Analysis of Artifacts from Marshall's Blackboard Outcomes Repository

Academic Year 2017 – 2018

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

Summer Assessment Support Staff: Mary Beth Reynolds and Tim Melvin (Office of Assessment), Doug Nichols (Academic Affairs Technical Support)

Executive Summary

Background

Recommendations from the 2017 Assessment Workgroup (Updates are in red).

Recommendations Concerning the General Process of Assignment Creation and Accurate Alignment to University Outcomes

We first noted that, beginning with academic year 2016-2017, faculty were asked to develop assignments that aligned to the outcomes as stated in Marshall University's Baccalaureate Degree Profile (BDP). We abandoned the former practice of asking instructors to indicate which performance level on the rubric they used when creating assignments. The Assessment Workgroup began the process of redeveloping rubrics for each of the BDP outcomes so that performance levels now specify *how well* each student demonstrates mastery of the university's *outcomes*, not whether the student achieves progressively more complex outcomes. Outlined below are concerns and recommendations from the Assessment Workgroup.

The transition from our former General Education Assessment Repository to Blackboard for purposes of assessment is off to a good start; however, the Summer Assessment Workgroup made the following recommendations to improve faculty understanding of this process.

1. Staff from the Assessment Office and the MU Online Design Center should schedule meetings with small groups of faculty to discuss, demonstrate, and answer questions about the process of creating assignments in Blackboard's Assignment Module, aligning those assignments to one (or more) of Marshall's BDP outcomes, and having students submit their assignment artifacts using the Blackboard Assignment Module. The Workgroup recommended that this process begin with the staff requesting to be on the schedule of a Chairs' meeting and then following this up with visits to the faculty in as many departments as possible. Kristen Huff and Mary Beth Reynolds

attended a chairs' meeting to discuss the importance of aligning significant course assignments with BDP outcomes and to demonstrate how to do this. This meeting also included a request to begin including capstone projects in this upload. Printed step-by-step directions were disseminated and these directions are available online through Blackboard. Following this meeting, Kristen and Mary Beth met with Chairs of each undergraduate college except the College of Information Technology and Engineering. Kristen Huff regularly meets with faculty to demonstrate this process.

2. In meetings with faculty, Assessment and Design Center staff should emphasize the importance of the inclusion of assignment instructions in Blackboard that explain in some detail how the assignment addresses the BDP outcome to which the faculty member aligned it. If the assignment is meant to address some (but not all) traits of the outcome, the assignment instructions should include the traits that are addressed. The importance of including these directions is included in assignment alignment and artifact upload directions; however, this year's project suggested that some faculty may think they are including instructions (because they include them in Blackboard Learn), but do not actually upload them so that assessors can see them in Blackboard Outcomes. We must make these instructions more explicit.
3. All assignment artifacts that students submit to the Blackboard assignment module for purposes of assessing Marshall's BDP should include *process statements* (aka reflection papers). In other words, each student should describe the *process* s/he used to complete the assignment. This reflection on the process should clearly explain how the assignment helped the student achieve the BDP outcome to which the assignment was aligned. This recommendation has not been accomplished.

Recommendations Concerning the Blackboard Outcomes Assessment Tool

The following items are issues that we will ask Blackboard to address; however, we understand that Blackboard is a large company with many clients and must prioritize improvements to the product. So, while we are hopeful that many of our concerns will be addressed, we realize that addressing them all may take some time.

1. During our assessment cycle, each assessor's artifact queue disappeared upon completion of scoring. This was problematic when score disagreements between the two raters needed discussion. Blackboard has not addressed this issue. The only way for assessors to access their queues after they have completed their assessments is to enter through a link in their original notification emails. If they delete these emails, they cannot reenter their queues. However, we made this known to assessors for the capstone project that followed and they were able to access their queues using this process for reconciliations during that project.
2. We use an assessment process where each artifact is reviewed by two independent reviewers. The random reviewer assignment process that Blackboard uses is too simplistic. We had a total of nine reviewers and each of the nine reviewers had only two review partners for the Blackboard artifact reviews, whereas for the non-Blackboard part of our assessment process, we were able to pair each of our nine reviewers with each of the other eight people on the Assessment team. Blackboard has not addressed this issue. Mary Beth talked with a Blackboard representative on June 11 and the Blackboard representative said she would have some of the technical staff discuss this issue with Mary Beth.

3. Blackboard does not accommodate a third reader for those artifacts for which the original two readers cannot agree on a final score. We had to complete the third reader process this year outside of the Blackboard platform. Blackboard also has not fixed this issue, but it is lower priority than the first two.
4. One of our team members noted that, in Blackboard Learn, course instructors can evaluate student work by having an artifact and rubric next to each other on the computer screen. This was not possible for assessors using Blackboard Outcomes. I am not sure if this issue was addressed.
5. Course names were visible to assessors. We would prefer that course information not be visible to assessors. Course names are still visible.
6. Artifacts did not have unique identifiers in the data download. Rather, each student had an anonymized identifier. Unfortunately, in one project, we had several students who had more than one assignment artifact in our assessment pool. While we were able to make sure we coded each assignment correctly, it took some time and checking to do this. Artifacts are still identified by anonymized student identifier, rather than by unique artifact identifier.
7. Some of the comment columns contained an excessive amount of HTML code, making the comments almost impossible to read. While there was some HTML code in comment columns this year, it was not excessive.
8. We had several other technical questions which we will send to Blackboard.

Recommendations Concerning the Potential Use of Value Rubrics Developed by the American Association of Colleges and Universities (AAC&U)

There was discussion about the potential benefits of using rubrics created and validated by the American Association of Colleges and Universities (AAC&). These AAC&U Value Rubrics have been tested and used widely throughout the United States. The Assessment Workgroup conducted a pilot in which they scored a very small sample of capstone project artifacts using the AAC&U's *Critical Thinking* and *Written Communication* Value rubrics. The group found these rubrics easy to use and their scoring resulted in very few scores of N/A. The Workgroup decided to extend this pilot project to next year's assessment. The pilot will work as outlined below.

1. Course instructors will continue to create assignments using the Assignment Module in Blackboard. Instructors will align the assignment to the appropriate BDP outcome (or outcomes). Students will submit assignment artifacts using the Blackboard Assignment Module.
2. Prior to time for the Summer Assessment Workgroup to begin its work in May 2018, we will again create collections for artifacts aligned to the same outcomes we assessed in May 2017 (*Creative, Inquiry-Based*, and *Quantitative Thinking*).
3. In May/June 2018, assessors will score <u>each</u> artifact with two rubrics as follows – *Creative Thinking* (Marshall's *Creative Thinking* rubric and the AAC&U *Creative Thinking* Value rubric); *Inquiry-Based Thinking* (Marshall's *Inquiry-Based Thinking* rubric and the AAC&U *Critical Thinking* Value rubric); *Quantitative Thinking* (Marshall's Quantitative Thinking rubric and the AAC&U's *Quantitative Literacy* Value rubric). This procedure will help us to address the following issues that emerged during our discussions:

1. There was concern that, in an effort to create an outcome for *Creative Thinking* that would include all disciplines, we may have made it more difficult for programs in the traditional fine arts disciplines to create assignments that align to Marshall's outcome and rubric. Although the

three traits of Marshall's *Creative Thinking* rubric had similar numbers of usable scores, we noted that more artifacts were judged not to align with the outcome at all than was the case for the other two outcomes we assessed this year.

2. For *Inquiry-Based Thinking*, there is a concern that Marshall's outcome and rubric are geared too specifically to traditional scientific fields and are not as applicable as they should be to assignments from fields in the liberal, visual, and performing arts. We believe that the AAC&U's *Critical Thinking* Value Rubric may be more applicable to <u>all</u> fields of study.
3. For *Quantitative Thinking*, there was concern that very few assignment artifacts aligned to two of Marshall's outcome traits (visual representation and statistics). There is a greater difference between Marshall's *Quantitative Thinking* rubric and the AAC&U's *Quantitative Literacy* Value Rubric than between the other Marshall and AAC&U cognates and using both rubrics in next year's assessment has the potential to help us determine which works better for our instructors and students.

We followed-up on this suggestion and scored each artifact in our sample aligned to *Creative*, *Inquiry-Based*, and *Quantitative Thinking* with both their MU rubric and corresponding AAC&U Value rubric. We will discuss results at greater length later in the report.

Longitudinal Analysis

For the initial assessment of artifacts uploaded to GEAR (summer 2013), all artifacts assessed were drawn from the university's First Year Seminar in Critical Thinking (FYS) course and we used these artifacts to assess all nine university outcomes. Mean performance across students ranged from a low of 0 for *Intercultural Thinking* (communication with other cultures) to a high of 1.24 for *Communication Fluency* (design/organization and diction). However, since artifacts were spread among so many outcomes, many traits had very small numbers (9 for communication with other cultures as compared to 24 for design/organization and 23 for diction). Other than the fact that all students included in the 2013 sample were freshmen, low means can be attributed to the fact that we had not yet settled on a score for misaligned artifacts, defaulting many of the scores to 0.

The second assessment of artifacts uploaded to GEAR (summer 2014) also included all nine outcomes, but we included artifacts from *Multicultural, International, Service Learning,* and *Writing Intensive* courses, in addition to those from FYS. The sample, however, continued to be skewed toward artifacts from lower level courses with freshman being the modal class rank for student artifacts in our sample. We decided to assign special codes to artifacts we felt to be misaligned to the outcomes or in cases of student upload or other technical issues that prevented assessment. This allowed us to see which outcomes/traits resulted in the greatest amount of confusion during the outcome/trait alignment process and resulted in recommendations to make sure instructors uploaded assignment instructions, specified the <u>primary</u> outcome to which their assignment aligned, and identified the performance level to which the assignment was written. Due to assessing all nine university outcomes again in 2014, we continued to have small numbers of artifacts aligned to each outcome, which led to the recommendation that we choose only three outcomes to assess in 2015, three more in 2016, and the last three in 2017 and continue to assess on a three-year cycle.

The third assessment of artifacts uploaded to GEAR (summer 2015) consisted of an in-depth assessment of artifacts that instructors aligned to the following outcomes as primary: *Intercultural Thinking* (due to sampling error, five of the alignments for *Intercultural Thinking* were secondary), *Ethical and Civic Thinking,* and *Communication Fluency*. One hundred eight artifacts were included for each outcome, resulting in a total of 324 artifacts. This sample resulted in higher numbers for each outcome trait. Results from summer 2015 suggested a need to redesign rubrics to be continuous, rather than categorical, in nature.

