Analysis of Artifacts from Marshall's Senior Capstone Courses
Academic Year 2021 – 2022

Summer Assessment Team Members: Marie Archambault, Clinton Brown, Kim DeTardo-Bora, Robert Ellison, Victor Fet, Marty Laubach, and Anita Walz

Summer Assessment Support Staff: Mary Beth Reynolds, Adam Russell, and Chris Sochor

Executive Summary

Recommendations from the 2021 Summer Assessment Team

The Sumer Assessment Team made the following recommendations:
1. As mentioned earlier in this report, we noted alignment between four of the traits of AAC&U's *Critical Thinking* rubric with five of the traits of its *Inquiry and Analysis* Value rubric. This led us to combine the rubrics, resulting in a single rubric that applied to culminating artifacts from courses in the humanities with those from the social and natural sciences. However, in developing this new rubric, we omitted influence of context and assumptions (a trait from AAC&U's Critical Thinking Value rubric, which had been a significant weakness in capstone artifacts assessed during the previous three review cycles) and limitations and implications (a trait from AAC&U's Inquiry and Analysis Value rubric). We note that these traits do not appear as part of any of Marshall University's Baccalaureate Degree Profile outcomes. That said, since students are introduced to considering the influence of context and assumptions during their First Year Seminar in Critical Thinking, we recommend consideration of this element of critical thinking be incorporated into discussions of potential revisions to the Baccalaureate

Degree Profile. We had no discussions about revisions to the Baccalaureate Degree Profile during 2021-2022; these discussions remain on our "to do" list.

2. That the Office of Assessment and Quality Initiatives continue to provide and distribute shorter reports in more digestible formats. We recommend that these reports be disseminated campus-wide through the Assessment Newsletter and shared with the Faculty Senate. We will follow up this year.

3. That we work with the Center for Teaching and Learning to form a committee of faculty to determine the most appropriate rubrics to use for assessment of capstone projects and to evaluate whether there is a need for modifications to some of our existing Baccalaureate Degree Profile outcomes (refer to recommendations from the Baccalaureate Degree Profile Assessment Report). Refer to comment in response to recommendation # 1.

4. That we continue to work closely with the Online Design Center. As more faculty use Blackboard, the Design Center staff are in a unique position to help faculty make appropriate assignment alignments that make student artifacts accessible for university-wide assessment. We are very pleased with the support we have received from the Online Design Center!

Background

In June 2017 the Assessment Team 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 Office of Assessment and Quality Initiatives encourage degree programs' capstone instructors to align their capstone assignments to the "Capstone Critical Thinking" outcome in Blackboard <u>and</u> to require students to submit their final projects using Blackboard's assignment module.   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 had the potential to allow us to evaluate a truly random sample of artifacts from multiple degree programs and to apply validated rubrics to assess work that students complete as part of their degree programs. Staff from the Office of Assessment and Quality Initiatives and the Online Design Center met with chairs and deans in most of Marshall's academic colleges during academic year 2017-2018 to ask that they encourage capstone instructors to follow the instructions outlined above. This year marks our fifth summer (since the initial pilot project) to assess senior capstone projects. The number of senior capstone artifacts submitted during academic year 2021-2022 was 199 from twelve academic disciplines. After reviewing sample artifacts from two disciplines, the Summer Assessment Team determined that, in one case, they did not all align to the AAC&U rubrics we planned to use for evaluation. In the second case, they files were such that they could not be easily accessed for evaluation. Elimination of artifacts from these disciplines reduced the number of usable artifacts to 167 from ten disciplines. From these, we sampled 154 artifacts for assessment. These artifacts came from disciplines within the Colleges of Liberal Arts, Business, Engineering & Computer Sciences, Health Professions, and Science.

Procedures for 2022 Assessment

General Procedures

Eight faculty representing the Colleges of Business, Liberal Arts, and Science served as the assessment team for this project. They evaluated each capstone artifact using a rubric to evaluate *Critical Thinking* that was modified from AAC&U's *Critical Thinking* and *Inquiry and Analysis* value rubrics (please refer to supporting documentation at the end of this report to view the resulting rubric). We used AAC&U's *Written Communication* Value rubric or evaluate students' writing skills. This project was coordinated by the Office of Assessment and Quality Initiatives.

