UNIVERSITY OF TEXAS ARLINGTON

INSTITUTIONAL EFFECTIVENESS AND REPORTING

MEASURING WRITTEN COMMUNICATION USING AAC&U VALUE RUBRICS AT THE UNIVERSITY OF TEXAS ARLINGTON

Spring 2024 Report

Measuring Written Communication, Spring 2024 Report

Written communication is a fundamental skill that enables individuals to effectively convey ideas, arguments, and information across various contexts, including academic, professional, and personal settings (Graham, 2018). It is a critical component of higher education, as it enhances students' ability to structure coherent arguments, engage in critical analysis, and adapt their writing to different audiences and purposes (Bean, 2011). Recognizing its importance, the Texas Higher Education Coordinating Board (THECB) has designated Written Communication as one of the six core objectives within the Texas Core Curriculum (TCC), ensuring that undergraduate students develop proficiency in expressing ideas clearly, concisely, and effectively (THECB, 2019). This objective is integrated across multiple Foundational Component Areas (FCAs), reinforcing the role of writing in supporting academic and professional success.

At The University of Texas at Arlington (UTA), the assessment of Written Communication is conducted as part of a structured, multi-year evaluation cycle that measures student proficiency across all six TCC core objectives. This systematic approach aims to provide consistent, reliable data to inform curricular improvements while minimizing the assessment burden on academic departments. The assessment process involves collecting student work samples from embedded assignments in core curriculum courses, which are then evaluated using the Written Communication VALUE Rubric developed by the Association of American Colleges and Universities (AAC&U, 2009). This rubric provides a standardized framework for assessing students' writing skills, including content development, organization, clarity, genre and disciplinary conventions, and control of syntax and mechanics (Rhodes, 2010).

This report presents findings from the assessment of Written Communication conducted at UTA, focusing on student work samples collected from designated FCA courses. The insights gained from this analysis serve as a foundation for faculty-driven discussions to enhance student learning outcomes within the core curriculum. Research indicates that strong written communication skills are directly linked to students' academic achievement and professional readiness, highlighting the necessity of targeted instructional interventions and continuous assessment (Graham & Perin, 2007). By prioritizing developing and assessing written communication skills, UTA reaffirms its commitment to fostering an educational environment that equips students with the essential competencies needed for academic, professional, and civic life.

Method

Participants

The project gathered evidence of written *communication* within a representative sample of undergraduates at UTA. Data were collected from multiple sections of high-enrollment core curriculum courses in English; this course is offered every fall and spring semesters: ENGL 2329 – American Literature. The total population for this report consists of 238 students. The gender distribution is predominantly female, with 74.4% (n=177) identifying as female and 25.6% (n=61) identifying as male. The racial and ethnic breakdown of the student population reveals a diverse composition. The largest racial/ethnic groups are White (30.7%) and Hispanic/Latino (29.4%), followed by Black/African American students (16.4%) and Asian students (14.7%). Additionally, 5.0% of students identify as having multiple ethnicities, while 2.9% are classified as international students. American Indian/Alaska Native students and those with unspecified race/ethnicity each constitute 0.4% of the population.

A nearly even split exists between first-generation students (50.4%) and non-first-generation students (49.6%). Additionally, 58.4% (n=139) of the population consists of transfer students, while 41.6% (n=99) are non-transfer students (see table 1).

Table 1: Student Demographics

Categorical Information	N	%
Gender		
Female	177	74.4%
Male	61	25.6%
Racial/Ethnic Description		
White	73	30.7%
Hispanic/Latino	70	29.4%
Black/African American	39	16.4%
Asian	35	14.7%
Multiple Ethnicities	12	5.0%
Foreign/International	7	2.9%
American Indian/Alaska Native	1	0.4%
Not Specified	1	0.4%
First Generation Student		
First Generation	120	50.4%
Non-First Generation	118	49.6%

Most students fall within upper academic levels, with 41.2% (n=98) being seniors and 29.8%

(n=71) classified as juniors. Sophomores make up 22.3% (n=53), while fifth-year students and freshmen represent 3.8% (n=9) and 2.9% (n=7), respectively. Typically, freshmen and sophomore-level students represent a majority, but TCC course rosters also contain upper-division and transfer students who must meet the TCC graduation criteria.