Finally, assessment data from 2013-2014 and 2014-2015 showed that Marshall's students improved their writing skills as they moved through the curriculum and, specifically, as they passed from 100/200 level writing intensive courses to 300/400 level writing intensive courses. Analyses in 2016-2017 showed that students generally showed stronger performance in 300/400 level courses than in 100/200 level courses in *Inquiry-Based Thinking.*

Procedures for 2018 Assessment

General Procedures

In summer 2018 we evaluated student artifacts (as we did in summer 2017) produced in response to course assignments aligned to *Creative*, *Inquiry-Based*, and *Quantitative Thinking*. In May 2018 a group of nine faculty representing several academic colleges from across the university evaluated a sample of these artifacts using Marshall's outcome specific rubrics and each outcome's corresponding rubric from the American Association of Colleges and Universities (AAC&U Value rubrics). Specifically, AAC&U Value rubrics used were *Creative Thinking* for Marshall's *Creative Thinking* outcome; *Critical Thinking* for Marshall's *Inquiry-Based Thinking* outcome; and *Quantitative Literacy* for Marshall's *Quantitative Thinking* outcome. These rubrics are included in the supporting documentation. Our sample initially consisted of 324 artifacts, 108 per outcome. However, during scoring we discovered that 5 artifacts (1 aligned to *Creative*, 3 to *Inquiry-Based,* and 1 to *Quantitative Thinking*) were not able to be opened or otherwise accessed by the reviewers for scoring. This reduced the number of usable artifacts to 319 (107 *Creative*, 105 *Inquiry-Based*, and 107 *Quantitative Thinking*). Reviewers further determined that 60 artifacts (18 *Creative*, 12 *Inquiry-Based*, and 30 *Quantitative Thinking*) were misaligned with all of the traits of the outcomes to which they had been tagged on the MU rubrics and 48 (11 *Creative*, 10 *Inquiry-Based*, and 27 *Quantitative Thinking*) were misaligned with all of the traits of the AAC&U rubrics. This reduced the number of scorable artifacts to 259 for the MU rubric (89 *Creative Thinking*, 93 for *Inquiry-Based Thinking*, and 77 for *Quantitative Thinking*) and 271 for the AAC&U rubric (96 *Creative*, 95 for *Inquiry-Based*, and 80 for *Quantitative Thinking*). Each artifact was read by two independent reviewers (to arrive either at scores or to agreements of nonalignment for specific traits of each outcome). This project was coordinated by the Office of Assessment.

Scoring Procedures

Evaluators assessed each artifact using the following scale:

Special Scoring Codes	
Score	**Explanation**
N/A	In the opinion of the evaluator, the artifact was misaligned with the outcome/trait to which the instructor had tagged it.
Error	The student did not upload the correct assignment or there was a technical problem with the upload that prevented the artifact from being opened or assessed.
Regular Scoring Codes These codes were given to artifacts that, in the opinion of the evaluator, were aligned with appropriate outcomes/traits and contained enough information to allow assessment.	
0	The artifact did not demonstrate the minimum level of performance expected at Level 1.
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

Two hundred sixty-seven (267; 82%) of the artifacts in our sample were drawn from courses at the 100/200 level, with the remaining 57 (18%) drawn from courses at the 300/400 level.

Results and Analysis

One challenge in reporting results of Blackboard assessment is that, although we assessed 324 artifacts (each of which was aligned to one of the BDP outcomes assessed this year), results were analyzed using two rubrics for each outcome and each was analyzed by outcome trait. The total number of traits across the three outcomes was 11 for the MU rubrics (4 each for *Inquiry-Based* and *Quantitative Thinking*, and 3 for *Creative Thinking*) and 17 for the AAC&U rubrics (6 each for *Creative Thinking* and *Quantitative Literacy* and 5 for *Critical Thinking)*. As mentioned previously, 5 artifacts were not able to be assessed due to upload or artifact file error, reducing the number of readable artifacts to 319. Of those, assessors agreed that 60 did not align to any trait of the outcomes to which they were tagged (using the MU rubrics) and 48 did not align with to any trait on the AAC&U Value rubrics. This left 259 scorable artifacts (for the MU rubric) and 271 (for the AAC&U Value rubrics). However, not all of those artifacts aligned to <u>every</u> trait of the outcomes to which they were tagged. A perusal of our supporting

documentation shows that the artifacts evaluated by the Assessment Workgroup tagged to a total of 725 traits using the MU rubrics (232 for *Creative*, 296 for *Inquiry-Based*, and 197 for *Quantitative Thinking*), all of which were used in calculating means. Using the AAC&U Value rubrics, artifacts tagged to a total of 1,248 traits (472 for *Creative Thinking*, 393 for *Critical (Inquiry-Based) Thinking*, and 383 for *Quantitative Literacy*). As can be seen in the chart below, AAC&U's *Creative Thinking* rubric traits solving problems and embracing contradictions are aligned to less often than its other traits; MU's *Creative Thinking* rubric trait ambiguities and possibilities is aligned to less often than the other two traits. For *Inquiry-Based Thinking* (we used AAC&U's *Critical Thinking* rubric for this analysis), the trait influence of context and assumptions was aligned to least often, whereas for the MU rubric the trait problem/question was aligned to least often. There were visible differences when considering the artifacts aligning to each of the MU rubric traits for *Quantitative Thinking*, with far fewer aligning to estimation, visual representation, or statistics than to context; for the AAC&U rubric, fewer artifacts aligned to the assumptions trait than to the others.

Outcome	Trait (MU rubric)	Total Traits Aligned	Trait (AAC&U rubric)	Total Traits Aligned
Creative Thinking	Ambiguities and Possibilities	73	Acquiring Competencies	82
			Taking Risks	83
	Risk Taking	80	Solving Problems	70
			Embracing Contradictions	67
	Synthesis and Innovation	79	Innovative Thinking	84
			Connecting, Synthesizing, Transforming	86
Total for Creative Thinking		**232**		**472**
			Please note that AAC&U Critical Thinking rubric was used	
Inquiry-Based Thinking	Problem/Question	59	Explanation of Issues	75
	Research of Existing Knowledge	76	Evidence	87
	Data Collection and Analysis	76	Influence of Context and Assumptions	60
	Conclusions	85	Student's Position	82
			Conclusions and Related Outcomes	89
Total for Inquiry-Based Thinking		**296**		**393**
Quantitative Thinking	Context	75	Interpretation	66
			Representation	67
	Estimation	45	Calculation	62
			Application/Analysis	70
	Visual Representation	46	Assumptions	45
	Statistics	31	Communication	73

Total for Quantitative Thinking		**197**		**383**
Totals		**725**		**1,248**

Based on course level, *Creative Thinking* means did not differ significantly for any trait on the MU rubric. However, students enrolled in courses at the 300/400 levels had significantly higher means for the AAC&U *Creative Thinking* rubric on two traits (acquiring competencies and taking risks) than did students enrolled in 100/200 level courses.

Students enrolled in courses at the 300/400 levels had significantly higher means for all traits of MU's *Inquiry-Based Thinking* rubric than did students enrolled in 100/200 level courses. For the AAC&U *Critical Thinking* rubric means at the 300/400 level were significantly higher for all traits *except* evidence and context/assumptions. Course level performance was not significant for any trait on either the MU or AAC&U *Quantitative Thinking/Literacy* rubrics.

Overall results showed mean performance for traits to range from 1.35 (MU *Inquiry-Based Thinking*: research of existing knowledge) to 2.35 (AAC&U *Quantitative Literacy:* calculations). Mean performance for artifacts uploaded from 100/200 level courses ranged from 1.19 (MU *Inquiry-Based Thinking*: research of existing knowledge) to 2.14 (AAC&U *Quantitative Literacy:* communication) and for 300/400 level courses from 1.21 (MU *Creative Thinking:* ambiguities and possibilities) to 2.74 (AAC&U *Critical Thinking:* explanation of issues). Although there does not appear to be an overall strength for our students, their weakest performance was in *Creative Thinking*, where no overall means reached 2.0 on either MU or AAC&U rubrics. We also noted that mean scores for *Quantitative Thinking* were more likely to reach 2.0 or higher using the AAC&U than the MU rubric and that the *Quantitative Thinking* artifacts were more often aligned to the traits of the AAC&U *Quantitative Literacy* rubric than to those of the MU rubric.

Results for Course Type

Analyzing results by course type posed several challenges. First, the only course type that is unique (i.e. can have only one course type attribute) is First Year Seminar in Critical Thinking (FYS). Courses can have the other attributes (Critical Thinking [CT], Multicultural [MC], International [INT], Writing Intensive [WI], Service Learning [SL], and Core II) in combination (and many do). So, when analyzing results by course type, we included all courses with the attribute we wanted to assess; this resulted in some courses being included in the analysis for more than one course type. Because MC and INT courses are asked to create assignments whose primary alignment is to *Intercultural Thinking* (an outcome we did not assess this cycle), were did not conduct analyses of these course types. SL courses (which align to *Ethical and Civic Thinking*) were not included in our sample this year.

Critical Thinking (CT) Courses

CT courses in the assessment sample included those that aligned to each of the outcomes assessed. First, we present results for *Creative Thinking*, assessed using the MU and AAC&U Value rubrics. All CT courses are at the 100/200 level. Results are below:

Creative Thinking (MU rubric)			**Creative Thinking (AAC&U rubric)**		
Trait	Number	Mean Score	**Trait**	Number	Mean Score
			Acquiring Competencies	38	1.58
Ambiguities and Possibilities	33	1.53	**Raking Risks**	37	1.46
			Solving Problems	32	1.45
Risk Taking	40	1.65	**Contradictions**	36	1.49
			Innovation	37	1.54
Synthesis and Innovation	35	1.27	**Connecting and Synthesizing**	37	1.37

These results must be interpreted with caution; however, overall means were 1.49 for the MU rubric and 1.48 for the AAC&U rubric.

Critical Thinking (CT) Courses

CT courses in the assessment sample included those that aligned to each of the outcomes assessed. Second, we present results for *Inquiry-Based Thinking*, assessed using the MU rubric and AAC&U *Critical Thinking* Value rubric. All CT courses are at the 100/200 level. Results are below:

Inquiry-Based Thinking (MU rubric)			**Critical Thinking (AAC&U rubric)**		
Trait	Number	Mean Score	**Trait**	Number	Mean Score
Problem/Question	21	1.69	**Explanation of Issues**	33	2.03
Research of Existing Knowledge	33	1.26	**Evidence**	38	1.9
Data Collection and Analysis	35	1.9	**Influence of Context and Assumptions**	28	1.39
Conclusions	37	2.23	**Student's Position**	39	1.41
			Conclusions and Related Outcomes	41	1.78

These results must be interpreted with caution; however, overall means were 1.79 for the MU rubric and 1.51 for the AAC&U rubric.