Scoring Procedures

Evaluators assessed each artifact using the following scale:

Scoring Codes	
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 154 artifacts assessed, 33 were from the Lewis College of Business, 36 from the College of Health Professions, 42 from the College of Liberal Arts, 37 from the College of Science, and 6 from the College of Engineering & Computer Sciences. One capstone assignment did not require students to complete work that aligned with the last two traits of the *Critical Thinking* rubric we used, leaving 122 artifacts aligning to those traits.

Results and Analysis

One challenge in reporting results of the capstone assessment is that, although we assessed 154 artifacts for *Critical Thinking* and for *Written Communication*, each was analyzed by rubric outcome trait. The total number of traits across the two outcome rubrics was nine (four for *Critical Thinking* and five for *Written Communication*), potentially resulting in a total of 552 total trait scores for *Critical Thinking* and 770 for *Written Communication*. The chart below provides the total scorable traits for each outcome, along with mean scores, standard deviations, and frequency counts.

Outcome	Trait (AAC&U adapted rubric [CT] and rubric [WC])	Total Traits Aligned	Mean Score (SD)	Number of Students Scoring 2.5 – 4	Number of Students Scoring 3.5 – 4
Critical Thinking	Issues/Topic	154	2.83 (0.62)	130 (84%)	42 (27%)
	Evidence/Existing Knowledge	154	2.82 (0.54)	131 (85%)	35 (23%)
	Position/Analysis	122	2.87 (0.53)	111 (91%)	28 (23%)
	Conclusions	122	2.85 (0.54)	109 (89%)	31 (25%)
Total for Critical Thinking		**552**		**481 (87%)**	**136 (25%)**
Written Communication	Context/Purpose	154	2.94 (0.52)	143 (93%)	39 (25%)
	Content	154	2.88 (0.62)	133 (86%)	48 (31%)
	Genre	154	2.80 (0.57)	135 (88%)	30 (19%)
	Evidence	154	2.88 (0.55)	132 (86%)	42 (27%)
	Syntax/Mechanics	154	2.74 (0.48)	133 (86%)	10 (6%)
Total for Written Communication		**770**		**676 (88%)**	**169 (22%)**

A series of paired-samples *t-tests* were used to determine whether there were significant differences among trait means for each outcome. We used Bonferroni adjusted alpha levels of .005 (for Written Communication) and .008 (for Critical Thinking) to control for Type 1 error. These analyses showed the following results:

Critical Thinking: Pairwise comparisons showed no significant differences between any pairs of means.

Written Communication: The mean for context and purpose of writing was significantly higher than those for genre and disciplinary conventions and control of syntax and mechanics. The mean for content development was significantly higher than the mean for control of syntax and mechanics. The mean for sources and evidence was significantly higher than that for control of syntax and mechanics.

Conclusion

We are pleased to report that 87% of artifacts in our sample scored within the range of 2.5 to 4.0 for *Critical Thinking*, with 88% scoring in this range for *Written Communication*. This was an improvement over last year's sample, where percentages in this range were 72% for *Critical Thinking* (15% increase) and 84% for *Written Communication* (4% increase). Perhaps even more impressive was that 25% of artifacts in this year's sample scored in the range of 3.5 to 4.0 for *Critical Thinking*, with 22% scoring in this range for *Written Communication*. These

percentages compared to only 7% scoring in this range for *Critical Thinking* (18% improvement) and 11% for *Written Communication* (11% improvement) in last year's sample.   We note that the most significant improvement over last year's sample came at the highest possible range, as all students scoring in the 3.5 to 4.0 range received at least one score of 4.0 from reviewers.

There were no significant differences among means for the four traits of the *Critical Thinking* rubric.

Within *Written Communication*, control of syntax and mechanics emerged as the only relative weakness, being significantly lower than means for three other rubric traits – context and purpose of writing, content development, and sources and evidence. We note that control of syntax and mechanics has consistently been a relative weakness since the inception of this assessment in 2018.

Recommendations from the 2022 Summer Assessment Team

The Sumer Assessment Team made the following recommendations:
1. That we determine, in the context of the Baccalaureate Degree Profile and with the goal of comparing performance of Marshall's seniors to a national norm, that we consider the minimally important outcomes (and associated rubric traits) that instructors should use when designing capstone experiences for seniors. Although we realize and respect disciplinary differences in emphasis, we suggest that these experiences also should extend the work begun in First Year Seminar in Critical Thinking and, at minimum, should address critical thinking and communication skills.   Analyzing the alignment of degree program outcomes to those of Marshall's Baccalaureate Degree Profile would be a good first step for this recommendation.
2. Remind each program to work with the Online Design Center to align capstone projects in Blackboard to make them available for the Summer Assessment Team's sample.

