Analysis of Artifacts from Marshall's Senior Capstone Courses

Academic Year 2018 – 2019

Summer Assessment Team Members: Marie Archambault, Cam Brammer, Kim DeTardo-Bora, Robert Ellison, Victor Fet, Marty Laubach, Joan St. Germain, and Mary Welch

Summer Assessment Support Staff: Mary Beth Reynolds and Chris Sochor

Executive Summary

Background

In June 2017 the Assessment Workgroup conducted a pilot assessment in which they scored a small sample of capstone project artifacts using the American Association of Colleges and Universities' (AAC&U's) *Critical Thinking* and *Written Communication* Value rubrics. Given the difficulty we have experienced over the years in drawing representative samples of seniors to complete either the *Collegiate Learning Assessment (CLA+)* or Marshall's Senior Assessment, we recommended that staff from the Assessment Office and Quality Initiatives encourage degree programs ask capstone instructors to align their capstone assignments to the "Capstone Critical Thinking" outcome in Blackboard and require students to submit their final projects using the assignment module in Blackboard. We recommended that these discussions be incorporated into larger discussions regarding the process of creating assignments in Blackboard and aligning them to appropriate outcomes of Marshall's Baccalaureate Degree Profile (BDP). We felt that this has the potential to allow us to evaluate a truly random sample of artifacts from multiple degree programs and apply validated rubrics to assess work that students complete as part of their degree programs. Staff from the Office of Assessment and the Online Design Center met with chairs and deans in all colleges except the College of Information Technology and

Engineering during academic year 2017-2018 to ask that they encourage capstone instructors to follow the instructions outlined above. During academic year 2018-2019, 164 non-duplicated artifacts from the Colleges of Arts and Media, Liberal Arts, Business, Health Professions, and Science submitted capstone artifacts for assessment. The Assessment Team evaluated 160 of these artifacts.

Procedures for 2019 Assessment

General Procedures

Eight faculty representing the Colleges of Arts and Media, Business, Liberal Arts, and Science served as the assessment team for this project. They evaluated capstone artifacts using the AAC&U's *Critical Thinking* and *Written Communication* Value rubrics. These rubrics are included in the supporting documentation. Our sample consisted of 160 artifacts. The assessors judged that five of the 160 artifacts did not align to the AAC&U *Critical Thinking* rubric, reducing the number of artifacts in that analysis to 155, while all 160 artifacts were evaluated using the AAC&U's *Written Communication* rubric. This project was coordinated by the Office of Assessment and Quality Initiatives.

Scoring Procedures

Evaluators assessed each artifact using the following scale:

Scoring Codes	
N/A	In the judgment of the evaluators, the artifact did not align with the specific trait of the outcome being assessed.
1	The artifact demonstrated Level 1 performance.
2	The artifact demonstrated Level 2 performance.
3	The artifact demonstrated Level 3 performance.
4	The artifact demonstrated Level 4 performance.

Please see the supporting information that follows this summary for a detailed explanation of scoring procedures.

General Information about the Sample

Of the 155 artifacts assessed for critical thinking 17 were from the College of Arts and Media, 37 from the Lewis College of Business, 5 from the College of Health Professions, 38 from the College of Liberal Arts, and 58 from the College of Science. Of the 160 artifacts assess for written communication 22 were from the College of Arts and Media, 37 from the Lewis College of Business, 5 from the College of Health Professions, 38 from the College of Liberal Arts, and 55 from the College of Science.

Results and Analysis

One challenge in reporting results of the capstone assessment is that, although we assessed 155 artifacts for *Critical Thinking* and 160 for *Written Communication*, each was analyzed by outcome trait. The total number of traits across the two outcomes was 10 (5 each for *Critical Thinking* and for *Written Communication*), potentially resulting in a total of 775 total trait scores for *Critical Thinking* and 800 for *Written Communication*. However, in some instances, evaluators determined that some traits were not addressed in specific student artifacts. The chart below provides the total scorable traits for each outcome, along with mean scores, standard deviations, and frequency counts.