Critical Thinking (CT) Courses

CT courses in the assessment sample included those that aligned to each of the outcomes assessed. Third, we present results for *Quantitative Thinking*, assessed using the MU rubric and AAC&U *Quantitative Literacy* Value rubric. All CT courses are at the 100/200 level. Results are below:

Quantitative Thinking (MU rubric)			**Quantitative Literacy (AAC&U rubric)**		
Trait	Number	Mean Score	**Trait**	Number	Mean Score
			Interpretation	55	2.03
Context	62	1.46	**Representation**	56	2.00
Estimation	36	1.42	**Calculation**	51	2.38
			Application/Analysis	57	1.76
Visual Representation	37	1.49	**Assumptions**	37	1.49
Statistics	23	1.61	**Communication**	60	2.18

These results must be interpreted with caution; however, overall means were 1.48 for the MU rubric and 2.00 for the AAC&U rubric.

Core II Courses

Core II courses in the assessment sample included those that aligned to each of the outcomes assessed: *Creative Thinking, Inquiry-Based Thinking,* and *Quantitative Thinking.* First, we present results for *Creative Thinking*, assessed using the MU and AAC&U Value rubrics. All Core II courses are at the 100/200 level. Results are below:

Creative Thinking (MU rubric)			**Creative Thinking (AAC&U rubric)**		
Trait	Number	Mean Score	**Trait**	Number	Mean Score
			Acquiring Competencies	44	1.71
Ambiguities and Possibilities	37	1.45	**Raking Risks**	43	1.56
			Solving Problems	37	1.46
Risk Taking	46	1.72	**Contradictions**	37	1.53
			Innovation	43	1.61
Synthesis and Innovation	39	1.35	**Connecting and Synthesizing**	43	1.43

These results must be interpreted with caution; however, overall means were 1.52 for the MU rubric and 1.55 for the AAC&U rubric.

Core II Courses

Core II courses in the assessment sample included those that aligned to each of the outcomes assessed. Second, we present results for *Inquiry-Based Thinking*, assessed using the MU rubric and AAC&U *Critical Thinking* Value rubric. All CT courses are at the 100/200 level. Results are below:

Inquiry-Based Thinking (MU rubric)			**Critical Thinking (AAC&U rubric)**		
Trait	Number	Mean Score	**Trait**	Number	Mean Score
Problem/Question	19	1.68	**Explanation of Issues**	27	1.85
Research of Existing Knowledge	32	1.25	**Evidence**	36	1.79
Data Collection and Analysis	31	1.82	**Influence of Context and Assumptions**	28	1.45
Conclusions	33	2.12	**Student's Position**	34	1.34
			Conclusions and Related Outcomes	38	1.66

These results must be interpreted with caution; however, overall means were 1.72 for the MU rubric and 1.63 for the AAC&U rubric.

Core II Courses

Core II courses in the assessment sample included those that aligned to each of the outcomes assessed. Third, we present results for *Quantitative Thinking*, assessed using the MU rubric and AAC&U *Quantitative Literacy* Value rubric. All CT courses are at the 100/200 level. Results are below:

Quantitative Thinking (MU rubric)			**Quantitative Literacy (AAC&U rubric)**		
Trait	Number	Mean Score	**Trait**	Number	Mean Score
			Interpretation	53	2.02
Context	60	1.46	**Representation**	54	1.98
Estimation	36	1.42	**Calculation**	49	2.37
			Application/Analysis	55	1.74
Visual Representation	36	1.5	**Assumptions**	36	1.51
Statistics	23	1.59	**Communication**	58	2.20

These results must be interpreted with caution; however, overall means were 1.48 for the MU rubric and 1.99 for the AAC&U rubric.

First Year Seminar in Critical Thinking (FYS) Courses

FYS courses in the assessment sample included those that aligned to *Creative* and *Inquiry-Based* Thinking. First, we present results for *Creative Thinking*, assessed using the MU and AAC&U Value rubrics. All FYS courses are at the 100-level. Results are below:

Creative Thinking (MU rubric)			**Creative Thinking (AAC&U rubric)**		
Trait	Number	Mean Score	**Trait**	Number	Mean Score
			Acquiring Competencies	17	1.21
Ambiguities and Possibilities	16	1.38	**Raking Risks**	20	1.25
			Solving Problems	16	1.31
Risk Taking	18	1.25	**Contradictions**	16	1.38
			Innovation	19	1.45
Synthesis and Innovation	18	1.42	**Connecting and Synthesizing**	20	1.23

These results must be interpreted with caution; however, overall means were 1.35 for the MU rubric and 1.30 for the AAC&U rubric.

First Year Seminar in Critical Thinking (FYS) Courses

FYS courses in the assessment sample included those that aligned to *Creative* and *Inquiry-Based* Thinking. Second, we present results for *Inquiry-Based Thinking*, assessed using the MU rubric and the AAC&U *Critical Thinking* Value rubric. All FYS courses are at the 100-level. Results are below:

Inquiry-Based Thinking (MU rubric)			**Critical Thinking (AAC&U rubric)**		
Trait	**Number**	**Mean Score**	**Trait**	**Number**	**Mean Score**
Problem/Question	18	1.56	**Explanation of Issues**	22	1.75
Research of Existing Knowledge	18	0.97	**Evidence**	23	1.33
Data Collection and Analysis	18	1.19	**Influence of Context and Assumptions**	17	1.09
Conclusions	22	1.66	**Student's Position**	22	1.18
			Conclusions and Related Outcomes	22	1.43

These results must be interpreted with caution; however, overall means were 1.36 for the MU rubric and 1.37 for the AAC&U rubric.

Writing Intensive (WI) Courses

WI courses in the assessment sample aligned to all outcomes assessed: *Creative Thinking, Inquiry-Based Thinking,* and *Quantitative Thinking.* Results are given below by course level for *Creative Thinking*, using both the MU and AAC&U rubrics:

Creative Thinking (MU rubric)				**Creative Thinking (AAC&U rubric)**			
Trait	**Course Level**	**Number**	**Mean Score**	**Trait**	**Course Level**	**Number**	**Mean Score**
				Acquiring Competencies	100/200	31	1.69
					300/400	14	2.29
Ambiguities and Possibilities	100/200	31	1.52	**Raking Risks**	100/200	35	1.54
	300/400	12	1.21		300/400	12	1.92
				Solving Problems	100/200	30	1.53
					300/400	11	1.73
Risk Taking	100/200	32	1.78	**Contradictions**	100/200	24	1.67
	300/400	13	1.54		300/400	10	1.65
				Innovation	100/200	34	1.66
					300/400	15	1.83
Synthesis and Innovation	100/200	31	1.47	**Connecting and Synthesizing**	100/200	34	1.47
	300/400	15	1.53		300/400	14	1.68

These results must be interpreted with caution; however, overall means for the MU rubric were 1.59 (100/200 level) and 1.44 (300/400 level). Overall means for the AAC&U rubric were 1.59 (100/200 level) and 1.86 (300/400 level). While students in 300/400 level courses had a lower mean than those in 100/200 level courses using the MU rubric, the opposite was true using the AAC&U Value rubric.

WI results are given below by course level for *Inquiry-Based Thinking*:

Creative Thinking (MU rubric)				**Creative Thinking (AAC&U rubric)**			
Trait	**Course Level**	**Number**	**Mean Score**	**Trait**	**Course Level**	**Number**	**Mean Score**
Problem/Question	100/200	14	1.64	**Explanation of Issues**	100/200	16	1.78
	300/400	16	2.13		300/400	16	2.72
Research of Existing Knowledge	100/200	15	1.33	**Evidence**	100/200	18	1.94
	300/400	17	1.97		300/400	16	2.03
Data Collection and Analysis	100/200	15	1.93	**Influence of Context and Assumptions**	100/200	16	1.28
	300/400	16	2.28		300/400	11	1.86
Conclusions	100/200	17	2.21	**Student's Position**	100/200	17	1.27
	300/400	18	2.53		300/400	14	1.82

				Conclusions and Related Outcomes	100/200	19	1.66
					300/400	17	2.35

These results must be interpreted with caution; however, overall means for the MU rubric were 1.79 (100/200 level) and 2.23 (300/400 level). Overall means for the AAC&U rubric were 1.59 (100/200 level) and 2.19 (300/400 level). Students in 300/400 level courses had higher means than students in 100/200 level courses using both rubrics.

WI results are given below by course level for *Quantitative Thinking*:

Creative Thinking (MU rubric)				**Creative Thinking (AAC&U rubric)**			
Trait	**Course Level**	**Number**	**Mean Score**	**Trait**	**Course Level**	**Number**	**Mean Score**
				Interpretation	100/200	17	1.74
					300/400	6	2.42
Context	100/200	18	1.17	**Representation**	100/200	17	1.74
	300/400	7	1.71		300/400	6	2.42
Estimation	100/200	7	1.29	**Calculation**	100/200	10	2.35
	300/400	5	1.2		300/400	6	1.92
				Application/Analysis	100/200	17	1.59
					300/400	7	1.57
Visual Representation	100/200	12	1.33	**Assumptions**	100/200	11	1.32
	300/400	6	1.67		300/400	5	1.3
Statistics	100/200	6	1.33	**Conclusions**	100/200	18	2.06
	300/400	6	1.42		300/400	7	2.21

These results must be interpreted with caution; however, overall means for the MU rubric were 1.26 (100/200 level) and 1.52 (300/400 level). Overall means for the AAC&U rubric were 1.79 (100/200 level) and 1.99 (300/400 level). While students in 300/400 level scored higher than students in 100/200 level courses using both rubrics, the number of students in this sample was quite small.

Conclusion

Comparison of the MU and AAC&U rubrics for *Creative Thinking* showed similar overall means between rubrics; however the AAC&U rubric showed greater differentiation between 100/200 and 300/400 level course performance, especially for the traits acquiring competencies and taking risks.

Comparison of the MU rubric for *Inquiry-Based Thinking* and the AAC&U rubric for *Critical Thinking* showed that overall means were similar between the two, with both showing growth between 100/200 and 300/400 level courses. With the exception of context/assumptions, more artifacts aligned to the traits of the AAC&U rubric than to the MU rubric.

Comparison of the MU rubric for *Quantitative Thinking* and the AAC&U rubric for *Quantitative Literacy* showed higher means for the AAC&U than for the MU rubric. Moreover, more artifacts aligned to the traits of the AAC&U than to those of the MU rubric, with two notable exceptions – the MU trait context (to which 75 artifacts aligned) and the AAC&U trait assumptions (to which only 45 artifacts aligned).

The highest overall mean score for 100/200 level courses was 1.98 using the AAC&U *Quantitative Literacy* rubric, while the highest overall mean score for 300/400 level courses was 2.32, also using the AAC&U *Quantitative Literacy* rubric. The three rubrics that showed the greatest growth between 100/200 and 300/400 level courses were the AAC&U *Critical Thinking* rubric (0.57 point increase), the MU *Inquiry-Based Thinking* rubric (0.54 point increase), and the AAC&U *Creative Thinking* rubric (0.41 point increase). The smallest difference was for the MU *Creative Thinking* rubric, which showed no increase between 100/200 and 300/400 level courses.

