

General Education Assessment Report – AY 2017-18

Author: Joseph F. van Gaalen, Ph.D., Director, 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. This report details the results of FSW's General Education assessment for AY 2017-

2018 which included the analysis of 'Research' and 'Investigate' from the new 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, Director of Assessment & Effectiveness, Academic Affairs (jfvanngaalen@fsw.edu; x16965).

2 RESEARCH (R)

The outcome of the 'Research' competency at FSW is that by completion of the general education requirements, students will be able to research and examine academic and non-academic information, resources, and evidence. 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 13 below depict achievement and inter-rater reliability for the 'Research' 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 by Keene State College (KSC) as a means of scoring achievement in this competency. Feedback from scorers regarding the KSC rubric for the 'Research' 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). The 1st, 2nd, 3rd, and 5th dimensions were called out by different scorers. Second, scorers noted that the KSC rubric addresses the communication area of research by asking scorers to reference specific dimensions of the college's separate communication competency rubric. Scorers felt that this was inconvenient and also could occasionally 'take the scorer out of focus' with the assignment and its purpose. And third, scorers commented positively regarding the rubric that dimensions did not appear too similar, and that there was little overlap across competencies. Ultimately, scorers felt the rubric was adequate, but not sure it is best suited for a wide variety of assignments. For this study, no modifications to the KSC rubric has been made. The LAC has already tended to the task of writing a new FSW 'Research' rubric based on the findings included in this report.

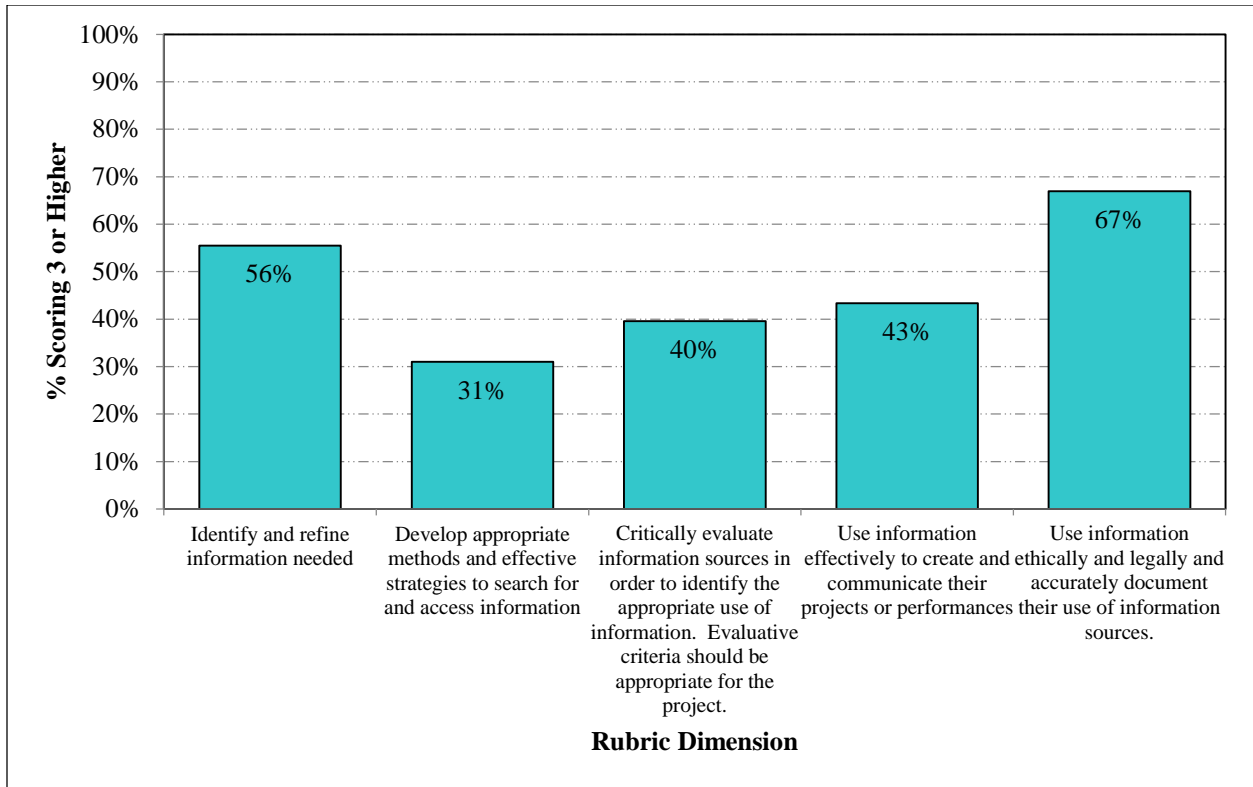


Figure 1. 'Research' achievement at 3 or higher across all rubric dimensions for 211 artifacts from 28 sampled course sections.

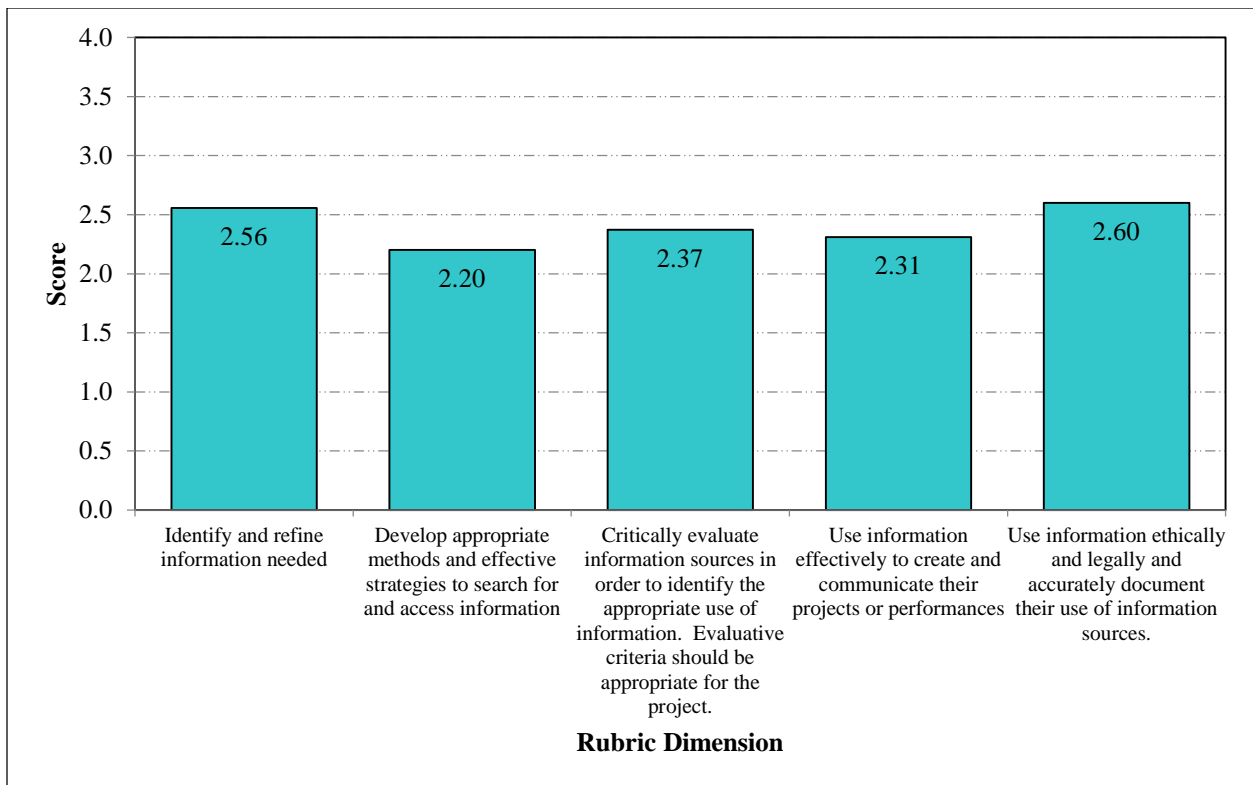


Figure 2. Mean score by rubric dimension for 'Research' for 211 artifacts from 28 sampled course sections.

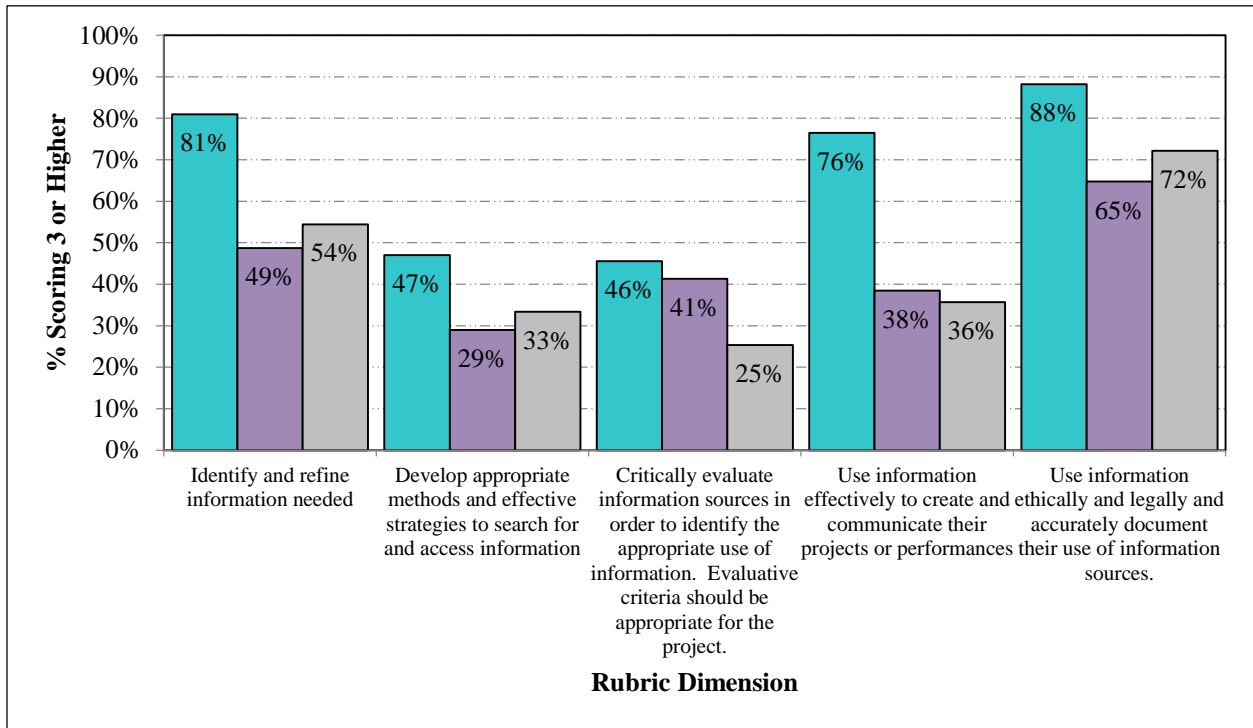


Figure 3. Comparison of 'Research' achievement by modality at 3 or higher across all rubric dimensions for 211 artifacts from 28 sampled course sections. Traditional (aqua), n=34, Online (purple), n=98, Dual Enrollment (concurrent) (gray), n=79.