The enrollment trends indicate that a significant portion of students began their studies recently, with 40.8% enrolling in 2021-2022 and 26.9% in 2022-2023. Enrollment drops significantly in earlier years, with 16.0% from 2020-2021 and only 5.9% from 2019-2020. Enrollment rates from years before 2019 represent a minimal portion of the population. Additionally, most students are enrolled full-time (68.5%), while 31.5% are part-time, which suggests that most students are committed to a full academic schedule.

Table 2: *Student Status at UT Arlington*

Categorical Information	N	%
Academic Level		
Senior	98	41.2%
Junior	71	29.8%
Sophomore	53	22.3%
Fifth Year	9	3.8%
Freshman	7	2.9%
Enrollment Year		
2021 - 2022	97	40.8%
2022 - 2023	64	26.9%
2020 - 2021	38	16.0%
2019 - 2020	14	5.9%
2018 and Prior	25	10.5%
Academic Load		
Enrolled Full-Time	163	68.5%
Enrolled Part-Time	75	31.5%
Transfer Student		
Transferred	139	58.4%
Non-Transferred	99	41.6%

*Eligibility as of Spring 2024

Students represented all nine UTA colleges and schools. The distribution of students across different colleges shows that the College of Nursing & Health Innovation comprises the largest share at 51.3% (n=122). Other colleges with notable representation include the College of Architecture, Planning & Public Affairs (10.9%), the Division of Student Success (8.0%), and the College of

Education (7.6%). The College of Liberal Arts (7.1%), College of Engineering (6.7%), and College of Science (6.7%) also have moderate representation. However, the College of Business (1.3%) and the School of Social Work (0.4%) have the slightest presence in this dataset.

Table 3: *Students by Colleges/Schools*

College/School	No. of Students	Percentage
College of Nursing & Health Innovation	122	51.3%
College of Architecture, Planning & Public Affairs	26	10.9%
Division of Student Success	19	8.0%
College of Education	18	7.6%
College of Liberal Arts	17	7.1%
College of Engineering	16	6.7%
College of Science	16	6.7%
College of Business	3	1.3%
School of Social Work	1	0.4%

Procedure

Faculty currently teaching undergraduate courses in the Language, Philosophy, and Culture Foundational Component Area (FCA) agreed to submit course Signature Assignments for this report. The syllabus for each core curriculum class at UT Arlington describes the Signature Assignment(s). Students enrolled in core courses complete the Signature Assignment(s) as they would complete other required coursework and assignments.

Written student work samples were obtained from multiple sections of the ENGL 2329 – American Literature course in the fall 2022 and spring 2023 semesters. Data were collected from a high-enrollment core English course offered every fall and spring semester: ENGL 2329 – American Literature. These were high-enrollment courses, with 200-300 students with different majors enrolled in this course every semester. The course focuses on literary texts written by American authors. The texts covered in the course focus on the roles that literacy, narrative, and textual production play in the making of American cultures. The course provides opportunities for students to demonstrate their understanding of American Literature and Culture in writing, making it a suitable sample course for evaluating written communication.

The samples submitted for this assessment process were ungraded, and the Office of Institutional Effectiveness and Reporting (IER) prepared them for rating. Preparation consisted of assigning the papers a coded tracking number and removing all personal identification information (e.g., the student's name, the faculty instructor's name) to prevent rater bias during the planned group

"Scoring Day" activities.

Assessment Instrument

The Signature Assignments were assessed using the Valid Assessment of Learning in Undergraduate Education (VALUE) Rubric for Written Communication (AAC&U 2009) developed by the Association of American Colleges and Universities (AAC&U). The rubric categorizes Written Communication into five dimensions: Context and Purpose, Organization and Structure, Content Development, Source and Evidence, and Control of Syntax and Mechanics. The rubric describes each dimension and uses a four-point scoring scale (see Appendix A). The rubric functions as a matrix that provides narrative descriptions of expected work quality and corresponding point values for scoring the five measures. The point values range from 1 to 4, with 1 indicating baseline performance (Benchmark-1), 2 indicating approaching milestone (Milestone-2), 3 indicating achieved milestone (Milestone-3), and 4 indicating the highest mastery (Capstone-4) of Written Communication. AAC&U, the authors of the rubric, permit zero ratings if the paper does not meet the minimum content or quality standards defined in the rubric. The attainment target (numerical ratings) was set at a score of 2 (Milestone-2). The attainment target was set above the benchmark following recommendations from AAC&U research (Greenhoot & Bernstein, 2012) and standard acceptance criteria in the assessment community.