Recommendations from the 2018 Assessment Workgroup

Recommendations Concerning the General Process of Assignment Creation and Accurate Alignment to University Outcomes

1. As it did last year, the Assessment Workgroup reiterated the need for instructors to include assignment instructions that clearly specify how their assignments align to the BDP outcomes to which they are tagged. Mary Beth will work with the MU Online Design Center staff to increase the number of faculty including these instructions during academic year 2018-2019.
2. The Assessment Workgroup recommended that, before beginning the evaluation of artifacts next year, the group take time to review each assignment with artifacts in the sample and collectively determine which outcome traits to which each assignment aligns. Reaching consensus before artifact assessment begins should help to reduce the number of disagreements regarding whether or not individual artifacts align with specific outcome traits. They also recommended that we discuss applications/interpretations of traits among the disciplines, especially context (quantitative), assumptions (critical thinking), ambiguities and risk taking (creative), and data collection (inquiry).
3. The Assessment Workgroup recommended having a discussion prior to scoring about the relevance of assignment instructions to assessors' interpretations of traits, as well as assessors' own assumptions.
4. The Assessment Workgroup recommended that we consider setting up the collections in Blackboard so that instructors can align their assignments with specific traits of an outcome.
5. The Assessment Workgroup echoed last year's recommendation that students include process papers with artifact uploads. The Assessment Office and Online Design Center staff will work on a communication strategy to ensure the best strategy for making this happen.

6. Due to lack of time, the Assessment Workgroup will discuss tentative recommendations to more carefully compare results of the paired rubrics used for this year's assessment. Basic findings and questions to be discussed include:
   - Overall, the AAC&U *Quantitative Literacy* Value rubric's traits aligned to more artifacts than did the MU *Quantitative Thinking* rubric. Additionally, using the same artifacts, mean scores were higher on the AAC&U than on the MU rubric. We note that all of the AAC&U rubrics have undergone extensive validation in institutions of higher education across the country.
   - The AAC&U *Creative Thinking* Value rubric was better able to show growth in performance between 100/200 and 300/400 level courses than was the MU *Creative Thinking* rubric. Also, we noted that 50% of artifacts uploaded by students from courses from the College of Arts and Media aligned to *Creative Thinking* were judged by assessors to be misaligned to all traits of the MU rubric, while only 28% were judged to be misaligned to all traits of the AAC&U rubric. Since students in the disciplines of this college produce creative works, we question the applicability of either rubric, but especially MU's *Creative Thinking* rubric to all disciplines within the university.
   - Although both the MU *Inquiry-Based Thinking* rubric and the AAC&U *Critical Thinking* Value rubric showed growth in performance between 100/200 and 300/400 level courses, more artifacts aligned to the traits of the AAC&U rubric than to those of the MU rubric. We note that the MU *Inquiry-Based Thinking* rubric is designed more specifically for scientific research, whereas the AAC&U *Critical Thinking* rubric might be applicable to a wider variety of disciplines. Although the numbers were not large, a greater percentage of FYS artifacts (19%) were judged to be misaligned with all traits of both the MU and AAC&U rubrics, while 11% of Liberal Arts artifacts were judged to be misaligned with all traits of the MU rubric (but only 5% of Liberal arts artifacts to all traits of the AAC&U rubric).

Recommendations Concerning the Blackboard Outcomes Assessment Tool

The following items are issues that we will ask Blackboard to address; however, we understand that Blackboard is a large company with many clients and must prioritize improvements to the product. So, while we are hopeful that many of our concerns will be addressed, we realize that addressing them all may take some time.

1. As we have always done, we used an assessment process where each artifact is independently reviewed by two reviewers. Blackboard has not developed an algorithm that will allow each assessor to be randomly paired with each other assessor during this process. Mary Beth will contact Blackboard again about this issue as we feel that the simplistic nature of having each person paired with the same two partners for all evaluations prejudices the objectivity of their evaluations. In other words, it becomes too easy for them to try to anticipate what their partners might think.
2. During the assessment process this year, the use of two rubrics made the time required to assess each artifact longer than last year. It appeared that, for some individuals, Blackboard "timed them out," and did not save their ratings. Mary Beth will contact Blackboard to try to resolve this issue.
3. During our assessment cycle, each assessor's artifact queue disappeared upon completion of scoring. Although we would like to have this issue corrected and Mary Beth will contact Blackboard about it, assessors can get into their queues for later reconciliation as long as they

save the email links that Blackboard generates for them at the beginning of the assessment process. We will be sure to emphasize saving these links until the end of the project during future assessment projects.

4. Artifacts did not have unique identifiers in the data download. Rather, each student had an anonymized identifier. Since the same students sometimes submit more than one artifact, we would prefer to have unique artifact identifiers. Although we will request the addition of unique artifact identifiers, we were able to identify students with multiple artifacts and create unique identifiers for each of their artifacts.

Supporting Documentation

General Education
Blackboard Artifact Assessment

Academic Year 2017 – 2018

Outcomes Assessed: MU Rubrics

Outcome	Abbreviation	Traits	Abbreviations
Creative Thinking	**CRT**	Ambiguities and Possibilities	A & P
		Risk Taking	Risk
		Synthesis/Innovation	Innovation
Inquiry-Based Thinking	**IBT**	Problem/Question	Question
		Research of Existing Knowledge	Knowledge
		Data Collection and Analysis	Analysis
		Conclusions	Conclusions
Quantitative Thinking	**QT**	Context	Context
		Estimation	Estimation
		Visual Representation	Visual
		Statistics	Statistics

Outcomes Assessed: AAC&U Rubrics

Outcome	Abbreviation	Traits	Abbreviations
Creative Thinking	**CRT**	Acquiring Competencies	Competencies
		Taking Risks	Risks
		Solving Problems	Solving Problems
		Embracing Contradictions	Contradictions
		Innovative Thinking	Innovation
		Connecting, Synthesizing, Transforming	Connecting
Inquiry-Based (Critical) Thinking	**IBT**	Explanation of Issues	Issues
		Evidence	Evidence
		Influence of Context and Assumptions	Context/Assumptions
		Student's Position	Position
		Conclusions and Related Outcomes	Conclusions
Quantitative Thinking/Literacy	**QT**	Interpretation	Interpretation
		Representation	Representation
		Calculation	Calculation
		Application/Analysis	Application/Analysis
		Assumptions	Assumptions
		Communication	Communication

Course Types

Course Type	Abbreviation
First Year Seminar in Critical Thinking	FYS
Critical Thinking	CT
Multicultural	MC
International	INT
Writing Intensive	WI
Service Learning	SL (Not included in this year's assessment)
Core II	Core II

Course Types in CRT, IBT, and QT Outcome Population

Each Course Counted Separately for Each Category

(i.e. sample *n* does not add to 324)

Course Type	Population *n*	Sample *n*	Percent
FYS	1,186	59	5.0%
CT	2,775	176	6.3%
MC	1,141	65	5.7%
INT	186	11	5.9%
WI	2,113	130	4.6%
Core II	2,851	184	6.5%
Total	**10,252**	**625**	**6.1%**

Population/Sample Comparisons for Marshall's Course Types by Course Level

Each Course Counted Separately for Each Category

(i.e. sample *n* does not add to 324)

Course Type	Course Level = 100/200			Course Level = 300/400		
	Population	Sample	Percent	Population	Sample	Percent
FYS	1,186	59	5.0%	0	0	0%
Critical Thinking	2,775	176	6.3%	0	0	0%
Multicultural	1,061	61	5.7%	80	4	5.0%
International	91	5	5.5%	95	6	6.3%
Writing Intensive	1,183	79	6.7%	930	51	5.5%
Core II	2,851	184	6.5%	0	0	0%
Total	**9,147**	**564**	**6.2%**	**1,105**	**61**	**5.5%**

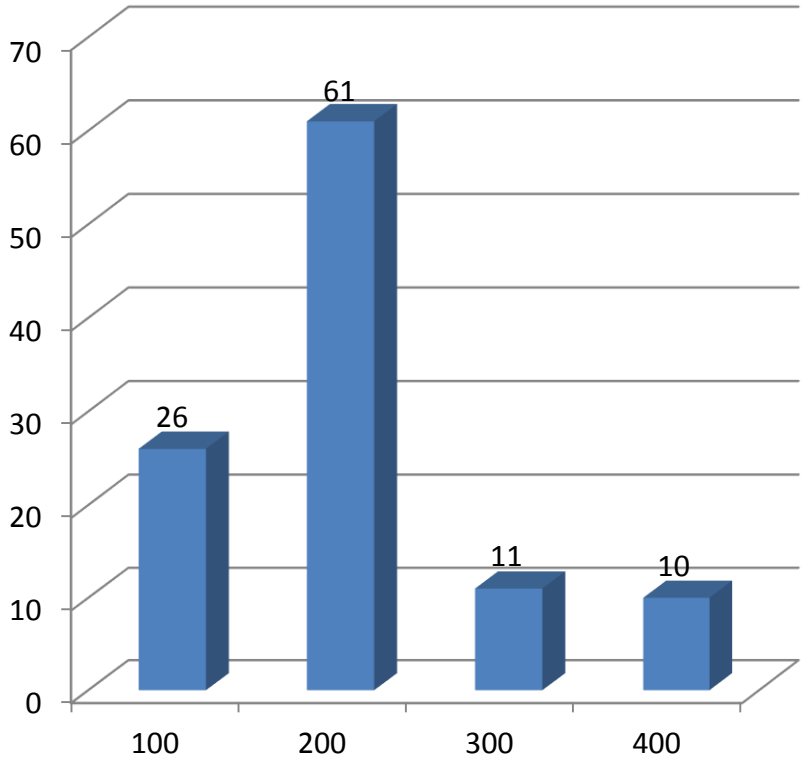
Population/Sample Comparisons for Marshall's Learning Outcomes by Course Level

Marshall Outcomes	Course Level = 100/200			Course Level = 300/400		
	Population	**Sample**	**Percent**	**Population**	**Sample**	**Percent**
Creative Thinking	1,129	87	7.7%	284	21	7.4%
Inquiry-Based Thinking	2,202	85	3.9%	519	23	4.4%
Quantitative Thinking	1,162	95	8.2%	222	13	5.9%
Total	**4,493**	**267**	**5.9%**	**1,025**	**57**	**5.6%**