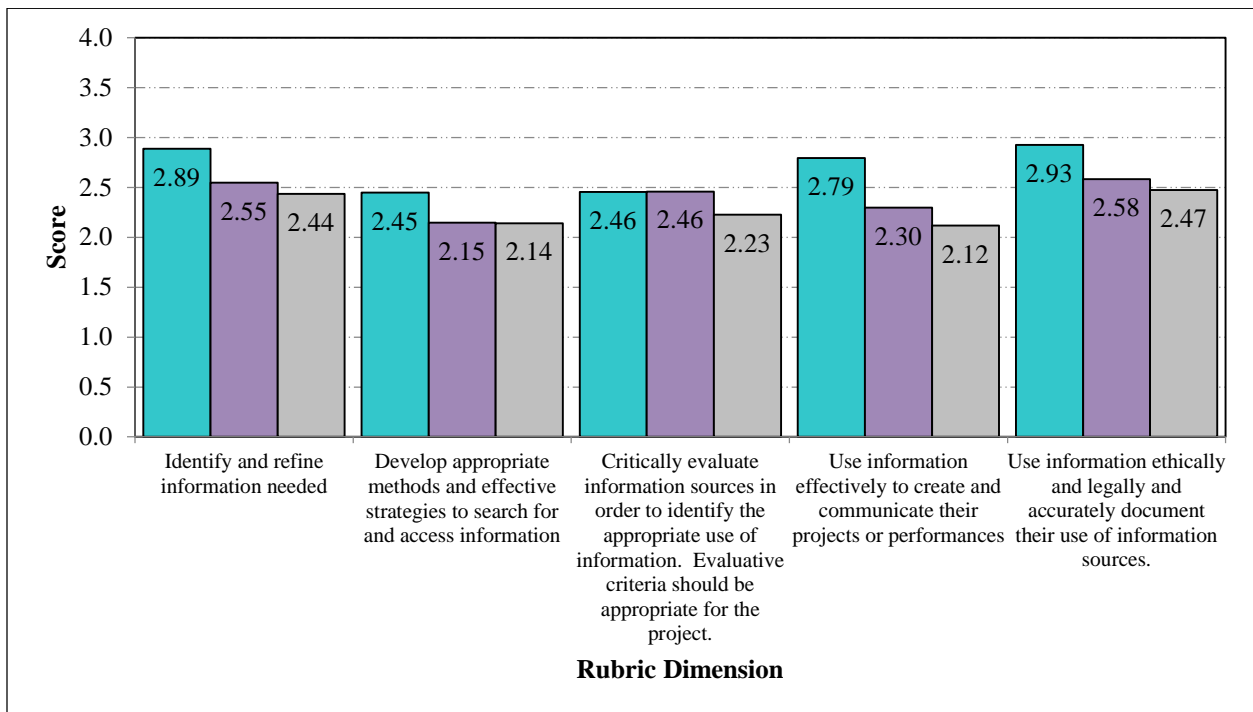


Figure 4. Mean score of 'Research' for each rubric dimension by modality at 3 or higher across all rubric dimensions for 211 artifacts from 28 sampled course sections. Traditional (aqua), n=34, Online (purple), n=98, Dual Enrollment (concurrent) (gray), n=79.

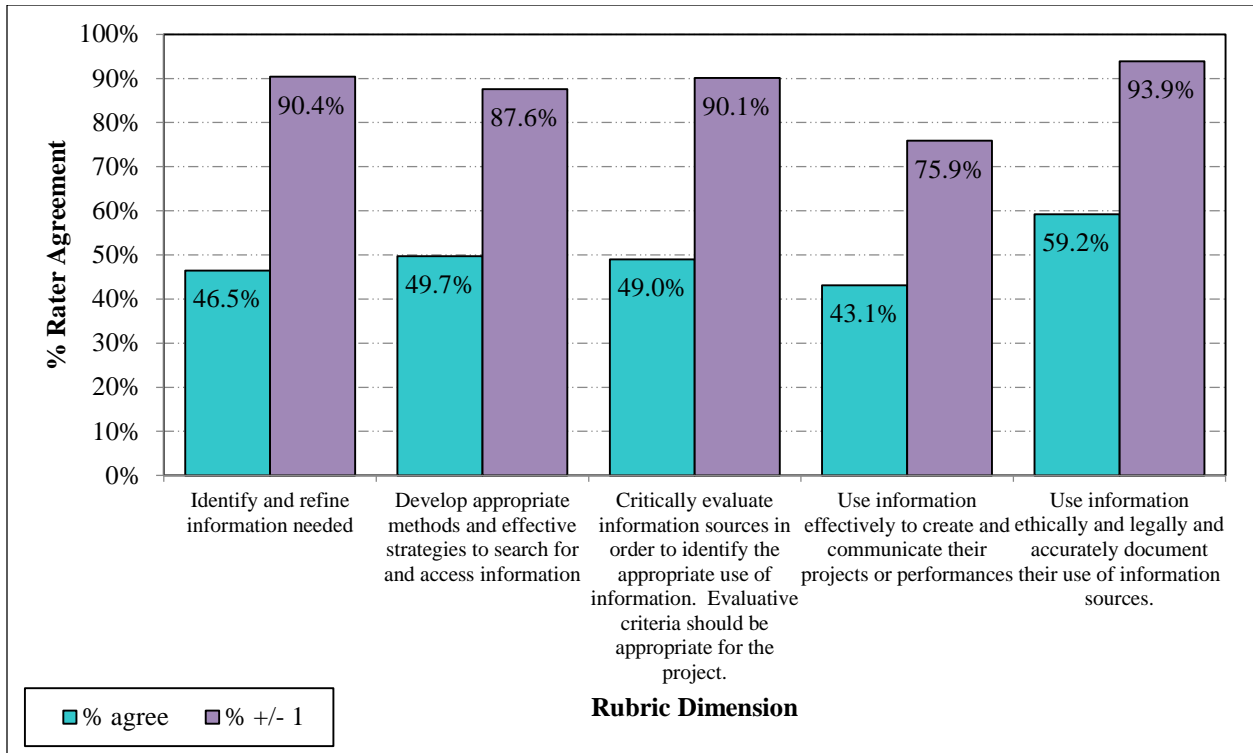


Figure 5. Inter-rater reliability (as %) for the 'Research' competency. Each artifact was scored by two scorers. Percentage (%) of agreement (aqua) is defined as cases where scores by each scorer were identical. Percentage (%) +/- 1 agreement (purple) is defined as cases where scores by each scorer were within 1 of each other.

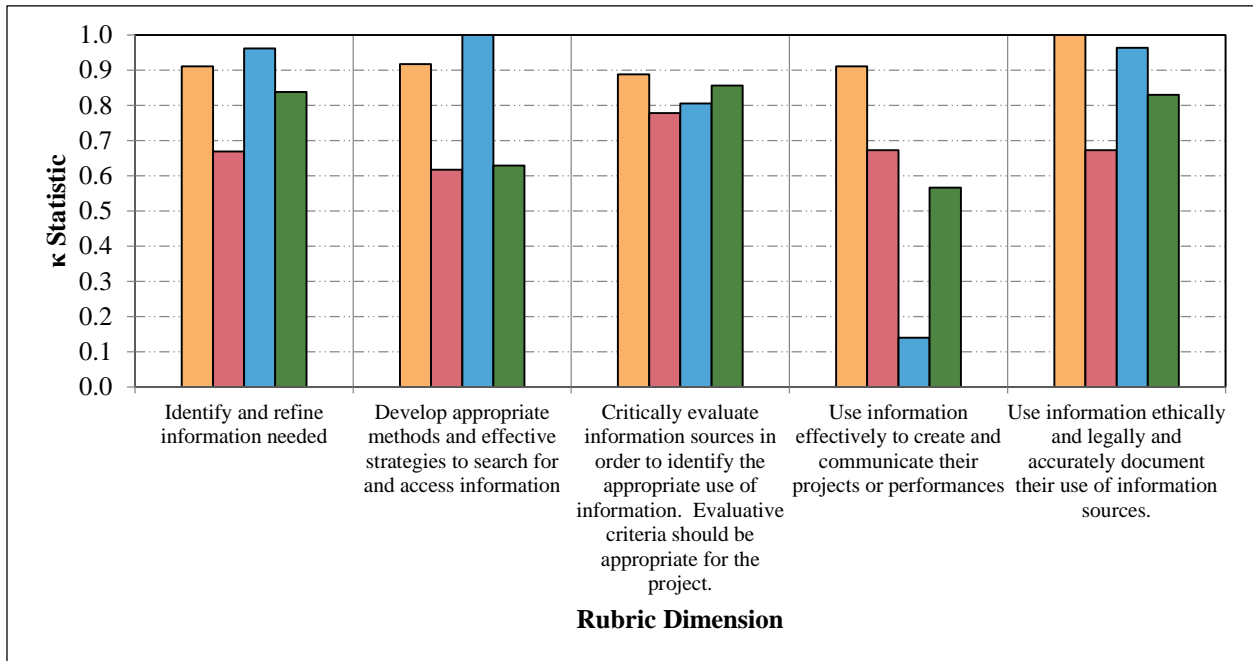


Figure 6. κ -statistic for scorer pairs of the 'Research' competency artifacts (Scorers 1A & 1B – orange, Scorers 2A & 2B – red, Scorers 3A and 3B – blue, Scorers 4A & 4B - green). The κ -statistic is evaluated for the percentage (%) +/- 1 agreement shown in Figure 4 above and takes into account the random chance that scorers would agree (Cohen, 1960; Gwet, 2002), thus the κ -statistic results will be slightly lower than a straightforward percentage of agreement. The κ -statistic was calculated for the 'research' study to be use as a guide to interpret percentage agreement in an academic sense.

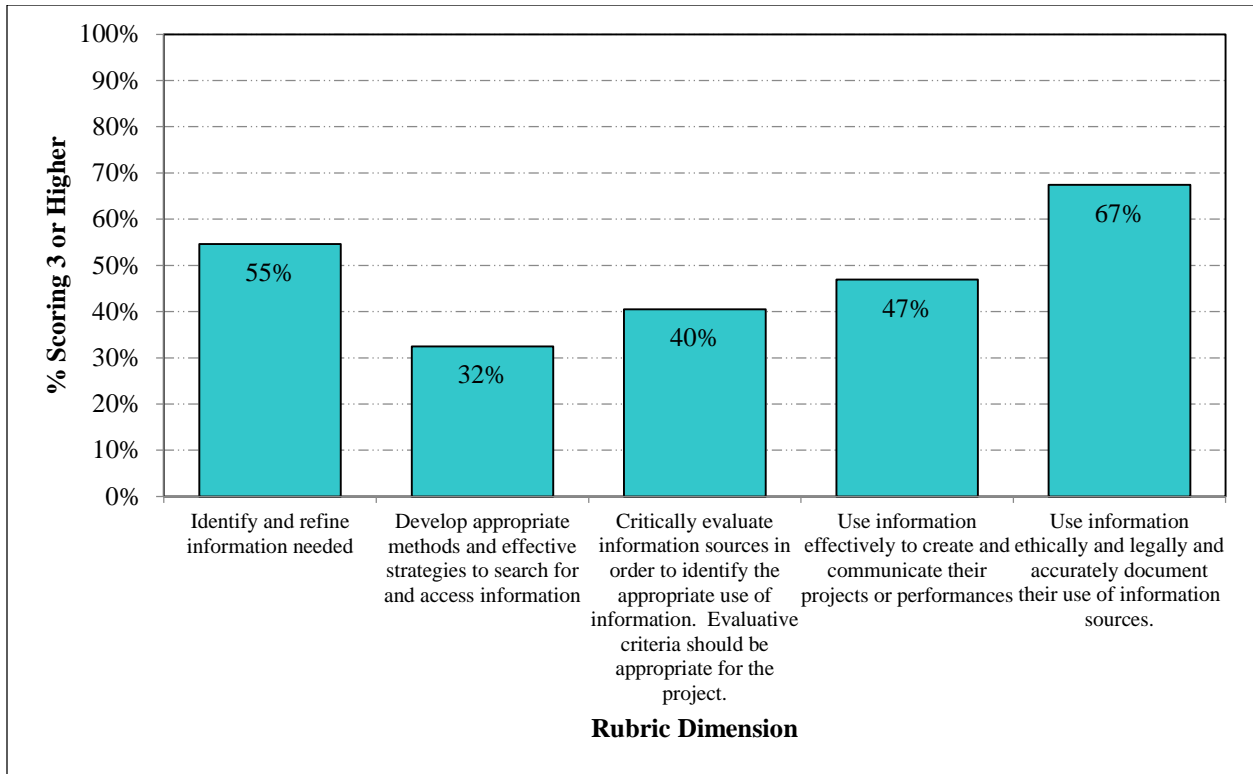


Figure 7. 'Research' achievement at 3 or higher across all rubric dimensions for AA courses only for 138 artifacts from 28 sampled course sections.

