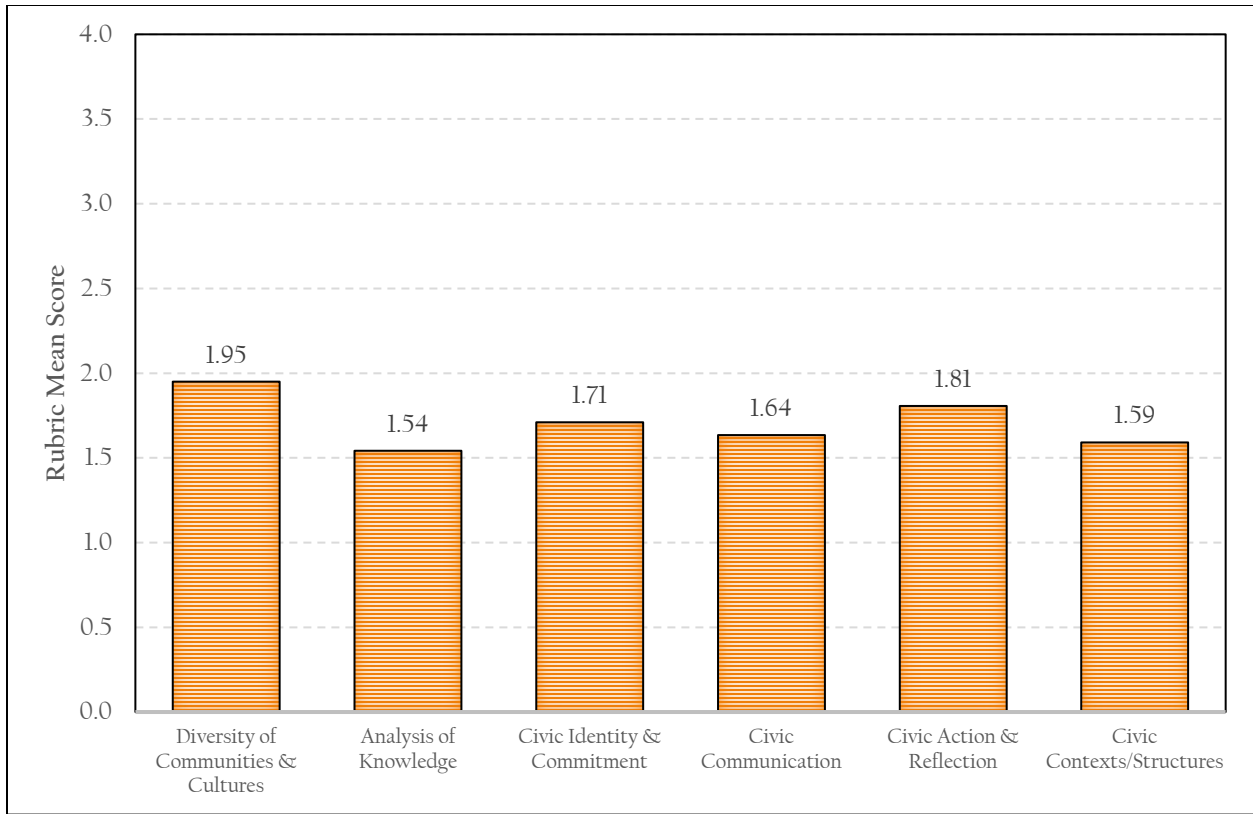
Figure 14. Mean score by rubric dimension for 'Engage' for 177 artifacts from 28 sampled course sections.
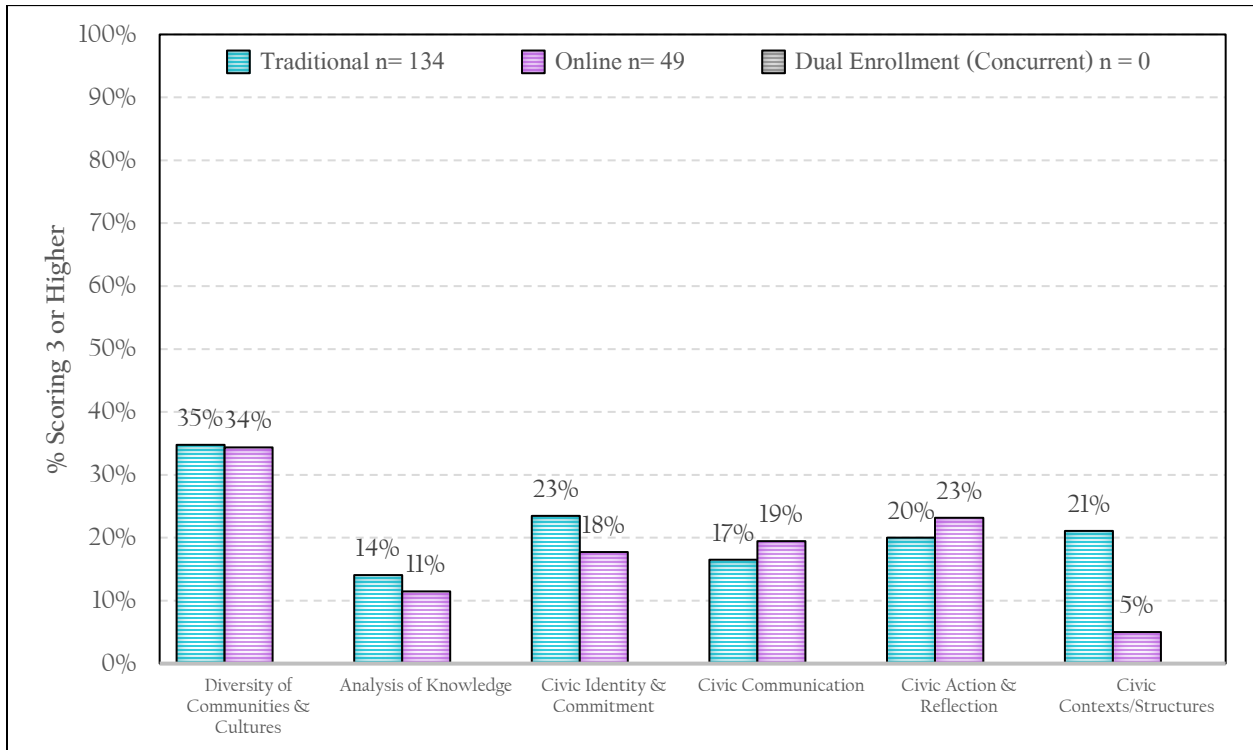

