Analysis of Artifacts from Marshall's Senior Capstone Courses
Academic Year 2022 – 2023

Summer Assessment Team Members: Marie Archambault, Clinton Brown, Robert Ellison, Victor Fet, Marty Laubach, Leslie Dawn Quick, and Anita Walz

Summer Assessment Support Staff: Mary Beth Reynolds, Adam Russell, Diana Adams, and Mary Welch

Executive Summary

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 <u>and</u> 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. This continues to be an ongoing discussion.
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. We need more work here, but there were a few new artifacts uploaded for this year's assessment.

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 sixth summer (since the initial pilot project) to assess senior capstone projects. The number of usable senior capstone artifacts submitted during academic year 2022-2023 was 197 from twelve academic 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, Science, and Arts and Media.

Procedures for 2023 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 in our sample, one was found to be corrupted, i.e., reviewers were not able to access it. This resulted in a sample of 153 scorable artifacts. Of these, 23 were from the Lewis College of Business, 26 from the College of Health Professions, 39 from the College of Liberal Arts, 63 from the College of Science, 1 from the College of Engineering & Computer Sciences, and 1 from the College of Arts and Media. One capstone assignment did not require students to complete work that aligned with the last two traits of the *Critical Thinking* rubric (student's position OR design process and analysis and conclusions and related outcomes OR conclusions), leaving 127 artifacts aligning to those traits.

Results and Analysis

One challenge in reporting results of the capstone assessment is that, although we assessed 153 artifacts for *Critical Thinking* and for *Written Communication*, we scored each artifact for four traits of *Critical Thinking* and five traits of *Written Communication*. This process had the potential of resulting in a total of 612 total trait scores for *Critical Thinking* and 765 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	153	2.78 (0.54)	128 (83%)	22 (14%)
	Evidence/ Existing Knowledge	153	2.57 (0.59)	112 (73%)	17 (11%)
	Position/ Analysis	127	2.67 (0.60)	102 (80%)	18 (14%)
	Conclusions	127	2.62 (0.58)	96 (76%)	15 (12%)
Total for Critical Thinking		**560**			
Written Communication	Context/ Purpose	153	2.92 (0.48)	139 (91%)	32 (21%)
	Content	153	2.76 (0.61)	126 (82%)	35 (23%)
	Genre	153	2.82 (0.57)	129 (84%)	31 (20%)
	Evidence	153	2.74 (0.55)	132 (86%)	23 (15%)
	Syntax/ Mechanics	153	2.81 (0.50)	139 (91%)	20 (13%)
Total for Written Communication		**765**			

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 that the mean for explanation of issues OR topic selection was significantly higher than the mean for all other traits. There were no significant differences between any other pairs of means.

Written Communication: The mean score for context and purpose of writing was significantly higher than those for content development and sources and evidence. There were no significant differences between any other pairs of means.

Frequency counts showed that the percentage of students scoring between 3.5 and 4.0 on each trait of the *Critical Thinking* rubric ranged from 11% (evidence OR existing knowledge) to 14% (explanation of

issues OR <u>topic selection</u> and <u>student's position</u> OR <u>design process and analysis</u>). Receiving a score in this range indicates that at least one reviewer gave the trait a score of "4," the highest score possible on the rubric used. The percentage of students scoring between 2.5 and 4.0 ranged from 73% (<u>evidence</u> OR <u>existing knowledge</u>) to 82% (<u>explanation of issues</u> OR <u>topic selection</u>). In this range, the artifacts that did not appear in the 3.5-4.0 range would have received at least one score of "3," which is generally considered an acceptable score for senior level work. Fewer than a quarter of the artifacts scored below this range.

Frequency counts showed that the percentage of students scoring between 3.5 and 4.0 on each trait of the *Written Communication* rubric ranged from 13% (<u>syntax/mechanics</u>) to 23% (<u>content development</u>). The percentage of students scoring between 2.5 and 4.0 ranged from 82% (<u>content development</u>) to 91% (<u>context and purpose for writing </u>and <u>syntax/mechanics</u>). Only between nine and 18% of students scored below this range on any trait.

Conclusion

Across all traits of *Critical Thinking*, on average 78% (as compared to 87% in spring 2022 and 72% in spring 2021) of students scored in the range of 2.5 to 4.0. Although there is room for improvement, we emphasize that scoring in this range indicates that at least one reviewer rendered scores of either "3" or "4." The consensus of the reviewers was that they considered a score of "3" to be perfectly acceptable for seniors, with a score of "4" reserved for truly outstanding work. The latter score (4) was given by at least one reviewer to approximately 13% of our sample, which was down from 25% in 2022. We note that our results suggest that <u>explanation of issues </u>OR <u>topic selection</u> is a significant strength within the *Critical Thinking* outcome.