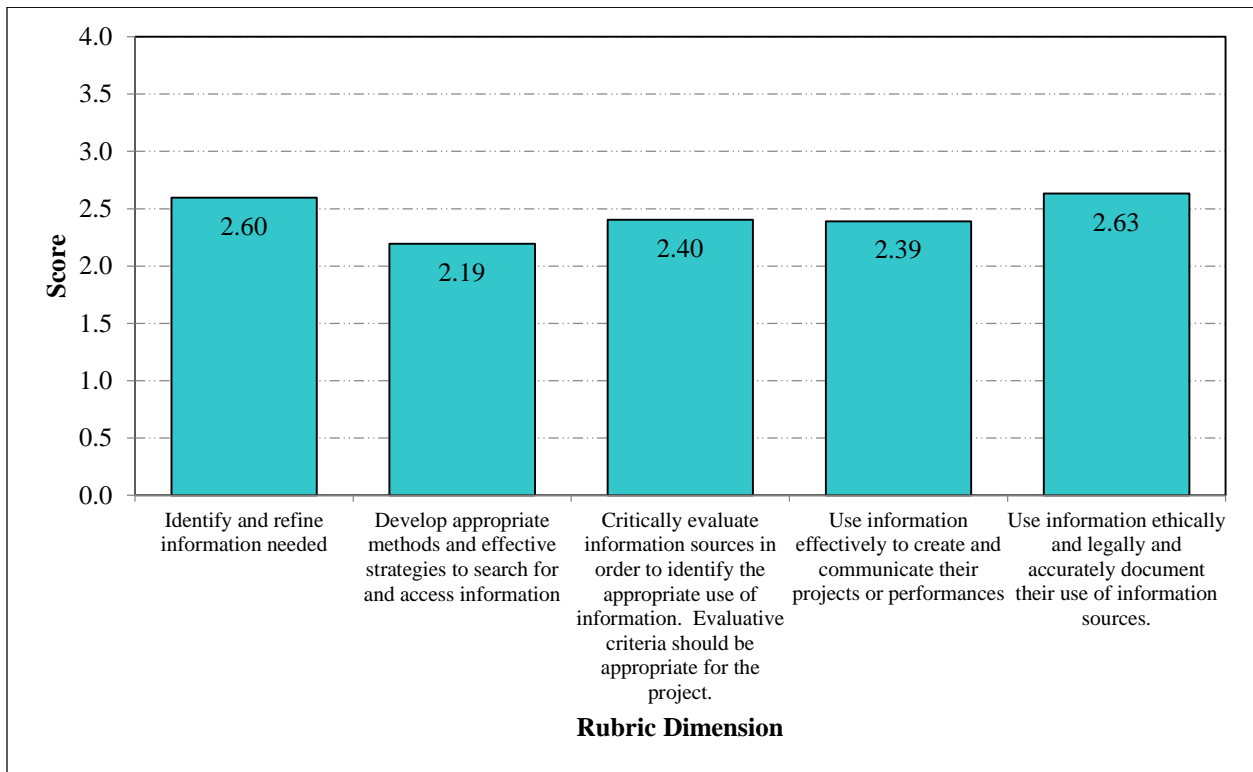


Figure 8. Mean score by rubric dimension for 'Research' for AA courses only for 138 artifacts from 28 sampled course sections.

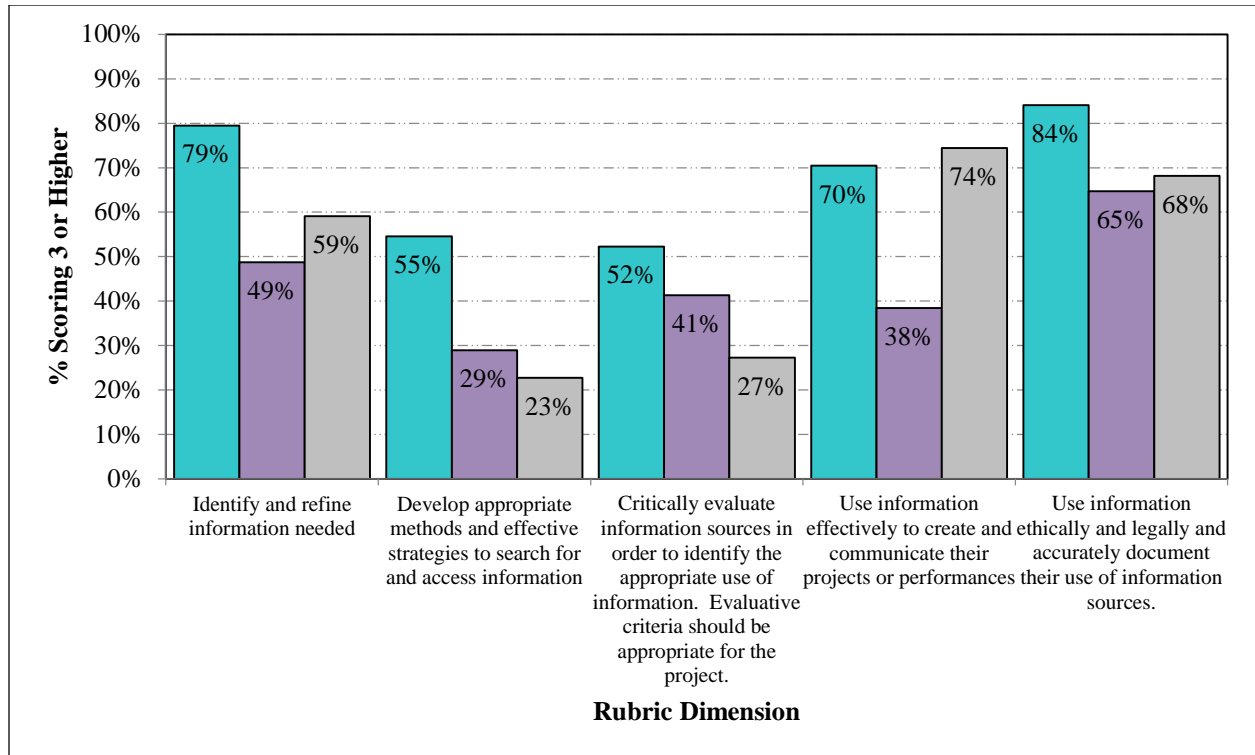


Figure 9. 'Research' achievement at 3 or higher across all rubric dimensions for AA courses only for 138 artifacts from 28 sampled course sections. Traditional (aqua), n=20, Online (purple), n=98, Dual Enrollment (concurrent) (gray), n=22.

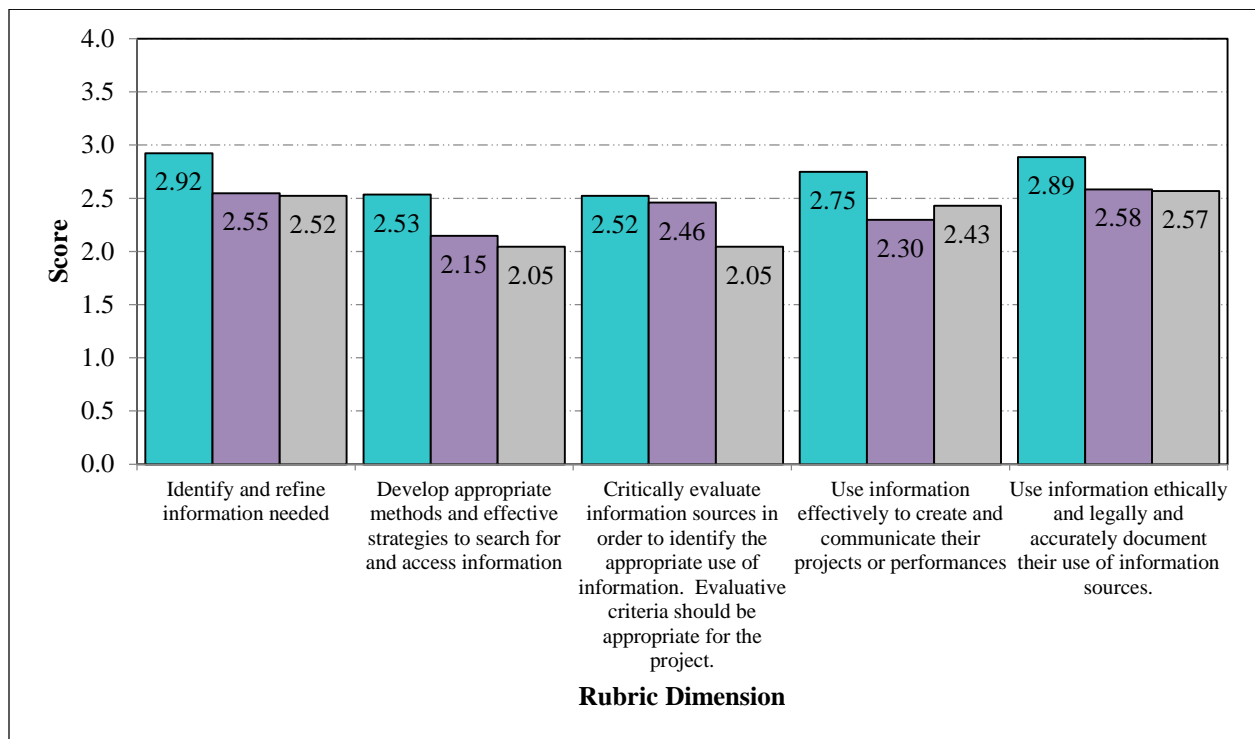


Figure 10. Mean score of 'Research' for each rubric dimension by modality at 3 or higher across all rubric dimensions for AA courses only for 138 artifacts from 28 sampled course sections. Traditional (aqua), n=20, Online (purple), n=98, Dual Enrollment (concurrent) (gray), n=22.

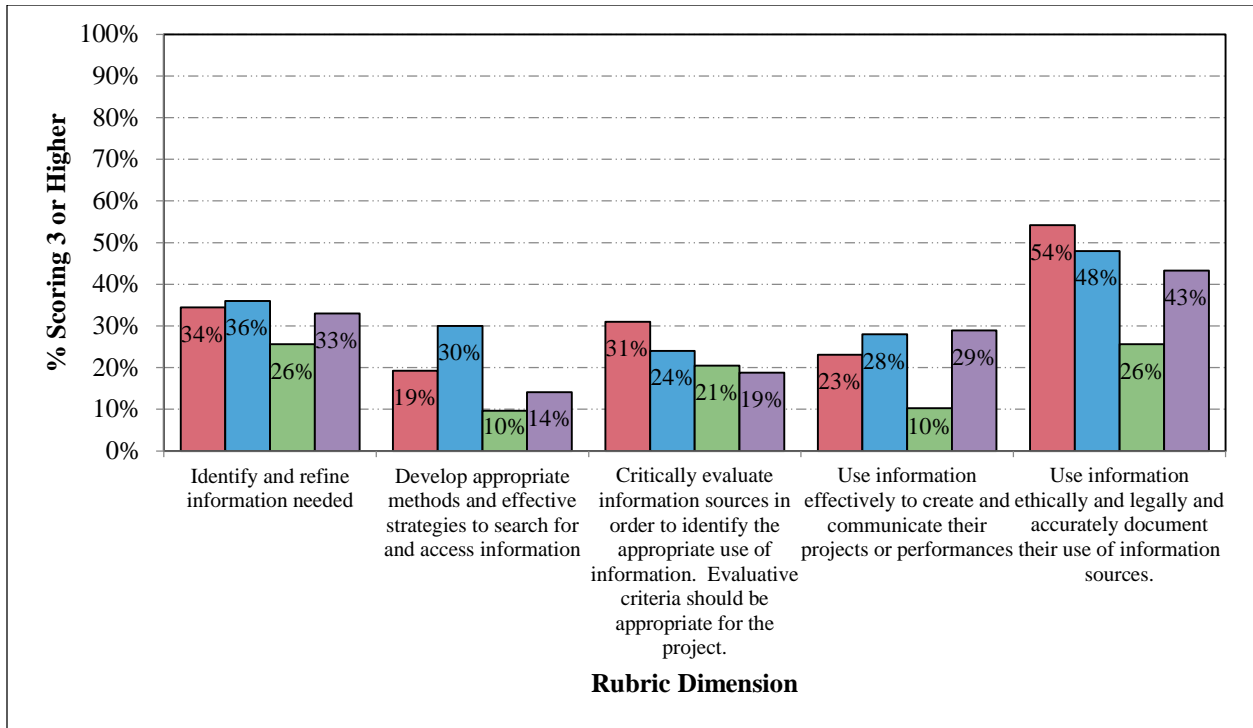


Figure 11. Comparison of 'Research' achievement at 3 or higher across all rubric dimensions for 114 artifacts in which credit information could be matched to artifact score. 0-15 credits earned (red) n=22, 16-30 credits earned (blue) n=13, 31-60 credits earned (green) n=22, > 60 credits earned (purple) n=57. *Credits earned based on number of credits earned entering fall 2017 term.