Outcome	Trait (AAC&U rubric)	Total Traits Aligned	Mean Score (SD)	Number of Students Scoring 2.5 – 4	Number of Students Scoring 3.5 – 4
Critical Thinking	Explanation of Issues	155	2.81 (0.57)	132 (85%)	27 (17%)
	Evidence	149	2.57 (0.59)	108 (72%)	14 (9%)
	Influence of Context and Assumptions	153	2.48 (0.55)	105 (68%)	8 (5%)
	Student's Position	154	2.58 (0.60)	119 (77%)	13 (8%)
	Conclusions and Related Outcomes	155	2.63 (0.62)	112 (72%)	20 (13%)
Total for Critical Thinking		**766**		**576 (75%)**	**82 (11%)**
Written Communication	Context of and Purpose for Writing	160	2.89 (0.61)	137 (86%)	43 (27%)
	Content Development	160	2.75 (0.65)	128 (81%)	36 (23%)
	Genre and Disciplinary Conventions	160	2.78 (0.63)	131 (82%)	33 (21%)
	Sources and Evidence	150	2.80 (0.66)	125 (84%)	34 (23%)
	Control of Syntax and Mechanics	160	2.60 (0.60)	118 (74%)	21 (13%)
Total for Written Communication		**790**		**639 (81%)**	**167 (21%)**

A series of paired-samples *t-tests*, using a Bonferroni adjusted alpha level of .005 to control for Type 1 error, showed that, among the traits of *Critical Thinking*, explanation of issues emerged as a strength, being significantly higher than means for all other traits. The mean score for the trait influence of context and assumptions was significantly lower than the means for explanation of issues, student's position, and conclusions

and related outcomes. Frequency counts mirror these outcomes, with 85% of artifacts receiving scores of 2.5 or higher on explanation of issues. That said, the lowest number of artifacts scoring 2.5 or higher was 68% on influence of context and assumptions (which also was the trait that several artifacts did not appear to address at all), with the average number of artifacts scoring 2.5 or higher across traits being 75%.

A series of paired samples *t-tests*, again using a Bonferroni adjusted alpha level of .005 to control for Type I error, showed that, among the traits of *Written Communication*, context and purpose of writing was significantly higher than all other traits, while genre and disciplinary conventions and sources and evidence were each significantly higher than control of syntax and mechanics.

Conclusion

These results provide evidence that the majority of artifacts in this sample of Marshall's capstone courses achieved expected levels of performance in *Critical Thinking* and in *Written Communication*, with 75% and 81% of artifacts scoring between 2.5 and 4.0 for these outcomes.

Within *Critical Thinking*, explanation of issues emerged as a relative strength, while influence of context and assumptions emerged as relative weaknesses. The Assessment Team suggested that we consider expanding the number of AAC&U Value rubrics we used for future capstone assessments. They suggested, for example, that the AAC&U rubrics for *Inquiry and Analysis* or for *Problem-Solving* might fit capstone projects in some disciplines better than the AAC&U rubric for *Critical Thinking*.

Within *Written Communication*, context of and purpose of writing emerged as a relative strength, while control of syntax and mechanics was a relative weakness.

The Assessment Team recommended that the Office of Assessment and Quality Initiatives share these results with the faculty senate and reach out again this year to academic departments to ask that more programs submit capstone work for assessment. In doing so, the team recommended that we provide them with a choice of which AAC&U Value rubric (*Critical Thinking*, *Inquiry and Analysis*, or *Problem-Solving*) best aligns with their students' projects.

MARSHALL UNIVERSITY®

Supporting Documentation

Capstone
Artifact Assessment

Academic Year 2018 – 2019

Outcomes Assessed: AAC&U Rubrics

Outcome	Abbreviation	Traits	Abbreviations
Critical Thinking	**CT**	Explanation of Issues	Issues
		Evidence	Evidence
		Influence of Context and Assumptions	Context/Assumptions
		Student's Position	Position
		Conclusions and Related Outcomes	Conclusions
Written Communication	**WC**	Context and Purpose of Writing	Context
		Content Development	Content
		Genre and Disciplinary Conventions	Genre
		Sources and Evidence	Evidence
		Control of Syntax and Mechanics	Syntax/Mechanics

Review Procedures

- Each artifact had two independent raters and usable scores on the 1 – 4 scale were determined in the following manner:
  - If raters assigned the same score, that became the score for the artifact.
  - If raters' scores differed by one point, e.g. Rater 1 assigned a score of 1 and Rater 2 a score of 2, the final score was the mean, i.e. 1.5.
  - If raters' scores differed by more than one point, e.g. Rater 1 assigned a score of 1 and Rater 2 a score of 3, the raters met to discuss the rationale for their scores to see if they could agree on a score or, at minimum, scores that differed by no more than one point.
  - If raters' scores differed by more than one point and, after discussion, they were not able to resolve the differences, a third rater was assigned to review the artifact. (For this review, all raters were able to come to agreement, so third raters were not needed).