All raters assigned a score to each of the five dimensions in the rubric for each student's work sample. Higher values indicate more evidence of Written Communication in student work and vice versa. Raters were advised to use zero per AAC&U recommendations if any dimension is absent in student work.

Raters, Rater Calibration, and Scoring

Raters scored the student writing samples during a scheduled scoring day, and each paper was reviewed twice (by two separate raters) in a group setting. A third "tiebreaker" rating was obtained when ratings diverged by more than one rating interval on a single dimension. In these cases, the mean score of three ratings was used as a final score. Twelve faculty members and professional staff with advanced degrees served as raters for the scoring session.

The scoring day began with an orientation and description of the rating process. A qualified UTA facilitator led the raters through reviewing the rubric and discussing the rating dimensions and scale designed to calibrate the raters' understanding and use of the rubric in the rating process. Then,

the entire group read and rated one practice anchor paper, which was chosen beforehand by the facilitator. Following the sample paper review, the facilitator led a discussion among all raters using the anchor paper to reach a common understanding of the Written Communication dimensions and to find exemplar indicators within the paper for the rubric levels of mastery. Following the completion of the calibration activity, formal review and rating of the de-identified student papers began. During the formal review and rating of papers, raters read each paper and assigned scores for each dimension on the rubric using the four-point scale (plus the available "zero" rating). If the values of the skill measure scores for a paper from the two raters were identical or within a one-point difference, then the two scores were considered in agreement and averaged. For example, if Rater A scored the Content Development measure with a value of 2 and Rater B scored the same measure with a value of 3, then the rating was considered in agreement, and scores for that dimension were averaged, resulting in a score value of 2.5. If the scores from the two raters differed by over two points, a third rater was assigned the paper, and then the three scores were averaged together to determine the final score in such cases. For example, if Rater A scored the Content Development measure with a value of 1 and Rater B scored the same measure with a value of 3, the rating was not in agreement, and a third rater was asked to read and score the paper.

Analysis and Results

Inter-rater Agreement and reliability

Once each paper had been rated twice, the IER staff collected the rating sheets, entered the rating scores into a spreadsheet, and analyzed them to determine agreement. Each score was calculated as the average of the two raters' scores if the values assigned by the raters differed by one point or less. The agreement percentages among raters across different dimensions of the Written Communication VALUE Rubric demonstrate high consistency in scoring. The Context and Purpose dimension achieved the highest agreement at 93%, indicating strong alignment between raters in evaluating how well students address the intended purpose within the given context. Organization and Structure followed closely at 92%, showing that raters largely concurred on students' ability to organize and present their work logically. Content Development and Source and Evidence had a slightly lower agreement at 87%, suggesting a moderate level of subjectivity or variation in interpretation regarding how students develop their arguments and incorporate supporting evidence. Control of Syntax and Mechanics showed 88% agreement, indicating relative consistency in

assessing grammatical accuracy, clarity, and writing mechanics.

Table 4: *Agreement Percentages Among Raters*

Dimension (Written Communication VALUE Rubric)	Percentages
Context and Purpose	93%
Organization and Structure	92%
Content Development,	87%
Source and Evidence	87%
Control of Syntax and Mechanics	88%

Note: If values assigned by the raters differed by the rating interval of one point or less, it was counted as agreement. The agreement percentage was computed by dividing the number of agreements by the total number of ratings.

Apart from the simple percentage agreements, researchers widely measure the reliability of rating agreements between different raters to eliminate chance agreements. All raters who participated in the scoring process had advanced degrees and work experience, and attended the same training just before the scoring session. Hence, the probability of chance agreement was very low, but inter-rater agreement was computed to follow best research practices. Inter-rater reliability is the consistency among raters when scoring the same subjects independently. The extent to which different raters agree on their judgments establishes the validity and credibility of measurements or ratings.

The inter-rater agreement was determined to check the consistency level of the rating by calculating the Intraclass Correlation Coefficient (ICC). ICC values reflect the reliability of scores assigned by different raters across various dimensions of the Written Communication VALUE Rubric. ICC is commonly used to assess the degree of agreement among raters beyond chance. High ICC values indicate more reliability between rater scores. Commonly accepted guidelines were used to interpret the ICC results. These suggest that the range of 0.40 to 0.74 is considered fair to good inter-rater agreement, with results above 0.74 classified as excellent inter-rater agreement and results lower than 0.40 considered poor inter-rater agreement (Fleiss, 1986; Shrout & Fleiss, 1979). ICC values for the Written Communication scoring session are presented in Table 4. Context and Purpose has the highest ICC at 0.71, indicating good agreement among raters. Organization and Structure follows with 0.64, which still falls within the fair to good agreement range. Content Development, Source and Evidence, and Control of Syntax and Mechanics all fall between 0.52–0.54, suggesting moderate (fair to good) agreement but lower than ideal reliability.