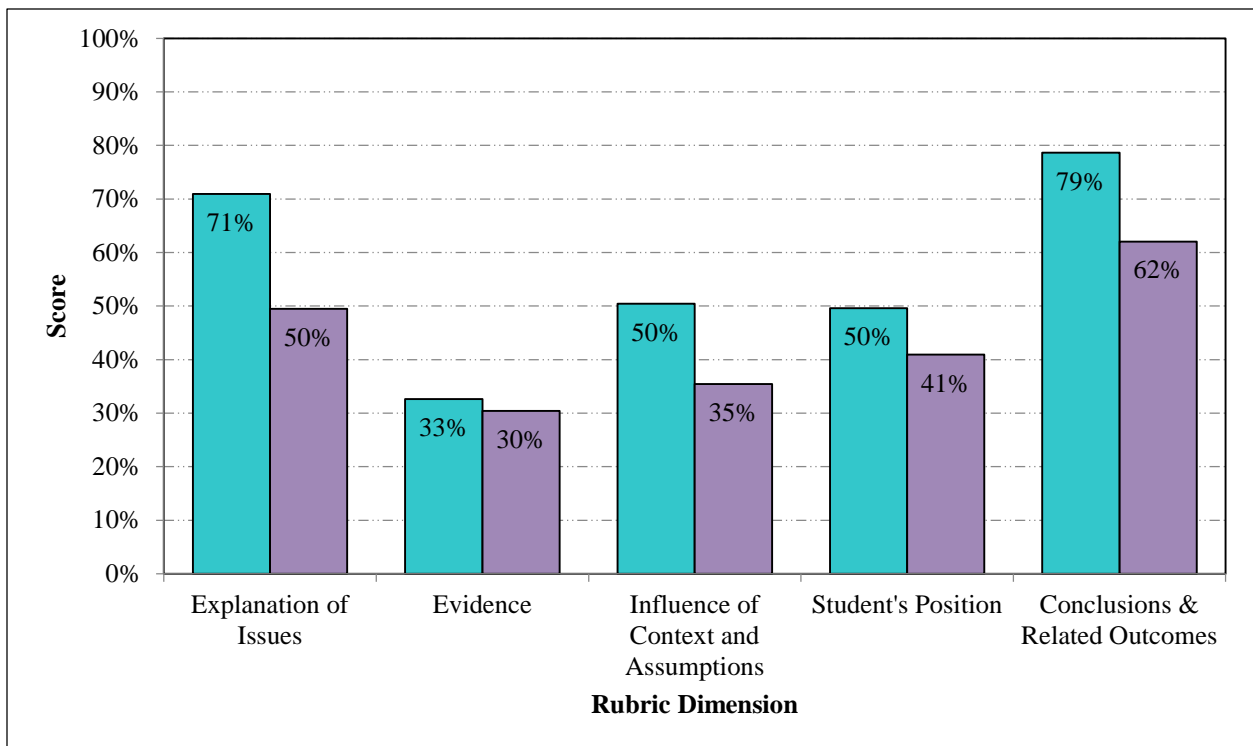


Figure 12. Comparison of 'Research' achievement at 3 or higher across all rubric dimensions based on courses with pre-requisites. Courses with pre-requisites (aqua, n=59), courses without pre-requisites (purple, n=184).

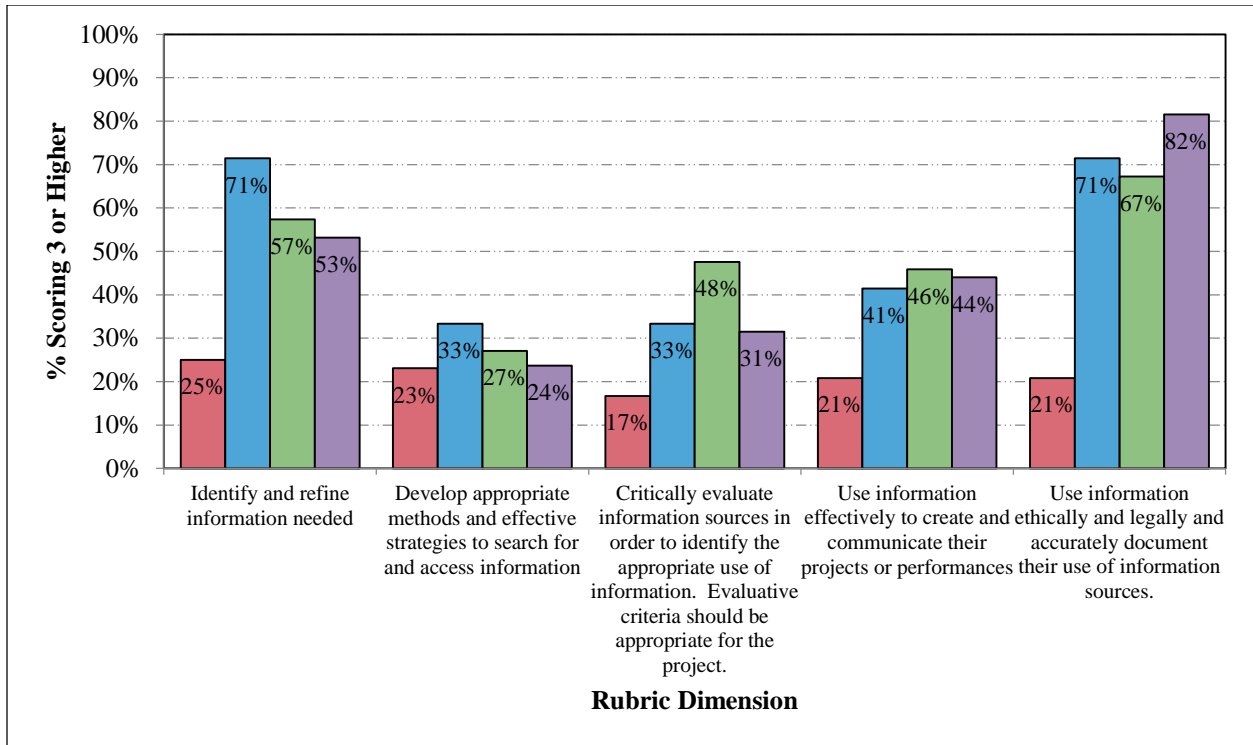


Figure 13. Comparison of 'Research' achievement at 3 or higher across all rubric dimensions based on GPA. GPA < 2.5 (red) n=17, GPA 2.5-3.0 (blue) n=24, GPA 3.0-3.5 (green) n=34, GPA ≥ 3.5 (purple) n=35. *Credits earned based on number of credits earned entering fall 2017 term.

3 INVESTIGATE (I)

The outcome of the 'Investigate' competency at FSW is that by completion of the general education requirements, students will be able to investigate and engage in the transdisciplinary applications of research, learning, and knowledge. 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 14 through 26 below depict achievement and inter-rater reliability for the 'Investigate' 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 'Investigate' competency included three main trends regarding rubric suitability. First, multiple scorers noted poor alignment between the assignments being scored and the rubric dimensions. Second, scorers noted that the first and fifth rubric dimensions appear to cover some of the same ground. And third, very few assignments appear to really get at the heart of the concept of transdisciplinary studies, a core component of the 'Investigate' competency. Assignments often made an attempt at addressing transdisciplinary thinking, which is not an easy task, but didn't fully achieve this goal.

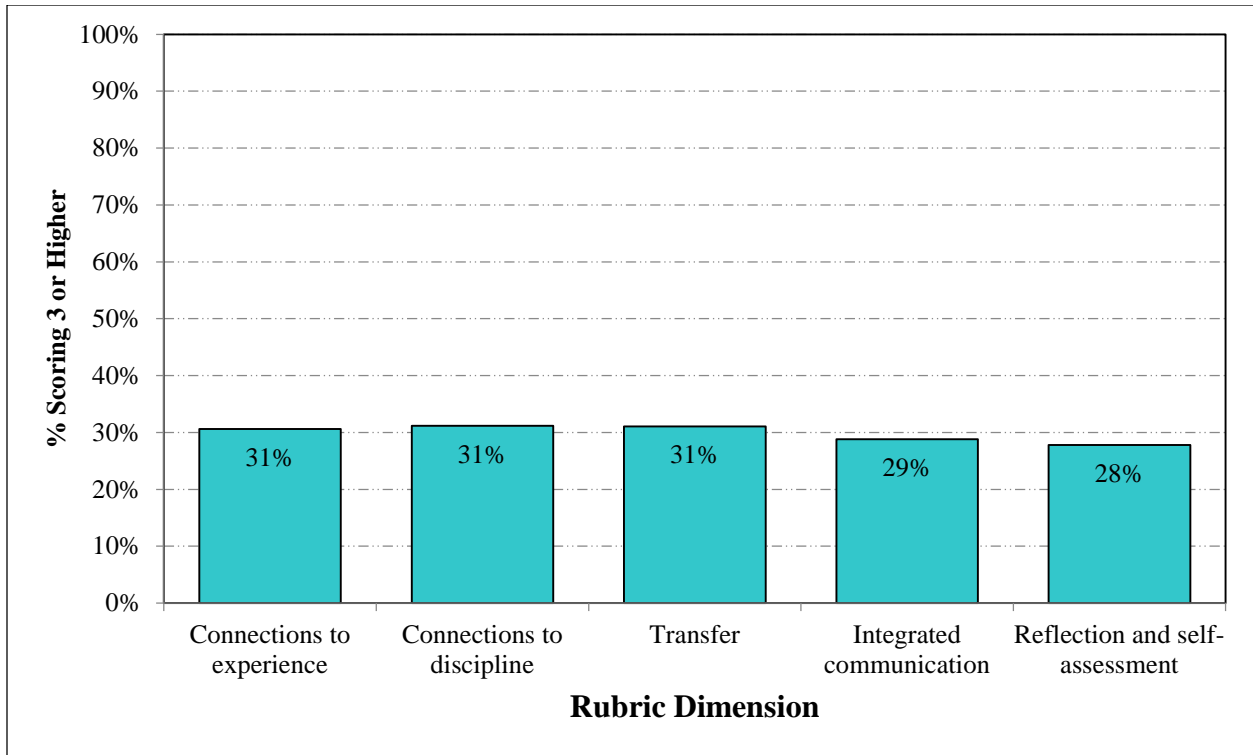


Figure 14. 'Investigate' achievement at 3 or higher across all rubric dimensions for 172 artifacts from 25 sampled course sections.