Interrater Reliability

- We conducted interrater reliability analyses using the Cohen's Kappa statistical procedure. In so doing, we used the following rules, similar to those suggested Stellmack, Kohneim-Kalkstein, Manor, Massey, & Schmitz (2009):
  - Since our scoring procedure was to average final scores between two raters when scores differed by only one point, we used that averaged score (e.g. 1.5) as the score for both raters, counting it as an agreement in the interrater reliability analysis.
  - For scores that were two or more points apart, the original score of each reviewer was used in the analysis. Therefore, these scores were counted as disagreements.

Artifacts Excluded from Analysis of Means Due to Issues that prevented assessors from evaluating the artifacts.

Outcome	Total Artifacts	Total Artifacts Eliminated due to Misalignment with Rubric	Total Used for Analysis
Critical Thinking	160	5	155
Written Communication	160	0	160

Critical Thinking AAC&U Value Rubric

AAC&U Critical Thinking Value Rubric

Traits	N/A	Level 1	Level 2	Level 3	Level 4
Explanation of issues	Does not apply to this assignment.	Issue/problem to be considered critically is stated without clarification or description.	Issue/problem to be considered critically is stated but description leaves some terms undefined, ambiguities unexplored, boundaries undetermined, and/or backgrounds unknown.	Issue/problem to be considered critically is stated, described, and clarified so that understanding is not seriously impeded by omissions.	Issue/problem to be considered critically is stated clearly and described comprehensively, delivering all relevant information necessary for full understanding.
Evidence ***Selecting and using information to investigate a point of view or conclusion***	Does not apply to this assignment.	Information is taken from source(s) without any interpretation/evaluation. Viewpoints of experts are taken as fact, without question.	Information is taken from source(s) with some interpretation/evaluation, but not enough to develop a coherent analysis or synthesis. Viewpoints of experts are taken as mostly fact, with little questioning.	Information is taken from source(s) with enough interpretation/evaluation to develop a coherent analysis or synthesis. Viewpoints of experts are subject to questioning.	Information is taken from source(s) with enough interpretation/evaluation to develop a comprehensive analysis or synthesis. Viewpoints of experts are questioned thoroughly.
Influence of context and assumptions	Does not apply to this assignment.	Shows an emerging awareness of present assumptions (sometimes labels assertions as assumptions). Begins to identify some contexts when presenting a position.	Questions some assumptions. Identifies several relevant contexts when presenting a position. May be more aware of others' assumptions than one's own (or vice versa).	Identifies own and others' assumptions and several relevant contexts when presenting a position.	Thoroughly (systematically and methodically) analyzes own and others' assumptions and carefully evaluates the relevance of contexts when presenting a position.
Student's position (perspective, thesis/hypothesis)	Does not apply to this assignment.	Specific position (perspective, thesis/hypothesis) is stated, but is simplistic and obvious.	Specific position (perspective, thesis/hypothesis) acknowledges different sides of an issue.	Specific position (perspective, thesis/hypothesis) takes into account the complexities of an issue. Others' points of view are acknowledged within position (perspective, thesis/hypothesis).	Specific position (perspective, thesis/hypothesis) is imaginative, taking into account the complexities of an issue. Limits of position (perspective, thesis/hypothesis) are acknowledged. Others' points of view are synthesized within position (perspective, thesis/hypothesis).
Conclusions and related outcomes (implications and consequences)	Does not apply to this assignment.	Conclusion is inconsistently tied to some of the information discussed; related outcomes (consequences and implications) are oversimplified.	Conclusion is logically tied to information (because information is chosen to fit the desired conclusion); some related outcomes (consequences and implications) are identified clearly.	Conclusion is logically tied to a range of information, including opposing viewpoints; related outcomes (consequences and implications) are identified clearly.	Conclusions and related outcomes (consequences and implications) are logical and reflect student's informed evaluation and ability to place evidence and perspectives discussed in priority order.