Figure 15. Comparison of 'Engage' achievement by modality at 3 or higher across all rubric dimensions for 177 artifacts from 28 sampled course sections. Traditional (aqua), n=134, Online (purple), n=49, Dual Enrollment (concurrent) (gray), n=0.
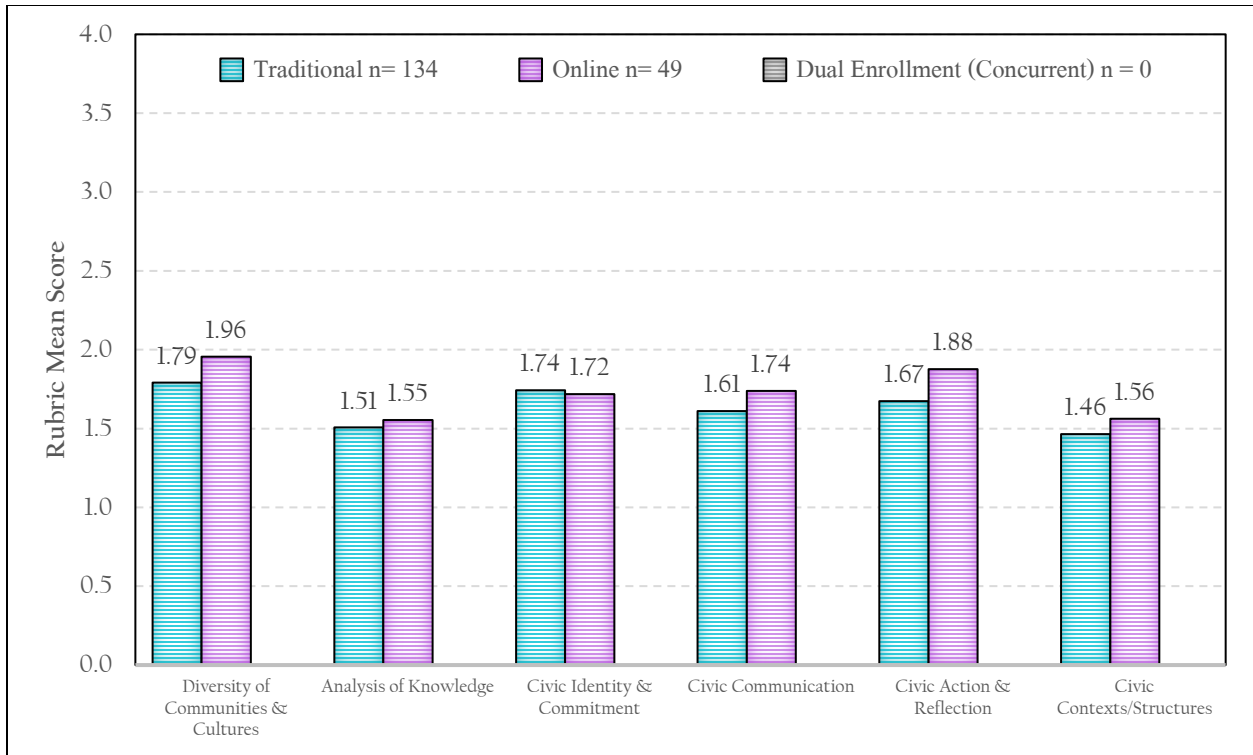
Figure 16. Mean score of 'Engage' for each rubric dimension by modality at 3 or higher across all rubric dimensions for 177 artifacts from 28 sampled course sections. Traditional (aqua), n=134, Online (purple), n=49, Dual Enrollment (concurrent) (gray), n=0.

Figure 17. Inter-rater reliability (as %) for the 'Engage' competency. Each artifact was scored by two scorers. Percentage (%) of agreement (dark beige) is defined as cases where scores by each scorer were identical. Percentage (%) +/- 1 agreement (light beige) is defined as cases where scores by each scorer were within 1 of each other. κ-statistic for the study exhibits similar results. Results are herein presented as percentages for reader convenience.

Figure 18. 'Engage' achievement at 3 or higher across all rubric dimensions for AA courses only for 148 artifacts from 25 sampled course sections.
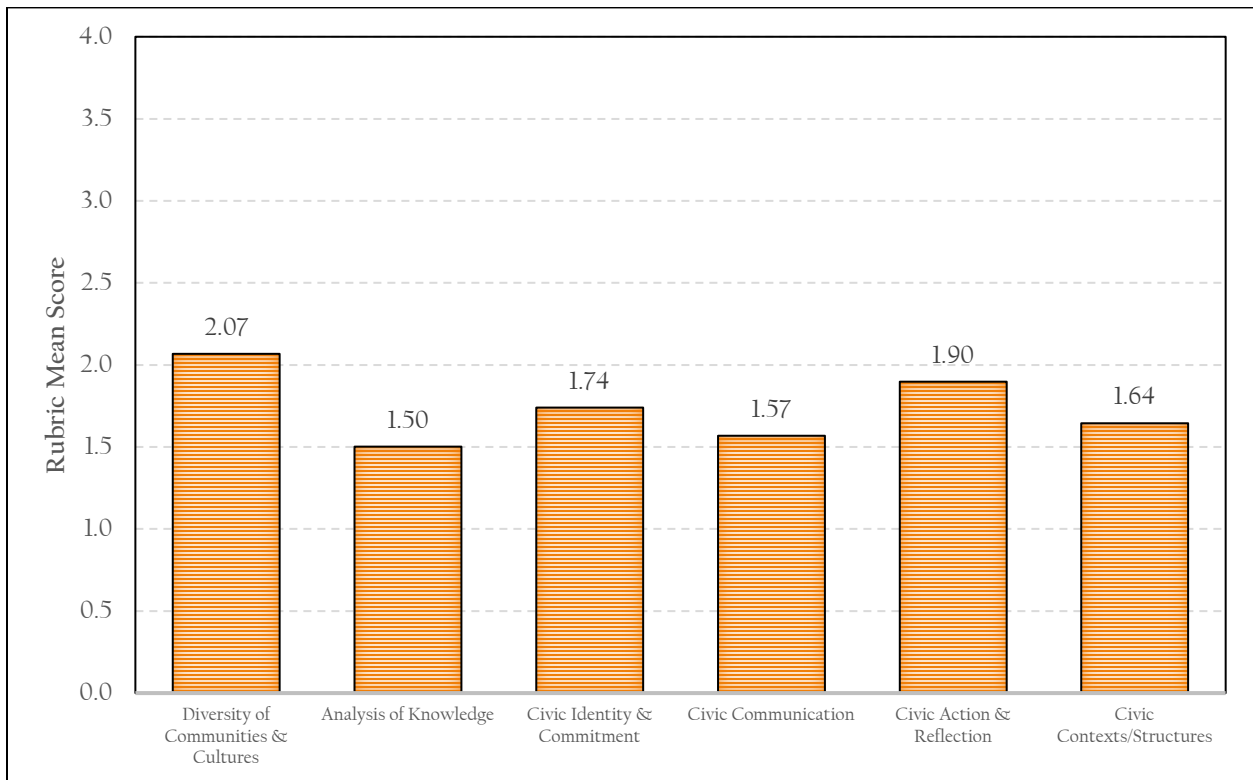
Figure 19. Mean score by rubric dimension for 'Engage' for AA courses only for 148 artifacts from 25 sampled course sections.

