# General Education Assessment Report - AY 2020-21 

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

## 1 INTRODUCTION

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.

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 each assessment, courses are randomly sampled from a list of courses which were identified by faculty as encompassing that competency.

A complete list of the years for analysis of each competency is shown below:
> AY 2015-16: Communication*
o *most closely associated with Communicate in the current competencies
> AY 2016-17: Critical Thinking*, Quantitative Reasoning*
o *most closely associated with Think and Evaluate in the current competencies
$>$ AY 2017-18: Research, Investigate
$>$ AY 2018-19: Visualize, Engage
$>$ AY 2019-20: Analyze, Research
> AY 2020-21: Communicate, Evaluate
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 Communicate (C) - Written

The outcome of the 'Communicate' competency at FSW is that by completion of the general education requirements, students will be able to communicate clearly in a variety of modes and media. The LAC will measure the percentage of artifacts scored a 3 or higher on the individual dimensions of the FSWspecific rubric. Figures 1 through 15 below depict achievement and inter-rater reliability for the 'Communicate (Written)' competency in college-wide, Associate of Arts (AA) cohorts, as well as valueadded studies as they relate to outcome goals and objectives, and longitudinal studies. For the study, the LAC utilizes an FSW-specific rubric developed by a selection of faculty representing various areas at the college (Figure 1).

### 2.1 Overall Achievement, Modality Comparison Study, \& Inter-Rater Reliability

| COMMUNICATE (Written) | Capstone (4) | Accomplished (3) | Developing (2) | Deficient (1) |
| :---: | :---: | :---: | :---: | :---: |
| Context \& Purpose of Writing | Synthesizes purpose, audience, and context to fulfill the objectives of the assignment. | Integrates purpose, audience, and context to fulfill the objectives of the assignment. | Understands purpose, audience, and context to fulfill the objectives of the assignment. | Shows limited understanding of purpose, audience, and context to fulfill the objectives of the assignment. |
| Content Development | Adapts appropriate, relevant, and compelling content to structure the writer's work and convey meaning. | Summarizes appropriate, relevant, and compelling content to structure the work and convey meaning. | Includes appropriate, relevant, and compelling content to structure the work and convey meaning. | Does not include appropriate, relevant, and compelling content to structure the work and convey meaning. |
| Genre \& Disciplinary Conventions | Executes an extensive use of conventions particular to a specific discipline and/or writing assignment which may include organization, content, presentation, formatting, and stylistic choices. | Demonstrates use of conventions particular to a specific discipline and/or writing assignment which may include organization, content, presentation, formatting, and stylistic choices. | Recognizes conventions particular to a specific discipline and/or writing assignment which may include organization, content, presentation, formatting, and stylistic choices. | Shows limited or no recognition of conventions particular to a specific discipline and/or writing assignment which may include organization, content, presentation, formatting, and stylistic choices. |
| Sources \& Evidence | Identifies and interprets information sources that are authoritative to the assignment and appropriate to the discipline. | Identifies but does not interpret (or interprets but does not identify) information sources that are authoritative to the assignment and appropriate to the discipline. | Includes some sources but demonstrates only a limited understanding of sources that are authoritative to the assignment and appropriate to the discipline. | Includes limited to no sources that are authoritative to the assignment and appropriate to the discipline. |
| Control of Syntax \& Mechanics | Communicates meaning clearly, accurately, and precisely to meet the assignment with few to no grammatical, mechanical, and syntactical errors. | Communicates meaning clearly, accurately, precisely, but has some grammatical, mechanical, and/or syntactical errors. | Communicates meaning with many grammatical, mechanical, and/or syntactical errors that affect clarity, accuracy, and/or precision. | Communicates with many grammatical, mechanical, and/or syntactical errors that impedes meaning. |

Figure 1. FSW-specific 'Communicate (Written)' rubric utilized in the study.


Figure 2. 'Communicate (Written)' achievement at 3 or higher across all rubric dimensions for 165 artifacts from 24 sampled course sections.