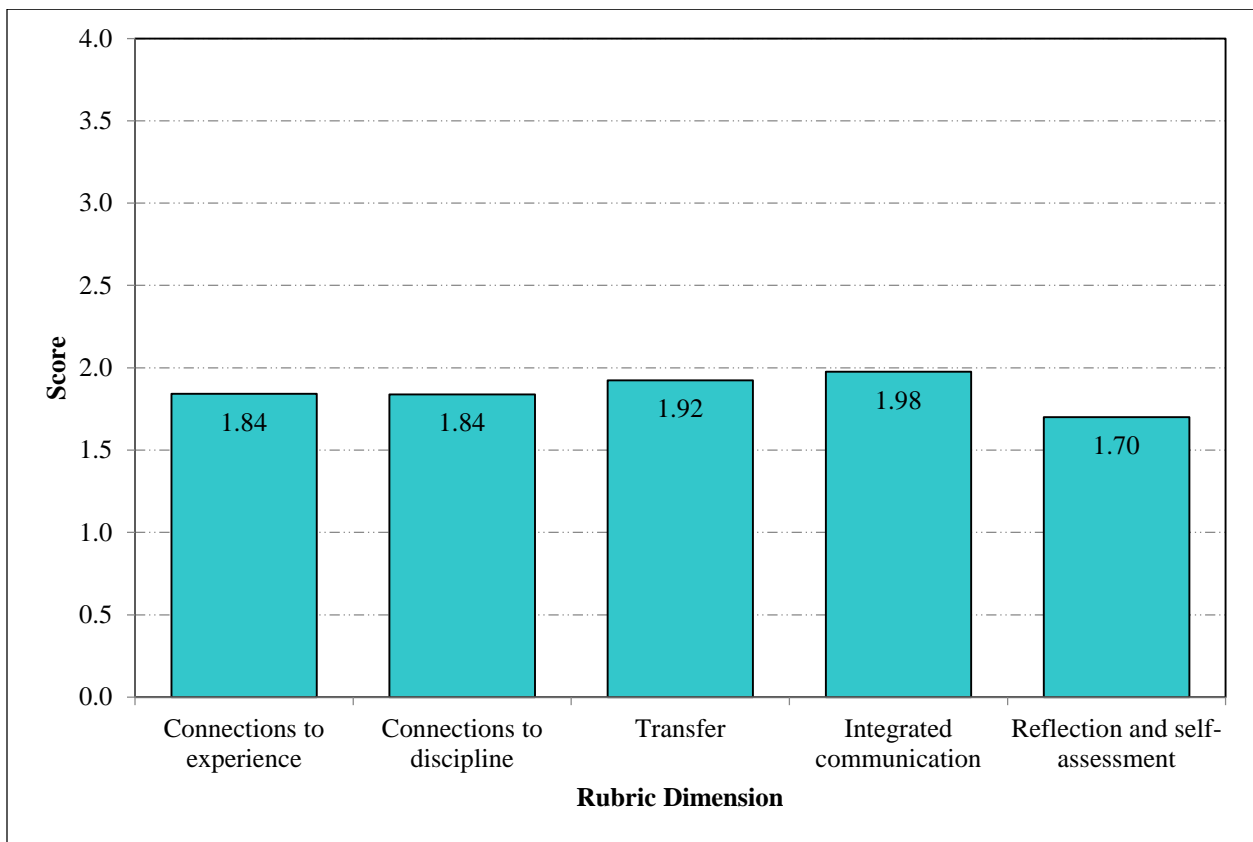


Figure 15. Mean score by rubric dimension for 'Investigate' for 172 artifacts from 25 sampled course sections.

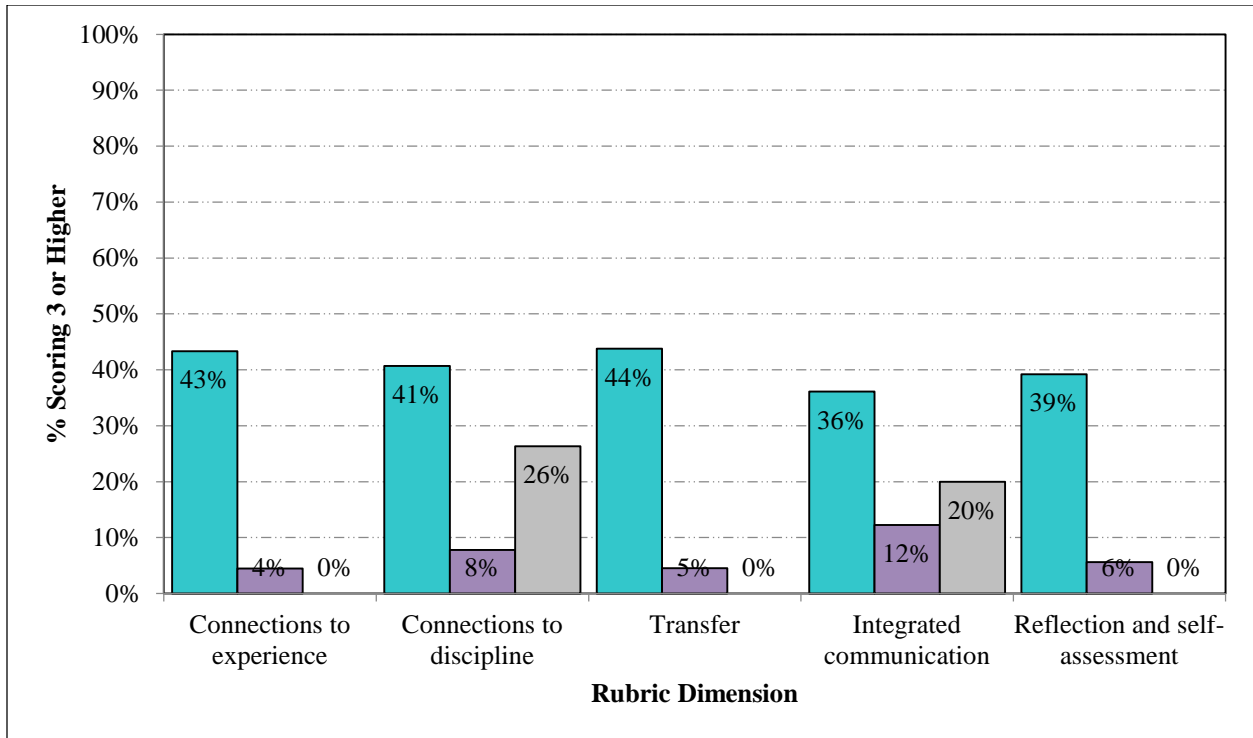


Figure 16. Comparison of 'Investigate' achievement by modality at 3 or higher across all rubric dimensions for 172 artifacts from 25 sampled course sections. Traditional (aqua), n=117, Online (purple), n=45, Dual Enrollment (concurrent) (gray), n=10.

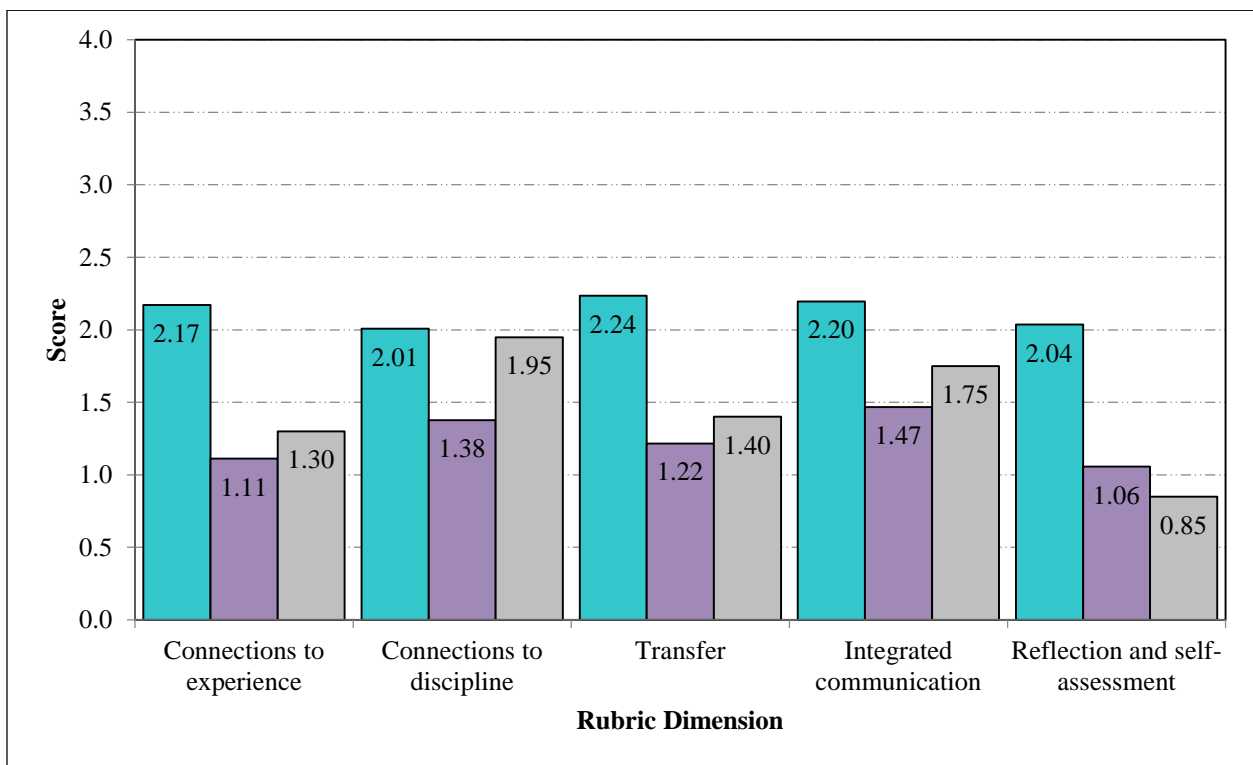


Figure 17. Mean score of 'Investigate' for each rubric dimension by modality at 3 or higher across all rubric dimensions for 172 artifacts from 25 sampled course sections. Traditional (aqua), n=117, Online (purple), n=45, Dual Enrollment (concurrent) (gray), n=10.

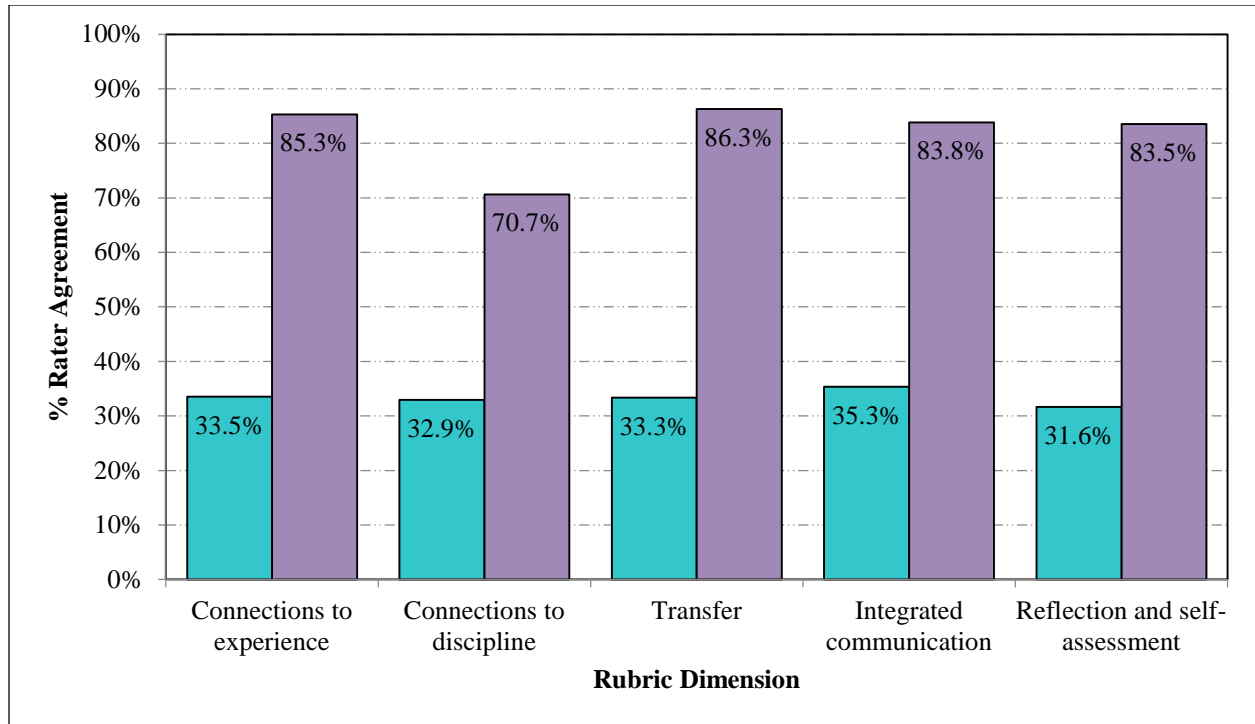


Figure 18. Inter-rater reliability (as %) for the 'Investigate' competency. Each artifact was scored by two scorers. Percentage (%) of agreement (aqua) is defined as cases where scores by each scorer were identical. Percentage (%) +/- 1 agreement (purple) is defined as cases where scores by each scorer were within 1 of each other.