Figure 20. 'Engage' achievement at 3 or higher across all rubric dimensions for AA courses only for 100 artifacts from 25 sampled course sections. Traditional (aqua), n=100, Online (purple), n=49, Dual Enrollment (concurrent) (gray), n=0.
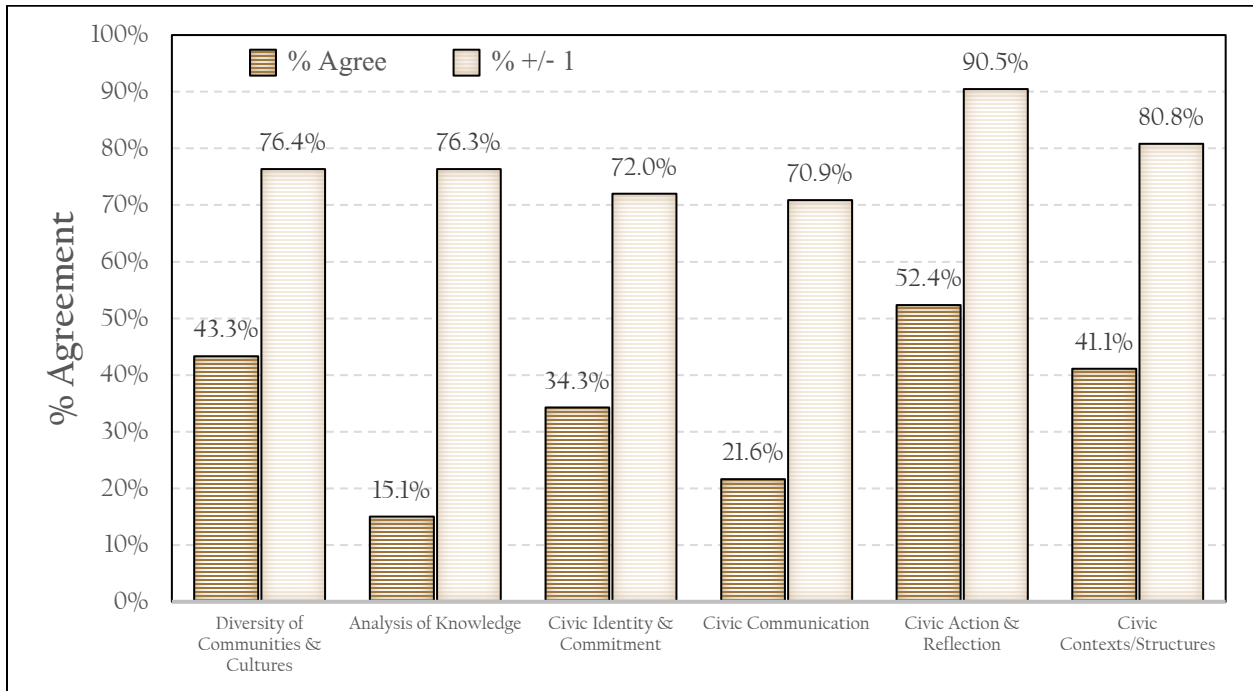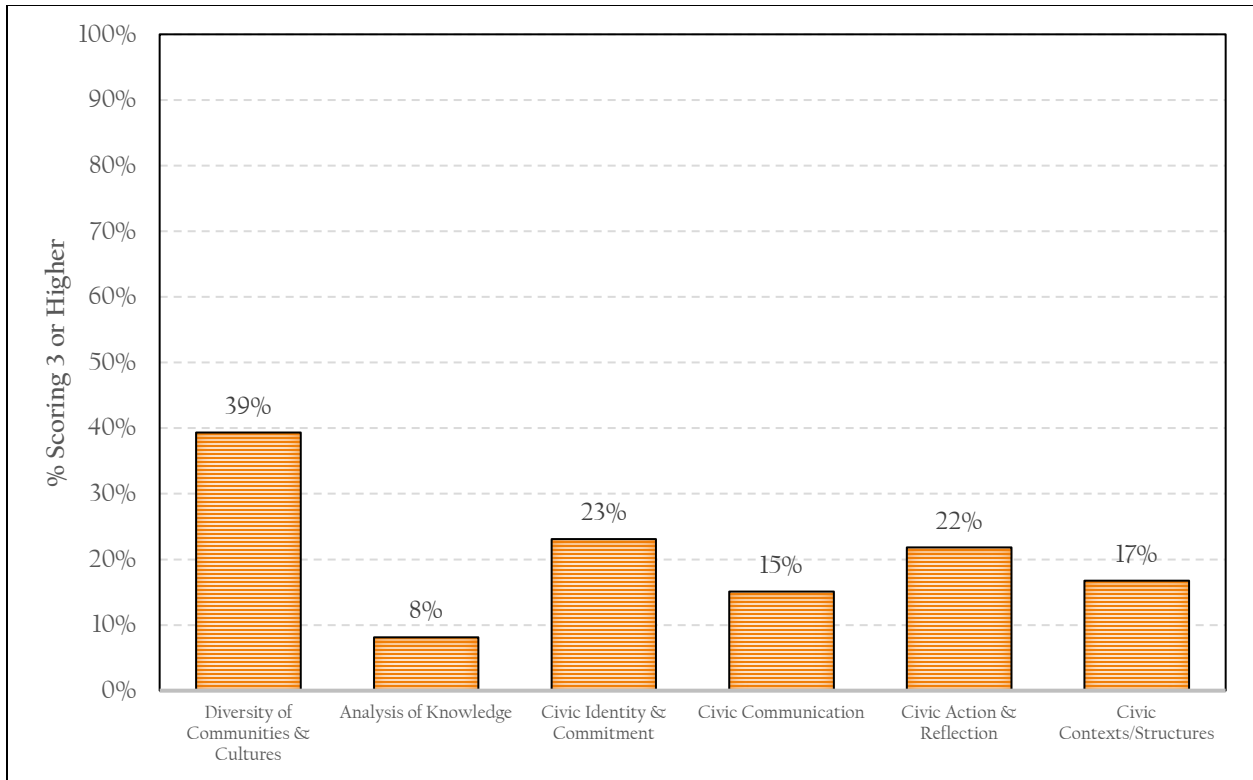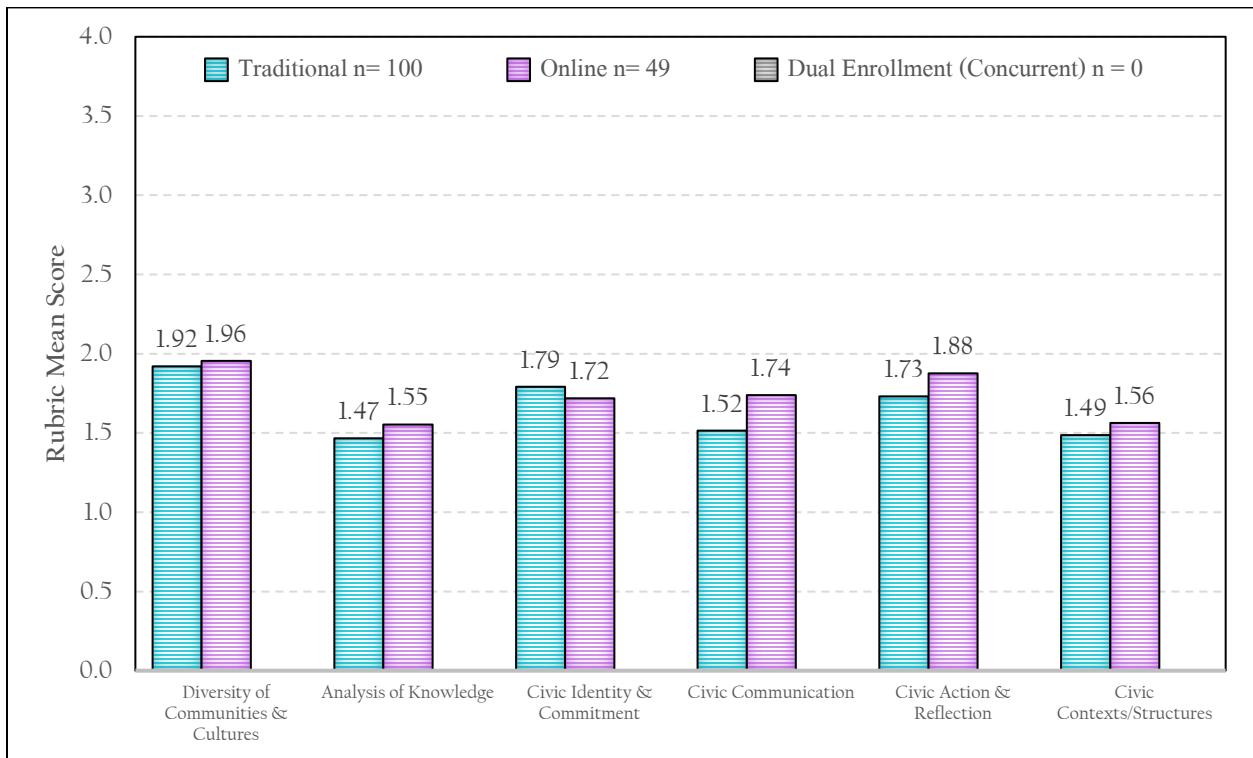