Figure 3. Mean score by rubric dimension for 'Communicate (Written)' for 165 artifacts from 24 sampled course sections.


Figure 4. Comparison of 'Communicate (Written)' achievement by modality at 3 or higher across all rubric dimensions for 165 artifacts from 24 sampled course sections. Traditional (red), $n=51$, Online (green), $n=82$, Dual Enrollment (concurrent) (beige), $n=27$.


Figure 5. Mean score of 'Communicate (Written)' for each rubric dimension by modality across all rubric dimensions for 165 artifacts from 24 sampled course sections. Traditional (red), $n=51$, Online (green), $n=82$, Dual Enrollment (concurrent) (beige), $n=27$.


Figure 6. Inter-rater reliability (as \%) for the 'Communicate (Written)' 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. k-statistic for the study exhibits similar results. Results are herein presented as percentages for reader convenience.
2.2 Results for A.A. General Studies Only


Figure 7. 'Communicate' achievement at 3 or higher across all rubric dimensions for $A A$ courses only for 132 artifacts from 18 sampled course sections.


Figure 8. Mean score by rubric dimension for 'Communicate (Written)' for AA courses only for 132 artifacts from 18 sampled course sections.


Figure 9. 'Communicate (Written)' achievement at 3 or higher across all rubric dimensions for AA courses only for 132 artifacts from 18 sampled course sections. Traditional (red), $n=38$, Online (green), $n=64$, Dual Enrollment (concurrent) (beige), $n=27$.


Figure 10. Mean score of 'Communicate (Written)' for each rubric dimension by modality across all rubric dimensions for $A A$ courses only for 132 artifacts from 18 sampled course sections. Traditional (red), $n=38$, Online (green), n=64, Dual Enrollment (concurrent) (beige), $n=27$.

### 2.3 Overall Value-added Studies



Figure 11. Comparison of mean score of 'Communicate (Written)' across all rubric dimensions for 165 artifacts in which credit information could be matched to artifact score. From light purple to dark, $0-15$ credits earned $n=98,16-30$ credits earned $n=14$, $31-45$ credits earned $n=11,46-60$ credits earned $n=9$, and $>60$ credits earned $n=33$. *Credits earned based on number of credits earned entering fall 2020 term.


Figure 12. Comparison of mean score of 'Communicate (Written)' across all rubric dimensions based on course success rates of students. From light beige to dark, students with $0-59 \% n=6,60-69 \% n=3,70-79 \% n=13,80-89 \% n=15$, and $90 \%$ or above $n=49$. *Note that inbound students would not have a success rate at FSW yet, which therefore limits sample size from the overall sample.


Figure 13. Comparison of mean score of 'Communicate (Written)' across all rubric dimensions based on GPA. From light green to dark, GPA < $2.0 n=6$, GPA 2.0-2.4 $n=15, G P A ~ 2.5-2.9 n=37, ~ G P A ~ 3.0-3.4 n=42, G P A \geq 3.5 n=40$. *GPA based on fall inbound GPA; first-time students would therefore not have an inbound FSW GPA, which limits sample size from the overall.


Figure 14. Comparison of mean score of 'Communicate (Written)' across all rubric dimensions based on First-Time-In-College (FTIC) status.

### 2.4 LONGITUDINAL STUDY

The AY 2020-21 assessment of the 'Communicate' competency is the first time this competency is measured using the new FSW-specific rubric. The school representatives used the Association of American Colleges and Universities (AAC \& U) Integrated Learning VALUE Rubric as a foundation for development ultimately adopting only the dimensions (in part) and achievement levels (4-3-2-1) with a 0 if no achievement is met. The rubric defines the fundamental criteria for each learning outcome and outline performance required to demonstrate levels of attainment through the use of Bloom's Taxonomic verbiage. Rubric achievement levels, in descending order: Capstone (4), Accomplished (3), Developing (2), and Deficient (1).