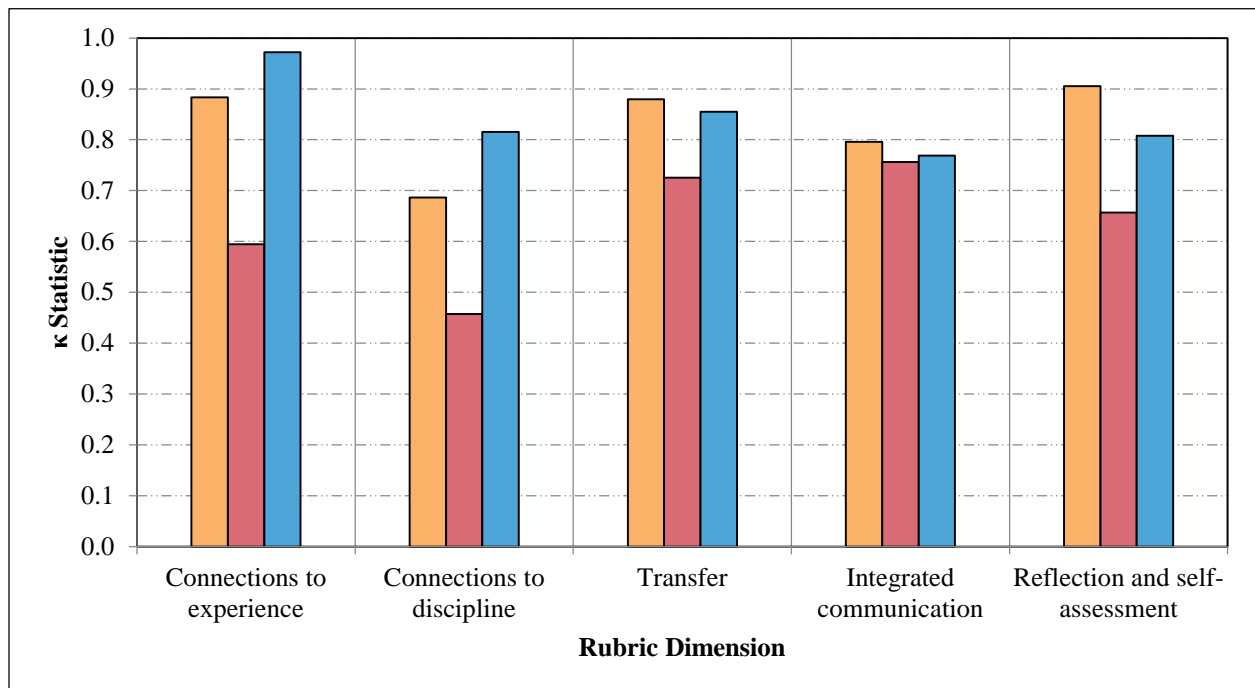


Figure 19. κ -statistic for scorer pairs of the 'Investigate' competency artifacts (Scorers 1A & 1B – orange, Scorers 2A & 2B – red, Scorers 3A and 3B – blue). The κ -statistic is evaluated for the percentage (%) +/- 1 agreement shown in Figure 4 above and takes into account the random chance that scorers would agree (Cohen, 1960; Gwet, 2002), thus the κ -statistic results will be slightly lower than a straightforward percentage of agreement. The κ -statistic was calculated for the 'research' study to be used as a guide to interpret percentage agreement in an academic sense.

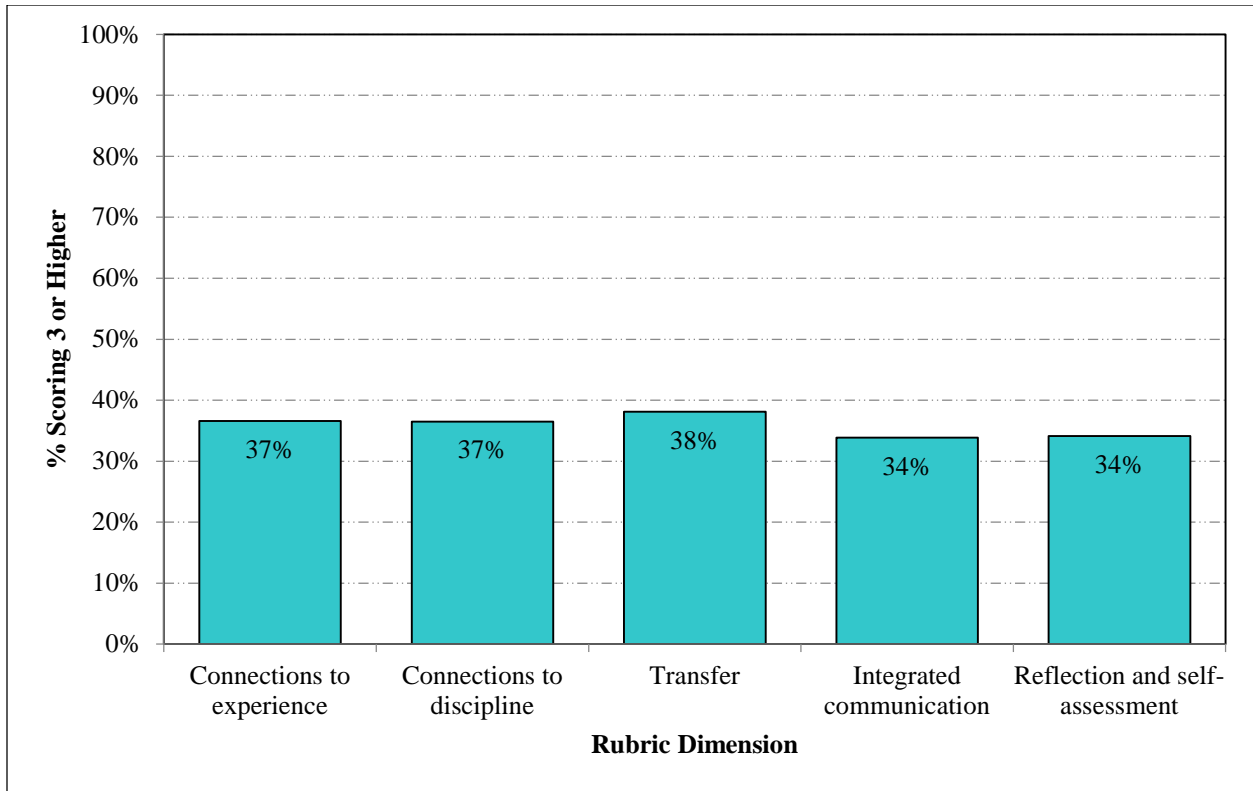


Figure 20. 'Investigate' achievement at 3 or higher across all rubric dimensions for AA courses only for 127 artifacts from 25 sampled course sections.

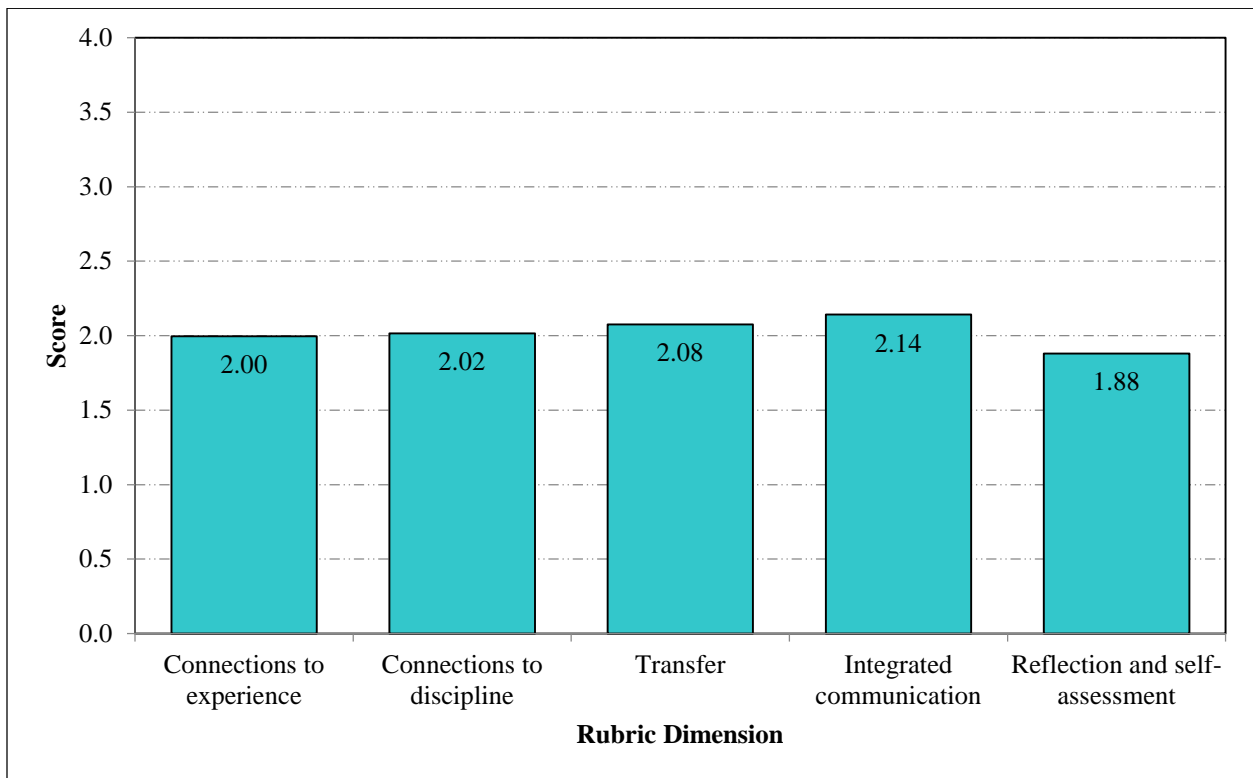


Figure 21. Mean score by rubric dimension for 'Investigate' for AA courses only for 127 artifacts from 25 sampled course sections.