Figure 21. Mean score of 'Engage' for each rubric dimension by modality at 3 or higher across all rubric dimensions for AA courses only for 100 artifacts from 25 sampled course sections. Traditional (aqua), n=100, Online (purple), n=49, Dual Enrollment (concurrent) (gray), n=0.
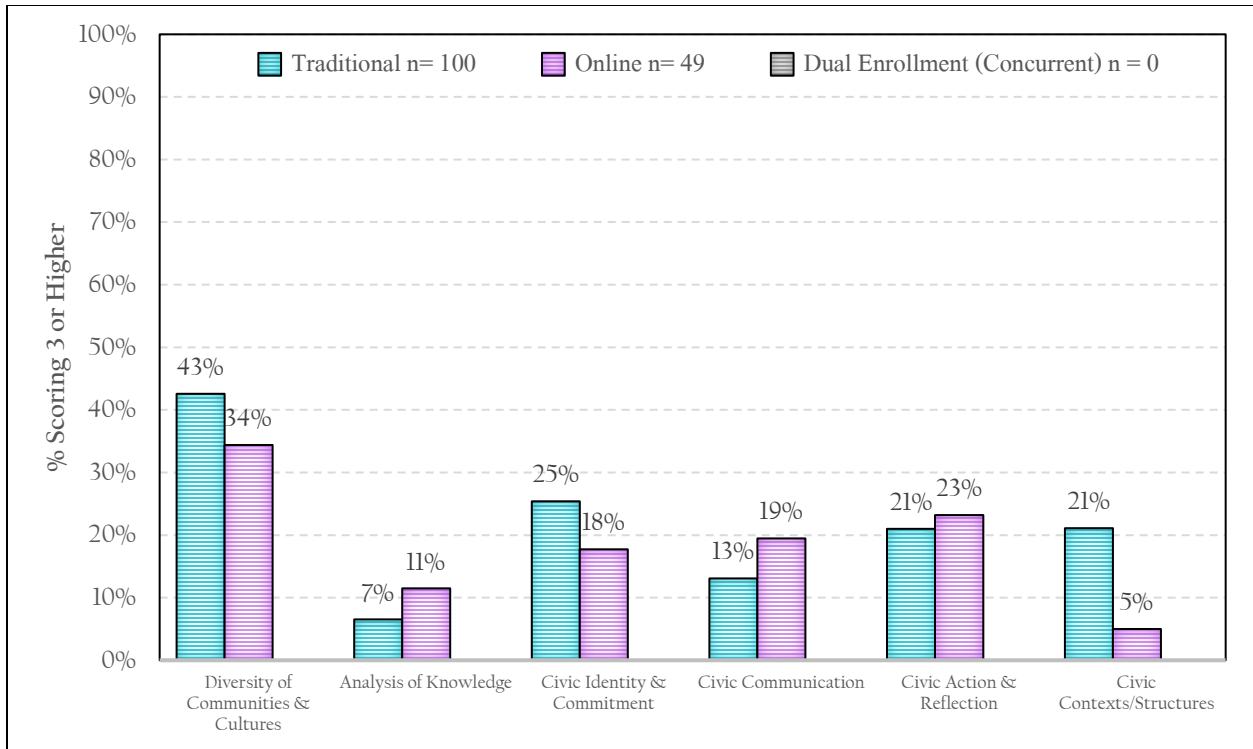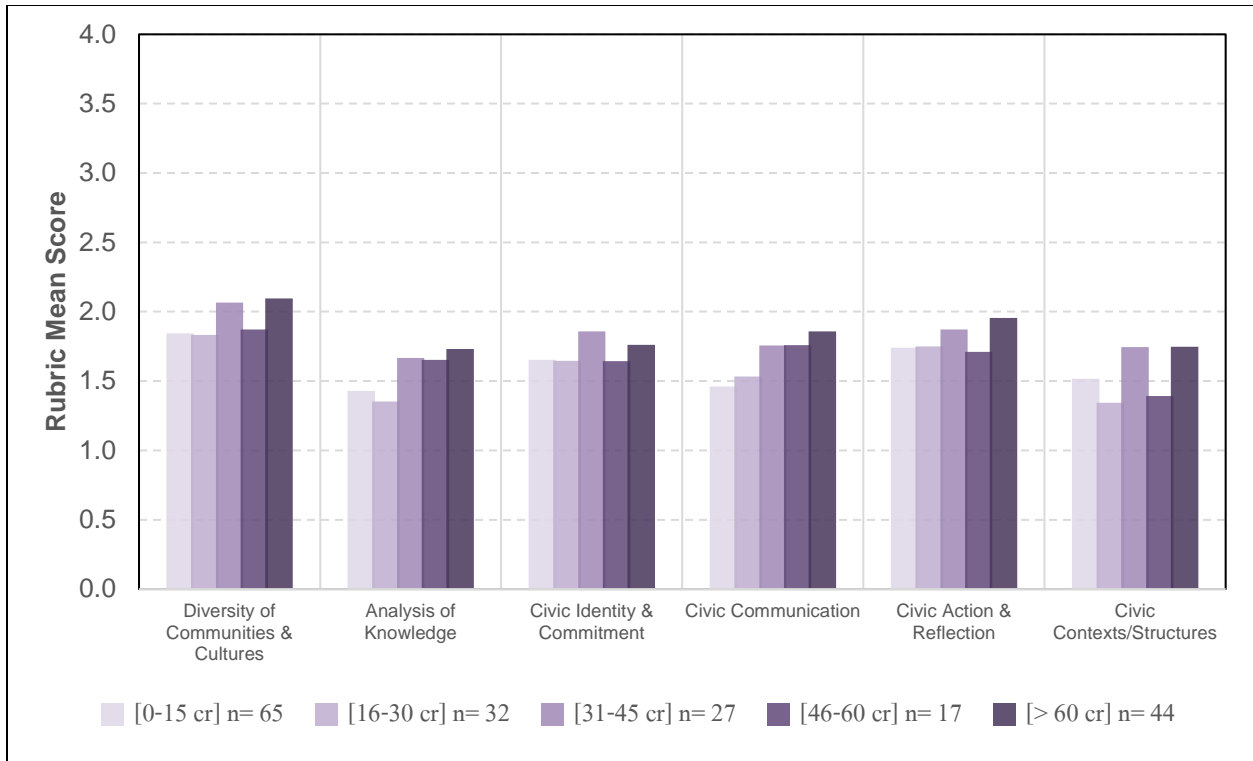
*Figure 22. Comparison of 'Engage' achievement at 3 or higher across all rubric dimensions for 185 artifacts in which credit information could be matched to at least one of the two artifact scorer's scores. From light purple to dark, 0-15 credits earned n=65, 16-30 credits earned n=32, 31-45 credits earned n=27, 46-60 credits earned n=17, and > 60 credits earned n=44. *Credits earned based on number of credits earned entering fall 2018 term.*