The intent of the rubric developers was to frame language such that the rubric is as inclusive as possible to all 'Communicate (Written)' assignments. Careful consideration was paid to providing descriptors detailed enough to score an artifact, yet to remain in general terms as much as possible to allow for application to a wide assortment of assignment types and styles. In order to increase clarity, action verbs were utilized in each achievement level description. The developers also attempted to place emphasis on dimensions being mutually exclusive, such that users of this rubric can elect to omit any dimension not required of a given assignment. To ensure that non-traditional assignments are scored properly, artifacts representing a variety of modes and media should be utilized during the 'Communicate (Written)' Rubric Calibration Sessions prior to the scoring process.

Because the AAC \& U VALUE Rubric was utilized as a foundation for the FSW-specific rubric, the rubric dimensions, while re-worked, are foundationally similar. To exploit this characteristic, this comparison study compares the results by dimension from the AY 2014-15 initial pilot study to the current AY 20202021 study.


Figure 15. Comparison of assessment of the Communicate (Written) competency over time. The AY 20-21 assessment is the first to utilize the FSW-specific rubric which utilized common dimensions to those in previous studies.

## 3 Communicate (C) - Oral

The outcome of the 'Communicate' competency at FSW is that by completion of the general education requirements, students will be able to communicate clearly in a variety of modes and media. The LAC will measure the percentage of artifacts scored a 3 or higher on the individual dimensions of the FSWspecific rubric. Figures 16 through 30 below depict achievement and inter-rater reliability for the 'Communicate (Written)' competency in college-wide, Associate of Arts (AA) cohorts, as well as valueadded studies as they relate to outcome goals and objectives, and longitudinal studies. For the study, the LAC utilizes an FSW-specific rubric developed by a selection of faculty representing various areas at the college (Figure 16). This is the first time the 'Communicate (Oral)' FSW-specific rubric was utilized for competency achievement assessment.

### 3.1 Overall Achievement, Modality Comparison Study, \& Inter-Rater Reliability

| COMMUNICATE (ORAL) | Capstone (4) | Accomplished (3) | Developing (2) | Deficient (1) |
| :---: | :---: | :---: | :---: | :---: |
| Introduction | Engages audience immediately and includes clear and inviting statement of purpose. | Acknowledges audience and includes statement of purpose. | Exhibits limited acknowledgment of audience and unclear statement of purpose. | Exhibits minimal audience acknowledgment and includes no statement of purpose. |
| Structure \& Transition | Organizes the speech appropriately and logically with effective transitions. | Organizes the speech with some internal logic and some effective transitions. | Organizes the speech with partial internal logic or partially effective transition. | Organizes the speech with minimal internal logic and transitions. |
| Supporting Evidence, Documentation, \& Presentation Media | Substantiates content with relevant credible support and/or presentation media, with the incorporation of oral citations if required by the assignment. | Substantiates content with some relevant and credible support and/or presentation media, and includes some oral citations if required by the assignment. | Substantiates the content with minimal support and/or presentation media with inconsistent oral citations as required by the assignment. | Includes minimal to no support and/or presentation media, and does not include oral citations if required by the assignment. |
| Delivery | Uses gestures, eye contact, vivid language, and voice effectively to add interest to speech with no oral fillers or non-verbal distractions. | Uses adequate gestures, eye contact, and language to add interest to presentation with minor reliance on notes and limited fillers and non-verbal distractions. | Uses ineffective eye contact, gestures, language, and voice, with heavy reliance on notes, multiple oral fillers, and non-verbal distractions. | Lacks appropriate gestures, eye contact, or voice and reads from notes only, with excessive use of oral fillers and non-verbal distraction. |
| Conclusion | Summarizes the main points and purpose and brings the presentation to a logical end. | Summarizes some of the main points and purpose and brings the presentation to an end. | Summarizes some of the main points or purpose bringing the presentation to an awkward end. | Ends presentation abruptly with no reference to main points or purpose. |

Figure 16. FSW-specific 'Communicate (Oral)' rubric utilized in the study.


Figure 17. 'Communicate (Oral)' achievement at 3 or higher across all rubric dimensions for 31 artifacts from 10 sampled course sections.


Figure 18. Mean score by rubric dimension for 'Communicate (Oral)' for 31 artifacts from 10 sampled course sections.