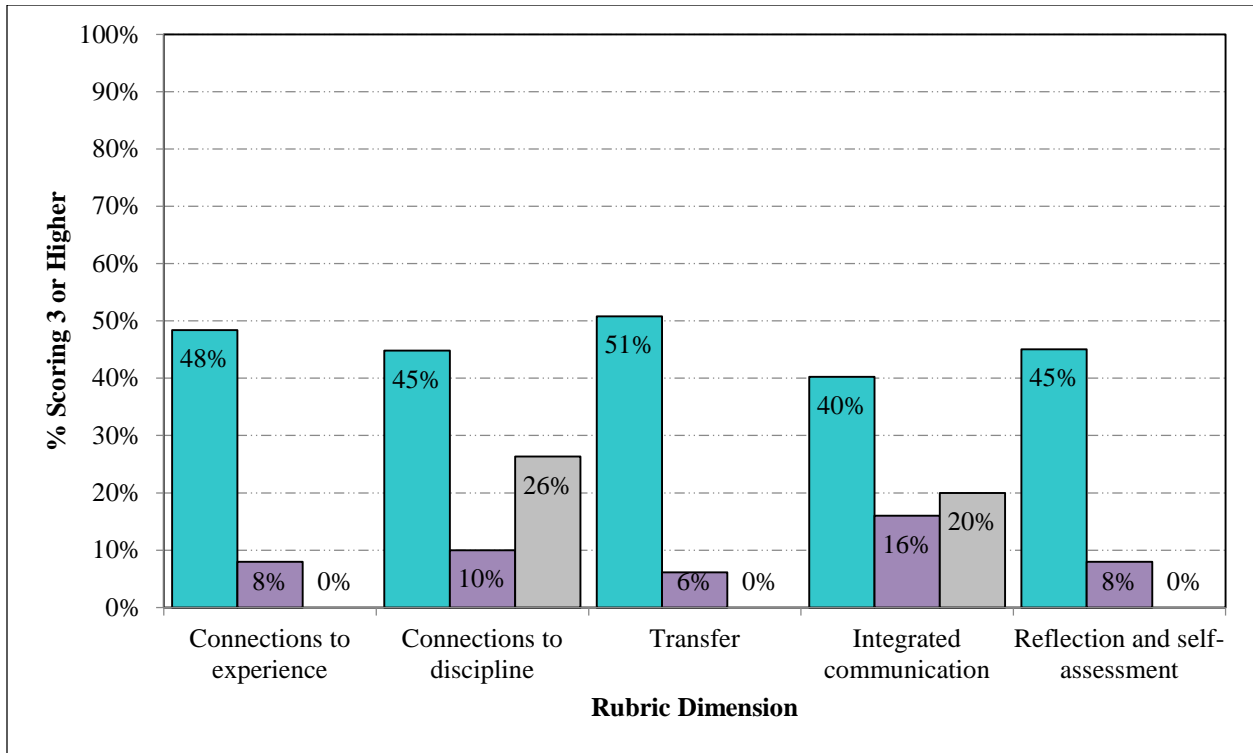


Figure 22. 'Investigate' achievement at 3 or higher across all rubric dimensions for AA courses only for 127 artifacts from 25 sampled course sections. Traditional (aqua), n=92, Online (purple), n=25, Dual Enrollment (concurrent) (gray), n=10.

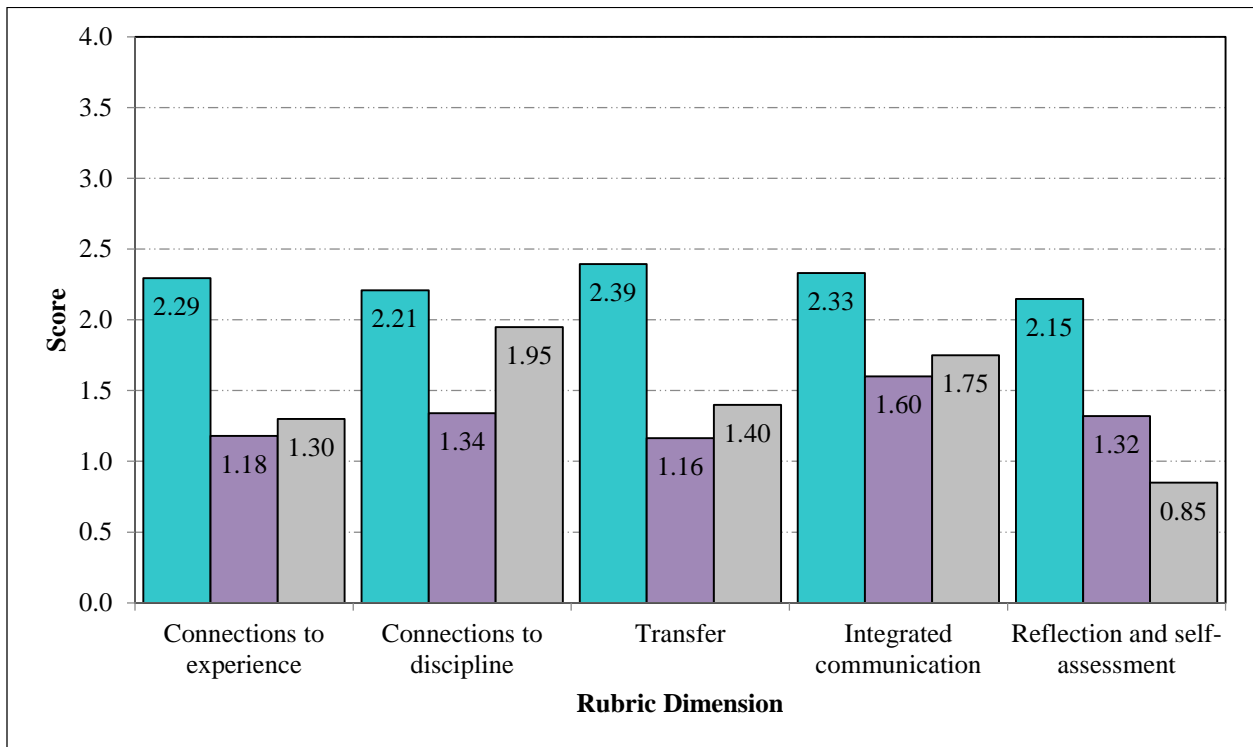


Figure 23. Mean score of 'Investigate' for each rubric dimension by modality at 3 or higher across all rubric dimensions for AA courses only for 127 artifacts from 25 sampled course sections. Traditional (aqua), n=92, Online (purple), n=25, Dual Enrollment (concurrent) (gray), n=10.

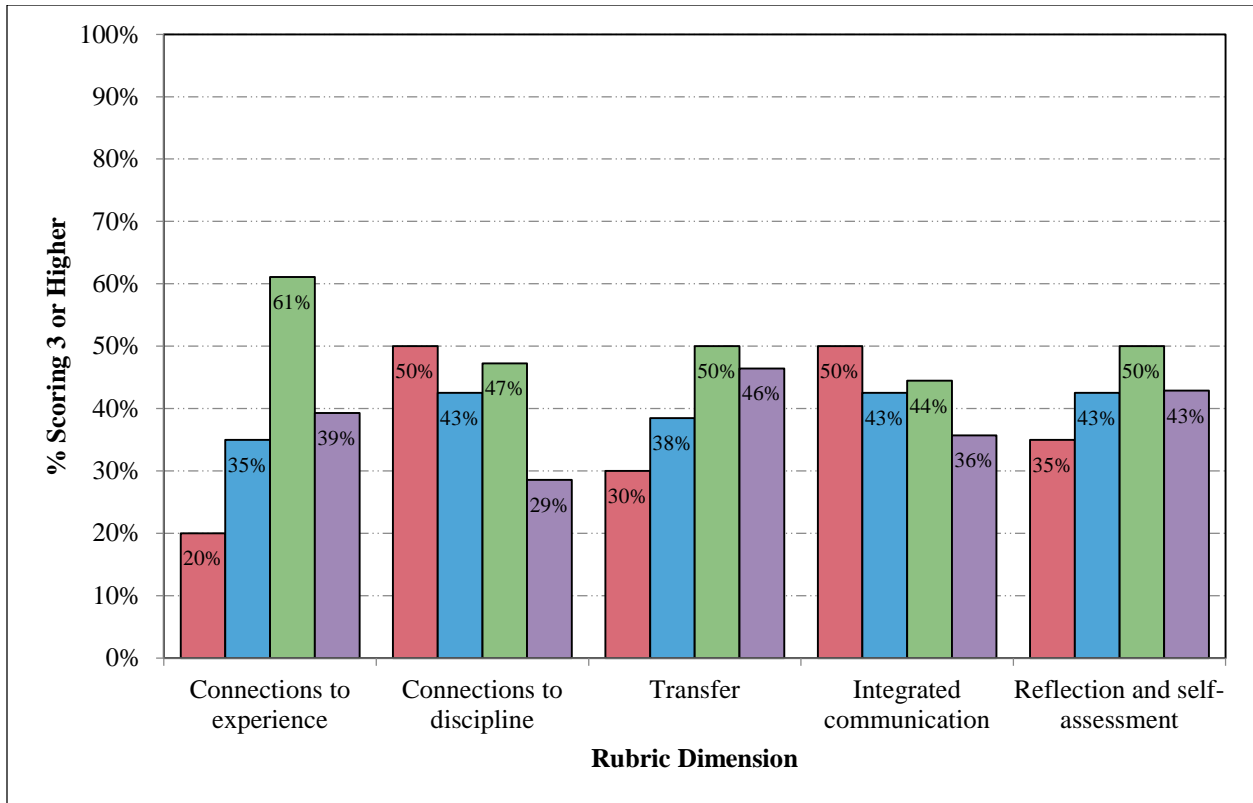


Figure 24. Comparison of 'Investigate' achievement at 3 or higher across all rubric dimensions for 108 artifacts in which credit information could be matched to artifact score. 0-15 credits earned (red) n=23, 16-30 credits earned (blue) n=28, 31-60 credits earned (green) n=23, > 60 credits earned (purple) n=34. *Credits earned based on number of credits earned entering fall 2017 term.

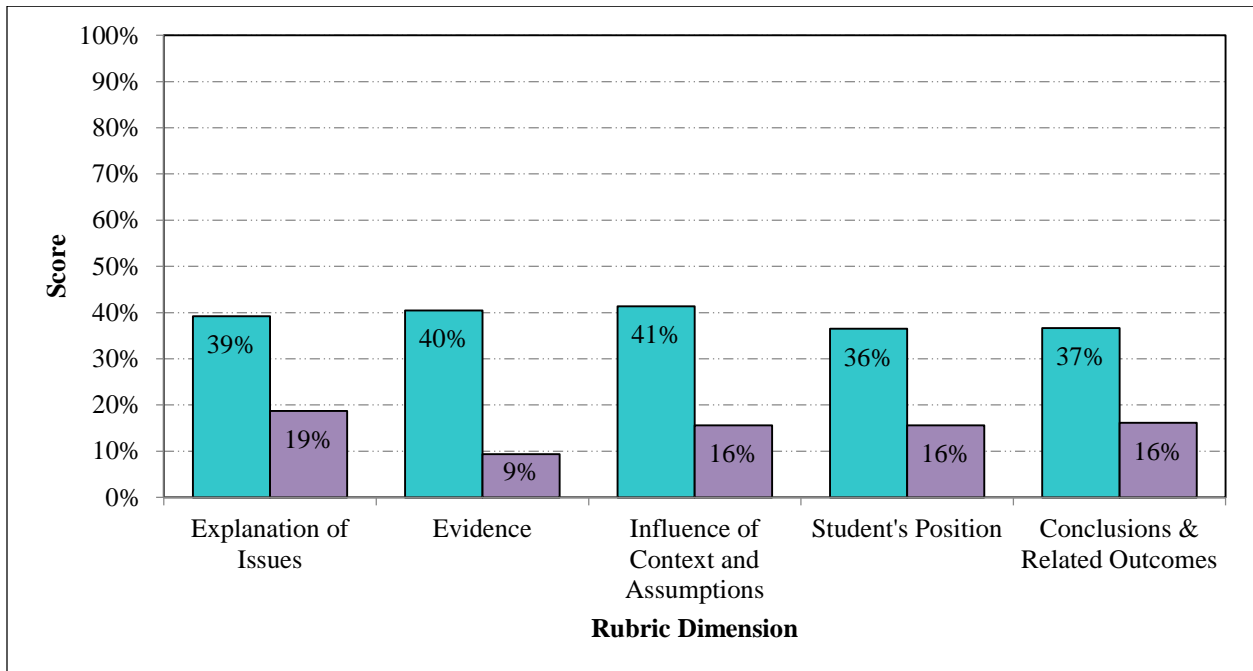


Figure 25. Comparison of 'Investigate' achievement at 3 or higher across all rubric dimensions based on courses with pre-requisites. Courses with pre-requisites (aqua, n=112), courses without pre-requisites (purple, n=131).