Table 5: *Intraclass Correlation Coefficient (Cronbach's Alpha)*

Dimension (Written Communication VALUE Rubric)	Coefficient
Context and Purpose	0.71
Organization and Structure	0.64
Content Development,	0.54
Source and Evidence	0.52
Control of Syntax and Mechanics	0.54

Note 1: *less than 0.40 = poor agreement; between .40 and .74 = fair to good agreement; greater than .74 = excellent agreement.*

Note 2: *The intra-class correlation coefficient (ICC) was calculated as a two-way random effects model. Values in this model type with random rater pairings are typically expected to be lower than those where rater pairings are fixed throughout the rating day.*

Students Performance

The final data set contains rating scores on the five dimensions, and all student papers (n = 238) were rated on the scoring day. The student performance data for the Written Communication measure reveals notable trends across different rubric dimensions. Among the five assessed dimensions, Context and Purpose received the highest average score (M = 2.50, SD = 0.74), indicating that students generally understood the objectives and relevance of their writing tasks. This dimension also demonstrated the most substantial rater agreement (93%) and the highest Intraclass Correlation Coefficient (ICC = 0.71), suggesting a more consistent evaluation process among raters. Conversely, Source and Evidence had the lowest student performance (M = 2.00, SD = 0.73), along with the lowest rater agreement (87%) and the weakest reliability (ICC = 0.52). This finding suggests that students struggled to integrate and support their arguments with appropriate evidence, and raters found this criterion more challenging to evaluate consistently. Content Development (M = 2.20, SD = 0.74) and Organization and Structure (M = 2.30, SD = 0.76) also showed moderate student performance, with corresponding ICC values of 0.54 and 0.64, respectively. These findings indicate that students need additional support in effectively developing their ideas and structuring their writing logically. The Control of Syntax and Mechanics dimension (M = 2.45, SD = 0.71) received a relatively higher mean score, reflecting students' competency in writing mechanics, though there remains room for improvement. The means and standard deviations of the analyzed data for each dimension are given in Table 6.

Table 6
Means for Written Communication Measure Scores

Measurement Dimensions	N	Mean	SD
Context and Purpose	238	2.50	0.74
Organization and Structure	238	2.30	0.76
Content Development,	238	2.20	0.74
Source and Evidence	238	2.00	0.73
Control of Syntax and Mechanics	238	2.45	0.71

Summary, Observations, and Limitations

This report analyzed student performance on the Written Communication core objective using the AAC&U VALUE rubric. The evaluation centered on three primary metrics: student mean scores, rater agreement, and scoring reliability. Among the five assessed dimensions, students demonstrated the highest mean score in Context and Purpose (M = 2.50, SD = 0.74), indicating strong proficiency in articulating the purpose and relevance of their writing. The lowest performance occurred in Source and Evidence (M = 2.00, SD = 0.73), highlighting student challenges in effectively integrating and citing evidence.

Rater agreement percentages were consistently high, ranging from 87% to 93%, suggesting strong consensus among raters. However, scoring reliability as measured by Intraclass Correlation Coefficients (ICCs) ranged from fair to good, with values between 0.52 (Source and Evidence) and 0.71 (Context and Purpose). The high rater agreement but lower ICC values, particularly for the dimensions Source and Evidence (0.52) and Content Development (0.54), suggest that while raters generally concurred, absolute scoring consistency could be improved. Additional rater training or rubric calibration may help address these discrepancies.

Student performance was notably weaker in Content Development (M = 2.20) and Source and Evidence (M = 2.00), dimensions associated with lower ICC values, indicating both instructional and assessment challenges. Conversely, the high scores and reliability observed in Context and Purpose suggest students effectively understood assignment expectations and that raters consistently evaluated this dimension.

The variability in ICC scores and reduced reliability in certain rubric dimensions highlight the need for additional rubric refinement or norming sessions. Improved training protocols and clearer rubric descriptors could enhance rater consistency, especially for criteria involving higher subjectivity. Additionally, because the current assessment represents only a single evaluation cycle, results may not fully capture student performance trends or long-term instructional effectiveness.