Figure 19. Comparison of 'Communicate (Oral)' achievement by modality at 3 or higher across all rubric dimensions for 31 artifacts from 10 sampled course sections. Traditional (red), $n=4$, Online (green), $n=0$, Dual Enrollment (concurrent) (gray), $n=27$.


Figure 20. Mean score of 'Communicate (Oral)' for each rubric dimension by modality across all rubric dimensions for 31 artifacts from 10 sampled course sections. Traditional (red), $n=4$, Online (green), $n=27$, Dual Enrollment (concurrent) (gray), $n=0$.


Figure 21. Inter-rater reliability (as \%) for the 'Communicate (Oral)' 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. k-statistic for the study exhibits similar results. Results are herein presented as percentages for reader convenience.

### 3.2 Results for A.A. General Studies Only



Figure 22. 'Communicate (Oral)' achievement at 3 or higher across all rubric dimensions for AA courses only for 31 artifacts from 10 sampled course sections.


Figure 23. Mean score by rubric dimension for 'Communicate (Oral)' for AA courses only for 31 artifacts from 10 sampled course sections.


Figure 24. 'Communicate (Oral)' achievement at 3 or higher across all rubric dimensions for AA courses only for 31 artifacts from 10 sampled course sections. Traditional (red), $n=4$, Online (green), $n=27$, Dual Enrollment (concurrent) (beige), $n=0$.


Figure 25. Mean score of 'Communicate (Oral)' for each rubric dimension by modality across all rubric dimensions for AA courses only for 31 artifacts from 10 sampled course sections. Traditional (red), $n=4$, Online (green), $n=27$, Dual Enrollment (concurrent) (beige), $n=0$.

### 3.3 Overall Value-added Studies



Figure 26. Comparison of mean score of 'Communicate (Oral)' across all rubric dimensions for 31 artifacts in which credit information could be matched to artifact score. From light purple to dark, 0-15 credits earned n=2, 16-30 credits earned n=14, $31-45$ credits earned $n=7,46-60$ credits earned $n=7$, and $>60$ credits earned $n=1$. *Credits earned based on number of credits earned entering fall 2020 term.


Figure 27. Comparison of mean score of 'Communicate (Oral)' across all rubric dimensions based on course success rates of students. From light beige to dark, students with $0-59 \% n=1,60-69 \% n=1,70-79 \% n=6,80-89 \% n=5$, and $90 \%$ or above $n=18$. *Note that inbound students would not have a success rate at FSW yet, which therefore limits sample size from the overall sample.


Figure 28. Comparison of mean score of 'Research' across all rubric dimensions based on GPA. From light green to dark, GPA < $2.0 n=0, G P A 2.0-2.4 n=7, G P A 2.5-2.9 n=6, G P A 3.0-3.4 n=7, G P A \geq 3.5 n=11$. *GPA based on fall inbound GPA; first-time students would therefore not have an inbound FSW GPA, which limits sample size from the overall.


Figure 29. Comparison of mean score of 'Communicate (Oral)' across all rubric dimensions based on First-Time-In-College (FTIC) status.

### 3.4 LONGITUDINAL STUDY

The AY 2020-21 assessment of the 'Communicate' competency is the first time this competency is measured using the new FSW-specific rubric. The school representatives used the Association of American Colleges and Universities (AAC \& U) Integrated Learning VALUE Rubric as a foundation for development ultimately adopting only the dimensions (in part) and achievement levels (4-3-2-1) with a 0 if no achievement is met.

The intent of the rubric developers was to frame language such that the rubric is as inclusive as possible to all 'Communicate (Oral)' assignments. Careful consideration was paid to providing descriptors detailed enough to score an artifact, yet to remain in general terms as much as possible to allow for application to a wide assortment of assignment types and styles. To increase clarity, action verbs were utilized in each achievement level description. The developers also attempted to place emphasis on dimensions being mutually exclusive, such that users of this rubric can elect to omit any dimension not required of a given assignment. To ensure that non-traditional assignments are scored properly, artifacts representing a variety of modes and media should be utilized during the 'Communicate (Oral)' Rubric Calibration Sessions prior to the scoring process.