Supporting Documentation

Capstone
Artifact Assessment

Academic Year 2021 – 2022

Outcomes Assessed: Modified AAC&U Rubrics

Outcome	Abbreviation	Traits	Abbreviations
Critical Thinking or Inquiry and Analysis	**CT**	Explanation of Issues OR Topic Selection	Issues
		Evidence OR Existing Knowledge	Evidence
		Student's Position OR Design Process and Analysis	Position
		Conclusions and Related Outcomes/Conclusions	Conclusions
Written Communication	**WC**	Context and Purpose of Writing	Context/Purpose
		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.

Critical Thinking Rubric

Modified from AAC&U Critical Thinking and Inquiry and Analysis Value Rubrics (Page 1)

This rubric was created using the Association of American Colleges and Universities (AAC&U) Critical Thinking VALUE Rubric. Retrieved from https://www.aacu.org/value-rubrics

AAC&U Value Rubrics (Critical Thinking and Inquiry/Analysis Combined)

Traits	Level 1	Level 2	Level 3	Level 4
Explanation of Issues OR Topic Selection	**Issue/problem** to be considered critically is stated without clarification or description. **OR** **Identifies** a topic that is far too general and wide-ranging as to be manageable and doable.	**Issue/problem** to be considered critically is stated, but description leaves some terms undefined, ambiguities unexplored, boundaries undetermined, and/or backgrounds unknown. **OR** **Identifies** a topic that, while manageable/doable, is too narrowly focused and leaves out relevant aspects of the topic.	**Issue/problem** to be considered critically is stated, described, and clarified so that understanding is not seriously impeded by omissions. **OR** **Identifies** a focused and manageable/doable topic that appropriately addresses relevant aspects of the topic.	**Issue/problem** to be considered critically is stated clearly and described comprehensively, delivering all relevant information necessary for full understanding. **OR** **Identifies** a creative, focused, and manageable topic that addresses potentially significant yet previously less-explored aspects of the topic.
Evidence (Selecting and using information to investigate a point of view or conclusion) OR Existing Knowledge, Research, and/or Views	Information is taken from sources without any interpretation/ evaluation. Viewpoints of experts are taken as fact, without question. **OR** **Presents** information from irrelevant sources representing limited points of view/approaches.	Information is taken from sources with some interpretation/ evaluation, but not enough to develop a coherent analysis of synthesis. Viewpoints of experts are taken as mostly fact, with little questioning. **OR** **Presents** information from relevant sources representing limited points of view/approaches.	Information is taken from sources with enough interpretation/ evaluation to develop a coherent analysis or synthesis. Viewpoints of experts are subject to questioning. **OR** **Presents** in-depth information from relevant sources representing various points of view/approaches.	Information is taken from sources with enough interpretation/ evaluation to develop a comprehensive analysis or synthesis. Viewpoints of experts are questioned thoroughly. **OR** **Synthesizes** in-depth information from relevant sources representing various points of view/approaches.
Student's Position (perspective, thesis/hypothesis) OR Design Process and Analysis	Specific position (perspective, thesis/hypothesis) is stated, but is simplistic and obvious. **OR** • Inquiry design demonstrates a misunderstanding of the methodology or theoretical framework. • Lists evidence, but it is not organized and/or is unrelated to the focus.	Specific position (perspective, thesis/hypothesis) acknowledges different sides of an issue. **OR** • Critical elements of the methodology or theoretical framework are missing, incorrectly developed, or unfocused. • Organizes evidence, but the organization is not effective in revealing	Specific position (perspective, thesis/hypothesis) takes into account the complexities of an issue. Others' points of view are acknowledged within position (perspective, thesis/hypothesis). **OR** • Critical elements of the methodology or theoretical framework are appropriately developed; however, more subtle	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). **OR** • All elements of the methodology or

Critical Thinking Rubric

Modified from AAC&U Critical Thinking and Inquiry and Analysis Value Rubrics (Page 2)