Written Communication AAC&U Value Rubric

AAC & U Written Communication Value Rubric

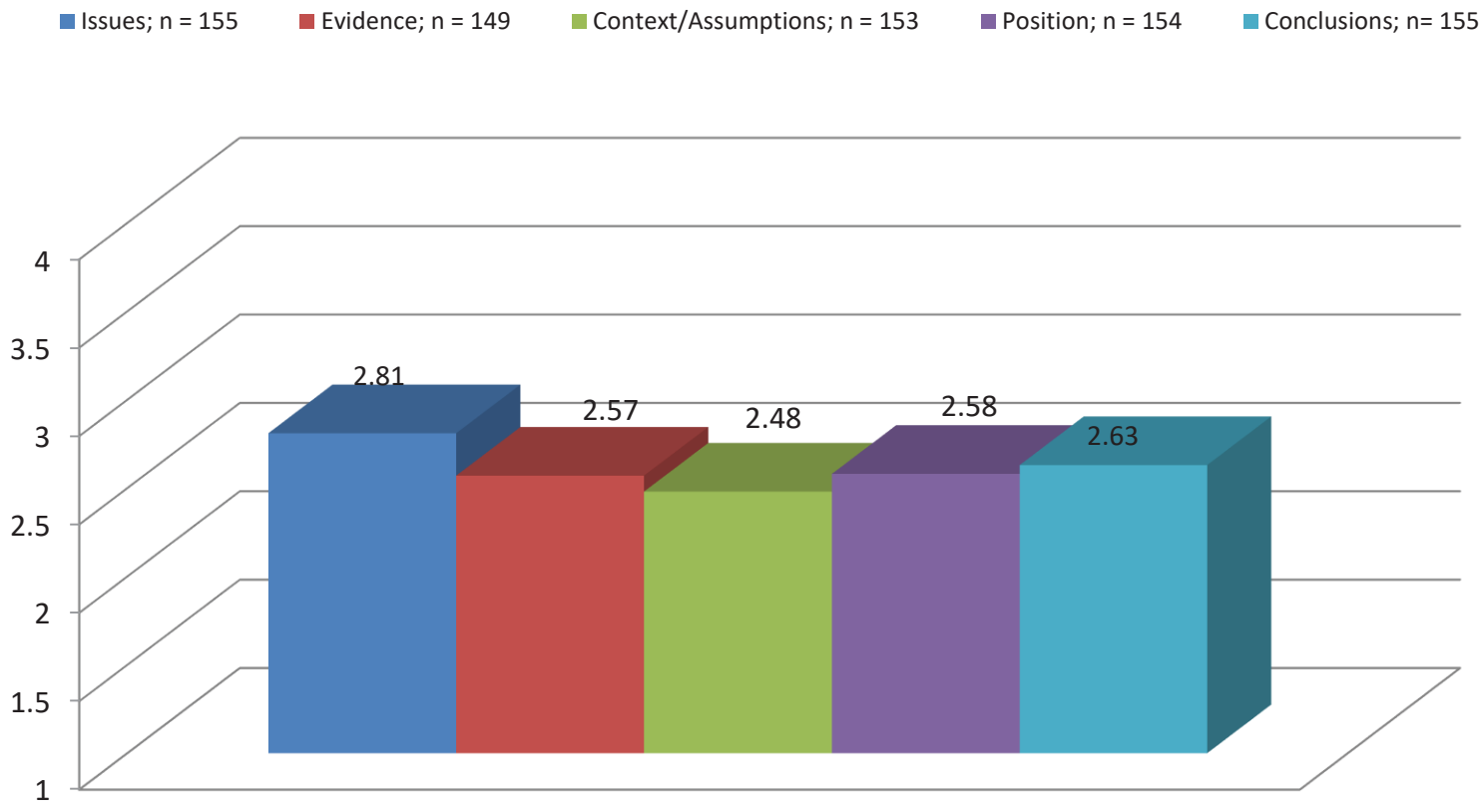
Traits	N/A	Level 1	Level 2	Level 3	Level 4
Context of and Purpose for Writing *Includes considerations of audience, purpose, and the circumstances surrounding the writing task(s).*	Does not apply to this assignment.	Demonstrates minimal attention to context, audience, purpose, and to the assigned tasks(s) (e.g., expectation of instructor or self as audience).	Demonstrates awareness of context, audience, purpose, and to the assigned tasks(s) (e.g., begins to show awareness of audience's perceptions and assumptions).	Demonstrates adequate consideration of context, audience, and purpose and a clear focus on the assigned task(s) (e.g., the task aligns with audience, purpose, and context).	Demonstrates a thorough understanding of context, audience, and purpose that is responsive to the assigned task(s) and focuses all elements of the work.
Content Development	Does not apply to this assignment.	Uses appropriate and relevant content to develop simple ideas in some parts of the work.	Uses appropriate and relevant content to develop and explore ideas through most of the work.	Uses appropriate, relevant, and compelling content to explore ideas within the context of the discipline and shape the whole work.	Uses appropriate, relevant, and compelling content to illustrate mastery of the subject, conveying the writer's understanding, and shaping the whole work.