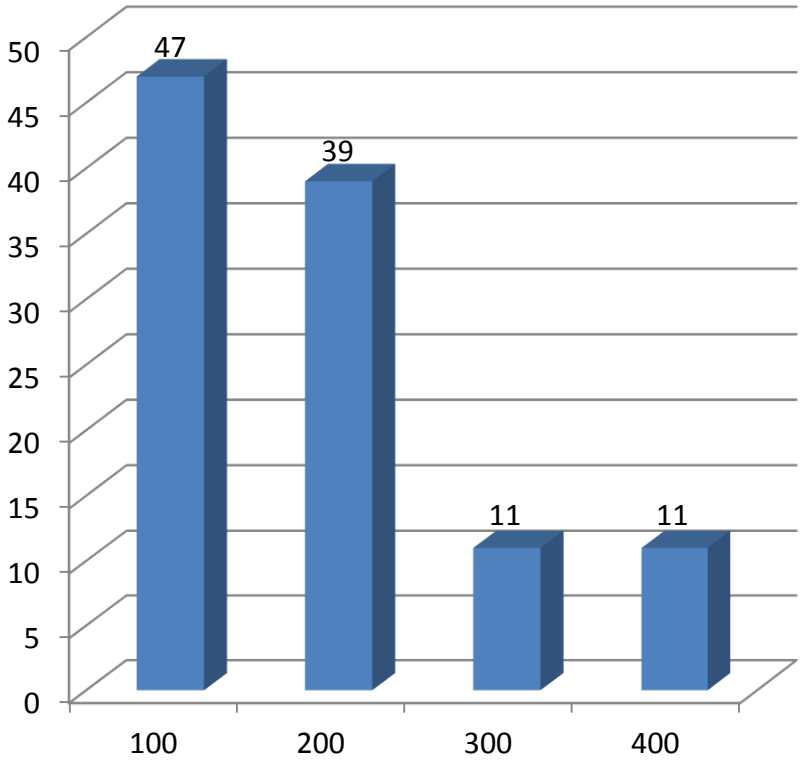
Sample Frequencies

Total # of artifacts assessed = 108 per outcome

Course Level Frequencies: Creative Thinking



Course Level Frequencies: Inquiry-Based Thinking

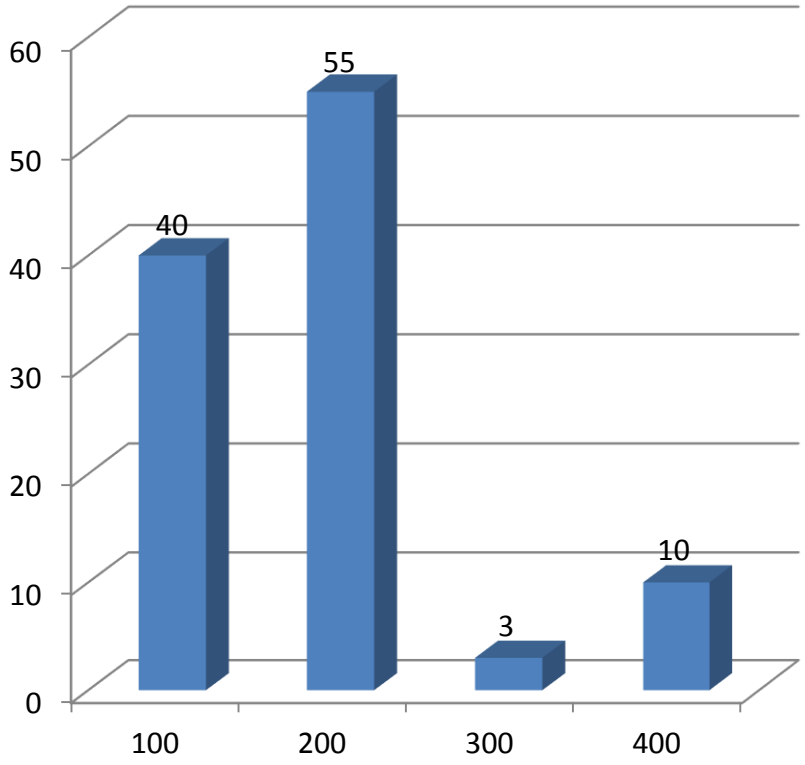


Sample Frequencies

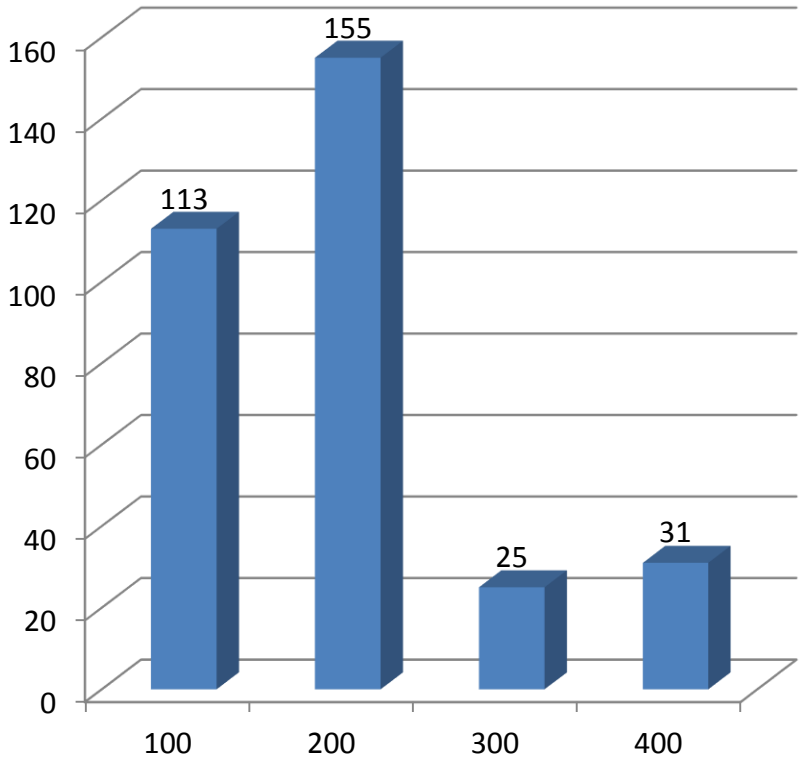
Total # of artifacts assessed = 108 per outcome

Total = 324

Course Level Frequencies: Quantitative Thinking



Course Level Frequencies: Total across the three outcomes



Review Procedures

- Each artifact had two independent raters and usable scores on the 0 – 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 or less, 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.

Review Procedures

- We also allowed reviewers to assign a score of *N/A (not applicable*) when they judged the assignment to not be aligned with the outcome trait, or a score of *Error* if there was a student upload error or other technical issue which prevented the reviewers from scoring the artifact. When one rater assigned a score of *N/A* and the second rater assigned a score of 0 – 4, they also met to discuss the rationale for their scores to see if they could agree on a score or, at a minimum, scores on the 0 – 4 scale that differed by not more than one point. If they could not agree, a third reader was assigned.

Third Readers for this Year's Review

We used two rubrics for each artifact, resulting in 648 rubrics (two sets for each of our 324 artifacts). Of these, we had 24 rubrics (with a total of 33 traits) that required a third review. In each case, the disagreement was between a score of either a score of *N/A* and a score on the 0 – 4 scale or between two scores on the 0 – 4 scale. All but four of the traits (which were eliminated from the analysis) were able to be reconciled with the third reader.

Artifacts Excluded from Analysis of Means Due to Inability to Assess or Misalignment with Tagged Outcomes

Outcome	Total Artifacts		Total Artifacts Eliminated Due to Error		Total Artifacts Eliminated due to Misalignment		Total Used for Analysis	
	MU	**AAC&U**	**MU**	**AAC&U**	**MU**	**AAC&U**	**MU**	**AAC&U**
Creative Thinking	108	108	1	1	18	11	89	96
Inquiry-Based Thinking	108	108	3	3	12	10	93	95
Quantitative Thinking	108	108	1	1	30	27	77	80
Total	**108**	**108**	**5**	**5**	**60**	**48**	**259**	**271**

Revised Creative Thinking MU Rubric

Creative Thinking: Students will **outline** multiple divergent solutions to a problem, **explore** and **develop** risky or controversial ideas, and **synthesize** ideas/expertise to **generate** innovations.

Traits: Performance Indicators/Performance Levels	Level 0	Level 1	Level 2	Level 3	Level 4
Ambiguities & Possibilities: **Outlines** multiple divergent solutions to a problem.	Does not **outline** solutions to a given problem.	**Outlines** a single solution to a problem, either feasible or infeasible.	**Outlines** some solutions, although not all might be divergent and/or feasible. (2 ideas)	**Outlines** multiple divergent and feasible solutions to a problem. (more than 2 ideas)	**Outlines** multiple divergent and feasible solutions to a problem and considers the potential pros and cons of each solution.
Risk Taking: **Explores** and **develops** risky or controversial ideas.	Does not **explore or develop** risky or controversial ideas.	**Explores,** but does not **develop** risky or controversial ideas.	**Explores** risky or controversial ideas and **develops** these ideas, but only in a superficial manner.	**Explores** risky or controversial ideas, and **develops** these ideas in some depth.	**Explores** risky or controversial ideas, and thoroughly **develops** these ideas.
Innovation: **Synthesizes** ideas/expertise to **generate** innovations.	Does not **synthesize** ideas/expertise or **generate** innovations.	**Demonstrates** rudimentary ability to **synthesize** ideas, but this synthesis does not result in innovations.	**Synthesizes** similar ideas and methods to **generate** an innovation.	**Synthesizes** divergent ideas and methods to **generate** an innovation.	**Synthesizes** by **elaborating or expanding** on divergent ideas and methods to **generate** an innovation.

Creative Thinking AAC&U Value Rubric

AAC&U Creative Thinking Value Rubric

Traits	Level 0	Level 1	Level 2	Level 3	Level 4
Acquiring Competencies *This step refers to acquiring strategies and skills within a particular domain.*	Does not meet level 1	Model: Successfully reproduces an appropriate exemplar.	Adapt: Successfully adapts an appropriate exemplar to his/her own specifications.	Create: Creates an entirely new object, solution or idea that is appropriate to the domain.	Reflect: Evaluates creative process and product using domain-appropriate criteria.
Taking Risks *May include personal risk (fear of embarrassment or rejection) or risk of failure in successfully completing assignment, i.e. going beyond original parameters of assignment, introducing new materials and forms, tackling controversial topics, advocating unpopular ideas or solutions.*	Does not meet level 1	Stays strictly within the guidelines of the assignment.	Considers new directions or approaches without going beyond the guidelines of the assignment.	Incorporates new directions or approaches to the assignment in the final product.	Actively seeks out and follows through on untested and potentially risky directions or approaches to the assignment in the final product.
Solving Problems	Does not meet level 1	Only a single approach is considered and is used to solve the problem.	Considers and rejects less acceptable approaches to solving problem.	Having selected from among alternatives, develops a logical, consistent plan to solve the problem.	Not only develops a logical, consistent plan to solve problem, but recognizes consequences of solution and can articulate reason for choosing solution.
Embracing Contradictions	Does not meet level 1	Acknowledges (mentions in passing) alternate, divergent, or contradictory perspectives or ideas.	Includes (recognizes the value of) alternate, divergent, or contradictory perspectives or ideas in a small way.	Incorporates alternate, divergent, or contradictory perspectives or ideas in a exploratory way.	Integrates alternate, divergent, or contradictory perspectives or ideas fully.
Innovative Thinking *Novelty or uniqueness (of idea, claim, question, form, etc.)*	Does not meet level 1	Reformulates a collection of available ideas.	Experiments with creating a novel or unique idea, question, format, or product.	Creates a novel or unique idea, question, format, or product.	Extends a novel or unique idea, question, format, or product to create new knowledge or knowledge that crosses boundaries.
Connecting, Synthesizing, Transforming	Does not meet level 1	Recognizes existing connections among ideas or solutions.	Connects ideas or solutions in novel ways.	Synthesizes ideas or solutions into a coherent whole.	Transforms ideas or solutions into entirely new forms.