Because the AAC \& U VALUE Rubric was utilized as a foundation for the FSW-specific rubric, the rubric dimensions, while re-worked, are foundationally similar. To exploit this characteristic, this comparison study compares the results by dimension from the AY 2015-16 initial pilot study to the current AY 20202021 study.


Figure 30. Comparison of assessment of the Communicate (Oral) competency over time. The AY 20-21 assessment is the first to utilize the FSW-specific rubric which utilized most common dimensions to those in previous studies.

## 4 Evaluate (Ev)

The outcome of the 'Evaluate' competency at FSW is that by completion of the general education requirements, students will be able to evaluate and utilize mathematical principles, technology, scientific and quantitative data. The FSW Learning Assessment Committee will measure the number of artifacts scored a 3 or higher on relevant dimensions of the rubric against results from previous studies. Figures 31 through 45 below depict achievement and inter-rater reliability for the 'Evaluate' competency in college-wide, Associate of Arts (AA) cohorts, as well as value-added studies as they relate to outcome goals and objectives, and longitudinal studies.

### 4.1 Overall Achievement, Modality Comparison Study, \& Inter-Rater Reliability

| Evaluate | Capstone (4) | Accomplished (3) | Developing (2) | Deficient (1) |
| :---: | :---: | :---: | :---: | :---: |
| Comprehension | Demonstrates purposeful and accurate understanding of the problem. | Demonstrates an understanding of the problem, but it is only partially accurate or insufficiently complex. | Demonstrates an understanding of the problem, but it is inaccurate and insufficiently complex. | Does not show a sufficient understanding of the problem. |
| Interpretation / Representation | Provides accurate explanation and/or conversion of relevant information. | Provides an explanation and/or conversion of relevant information with limited errors. | Provides an explanation and/or conversion of relevant information with many errors. | Provides an explanation and/or conversion of relevant information with excessive errors. |
| Application | Applies thorough and quantified judgment(s) and/or applies results in an insightful manner. | Applies quantified judgment(s) and/or applies results in an adequate manner. | Applies poorly quantified judgment(s) and/or applies results in an inadequate manner. | Does not apply quantified judgment(s) and/or does not apply results in a correct manner. |
| Inference / <br> Conclusion | Draws accurate inferences from results that exhibit awareness of supporting evidence. | Draws mostly accurate inferences from results that exhibit awareness of supporting evidence. | Draws partially accurate inferences from results that exhibit awareness of supporting evidence. | Draws inaccurate inferences from results that exhibit a lack of awareness of supporting evidence. |
| Communication | Establishes and promulgates results in a clear framework that is based on the purpose of the assignment. | Establishes results in a framework that is based on the purpose of the assignment. | Establishes results in a partially constructed framework that is based on the purpose of the assignment. | Establishes results in framework that is minimally based on the purpose of the assignment. |

Figure 31. FSW-specific 'Evaluate' rubric utilized in the study.


Figure 32. 'Evaluate' achievement at 3 or higher across all rubric dimensions for 146 artifacts from 20 sampled course sections.


Figure 33. Mean score by rubric dimension for 'Evaluate' for 146 artifacts from 20 sampled course sections.


Figure 34. Comparison of 'Evaluate' achievement by modality at 3 or higher across all rubric dimensions for 146 artifacts from 20 sampled course sections. Traditional (red), $n=57$, Online (green), $n=67$, Dual Enrollment (concurrent) (beige), $n=18$.


Figure 35. Mean score of 'Evaluate' for each rubric dimension by modality at 3 or higher across all rubric dimensions for 146 artifacts from 20 sampled course sections. Traditional (red), $n=57$, Online (green), $n=67$, Dual Enrollment (concurrent) (beige), $n=18$.