Genre and Disciplinary Conventions *Formal and informal rules inherent in the expectations for writing in particular forms and/or academic fields (please see glossary).*	Does not apply to this assignment.	Attempts to use a consistent system for basic organization and presentation.	Follows expectations appropriate to a specific discipline and/or writing task(s) for basic organization, content, and presentation	Demonstrates consistent use of important conventions particular to a specific discipline and/or writing task(s), including organization, content, presentation, and stylistic choices	Demonstrates detailed attention to and successful execution of a wide range of conventions particular to a specific discipline and/or writing task (s) including organization, content, presentation, formatting, and stylistic choices
Sources and Evidence	Does not apply to this assignment.	Demonstrates an attempt to use sources to support ideas in the writing.	Demonstrates an attempt to use credible and/or relevant sources to support ideas that are appropriate for the discipline and genre of the writing.	Demonstrates consistent use of credible, relevant sources to support ideas that are situated within the discipline and genre of the writing.	Demonstrates skillful use of high-quality, credible, relevant sources to develop ideas that are appropriate for the discipline and genre of the writing
Control of Syntax and Mechanics	Does not apply to this assignment.	Uses language that sometimes impedes meaning because of errors in usage.	Uses language that generally conveys meaning to readers with clarity, although writing may include some errors.	Uses straightforward language that generally conveys meaning to readers. The language in the portfolio has few errors.	Uses graceful language that skillfully communicates meaning to readers with clarity and fluency, and is virtually error-free.