Revised Inquiry-Based Thinking MU Rubric

Inquiry-Based Thinking: Students will **formulate** focused questions and/or hypotheses, **evaluate** existing knowledge, **collect** and **analyze** data, and **draw** justifiable conclusions.

Traits: Performance Indicators/Performance Levels	Level 0	Level 1	Level 2	Level 3	Level 4
Problem/Question: **Formulates** focused questions and/or hypotheses.	Does not **formulate** focused questions or hypotheses.	**Formulates** a question and/or hypothesis, but not one that is necessarily focused or manageable.	**Formulates** a question and/or hypothesis that is focused and manageable.	**Formulates** a question and/or hypothesis that is focused and manageable and addresses a potentially significant area of inquiry.	**Formulates** a focused, and manageable question and/or hypothesis that addresses significant yet less-explored aspects of the topic.
Research of Existing Knowledge: **Evaluates** existing knowledge.	Does not **evaluate** existing knowledge.	**Evaluates** some existing research relevant to the problem/question, but only includes those that support one side of an issue or includes information from some questionable sources.	**Evaluates** some existing research relevant to the problem/question from reputable sources. The review is balanced but not comprehensive.	**Uses** reputable sources to **conduct** a comprehensive evaluation of existing research relevant to the problem/question.	**Evaluates and synthesizes** in-depth relevant information from reputable sources representing various points of view/approaches.
Data Collection and Analysis: **Collects** and **analyzes** data.	Neither **collects** nor **analyzes** the data.	**Collects** but does not **analyze** the data.	**Collects** but incompletely **analyzes** the data.	Thoroughly **analyzes** the data.	Thoroughly **analyzes** and **synthesizes** the data.
Conclusions: **Draws** justifiable conclusions.	Does not **draw** conclusions.	Conclusions neither **address** the question and/or hypothesis nor are supported by the data.	Conclusions either **address** the question and/or hypothesis or are supported by the data.	Conclusions both **address** the question and/or hypothesis and are supported by the data.	**Fulfills level 3 plus suggests** how results might apply to other problems or inform future studies.

Critical Thinking AAC&U Value Rubric

Critical Thinking QAAC&U Value Rubric

Traits	Level 0	Level 1	Level 2	Level 3	Level 4
Explanation of issues	Does not meet Level 1	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 meet Level 1	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 meet Level 1	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 meet Level 1	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 meet Level 1	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.

Revised Quantitative Thinking MU Rubric

Quantitative Thinking: Students will **analyze** real-world problems quantitatively, **formulate** plausible estimates, **assess** the validity of visual representations of quantitative information, and **differentiate** valid from questionable statistical conclusions.

Traits: **Performance Indicators**/**Performance Levels**	**Level 0** **Can't do it**	**Level 1** **recognize**	**Level 2** **use**	**Level 3** **apply**	**Level 4** **create**
Context: Analyzes real-world problems quantitatively.	Does not **explain**, **report**, or **analyze** real-world problems quantitatively.	**Explains** and **reports** the problem within its context quantitatively. **Identifies** basic metrics to solve the problem.	**Level 1 plus uses** appropriate tools to analyze metrics to **solve** problems in a given context.	**Level 2 plus articulates** meanings of a quantitative analysis.	**Develops** metrics, **uses** appropriate tools, and **applies** solutions to **solve** novel problems.
Estimation: **Formulates** plausible estimates.	Does not **formulate** plausible estimates.	**Recognizes** applicable units, orders of magnitude and appropriate mathematical tools.	**Level 1 plus uses** tools to develop the estimate to solve a problem.	**Level 2 plus identifies** the limitations of the estimated solution as applied to a specified problem.	**Level 3 plus Identifies** how the purpose of the estimate affects the parameters of the solution.
Visual Representation: **Assesses** the validity of visual representations of quantitative information.	Does not **assess, summarize**, or **explain** the validity of visual representations of quantitative information.	**Identifies** which visual representation is applicable to construct the model of a problem.	**Level 1 plus constructs** the visual representation of a model appropriate to the problem.	**Level 2 plus recognizes** the limitations of how the visual representation of the model applies to the problem.	**Level 3 plus identifies, constructs, and recognizes** the applicability of the visual representation of the model of the problem.
Statistics: Differentiates valid from questionable statistical conclusions.	Is not able to explain basic statistical terms/concepts.	**Recognize** types of data appropriate to construct statistical models of the problem.	**Level 1 plus uses** appropriate statistical tools to **estimate** models of the problem.	**Level 2 plus interprets** and **applies** solutions derived from statistical models of problem.	**Independently develops** data, **uses** appropriate tools, **interprets** and **applies** results for a problem.

Quantitative Literacy AAC&U Value Rubric

Quantitative Literacy AAC&U Value Rubric

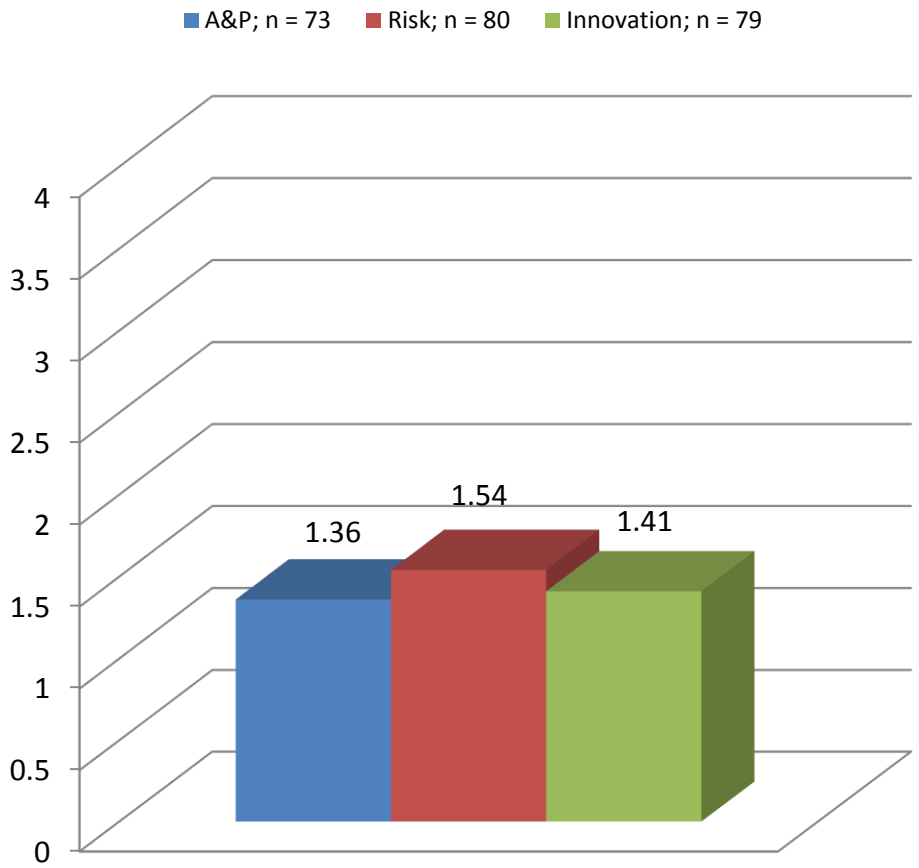
Traits	Level 0	Level 1	Level 2	Level 3	Level 4
Interpretation *Ability to explain information presented in mathematical forms (e.g., equations, graphs, diagrams, tables, words)*	Does not meet Level 1 Expectations	Attempts to explain information presented in mathematical forms, but draws incorrect conclusions about what the information means. *For example, attempts to explain the trend data shown in a graph, but will frequently misinterpret the nature of that trend, perhaps by confusing positive and negative trends.*	Provides somewhat accurate explanations of information presented in mathematical forms, but occasionally makes minor errors related to computations or units. *For instance, accurately explains trend data shown in a graph, but may miscalculate the slope of the trend line.*	Provides accurate explanations of information presented in mathematical forms. *For instance, accurately explains the trend data shown in a graph.*	Provides accurate explanations of information presented in mathematical forms. Makes appropriate inferences based on that information. *For example, accurately explains the trend data shown in a graph and makes reasonable predictions regarding what the data suggest about future events.*
Representation *Ability to convert relevant information into various mathematical forms (e.g., equations, graphs, diagrams, tables, words)*	Does not meet Level 1 Expectations	Completes conversion of information but resulting mathematical portrayal is inappropriate or inaccurate.	Completes conversion of information but resulting mathematical portrayal is only partially appropriate or accurate.	Competently converts relevant information into an appropriate and desired mathematical portrayal.	Skillfully converts relevant information into an insightful mathematical portrayal in a way that contributes to a further or deeper understanding.
Calculation	Does not meet Level 1 Expectations	Calculations are attempted but are both unsuccessful and are not comprehensive.	Calculations attempted are either unsuccessful or represent only a portion of the calculations required to comprehensively solve the problem.	Calculations attempted are essentially all successful and sufficiently comprehensive to solve the problem.	Calculations attempted are essentially all successful and sufficiently comprehensive to solve the problem. Calculations are also presented elegantly (clearly, concisely, etc.)
Application / Analysis *Ability to make judgments and draw appropriate conclusions based on the quantitative analysis of data, while recognizing the limits of this analysis*	Does not meet Level 1 Expectations	Uses the quantitative analysis of data as the basis for tentative, basic judgments, although is hesitant or uncertain about drawing conclusions from this work.	Uses the quantitative analysis of data as the basis for workmanlike (without inspiration or nuance, ordinary) judgments, drawing plausible conclusions from this work.	Uses the quantitative analysis of data as the basis for competent judgments, drawing reasonable and appropriately qualified conclusions from this work.	Uses the quantitative analysis of data as the basis for deep and thoughtful judgments, drawing insightful, carefully qualified conclusions from this work.
Assumptions *Ability to make and evaluate important assumptions in estimation, modeling, and data analysis*	Does not meet Level 1 Expectations	Attempts to describe assumptions.	Explicitly describes assumptions.	Explicitly describes assumptions and provides compelling rationale for why assumptions are appropriate.	Explicitly describes assumptions and provides compelling rationale for why each assumption is appropriate. Shows awareness that confidence in final conclusions is limited by the accuracy of the assumptions.
Communication *Expressing quantitative evidence in support of the argument or purpose of the work (in terms of what evidence is used and how it is formatted, presented, and contextualized)*	Does not meet Level 1 Expectations	Presents an argument for which quantitative evidence is pertinent, but does not provide adequate explicit numerical support. (May use quasi-quantitative words such as "many," "few," "increasing," "small," and the like in place of actual quantities.)	Uses quantitative information, but does not effectively connect it to the argument or purpose of the work.	Uses quantitative information in connection with the argument or purpose of the work, though data may be presented in a less than completely effective format or some parts of the explication may be uneven.	Uses quantitative information in connection with the argument or purpose of the work, presents it in an effective format, and explicates it with consistently high quality.