As has been the pattern over the years, students continued to have slightly higher scores on the traits of *Written Communication* than on those of *Critical Thinking*, with 91% of students scoring 2.5 or higher on <u>syntax/mechanics</u> and <u>context and purpose for writing</u>. These findings were consistent with our findings from 2022.

Recommendations from the 2023 Summer Assessment Team

The Sumer Assessment Team made the following recommendations:
1. There was discussion regarding whether there is any university mandate that senior capstone projects must address *Critical Thinking* and the consensus was that there is not. This assessment evolved because of Marshall's history of emphasizing critical thinking as the centerpiece of the Core Curriculum (approved in 2010 and updated in 2013). It also was implemented to extend the work the university had been doing to emphasize critical thinking by assessing it upon matriculation, as part of First Year Seminar in Critical Thinking, and with a special Critical Thinking Senior Assessment, more authentic for seniors. In other words, rather than ask all seniors to take a special critical thinking assessment (which we found nearly impossible to do), we felt that using a nationally normed *Critical Thinking* rubric to assess authentic work done in the major at the senior level would be more meaningful. After much discussion, the summer assessment team felt that a better approach would be to ask faculty to align senior-level work completed in the students' majors to appropriate Marshall University Baccalaureate Degree Profile (BDP) outcomes. They would continue to make these alignments in Blackboard. Since we would like them to align as many projects as

possible to as many BDP outcomes as possible, the alignments could include all appropriate senior level assignments.

2. To expedite the process mentioned above the Assessment Office will begin downloading and compiling all alignments programs have made between their program's student learning outcomes and the traits of each of the university's nine BDP outcomes. We note that, in the latter instance, each trait is worded as a student learning outcome. This process will enable us to examine more closely a) which programs have made these alignments and, 2) for those programs that have, to which outcomes/traits their program student learning outcomes most commonly align. This analysis will help to inform both our outreach to the programs concerning specific alignments we will ask them to make in Blackboard and changes the university might consider in its comprehensive review of the Core Curriculum, slated to be initiated in academic year 2023-2024.

Supporting Documentation

Capstone
Artifact Assessment

Academic Year 2022 – 2023

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.
  - A third reader was used for two artifacts in this sample.

Rules for arriving at final scores when there are three raters: these rules were followed for all assessments conducted.

- If the third rater's score agreed with one of the first two, the score with the two agreements was used.
- If the first two raters' scores were two points apart, e.g., 1 and 3 and the third rater's score was in the middle, e.g., 2, the third rater's score was used.
- If the first two raters' scores were two points apart, e.g., 1 and 3, and the third rater's scores was a "4", the two scores closer together were averaged, e.g., 3.5.
- If the first two raters' scores were three points apart, e.g., 1 and 4, the third rater's score was averaged with the closest other rater; e.g., if the third rater's score was 3, the final score was 3.5; if the third rater's score was 2, the final score was 1.5.

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.

AAC&U Rubric

■ Issues; n = 153   ■ Evidence; n = 153   ■ Position; n = 127   ■ Conclusions; n= 127

Critical Thinking

Number of artifacts scoring at each performance level

Trait/ Performance Level	Issues	Evidence	Position	Conclusions	Total
1.0	0	2 (1%)	0	0	2 (0%)
1.5 – 2.0	25 (16%)	39 (25%)	25 (20%)	31 (24%)	119 (21%)
2.5 – 3.0	106 (69%)	95 (62%)	84 (66%)	81 (64%)	367 (66%)
3.5 – 4.0	22 (14%)	17 (11%)	18 (14%)	15 (12%)	72 (13%)
Totals	153	153	127	127	560