Figure 23. Comparison of 'Engage' achievement at 3 or higher across all rubric dimensions based on course success rates of students. From light beige to dark, students with 0-59% n=10, 60-69% n=14, 70-79% n=14, 80-89% n=29, and 90% or above n=43.

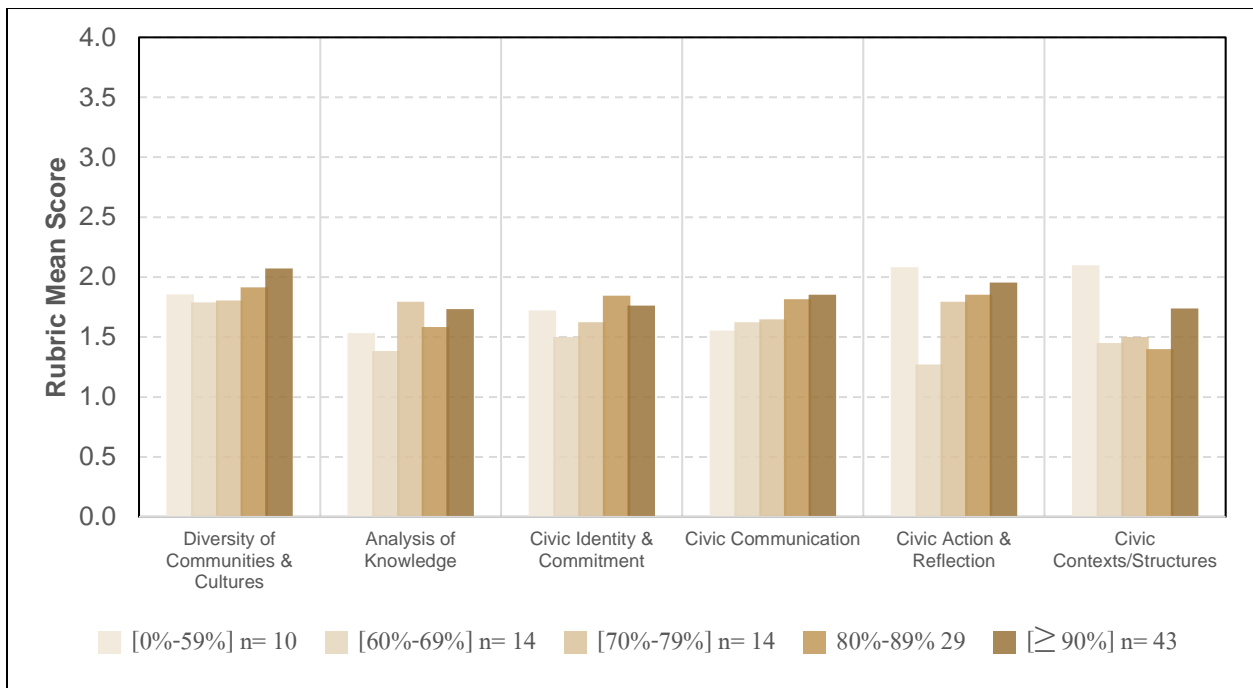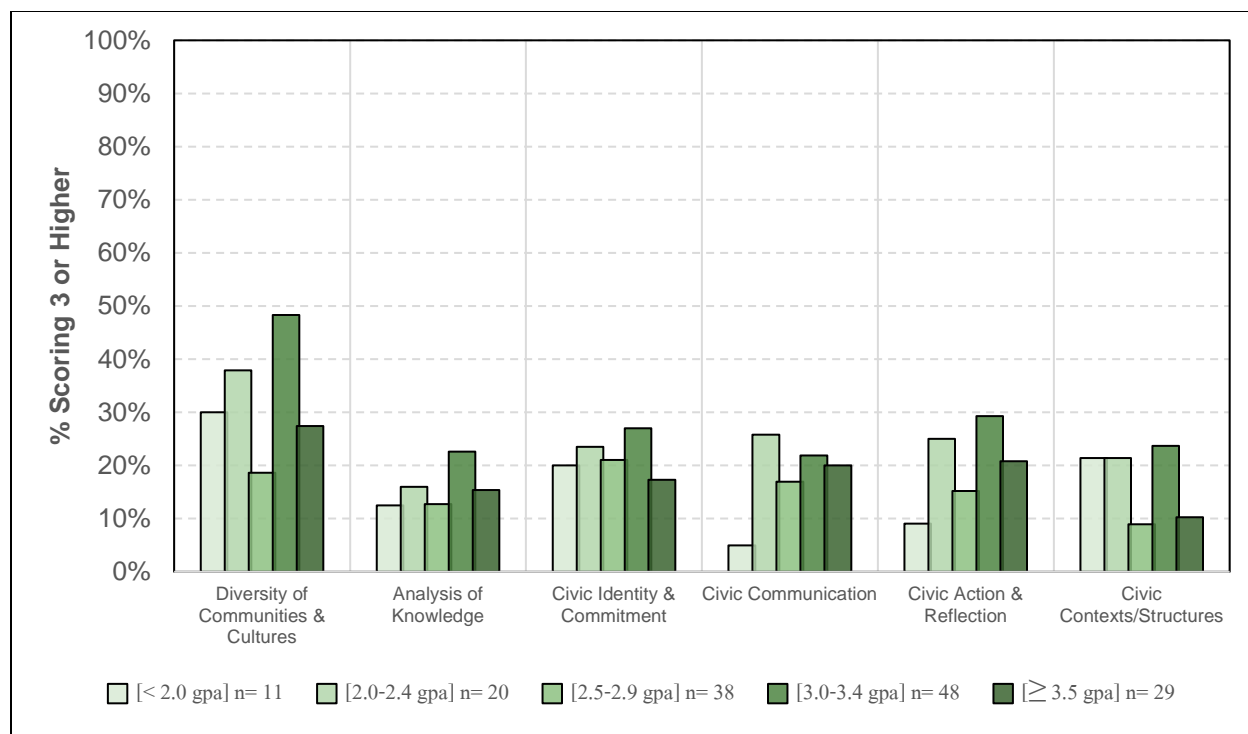*Figure 24. Comparison of 'Engage' achievement at 3 or higher across all rubric dimensions based on GPA. From light green to dark, GPA < 2.0 n=11, GPA 2.0-2.4 n=20, GPA 2.5-2.9 n=38, GPA 3.0-3.4 n=48, GPA ≥ 3.5 n=29. *Credits earned based on number of credits earned entering fall 2018 term.*

4   GENERAL LONGITUDINAL STUDY

In order to gain perspective into the results shared above, it can be valuable to look at generalized results from previous general education assessment studies at FSW. The new competencies are being studied in this assessment and no true longitudinal study can be completed. Therefore, instead of looking at a dimension by dimension comparison, it may be helpful to look at overall scores (combined average of rubric dimensions) from previous general education assessment studies with respect to the current results. Figures 25 and 26 provide these comparisons.