Figure 36. Inter-rater reliability (as \%) for the 'Evaluate' 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. $k$-statistic for the study exhibits similar results. Results are herein presented as percentages for reader convenience.
4.2 Results for A.A. General Studies Only


Figure 37. 'Evaluate' achievement at 3 or higher across all rubric dimensions for $A A$ courses only for 138 artifacts from 19 sampled course sections.


Figure 38. Mean score by rubric dimension for 'Evaluate' for AA courses only for 138 artifacts from 19 sampled course sections.


Figure 39. 'Evaluate' achievement at 3 or higher across all rubric dimensions for AA courses only for 138 artifacts from 19 sampled course sections. Traditional (red), $n=53$, Online (green), $n=67$, Dual Enrollment (concurrent) (beige), $n=18$.


Figure 40. Mean score of 'Evaluate' for each rubric dimension by modality across all rubric dimensions for AA courses only for 138 artifacts from 19 sampled course sections. Traditional (red), $n=53$, Online (green), $n=67$, Dual Enrollment (concurrent) (beige), $n=18$.

### 4.3 OVERALL Value-Added Studies



Figure 41. Comparison of mean score of 'Evaluate' across all rubric dimensions for 146 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 $n=59,16-30$ credits earned $n=20,31-45$ credits earned $n=18,46-60$ credits earned $n=13$, and $>60$ credits earned $n=18$. *Credits earned based on number of credits earned entering fall term.


Figure 42. Comparison of mean score of 'Evaluate' across all rubric dimensions based on course success rates of students. From light beige to dark, students with $0-59 \% n=6,60-69 \% n=5,70-79 \% n=11,80-89 \% n=12$, and $90 \%$ or above $n=61$. *Note that inbound students would not have a success rate at FSW yet, which therefore limits sample size from the overall sample.


Figure 43. Comparison of mean score of 'Evaluate' achievement across all rubric dimensions based on GPA. From light green to dark, GPA < $2.0 n=10, G P A$ 2.0-2.4 $n=11, G P A 2.5-2.9 n=30, G P A$ 3.0-3.4 $n=40, G P A \geq 3.5 n=37$. *GPA based on fall inbound GPA; first-time students would therefore not have an inbound FSW GPA, which limits sample size from the overall.


Figure 44. Comparison of mean score of 'Communicate (Oral)' across all rubric dimensions based on First-Time-In-College (FTIC) status.

### 4.4 LONGITUDINAL STUDY

The AY 2020-21 assessment of the 'Evaluate' competency is the first time this competency is measured using the new FSW-specific rubric. The school representatives used a previous FSW-specific rubric as a foundation for development ultimately adopting only the dimensions (in part) and achievement levels (4-3-2-1) with a 0 if no achievement is met.

The intent of the rubric developers was to frame language such that the rubric is as inclusive as possible to all 'Evaluate' assignments. Careful consideration was paid to providing descriptors detailed enough to score an artifact, yet to remain in general terms as much as possible to allow for application to a wide assortment of assignment types and styles. To increase clarity, action verbs were utilized in each achievement level description. The developers also attempted to place emphasis on dimensions being mutually exclusive, such that users of this rubric can elect to omit any dimension not required of a given assignment. To ensure that non-traditional assignments are scored properly, artifacts representing a variety of modes and media should be utilized during the 'Evaluate' Rubric Calibration Sessions prior to the scoring process.

Because the former FSW-specific rubric was utilized as a foundation for the FSW-specific rubric, the rubric dimensions, while re-worked, are foundationally similar. To exploit this characteristic, this comparison study compares the results by dimension to studies from previous years.


Figure 45. Comparison of assessment of the 'Evaluate' competency over time. The AY 20-21 assessment is the first to utilize the new FSW-specific rubric which utilized most common dimensions to those in previous studies.

## 5 Competency Overview: 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. Because a true longitudinal study is limited due to a transition in both competencies and rubrics utilized, 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 46 and 47 provide these comparisons.


Figure 46. Comparison of inter-rater reliability (percentage (\%) +/- 1 agreement) averaged across dimensions by each competency in FSW General Education Assessment cycle.


Figure 47. Comparison of achievement at 3 or higher averaged across dimensions by each competency in FSW General Education Assessment cycle.