Critical Thinking
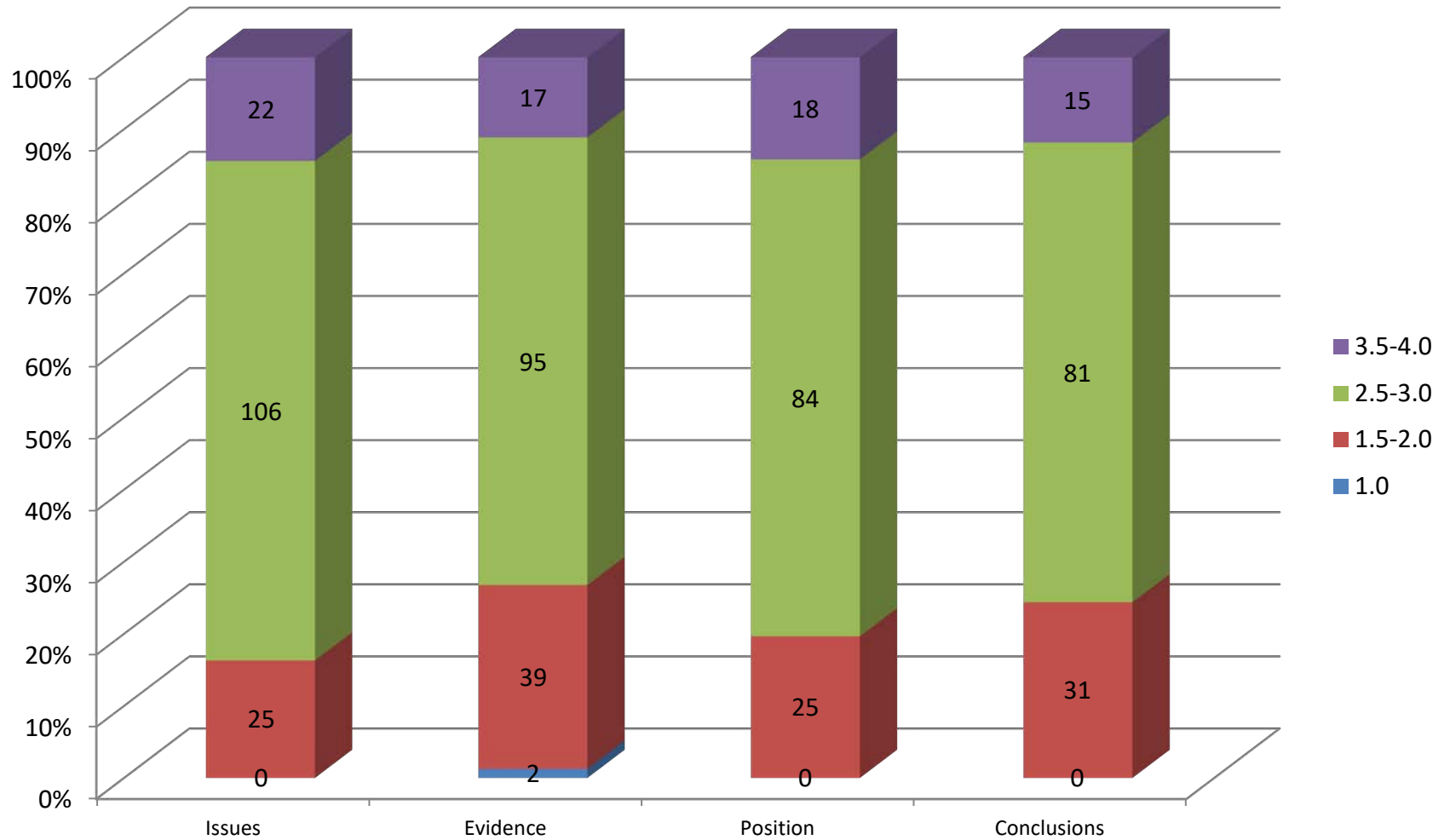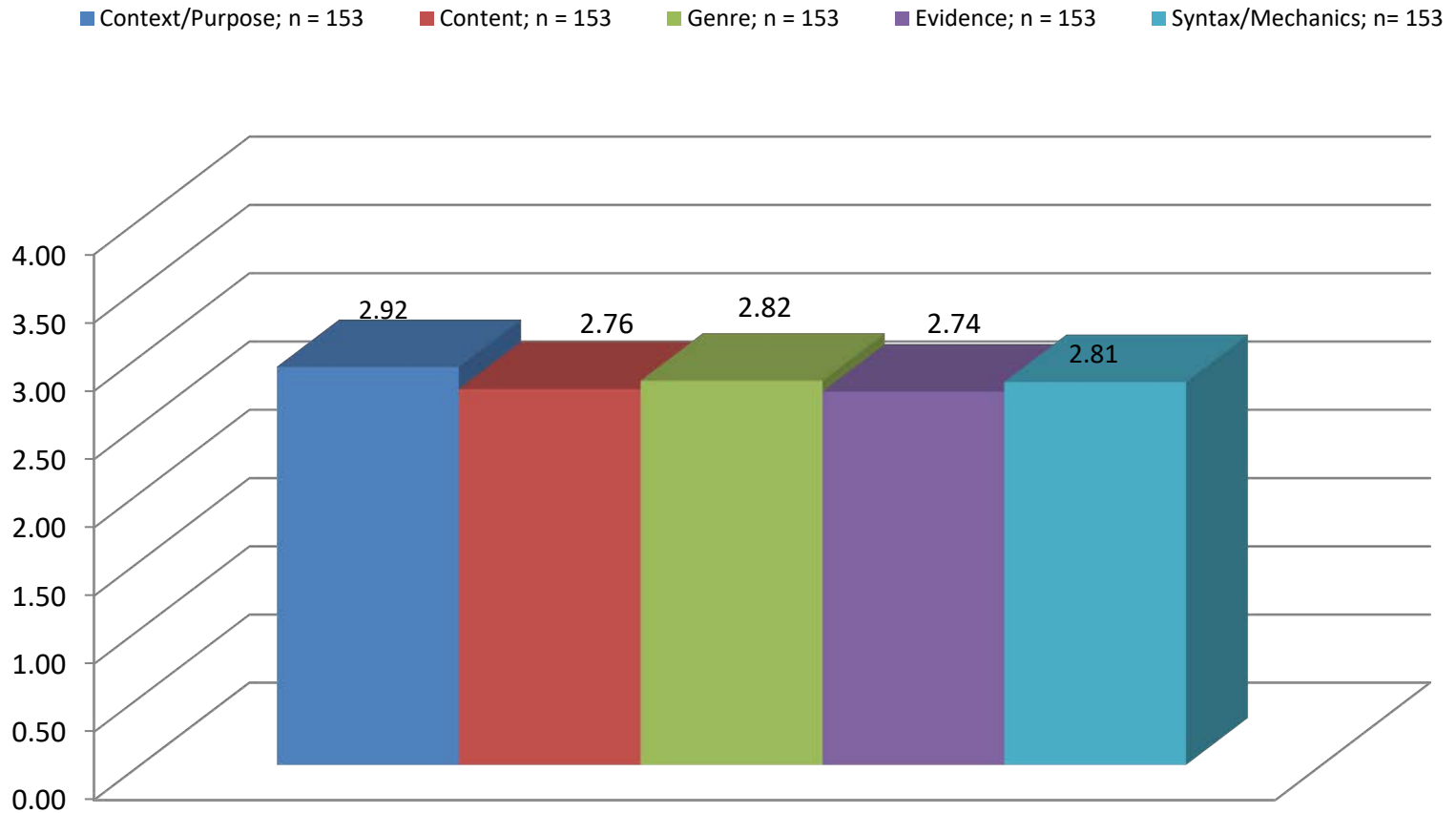
Critical Thinking
Inter-Rater Agreement Results