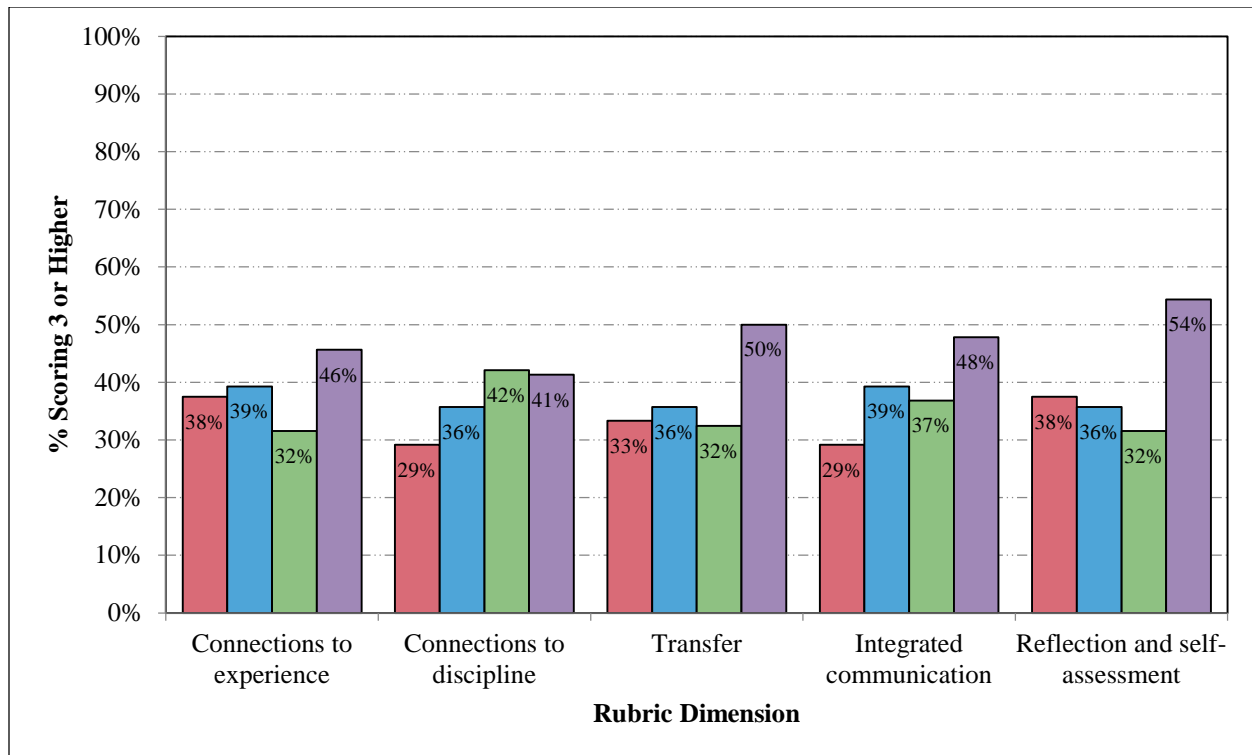


Figure 26. Comparison of 'Investigate' achievement at 3 or higher across all rubric dimensions based on GPA. GPA < 2.5 (red) n=17, GPA 2.5-3.0 (blue) n=24, GPA 3.0-3.5 (green) n=34, GPA ≥ 3.5 (purple) n=35. *Credits earned based on number of credits earned entering fall 2017 term.

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 27 and 28 provide these comparisons.

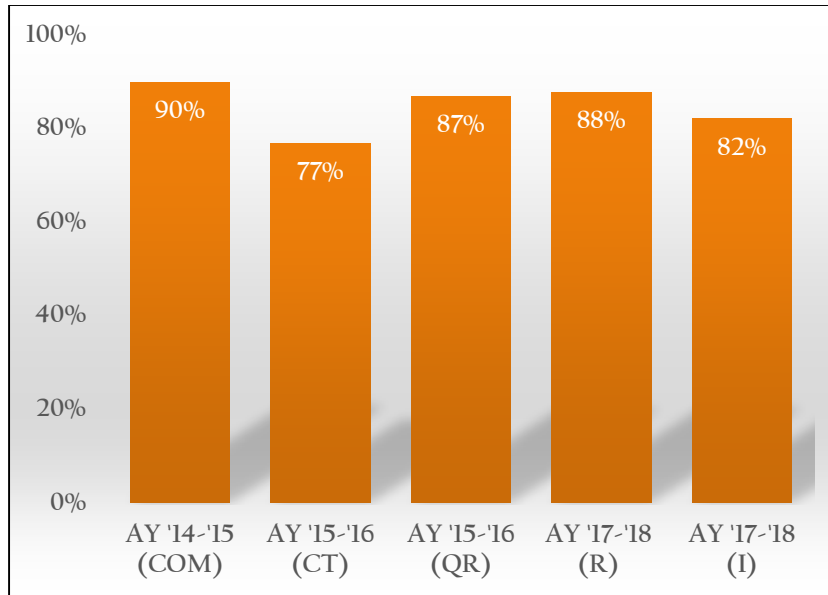


Figure 27. Comparison of inter-rater reliability (percentage (%) +/- 1 agreement) averaged across dimensions by each competency in FSW General Education Assessment cycle.

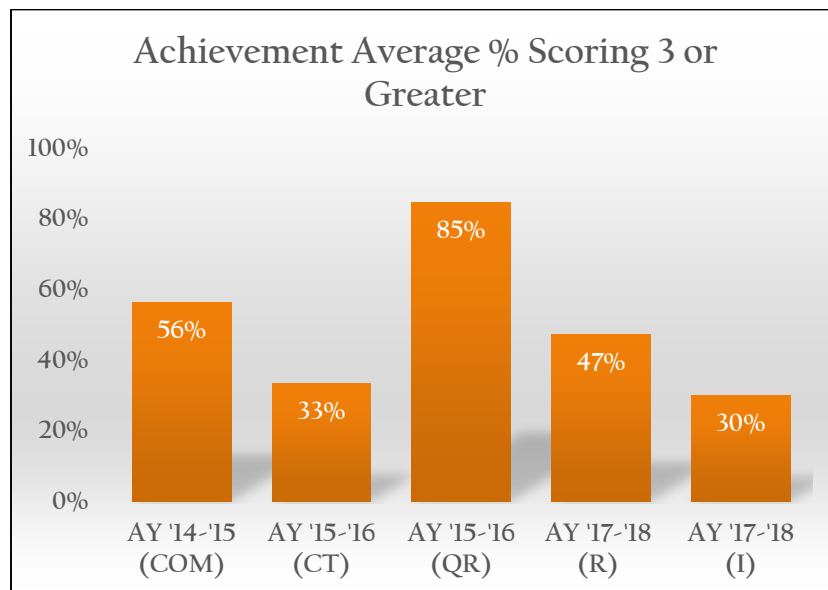


Figure 28. Comparison of achievement at 3 or higher averaged across dimensions by each competency in FSW General Education Assessment cycle.

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 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 2017-2018 which included the analysis of 'Research' and 'Investigate' 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, pre-requisites, and GPA.

A drilldown of 'Research' (R) results are as follows:

1. One of five rubric dimensions exhibit greater than 60% achievement at level '3'. The highest scored dimension is "Use information ethically and legally and accurately document their use of information sources" at 67% scoring '3' or higher.
2. Mean achievement levels for each of the five rubric dimensions of 'Research' range from 2.20 to 2.60 on a 4-point scale.
3. In a study comparing online, dual enrollment (concurrent), and traditional artifacts, the traditional modality exhibits the highest in all five dimensions. Results for three of five dimensions were statistically significantly different (traditional statistically significantly higher than both online or dual enrollment (concurrent)).
4. An inter-rater reliability study exhibits rubric scoring agreement ranging from 43% to 59% with a +/- 1 agreement ranging from 76% to 94%.
5. With respect to AA courses, one of five rubric dimensions exhibit greater than 60% achievement at level '3'. The highest scored dimension is "Use information ethically and legally and accurately document their use of information sources" at 67% 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 four of five dimensions. The dual enrollment modality exhibits the highest in the remaining dimension. There is a statistically significant difference in achievement between traditional and online in the "Identify and refine information needed" dimension, the "Use of information effectively to create and communicate their projects or performance" and "Use information ethically and legally and accurately document their use of information sources."

7. In a study comparing achievement at 3 or higher across rubric dimensions based on credits earned, achievement is higher for credits 0-30 compared with above 30 in four of five dimensions.
8. In a study comparing achievement at 3 or higher across rubric dimensions based on pre-requisites, in all five dimensions, artifacts stemming from courses which had pre-requisites exhibit a higher achievement.
9. In a study comparing achievement at 3 or higher based on GPA, achievement is lowest with the lowest GPA (< 2.5). Trends from other GPA cohorts (2.5-3, 3-3.5, ≥ 3.5) are less clear.
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). The 1st, 2nd, 3rd, and 5th dimensions were called out by different scorers. Scorers also noted that the KSC rubric addresses the communication area of research by asking scorers to reference specific dimensions of the college's separate communication competency rubric. Scorers felt that this was inconvenient and also could occasionally 'take the scorer out of focus' with the assignment and its purpose. Also, scorers commented positively regarding the rubric that dimensions did not appear too similar, and that there was little overlap across competencies.

A drilldown of 'Investigate' (I) results are as follows:

1. Zero of five rubric dimensions exhibit greater than 60% achievement at level '3' with percentages ranging from 28% to 31%.
2. Mean achievement levels for each of the five rubric dimensions of 'Investigate' range from 1.70 to 1.98 on a 4-point scale.
3. In a study comparing online, dual enrollment (concurrent), and traditional artifacts, the traditional modality exhibits the highest in all five dimensions. Differences from traditional to online are statistically significant in all dimensions.
4. An inter-rater reliability study exhibits rubric scoring agreement ranging from 32% to 35% with a +/- 1 agreement ranging from 71% to 86%.
5. With respect to AA courses, zero of five rubric dimensions exhibit greater than 60% achievement at level '3' with percentages ranging from 34% to 38%.
6. In a study comparing online, dual enrollment, and traditional artifacts, the traditional modality exhibits the highest in all five dimensions. The differences between traditional and online are statistically significantly different in all five dimensions.
7. In a study comparing achievement at 3 or higher across rubric dimensions based on credits earned, achievement is higher for credits 0-30 compared with above 30 in two of five dimensions.
8. In a study comparing achievement at 3 or higher across rubric dimensions based on pre-requisites, in all five dimensions, artifacts stemming from courses which had pre-requisites exhibit a higher achievement.
9. In a study comparing achievement at 3 or higher based on GPA, achievement is lowest with the lowest GPA (< 2.5) in three of five dimensions. Achievement is highest with the highest GPA (≥ 3.5) in four of five dimensions.
10. In a review of scorer feedback, multiple scorers noted poor alignment between the assignments being scored and the rubric dimensions. Also, scorers noted that the first and fifth rubric dimensions appear to cover some of the same ground. And lastly, very few assignments appear

to really get at the heart of the concept of transdisciplinary studies, a core component of the 'Investigate' competency. Assignments often made an attempt at addressing transdisciplinary thinking, which is not an easy task, but didn't fully achieve this goal.

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 'Research' and 'Investigate' exhibit results similar to those of past studies (88% and 82% compared with 77%, 87%, and 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 'Research' and 'Investigate' competencies exhibit the 3rd and 5th ranked achievement percentages of the past five assessments, respectively.

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.
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.