This rubric was created using the Association of American Colleges and Universities (AAC&U) Critical Thinking VALUE Rubric. Retrieved from https://www.aacu.org/value-rubrics

Page 2

Traits	Level 1	Level 2	Level 3	Level 4
		important patterns, differences, or similarities.	elements are ignored of unaccounted for. • Organizes evidence to reveal important patterns, differences, or similarities related to focus.	theoretical framework are skillfully developed. Appropriate methodology or theoretical frameworks may be synthesized from across disciplines or from relevant subdisciplines. • Organizes and synthesizes evidence to reveal insightful patterns, differences, or similarities related to focus.
Conclusions and related outcomes (implications and consequences) **OR** **Conclusions**	Conclusion is inconsistently tied to some of the information discussed; related outcomes (consequences and implications) are oversimplified. **OR** States an ambiguous, illogical, or unsupportable conclusion from Inquiry findings.	Conclusion is logically tied to information (because information is chosen to fit the desired conclusion); some related outcomes (consequences and implications) are identified clearly. **OR** States a general conclusion that, because it is so general, also applies beyond the scope of the inquiry findings.	Conclusion is logically tied to a range of information, including opposing viewpoints; related outcomes (consequences and implications) are identified clearly. **OR** States a conclusion focused solely on the Inquiry findings. The conclusion arises specifically from and responds specifically to the inquiry findings.	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. **OR** States a conclusion that is a logical extrapolation from the inquiry findings.

Written Communication AAC&U Value Rubric

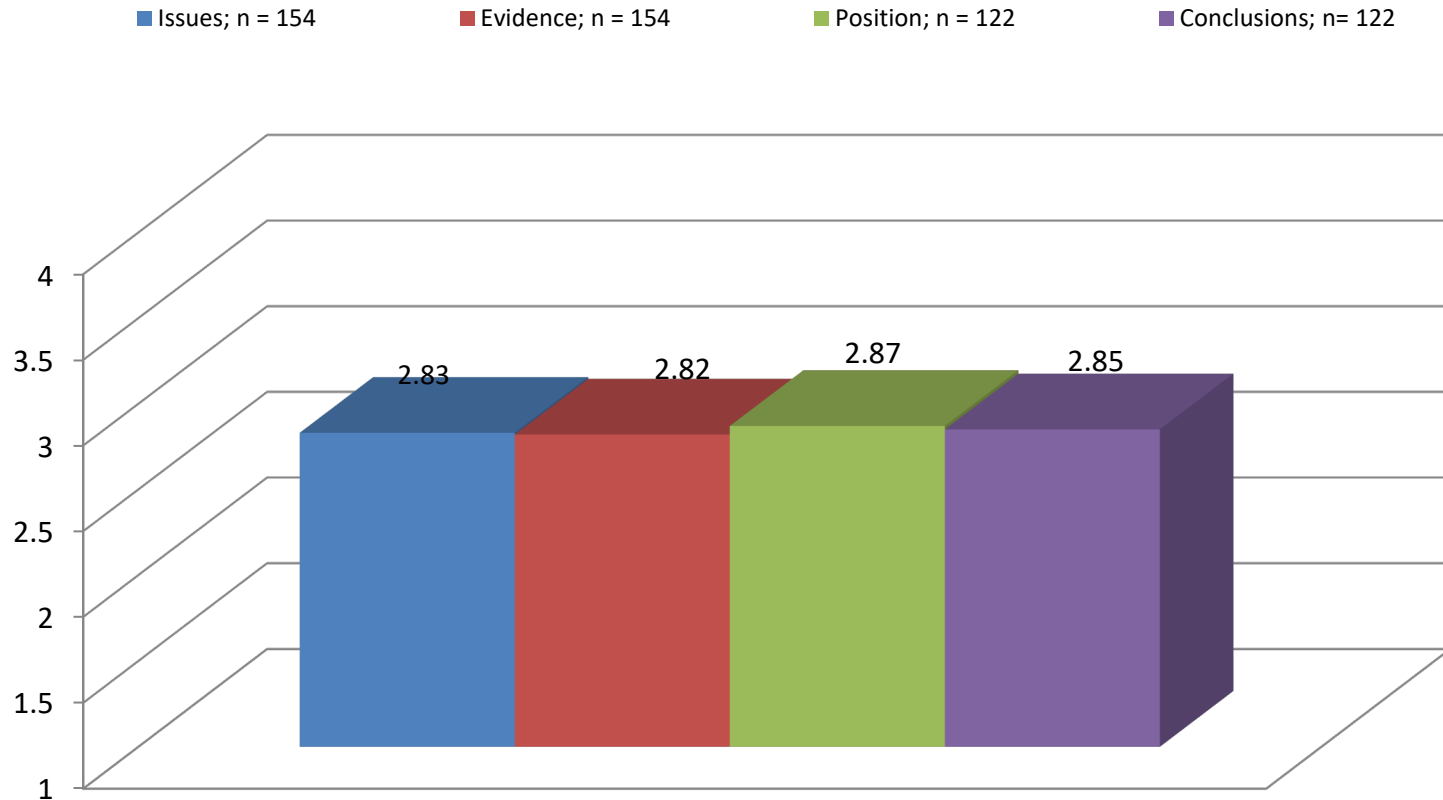**AAC & U Written Communication Value Rubric**