## 6 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.

In a review of this year's assessment and the professional development associated with it, several highlights are notable. (1) All rubrics are now available on the Assessment webpage (https://www.fsw.edu/facultystaff/assessment/genedcompetencies) and also in Canvas for any faculty to easily embend in their ow course pages for use with their assignments. (2) There are large differences in achievement between modalities. It is possible that the challenge of developing / transitioning courses to online in response to the pandemic may be related as a potential contribution to achievement levels. The challenge of introducing new modalities institution-wide in the Fall 2020 term might also be at play. (3) In general, feedback on the rubrics during use was very positive. During a Learning Assessment Committee meeting on May 3, 2021, it was determined that the only challenge is ensuring there are enough course experts to guide scorers. As a result, the committee developed a plan to have a 'lead' scorer as a representative discipline expert for scorers to consult, if necessary. (4)

General awareness and discussion of results is always important, so workshops on the competencies scored in AY 2020-21 are planned for AY 2021-22. (5) Finally, continued development of 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.

## 7 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 2020-2021 which included the analysis of 'Communicate (Written) and (Oral)' and 'Evaluate' from the 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 as well as longitudinal studies.

A drilldown of 'Communicate (Written)' (C) results are as follows:

1. Five of five rubric dimensions exhibit greater than $60 \%$ achievement at level ' 3 '. The highest scored dimension is "Documentation of Sources" at $60 \%$ scoring ' 3 ' or higher.
2. Mean achievement levels for each of the five rubric dimensions range from 2.83 to 3.12 on a 4point scale.
3. In a study comparing online, dual enrollment (concurrent), and traditional artifacts, the traditional modality exhibits the highest in 5 of 5 dimensions. Results for 3 of 5 dimensions were statistically significantly different.
4. An inter-rater reliability study exhibits rubric scoring agreement ranging from $40 \%$ to $53 \%$ with a +/- 1 agreement ranging from $89 \%$ to $94 \%$.
5. With respect to AA courses, five of five rubric dimensions exhibit greater than $60 \%$ achievement at level ' 3 '. The highest scored dimension is "Context \& Purpose of Writing" at $73 \%$ 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 4 of 5 dimensions. Concurrent dual enrollment exhibits the highest in 1 of 5 . Online exhibits the highest in 0 of 5 . Results for 3 of 5 dimensions were statistically significantly different.
7. In a study comparing achievement across rubric dimensions based on credits earned, in four of five dimensions, achievement generally increases with increasing credits earned. This is not as clear in the 'Sources and Evidence' dimension, although > 30 credits and higher all exhibit higher achievement than < 30 .
8. In a study comparing achievement across rubric dimensions based on success rates of students, achievement tends to increase with increasing success rates with some scatter to the data as a result of limited sample size.
9. In a study comparing achievement based on GPA, achievement increases with increasing GPA in all dimensions, although the increases tend to level off at higher GPA ranges.
10. In a study comparing achievement based on whether the student is a First-Time-In-College (FTIC) student, in all dimensions, FTIC students exhibit lower achievement than those that are non-FTIC.
11. In a longitudinal study, improvement is exhibited in 5 of 5 dimensions from the previous study of the competency.
A drilldown of 'Communicate (Oral)' (C) results are as follows:
12. Five of five rubric dimensions exhibit greater than $60 \%$ achievement at level ' 3 '. The highest scored dimension is "Documentation of Sources" at $60 \%$ scoring ' 3 ' or higher.
13. Mean achievement levels for each of the five rubric dimensions range from 3.03 to 3.58 on a 4point scale.
14. In a study comparing online, dual enrollment (concurrent), and traditional artifacts, the traditional modality exhibits the highest in 2 of 5 dimensions. The online modality exhibits the highest in 3 of 5 . The dual enrollment (concurrent) did not have any sample data. Results for 0 of 5 dimensions were statistically significantly different.
15. An inter-rater reliability study exhibits rubric scoring agreement ranging from $40 \%$ to $53 \%$ with a +/- 1 agreement ranging from $81 \%$ to $97 \%$.
16. With respect to AA courses, five of five rubric dimensions exhibit greater than $60 \%$ achievement at level ' 3 '. The highest scored dimension is "Structure \& Transition" at $97 \%$ scoring ' 3 ' or higher.
17. In a study comparing AA courses with online, dual enrollment, and traditional artifacts, the traditional modality exhibits the highest in 2 of 5 dimensions. The online modality exhibits the highest in 3 of 5 . The dual enrollment (concurrent) did not have any sample data. Results for 0 of 5 dimensions were statistically significantly different.
18. In a study comparing achievement across rubric dimensions based on credits earned, there is not any discernible trend across the five dimensions. In some cases, achievement declines with increasing credits and then rebounds. In others, achievement increases before declining again.
19. In a study comparing achievement across rubric dimensions based on success rates of students, achievement is again inconsistent similar to that of the credits earned study.
20. In a study comparing achievement based on GPA, achievement is again unclear with increasing GPA. In some cases there is no difference, whereas in others, declines exist.
21. In a study comparing achievement based on whether the student is a First-Time-In-College (FTIC) student, in all dimensions, FTIC students exhibit lower achievement than those that are non-FTIC.
22. In a longitudinal study, improvement is exhibited in all common dimension from the previous study of the competency.