Critical Thinking: Overall Analysis

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score.

Please note that, while there were 155 artifacts in this sample, not all aligned to every trait of the AAC&U Critical Thinking rubric.

AAC&U Rubric

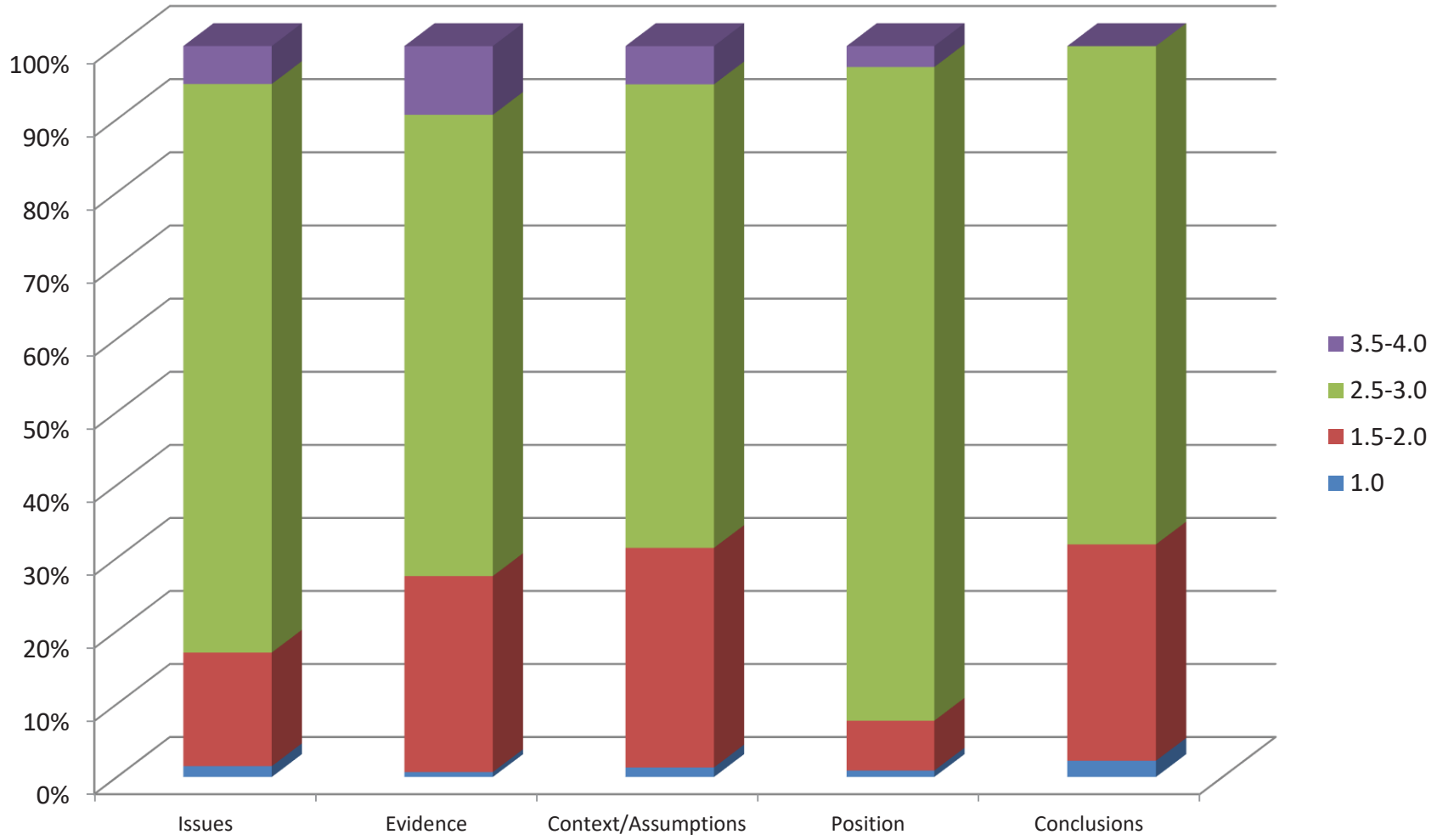


Critical Thinking

Number of artifacts (with usable scores) scoring at each performance level

Trait/ Performance Level	Issues	Evidence	Context/ Assumptions	Position	Conclusions	Total
1.0	2 (1%)	1 (1%)	2 (1%)	4 (3%)	3 (2%)	12 (2%)
1.5 – 2.0	21 (14%)	40 (27%)	46 (30%)	31 (20%)	40 (26%)	178 (23%)
2.5 – 3.0	105 (68%)	94 (63%)	97 (63%)	106 (69%)	92 (59%)	494 (64%)
3.5 – 4.0	27 (17%)	14 (9%)	8 (5%)	13 (8%)	20 (13%)	82 (11%)
Totals	155 (100%)	149 (100%)	153 (100%)	154 (100%)	155 (100%)	766 (100%)

Critical Thinking



Critical Thinking

Inter-Rater Agreement Results (Based on 160 artifacts assessed)