Trait/ Performance Level	Issues Cohen's Kappa (Liberal) = .868	Evidence Cohen's Kappa (Liberal) = .923	Position Cohen's Kappa (Liberal) = .907	Conclusions Cohen's Kappa (Liberal) = .938
Agree on Score	83 (54%)	68 (44%)	53 (42%)	61 (48%)
Difference = 1 point	55 (36%)	76 (50%)	65 (51%)	60 (47%)
Difference = 2 points	14 (9%)	8 (5%)	9 (7%)	6 (5%)
Difference = 3 points	1 (1%)	1 (1%)	0	0
Total	153	153	127	127

Written Communication: Overall Analysis

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

AAC&U Rubric

■ Context/Purpose; n = 153   ■ Content; n = 153   ■ Genre; n = 153   ■ Evidence; n = 153   ■ Syntax/Mechanics; n= 153

2.92   2.76   2.82   2.74   2.81

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	0	0	1 (1%)	2 (1%)	1 (1%)	4 (1%)
1.5 – 2.0	14 (9%)	27 (18%)	23 (15%)	19 (12%)	13 (8%)	96 (13%)
2.5 – 3.0	107 (70%)	91 (59%)	98 (64%)	109 (71%)	119 (78%)	524 (68%)
3.5 – 4 .0	32 (21%)	35 (23%)	31 (20%)	23 (15%)	20 (13%)	141 (18%)
Totals	153	153	153	153	153	765