A drilldown of 'Evaluate’ (Ev) results are as follows:

1. Five of five rubric dimensions exhibit greater than $60 \%$ achievement at level ' 3 ' with percentages ranging from $64 \%$ to $77 \%$.
2. Mean achievement levels for each of the five rubric dimensions range from 2.86 to 3.06 on a 4point scale.
3. In a study comparing online, dual enrollment (concurrent), and traditional artifacts, the concurrent modality exhibits the highest in 3 of 5 dimensions. Traditional artifacts exhibit the highest in 2 of 5 dimensions. Results for 3 of 5 dimensions were statistically significantly different.
4. An inter-rater reliability study exhibits rubric scoring agreement ranging from $48 \%$ to $62 \%$ with a +/- 1 agreement ranging from $88 \%$ to $97 \%$.
5. With respect to AA courses, five of five rubric dimensions exhibit greater than $60 \%$ achievement at level ' 3 ' with percentages ranging from $63 \%$ to $76 \%$.
6. In a study comparing online, dual enrollment, and traditional artifacts, the dual enrollment (concurrent), and traditional artifacts, the concurrent modality exhibits the highest in 3 of 5 dimensions. Traditional artifacts exhibit the highest in 2 of 5 dimensions. Results for 3 of 5 dimensions were statistically significantly different.
7. In a study comparing achievement across rubric dimensions based on credits earned, achievement exhibits no discernible difference with increasing credits earned.
8. In a study comparing achievement across rubric dimensions based on success rates of students, achievement exhibits no discernible difference with increasing success rates.
9. In a study comparing achievement at 3 or higher based on GPA, achievement exhibits no discernible difference with increasing GPA.
10. In a study comparing achievement based on whether the student is a First-Time-In-College (FTIC) student, in all dimensions, FTIC students exhibit lower achievement than those that are non-FTIC.
11. In a longitudinal study, improvement is exhibited in none of the common dimension from the previous study of the competency.

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 'Communicate' and 'Evaluate' exhibit results similar to those of past studies ( $92 \%$ and $93 \%$ compared with a range of $90 \%$ to $87 \%$ 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 'Communicate' and 'Evaluate' studies of AY 2020-21 exhibit the $3^{\text {rd }}$ and $2^{\text {nd }}$ highest achievement percentages of the past 11 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. The plan began in AY 2018-2019 and continues in AY 2020-2021.
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).
3. In general, feedback on the rubrics during use was very positive. During a Learning Assessment Committee meeting on May 3,2021 , it was determined that the only challenge is ensuring there are enough course experts to guide scorers. As a result, the committee developed a plan to have a 'lead' scorer as a representative discipline expert for scorers to consult, if necessary.

## 8 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?fDocumentld=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.