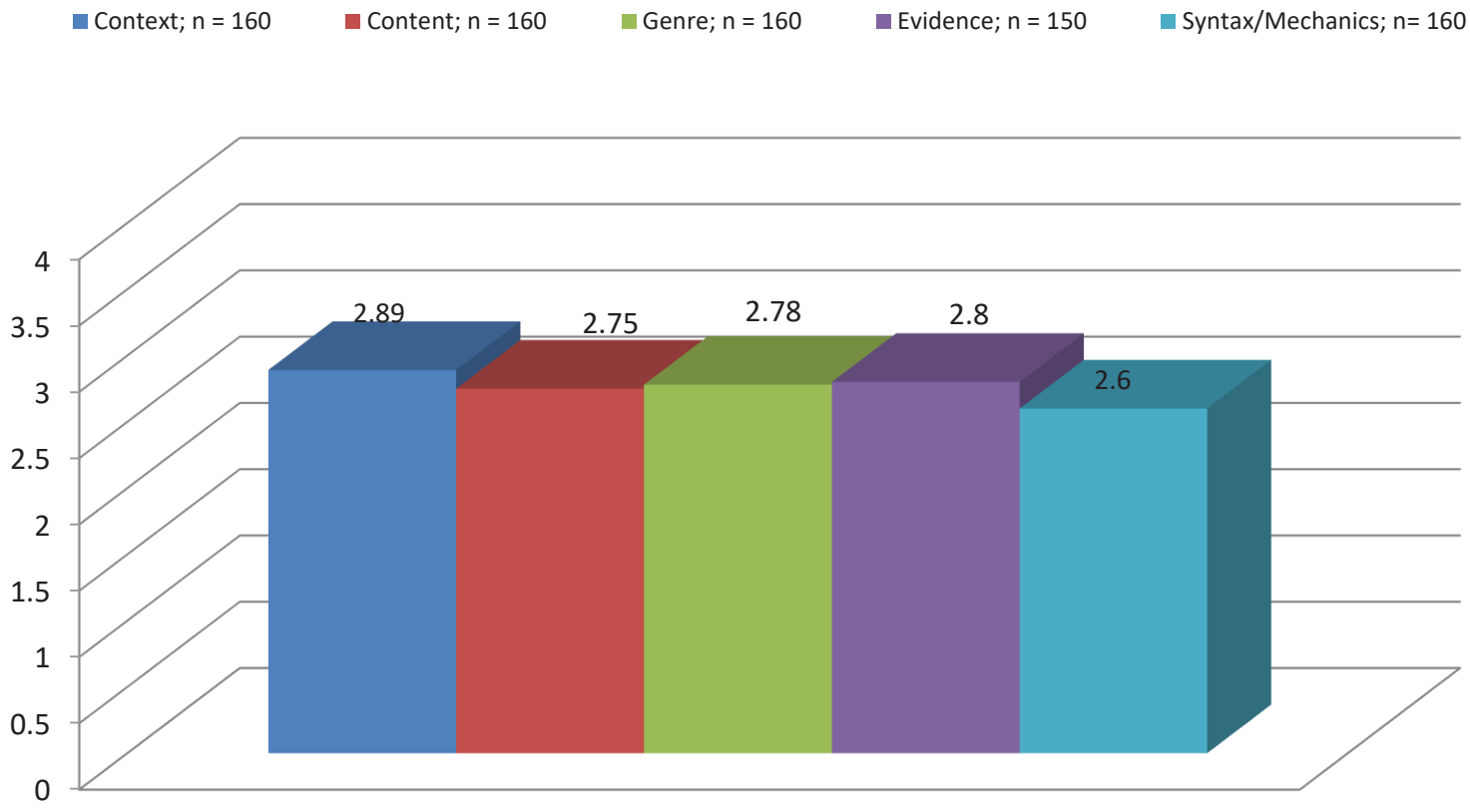
Trait/ Performance Level	Issues Kappa Liberal = .917	Evidence Kappa Liberal = .898	Context/Assumptions Kappa Liberal = .833	Position Kappa Liberal = .807	Conclusions Kappa Liberal = .920
Exact Agreement	93 (58%)	84 (53%)	74 (46%)	72 (45%)	92 (58%)
Difference = 1 point	57 (36%)	63 (39%)	65 (41%)	64 (40%)	58 (36%)
Difference = 2 points	10 (6%)	12 (8%)	16 (10%)	23 (14%)	9 (6%)
Difference = 3 points	0	1 (1%)	5 (3%)	1 (1%)	1 (1%)
Total	160 (100%)	160 (100%)	160 (100%)	160 (100%)	160 (100%)

Written Communication: Overall Analysis

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score.

Please note that, while there were 160 artifacts in this sample, not all aligned to every trait of the AAC&U Written Communication rubric.