While scorers found many assignments did not include this category and were not scored, going forward reliability will increase over time as the assignments and the rubrics speak to each other better. The more the FSW community embrace the competencies, the more the achievement should rise. In many cases scorers noted that there are cases where students demonstrate further achievement beyond the stated assignment guidelines. Therefore, some of this process is about embracing the achievements of both faculty and students via documentation.
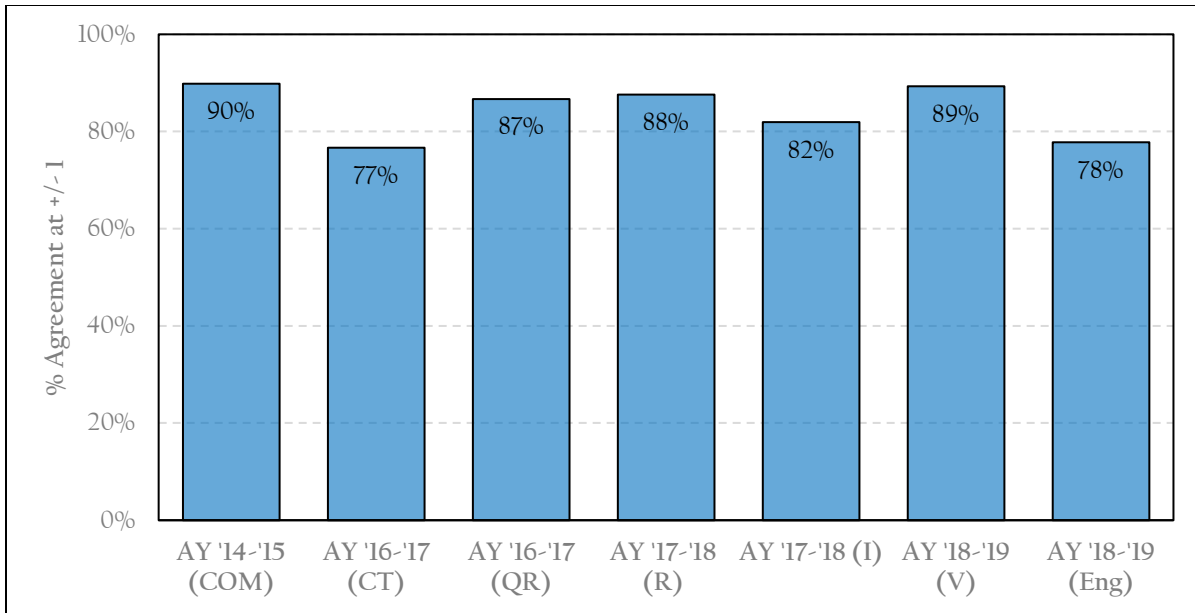
Figure 25. Comparison of inter-rater reliability (percentage (%) +/- 1 agreement) averaged across dimensions by each competency in FSW General Education Assessment cycle.
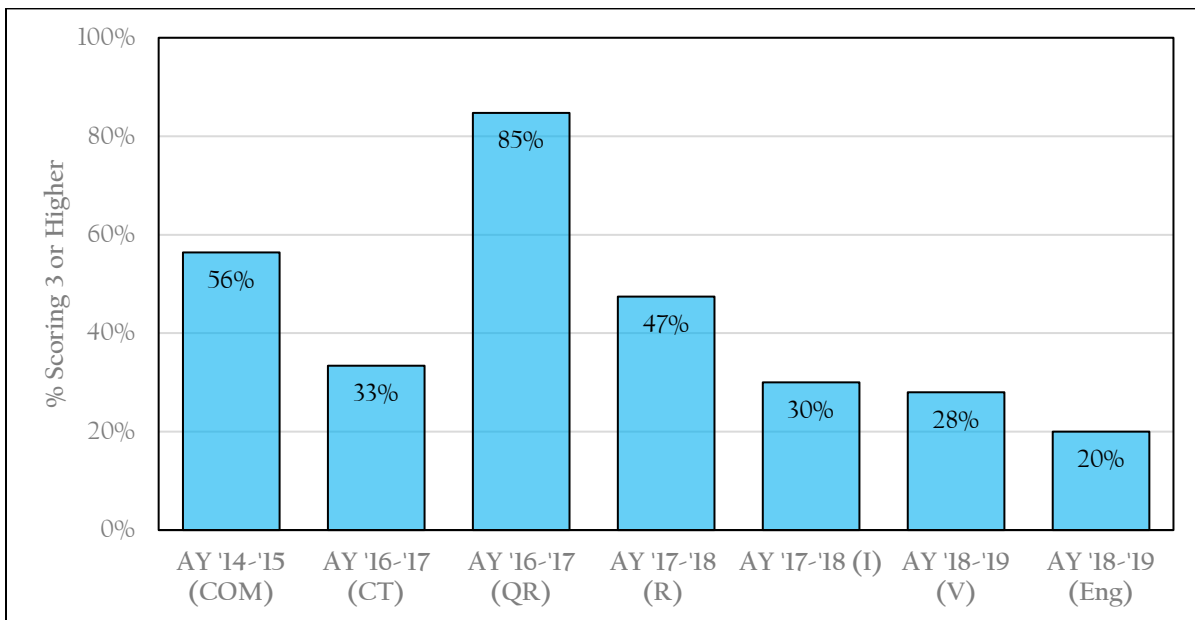
Figure 26. Comparison of achievement at 3 or higher averaged across dimensions by each competency in FSW General Education Assessment cycle.

5   PROFESSIONAL DEVELOPMENT PLANS

When reviewing general education assessment results, it is important to review assignments that are being assessed with respect to the rubric and the competency. Without a strong alignment between the task (competency) and the rubric/assignment, assessment measurements will always yield results more telling of the process and alignment rather than true achievement. This concept can be supported in

the work of Reeves (2006) in which the critical factors of learning are highlighted and assessment is one of eight major components.

Through a review of the results laid out above and discussions within the FSW Community two main plans emerged. The first plan is to develop a repository of ideal assignments that line up well with rubrics that would be available to FSW faculty may be a good way of alleviating some of the problems noted by scorers. The plan began in AY 2018-2019 and continues in AY 2019-2020. The second is to development assignment building workshops specific to the competency and bring them to departments that are rich in that competency as opposed to housing them at FSW's Teaching and Learning Center (TLC). In short, by bringing the training workshops to department meetings instead of asking faculty to encumber their schedules, the idea would be that support can be provided with a higher yield.

6 CONCLUSIONS

FSW's General Education Program was assessed through randomly sampled from a list of courses which were identified by faculty as encompassing that competency. The study details the results of FSW's General Education assessment for AY 2018-2019 which included the analysis of 'Visualize' and 'Engage' from the new C-R-E-A-T-I-V-E General Education competencies. Results also included these same outcomes with respect to courses included in the AA program and value-added studies based on credits earned, success rates of the student from which the artifacts are collected, and GPA from those students.