Future longitudinal assessments are recommended to better understand improvements over time.

In summary, this analysis underscores areas of strength and challenge in assessing Written Communication skills. While students excel in articulating the context and purpose of their writing, targeted instructional interventions and further rubric refinements are necessary to improve student proficiency in content development and the effective integration of evidence.

References

Association of American Colleges and Universities (2019). *VALUE Rubrics*. Retrieved from https://www.aacu.org/value-rubrics/

Association of American Colleges and Universities. (2009). *VALUE rubrics*: Written communication. Retrieved from https://www.aacu.org/value-rubrics

Bean, J. C. (2011). Engaging ideas: *The professor's guide to integrating writing, critical thinking, and active learning in the classroom* (2nd ed.). Jossey-Bass.

Graham, S. (2018). *A writer(s) within community model of writing. Educational Psychologist*, *53*(4), 258–277. https://doi.org/10.1080/00461520.2018.1536981

Fleiss J. L. (1986). *The design and analysis of clinical experiments*. New York: John Wiley & Sons.

Lederman, D. (2015). Are they learning? New effort aims to standardize faculty-driven review of student work. *Inside Higher Ed*. Retrieved from https://www.insidehighered.com/news/2015/09/25/new-effort-aims-standardize-faculty-driven-review-student-work

National Association of Colleges and Employers. (2018). *Job Outlook 2018*. Bethlehem, PA.

Rhodes, T. (Ed.). (2010). *Assessing outcomes and improving achievement: Tips and tools for using rubrics*. Washington, DC: Association of American Colleges and Universities.

Shrout, P., & Fleiss, J. L. (1979). Intraclass correlation: uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420 - 428.

Texas Higher Education Coordinating Board. (2019). Texas core curriculum Report]. http://board.thecb.state.tx.us/apps/TCC/

Appendix A: Written Communication VALUE Rubric

Communication Rubric

	Levels of Achievement			
Criteria	**Capstone 4**	**Milestone 3**	**Milestone 2**	**Benchmark 1**
Context & Purpose	Demonstrates a thorough understanding of context, audience, and purpose and a clear focus on the assigned task.	Demonstrates adequate consideration of context, audience, and purpose and is responsive to the assigned task	Demonstrates awareness of context, audience, and purpose and to the assigned task.	Demonstrates minimal attention to context, audience, purpose, and to assigned task.
Organization and Structure	Demonstrates detailed attention to successful organization, content presentation, formatting, and stylistic choices.	Demonstrates consistent use of organization, content presentation, formatting, and stylistic choices.	Follows expectations for basic organization, content presentation, formatting, and stylistic choices.	Attempts to use a consistent system for basic organization and presentation.
Content Development	Uses appropriate, relevant, and compelling content to illustrate mastery of the subject, conveying the writer's understanding, and shaping the whole work.	Uses appropriate, relevant, and compelling content to explore ideas within the context of the discipline and shape the whole work.	Uses appropriate and relevant content to develop and explore ideas through most of the work.	Uses appropriate and relevant content to develop simple ideas in some part of the work.
Sources and Evidence	Demonstrates skillful use of high quality, credible, relevant sources to develop ideas. Writer contextualizes sources and credits sources throughout the essay an in a works cited/bibliography page or other appropriate source documentation format.	Demonstrates consistent use of credible, relevant sources to support ideas. Writer clearly identifies sources in essay and in a works cited/bibliography page or other appropriate source documentation format.	Demonstrates an attempt to use credible and/or relevant sources to support ideas. Writer generally identifies sources in essay and in a works cited/bibliography page or other appropriate source documentation format.	Demonstrates a basic attempt to use sources to support ideas. Writer does not consistently credit borrowed material to its source in essay and/or in a works cited/bibliography page or other appropriate source documentation format.
Control of Syntax and Mechanics	Uses effective, virtually error-free, language that skillfully communicates meaning to readers with clarity and fluency.	Uses straightforward language with minimal errors to convey clear meaning to readers.	Uses language that conveys general meaning to readers although the language may contain some errors.	Uses language that sometimes impedes meaning because of errors in usage.

Adapted for The University of Texas at Arlington
from AAC&U's Written Communication VALUE Rubric
Last Revised 9/24/2014

UNIVERSITY OF TEXAS ARLINGTON

Association of American Colleges and Universities