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 184 artifacts in this sample aligned to *Critical Thinking* , seven artifacts aligned to <u>issues</u> and <u>evidence</u> only.

AAC&U Rubric

■ Issues; n = 154 ■ Evidence; n = 154 ■ Position; n = 122 ■ Conclusions; n= 122

Issues: 2.83 · Evidence: 2.82 · Position: 2.87 · Conclusions: 2.85

Critical Thinking

Number of artifacts scoring at each performance level

Trait/ Performance Level	Issues	Evidence	Position	Conclusions	Total
1.0	1 (1%)	0	1 (1%)	1 (1%)	3 (1%)
1.5 – 2.0	23 (15%)	23 (15%)	10 (8%)	12 (10%)	68 (12%)
2.5 – 3.0	88 (57%)	96 (62%)	83 (68%)	78 (64%)	345 (63%)
3.5 – 4.0	42 (27%)	35 (23%)	28 (23%)	31 (25%)	136 (25%)
Totals	154 (100%)	154 (100%)	122 (100%)	122 (100%)	552 (100%)

Critical Thinking
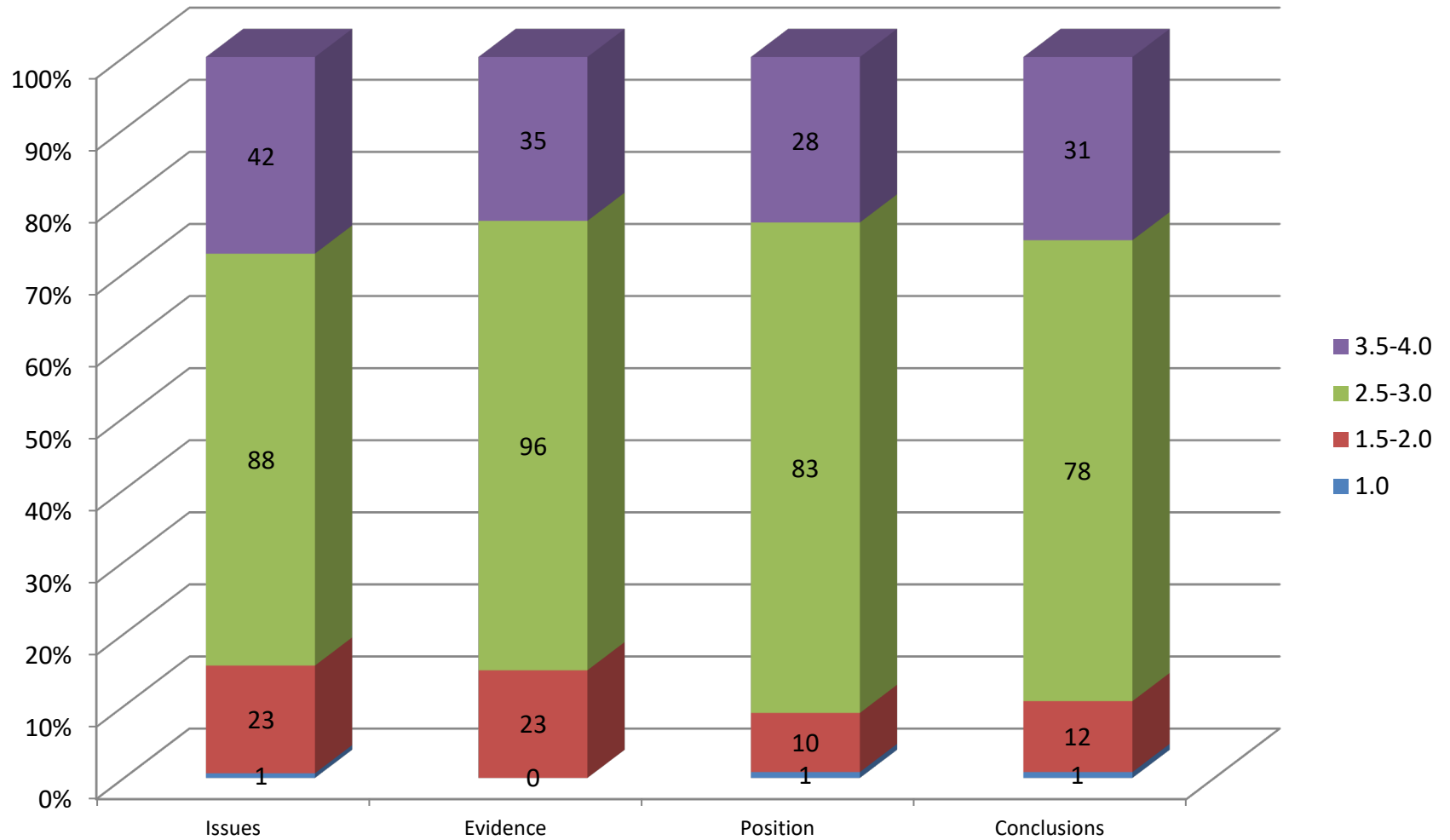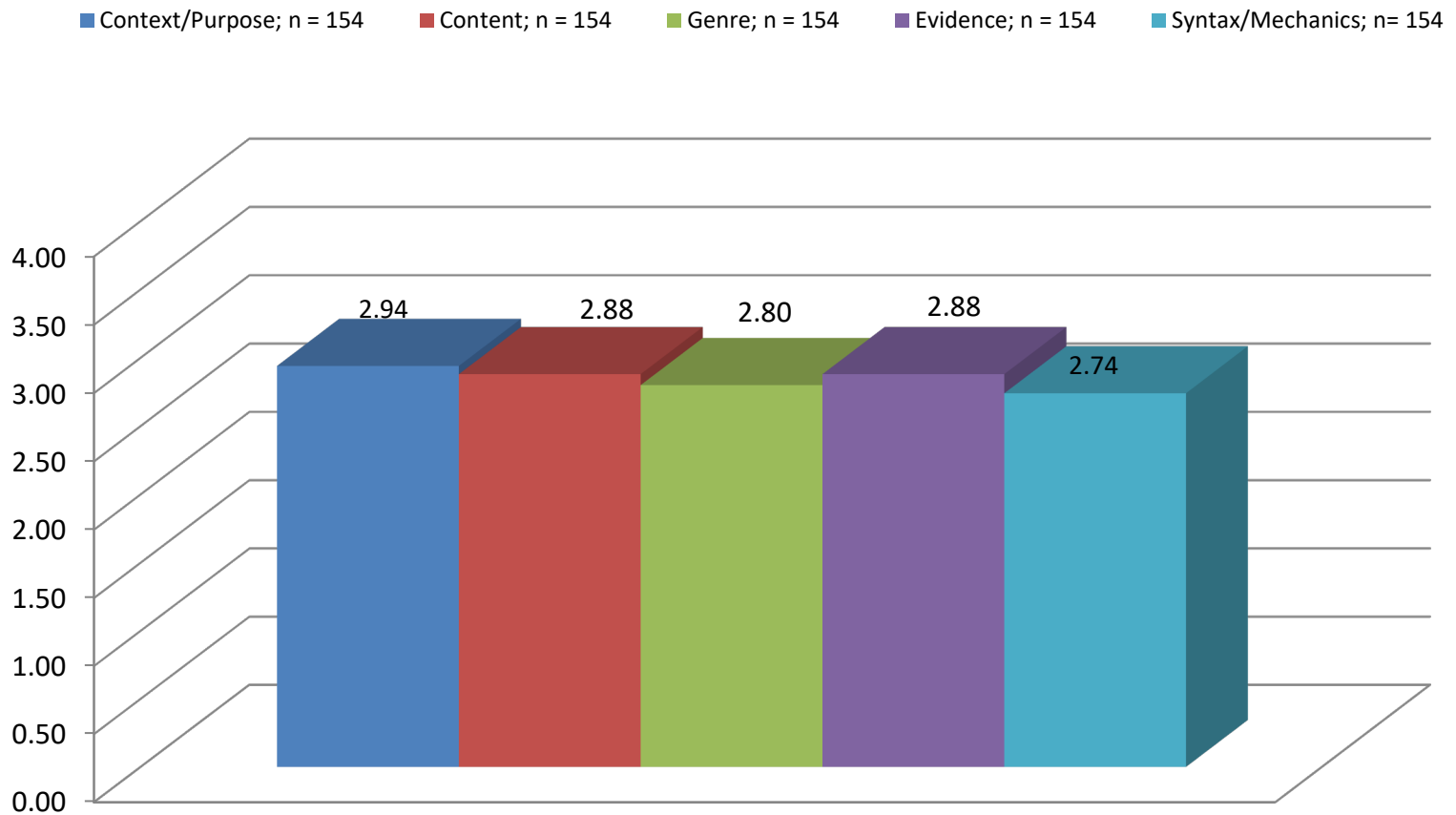
Critical Thinking
Inter-Rater Agreement Results