AAC&U Rubric

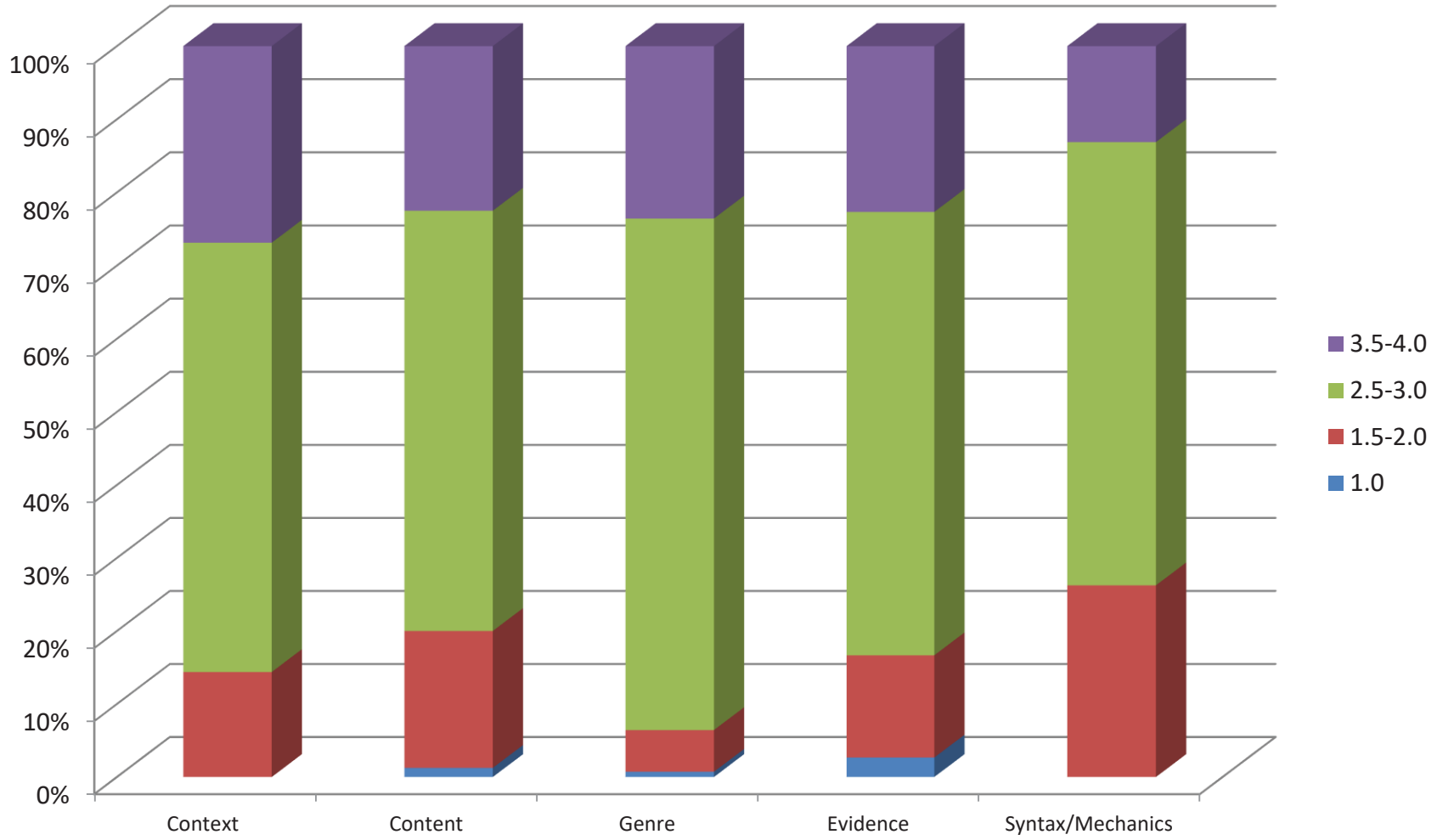


Written Communication

Number of artifacts (with usable scores) scoring at each performance level

Trait/ Performance Level	Context	Content	Genre	Evidence	Syntax/ Mechanics	Total
1.0	0	2 (1%)	1 (1%)	4 (3%)	0	7 (1%)
1.5 – 2.0	23 (14%)	30 (19%)	28 (18%)	21 (14%)	42 (26%)	144 (18%)
2.5 – 3.0	94 (59%)	92 (58%)	98 (61%)	91 (61%)	97 (61%)	472 (60%)
3.5 – 4 .0	43 (27%)	36 (23%)	33 (21%)	34 (23%)	21 (13%)	167 (21%)
Totals	160 (100%)	160 (100%)	160 (100%)	150 (100%)	160 (100%)	790 (100%)

Written Communication



Written Communication

Inter-Rater Agreement Results (Based on 160 artifacts assessed)

Trait/ Performance Level	Context Kappa Liberal = .919	Content Kappa Liberal = .890	Genre Kappa Liberal = .864	Evidence Kappa Liberal = .875	Syntax/Mechanics Kappa Liberal = .873
Exact Agreement	81 (51%)	69 (43%)	74 (46%)	80 (50%)	71 (44%)
Difference = 1 point	69 (43%)	77 (48%)	69 (43%)	64 (40%)	73 (46%)
Difference = 2 points	10 (6^)	14 (9%)	17 (11%)	16 (10%)	14 (9%)
Difference = 3 points	0	0	0	0	2 (1%)
Total	160 (100%)	160 (100%)	160 (100%)	160 (100%)	160 (100%)

Reference

Stellmack, M.A., Kohneim-Kalkstein, Y. L, Manor, J. E., Massey, A. R., & Schmitz, J. A. P. (2009). An assessment of reliability and validity of a rubric for grading APA-style introductions. *Teaching of Psychology*, *36*, 102-107.