A drilldown of 'Visualize' (V) results are as follows:
1. Zero of six rubric dimensions exhibit greater than 60% achievement at level '3'. The highest scored dimension is "Global Self-Awareness" at 40% scoring '3' or higher. A great many assignments were noted as difficult to score using the rubric, which may have influences results.
2. Mean achievement levels for each of the six rubric dimensions of 'Visualize' range from 1.65 to 2.19 on a 4-point scale.
3. In a study comparing online, dual enrollment (concurrent), and traditional artifacts, the traditional modality exhibits the highest in 1 of 6 dimensions. Concurrent dual enrollment exhibits the highest in 4 of 6. Online exhibits the highest in 1 of. Results for 1 of 6 dimensions were statistically significantly different (dual enrollment statistically significantly higher than both online or traditional).
4. An inter-rater reliability study exhibits rubric scoring agreement ranging from 35% to 53% with a +/- 1 agreement ranging from 86% to 93%.
5. With respect to AA courses, zero of six rubric dimensions exhibit greater than 60% achievement at level '3'. The highest scored dimension is "Global Self-Awareness" at 40% scoring '3' or higher.
6. In a study comparing AA courses with online, dual enrollment, and traditional artifacts, the traditional modality exhibits the highest in 1 of 6 dimensions. Concurrent dual enrollment exhibits the highest in 4 of 6. Online exhibits the highest in 1 of. Results for 1 of 6 dimensions were statistically significantly different (dual enrollment statistically significantly higher than both online or traditional).
7. In a study comparing achievement at 3 or higher across rubric dimensions based on credits earned, achievement is higher with increased credits earned in 2 of 6 cases. In the remaining

four cases, three exhibit similar results for both low and high credit completion, while one, "Applying Knowledge to Contemporary Global Contexts," exhibits a decline.

8. In a study comparing achievement at 3 or higher across rubric dimensions based on success rates of students, achievement is similar in all success rates in 4 of 6 dimensions. In the remaining two, "Global Self-Awareness" and "Applying Knowledge to Contemporary Global Contexts," a slight decline is present.

9. In a study comparing achievement at 3 or higher based on GPA, achievement does not exhibit a clear pattern across rubric dimensions. In four cases, the highest GPA exhibits the highest achievement. And in the case of "Personal and Social Responsibility," a clear trend exists where increased GPA is correlated with increased achievement (GPA < 2.0 = 17% scoring 3 or higher compared with ≥ 3.5 = 33%).

10. In a review of scorer feedback, multiple scorers noted that the rubric addresses areas the assignment does not call for (entire dimension can't be scored). All dimensions of the rubric were at one point or another called out in varying assignments as not being able to be assessed. Scorers noted that there is overlap in several areas that causes problems with scoring depending on the assignment. Elements of culture, diversity, and responsibility appear in varying dimension making interpretation troublesome. Also, scorers commented positively regarding the rubric that dimensions cover a lot of areas. So, while the rubric may overlap, it does appear to be versatile in the style of assignments. The LAC has already tended to the task of writing a new FSW 'Visualize' rubric based on the findings included in this report.

A drilldown of 'Engage' (Eng) results are as follows:

1. Zero of six rubric dimensions exhibit greater than 60% achievement at level '3' with percentages ranging from 13% to 35%.

2. Mean achievement levels for each of the six rubric dimensions of 'Engage' range from 1.54 to 1.95 on a 4-point scale.

3. In a study comparing online, dual enrollment (concurrent), and traditional artifacts, the traditional modality exhibits the highest in 4 of 6 dimensions. Differences from traditional to online are statistically significant in "Civic Contexts/Structures."

4. An inter-rater reliability study exhibits rubric scoring agreement ranging from 15% to 52% with a +/- 1 agreement ranging from 71% to 91%.

5. With respect to AA courses, zero of six rubric dimensions exhibit greater than 60% achievement at level '3' with percentages ranging from 8% to 39%.

6. In a study comparing online, dual enrollment, and traditional artifacts, the traditional modality exhibits the highest in 4 of 6 dimensions. Differences from traditional to online are statistically significant in "Civic Contexts/Structures."

7. In a study comparing achievement at 3 or higher across rubric dimensions based on credits earned, achievement exhibits a slight but steady increase with increased credits earned for all dimensions.

8. In a study comparing achievement at 3 or higher across rubric dimensions based on success rates of students, achievement exhibits increases with increased success rates in 6 of 6 cases, although in some cases the lowest success rate does not conform to this trend, likely a small sample size issue.

9. In a study comparing achievement at 3 or higher based on GPA, achievement does not exhibit a clear pattern across rubric dimensions.
10. In a review of scorer feedback, first, scorers noted some overlap between the Action/Reflection and Context/Structures dimension. Second, scorers noted that the scoring levels '1' and '2' in the 'Analysis of Knowledge' dimension are too similar. And third, very few assignments appear to capture the rubric entirely. This is not necessarily a problem as the assignment should be tied to the competency and not the rubric. The LAC has already tended to the task of writing a new FSW 'Engage' rubric based on the findings included in this report.

A drilldown of longitudinal studies are as follows:
1. In a comparison of inter-rater reliability (percentage (%) +/- 1 agreement) averaged across dimensions by each competency in FSW General Education Assessment cycle, both 'Visualize' and 'Engage' exhibit results similar to those of past studies (89% and 78% compared with a range of 77% to 90% in past studies).
2. In a comparison of achievement at 3 or higher averaged across dimensions by each competency in FSW General Education Assessment cycle, the 'Visualize' and 'Engage' competencies exhibit the 7th and 8th ranked achievement percentages of the past seven assessments, respectively. It is important to note, however, that the 5th, 6th, and 7th ranked achievement percentages are all within 5% points.

A drilldown of professional development plans:
1. To develop a repository of ideal assignments that line up well with rubrics that would be available to FSW faculty may be a good way of alleviating some of the problems noted by scorers. The plan began in AY 2018-2019 and continues in AY 2019-2020.
2. To development assignment building workshops specific to the competency and bring them to departments that are rich in that competency as opposed to housing them at FSW's Teaching and Learning Center (TLC).

7 REFERENCES

Braselton, M. "Using AAC&U Value Rubrics. Case Study: Midland College." https://www.aacu.org/VALUE/casestudies/midland.pdf

Cohen, J. 1960. A coefficient of agreement for nominal scales. Educational and Psychological Measurement, Vol. XX(1), 37-46.

Florida SouthWestern State College, 2014, General Education Assessment Subcommittee Meeting Minutes, https://docs.fsw.edu/docnew/view.php?fDocumentId=298910.

Gwet, K. 2002. Inter-rater reliability: Dependency on trait prevalence and marginal homogeneity. Series: Statistical methods for Inter-Rater Reliability Assessment, No. 2, 1-9.

Reeves, T.C. 2006. How do you know they are learning?: the importance of alignment in higher education. International Journal of Learning Technology, 2(4).

Rhodes, T. & Finley, A. (2013) Using the VALUE Rubrics for Improvement of Learning and Authentic Assessment. AAC & U.