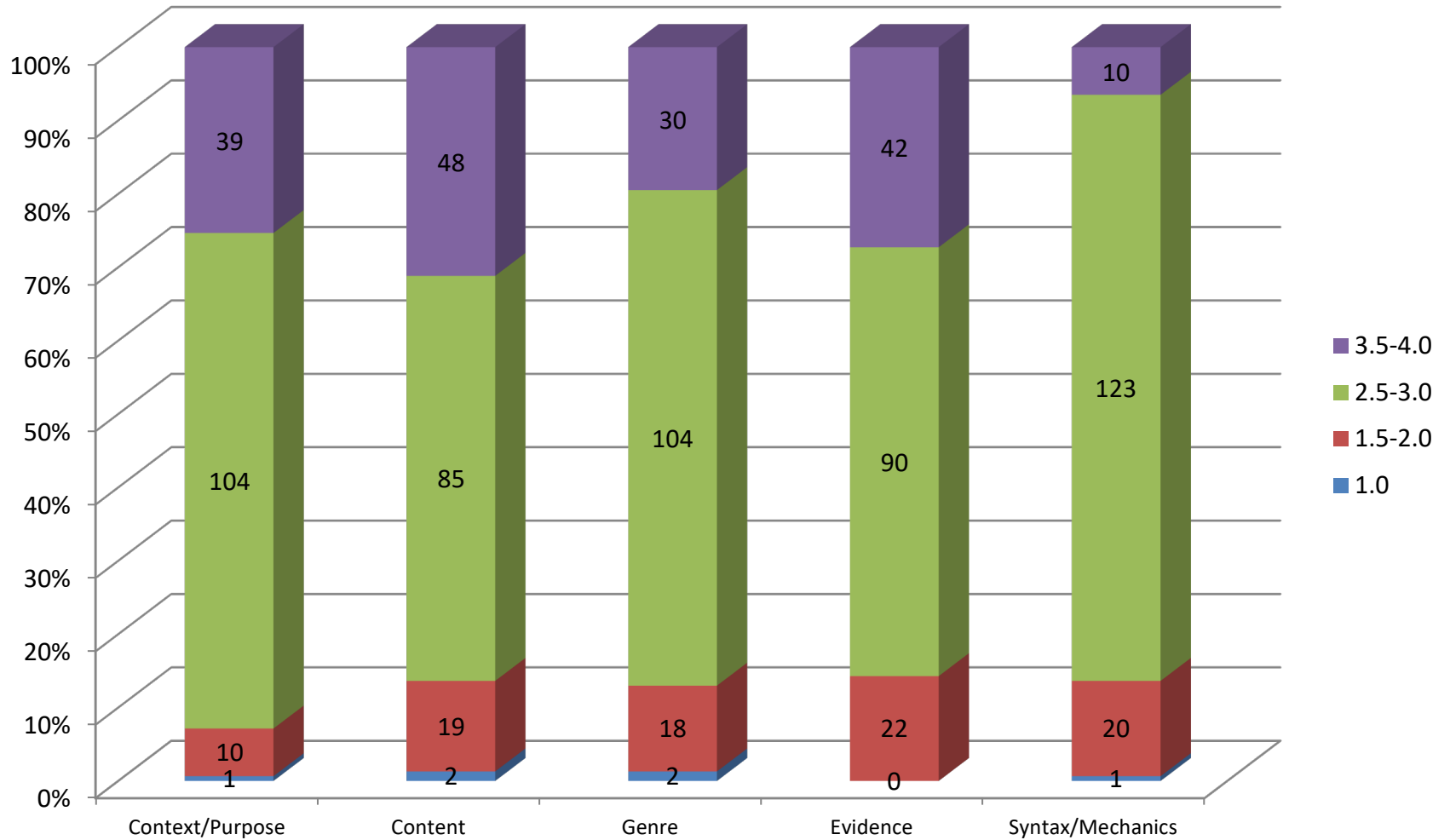| Trait/ Performance Level	Issues Cohen's Kappa (Liberal) = .925	Evidence Cohen's Kappa (Liberal) = .915	Position Cohen's Kappa (Liberal) = .930	Conclusions Cohen's Kappa (Liberal) = .945
Agree on Usable Score	67 (44%)	71 (46%)	55 (45%)	50 (41%)
Difference = 1 point	78 (51%)	73 (47%)	61 (50%)	67 (55%)
Difference = 2 points	9 (6%)	10 (6%)	6 (5%)	4 (3%)
Difference = 3 points	0	0	0	1 (1%)
Total	154 (100%)	154 (100%)	122 (100%)	122 (100%)