Written Communication
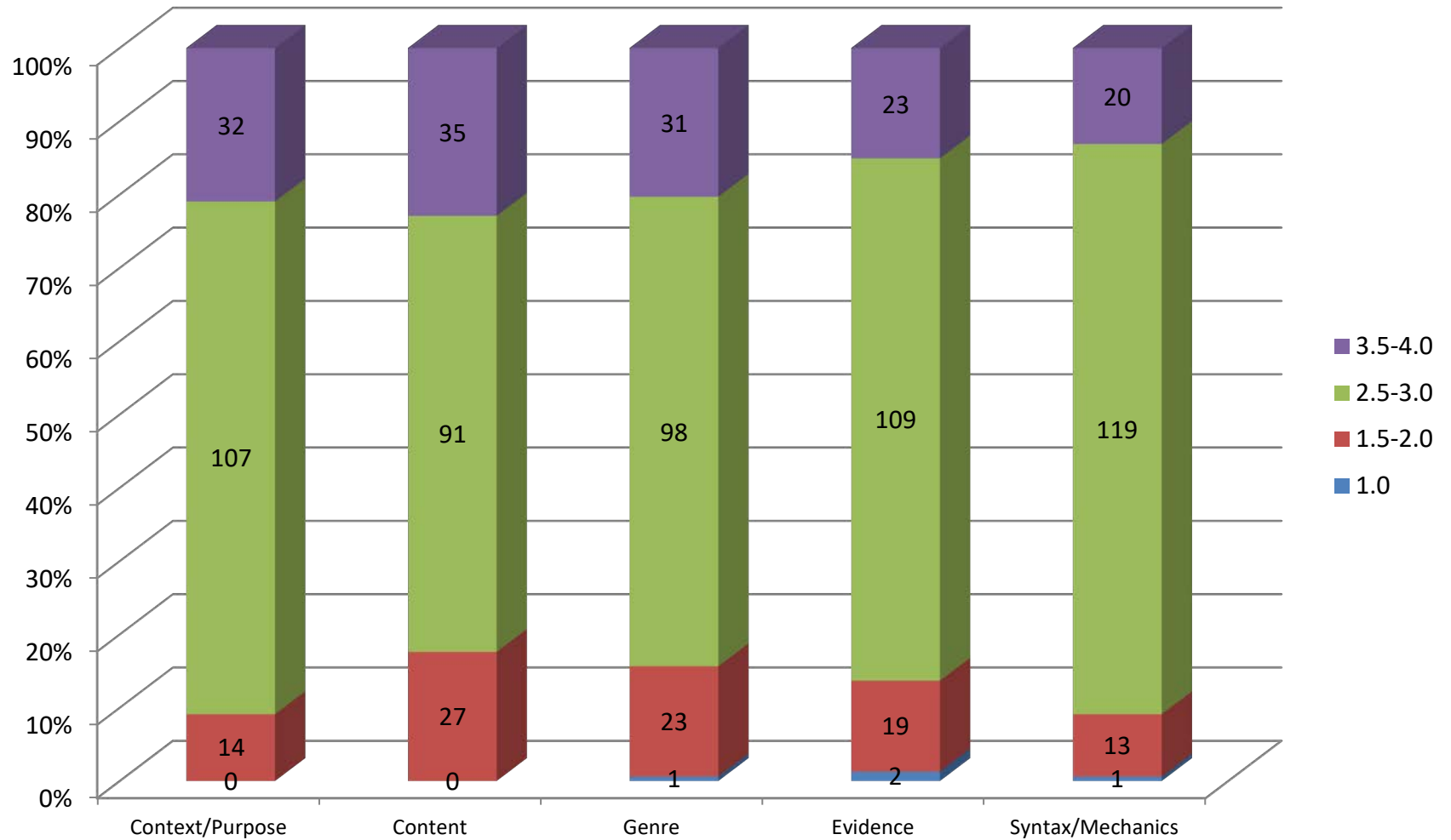
Written Communication

Inter-Rater Agreement Results

Trait/ Performance Level	Context/Purpose Cohen's Kappa (Liberal) = .925	Content Cohen's Kappa (Liberal) = .892	Genre Cohen's Kappa (Liberal) = .921	Evidence Cohen's Kappa (Liberal) = .955	Syntax/Mechanics Cohen's Kappa (Liberal) = .980
Agree on Score	87 (57%)	57 (37%)	81 (53%)	72 (47%)	86 (56%)
Difference = 1 point	58 (38%)	83 (54%)	63 (41%)	76 (50%)	65 (42%)
Difference = 2 points	8 (5%)	10 (7%)	9 (6%)	5 (3%)	2 (1%)
Difference = 3 points	0	3 (2%)	0	0	0
Total	**153**	**153**	**153**	**153**	**153**

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.