Creative Thinking: Overall Analysis

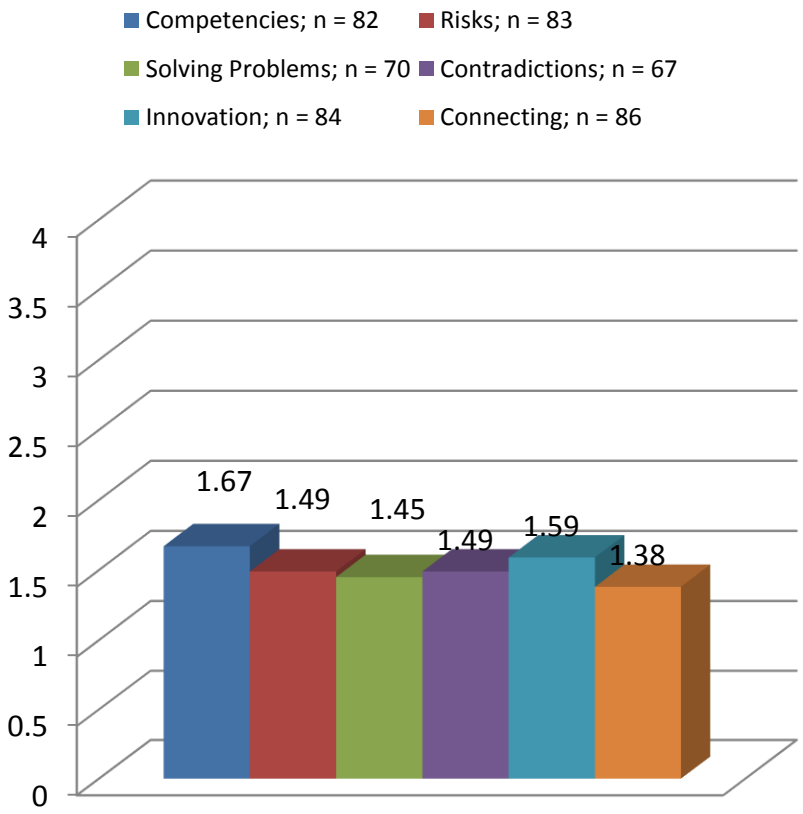
Mean Scores on a scale of 0 – 4, with 4 being the highest possible score

(Although there were 89 CRT artifacts in the MU rubric analysis and 96 in the AAC&U rubric analysis, not all artifacts aligned to every trait)

MU Rubric



AAC&U Rubric



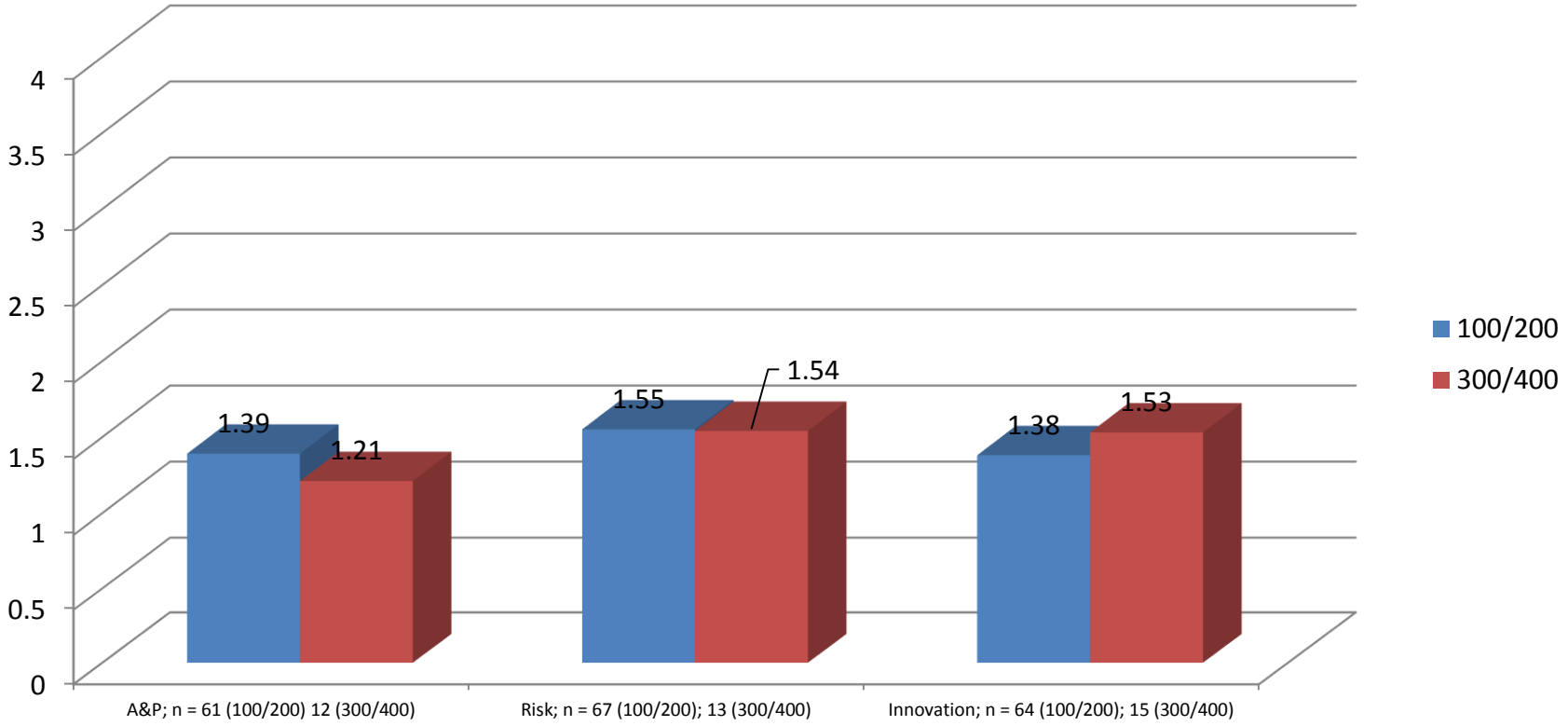
Creative Thinking: Course Level Analysis

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score

Mean differences for each trait at the course level were not significant.

Overall means for Creative Thinking were 1.44 for both 100/200 and 300/400 level courses.

MU Rubric

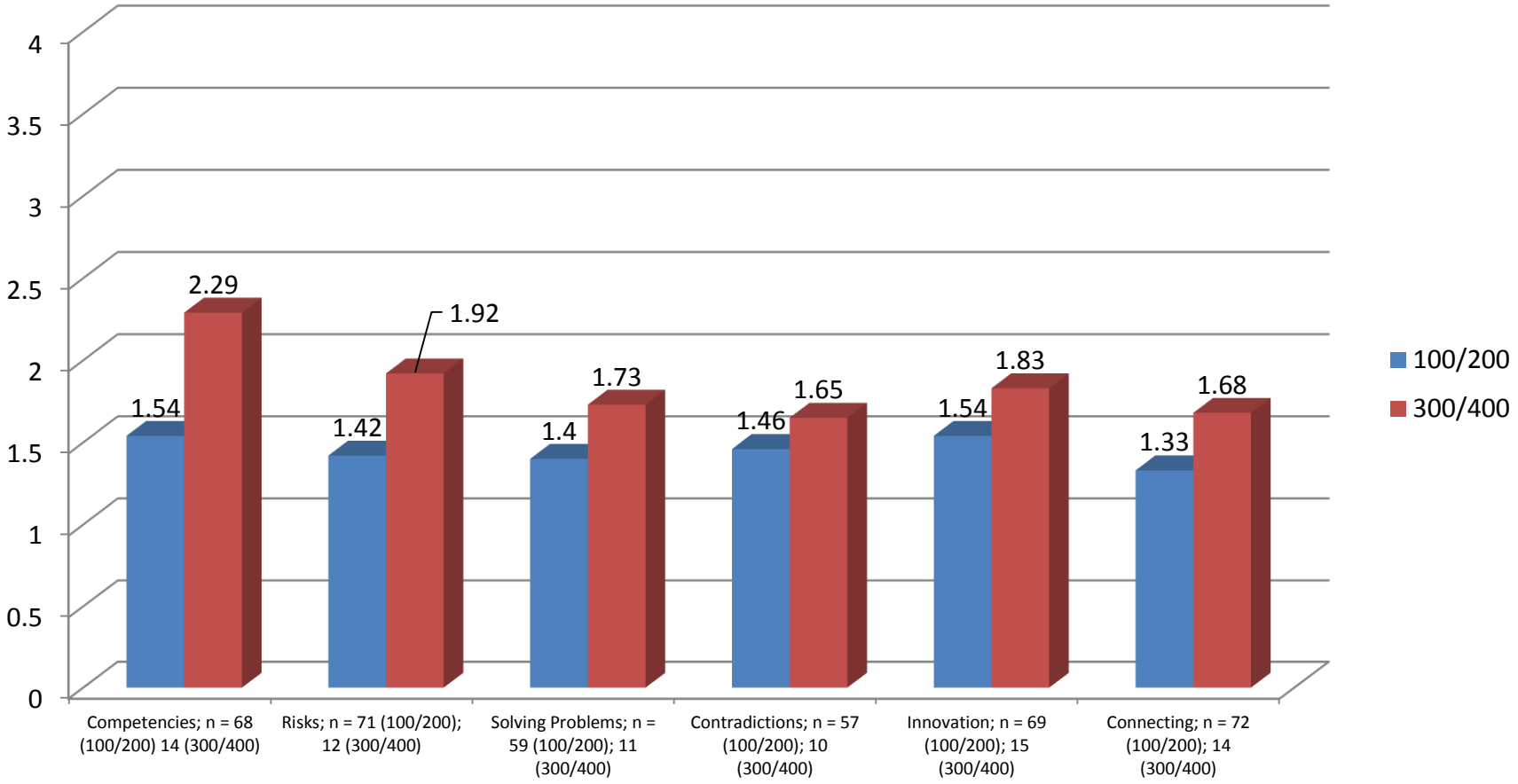


Creative Thinking: Course Level Analysis

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score

Mean scores for *competencies* and *risks* were significantly higher for 300/400 than for 100/200 level courses.
Overall means for Creative Thinking were 1.45 for 100/200 level courses and 1.86 for 300/400 level courses.