Written Communication: Overall Analysis

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

AAC&U Rubric

Legend: ■ Context/Purpose; n = 154 ■ Content; n = 154 ■ Genre; n = 154 ■ Evidence; n = 154 ■ Syntax/Mechanics; n= 154

Values: 2.94, 2.88, 2.80, 2.88, 2.74

Written Communication

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

Trait/ Performance Level	Context/ Purpose	Content	Genre	Evidence	Syntax/ Mechanics	Total
1.0	1 (1%)	2 (1%)	2 (1%)	0	1 (1%)	6 (1%)
1.5 – 2.0	10 (6%)	19 (12%)	18 (12%)	22 (14%)	20 (13%)	89 (12%)
2.5 – 3.0	104 (68%)	85 (55%)	104 (68%)	90 (58%)	123 (80%)	506 (66%)
3.5 – 4 .0	39 (25%)	48 (31%)	30 (19%)	42 (27%)	10 (6%)	169 (22%)
Totals	154 (100%)	154 (100%)	154 (100%)	154 (100%)	154 (100%)	770 (100%)

Written Communication
Inter-Rater Agreement Results

Trait/ Performance Level	Context/Purpose Cohen's Kappa (Liberal) = .912	Content Cohen's Kappa (Liberal) = .899	Genre Cohen's Kappa (Liberal) = .885	Evidence Cohen's Kappa (Liberal) = .921	Syntax/Mechanics Cohen's Kappa (Liberal) = .969
Agree on Usable Score	78 (51%)	61 (40%)	74 (48%)	75 (49%)	97 (63%)
Difference = 1 point	66 (43%)	79 (51%)	67 (44%)	70 (45%)	54 (35%)
Difference = 2 points	10 (6%)	14 (9%)	12 (8%)	9 (6%)	3 (2%)
Difference = 3 points	0	0	1 (1%)	0	0
Total	**154 (100%)**	**154 (100%)**	**154 (100%)**	**154 (100%)**	**154 (100%)**

References

Association of American Colleges and Universities (AAC&U). (2009). *Inquiry and analysis VALUE rubric.* Retrieved from https://www.aacu.org/value/rubrics/critical-thinking

Association of American Colleges and Universities (AAC&U). (2009). *Inquiry and analysis VALUE rubric.* Retrieved from https://www.aacu.org/value/rubrics/inquiry-analysis

Association of American Colleges and Universities (AAC&U). (2009). *Inquiry and analysis VALUE rubric.* Retrieved from https://www.aacu.org/value/rubrics/written-communication

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.