AAC&U Rubric

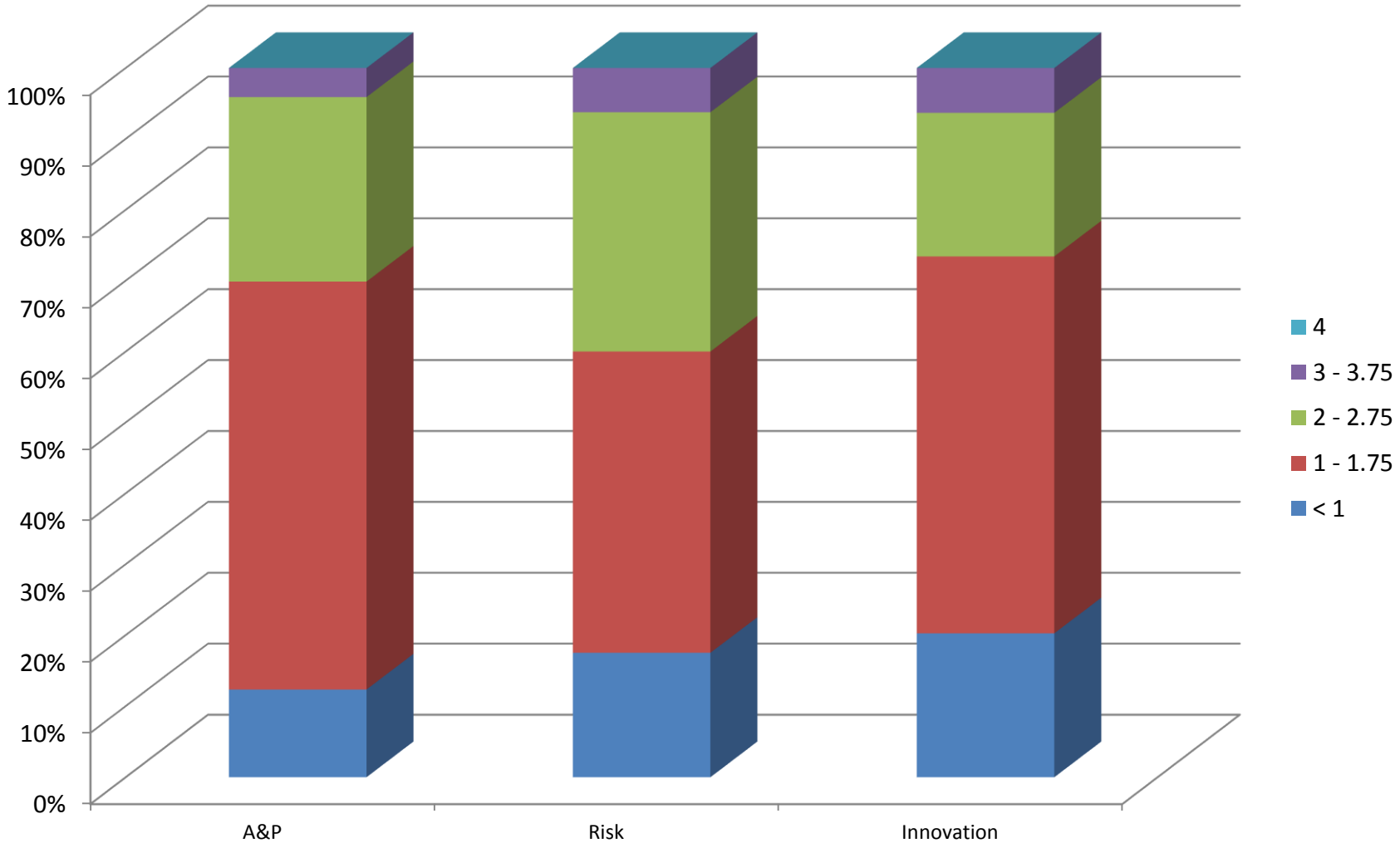


Creative Thinking (MU Rubric)

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

Trait/ Performance Level	A&P	Risk	Innovation	Total
< 1	9 (12%)	14 (18%)	16 (20%)	39 (17%)
1 – 1.75	42 (58%)	34 (43%)	42 (53%)	118 (51%)
2 – 2.75	19 (26%)	27 (34%)	16 (20%)	62 (27%)
3 – 3.75	3 (4%)	5 (6%)	5 (6%)	13 (6^)
4	0	0	0	0
Total Tags with Usable Scores	**73 (100%)**	**80 (100%)**	**79 (100%)**	**232 (100%)**

Creative Thinking (MU Rubric)

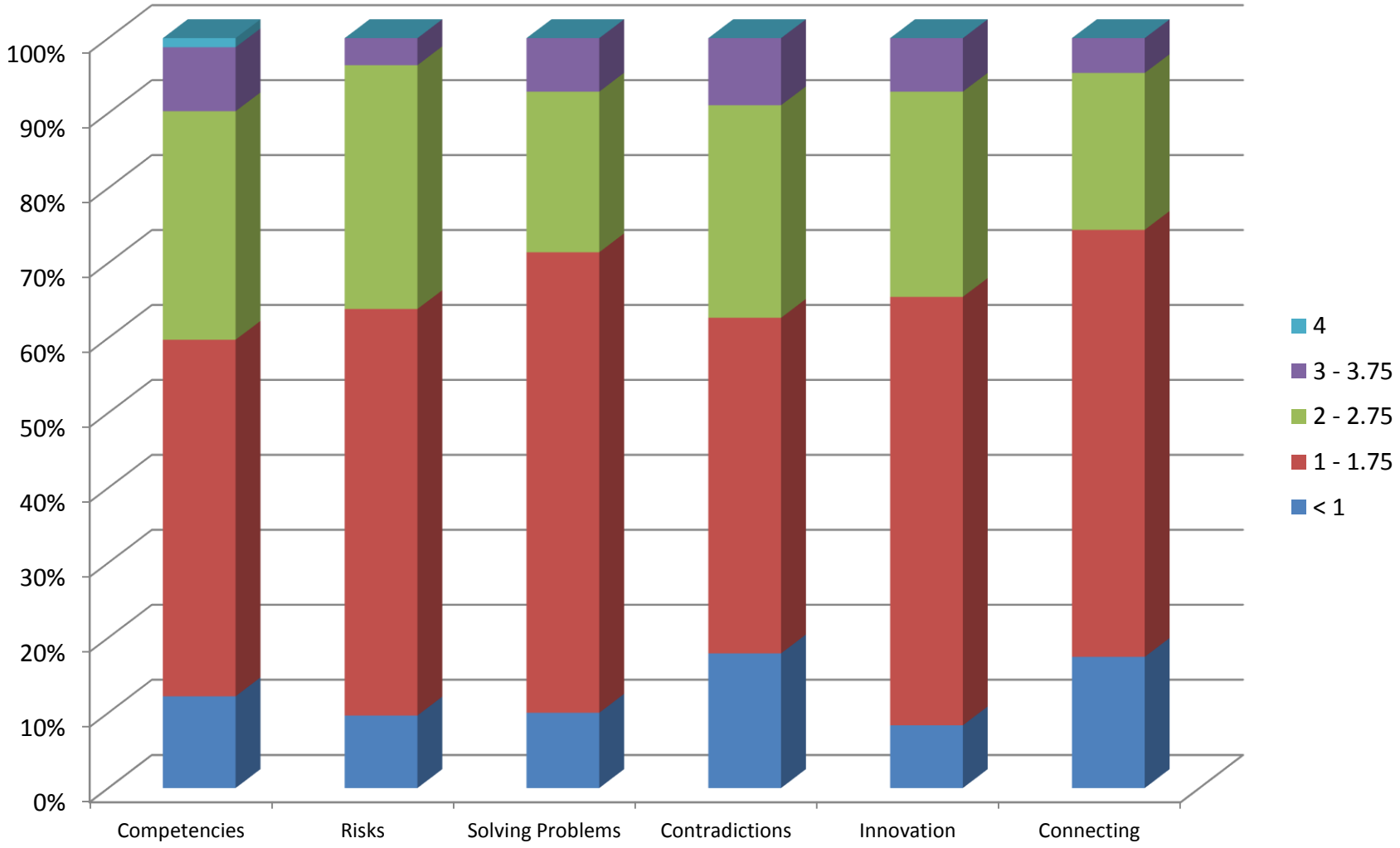


Creative Thinking (AAC&U Rubric)

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

Trait/ Performance Level	Acquiring Competencies	Taking Risks	Solving Problems	Contradictions	Innovation	Connecting	Total
< 1	10 (12%)	8 (10%)	7 (10%)	12 (18%)	7 (8%)	15 (17%)	59 (13%)
1 – 1.75	39 (48%)	45 (54%)	43 (61%)	30 (45%)	48 (57%)	49 (57%)	254 (54%)
2 – 2.75	25 (30%)	27 (33%)	15 (21%)	19 (28%)	23 (27%)	18 (21%)	127 (27%)
3 – 3.75	7 (9%)	3 (4%)	5 (7%)	6 (9%)	6 (7%)	4 (5%)	31 (7%)
4	1 (1%)	0	0	0	0	0	1 (0%)
Total Tags with Usable Scores	**82 (100%)**	**83 (100%)**	**70 (100%)**	**67 (100%)**	**84 (100%)**	**86 (100%)**	**472 (100%)**

Creative Thinking (AAC&U Rubric)



Creative Thinking (MU)

Inter-Rater Agreement Results

Trait/ Performance Level	A&P; Kappa = .307 (All scores); Kappa = .334 (Not aligned , unable to score, and one rater score missing excluded) Kappa Liberal = .548 (All Scores); Kappa Liberal = .822 (Exclusions Noted Above)	Risk; Kappa = .122 (All scores); Kappa = .062 (Not aligned , unable to score, and one rater score missing excluded) Kappa Liberal = .524 (All Scores); Kappa Liberal = .759 (Exclusions Noted Above)	Innovation; .125 (All scores); Kappa = .056 (Not aligned , unable to score, and one rater score missing excluded) Kappa Liberal = .516 (All Scores); Kappa Liberal = .764 (Exclusions Noted Above)
Agree on score	33 (31%)	20 (19%)	21 (19%)
Difference = 1 point or less	17 (16%)	31 (29%)	29 (27%)
Difference = 1.5 to 2 points	7 (6%)	11 (10%)	12 (11%)
Difference = > 2 points	1 (1%)	2 (2%)	0
Agree on Not Aligned	18 (17%)	12 (11%)	13 (12%)
Agree on Unable to Score due to error	1 (1%)	1 (1%)	1 (1%)
Score + Not Aligned	31 (29%)	31 (29%)	31 (29%)
Misaligned + Missing Second Rater Score	0	0	1 (1%)
Total	108 (100 %)	108 (100%)	108 (100%)

Creative Thinking (AAC&U)

Inter-Rater Agreement Results

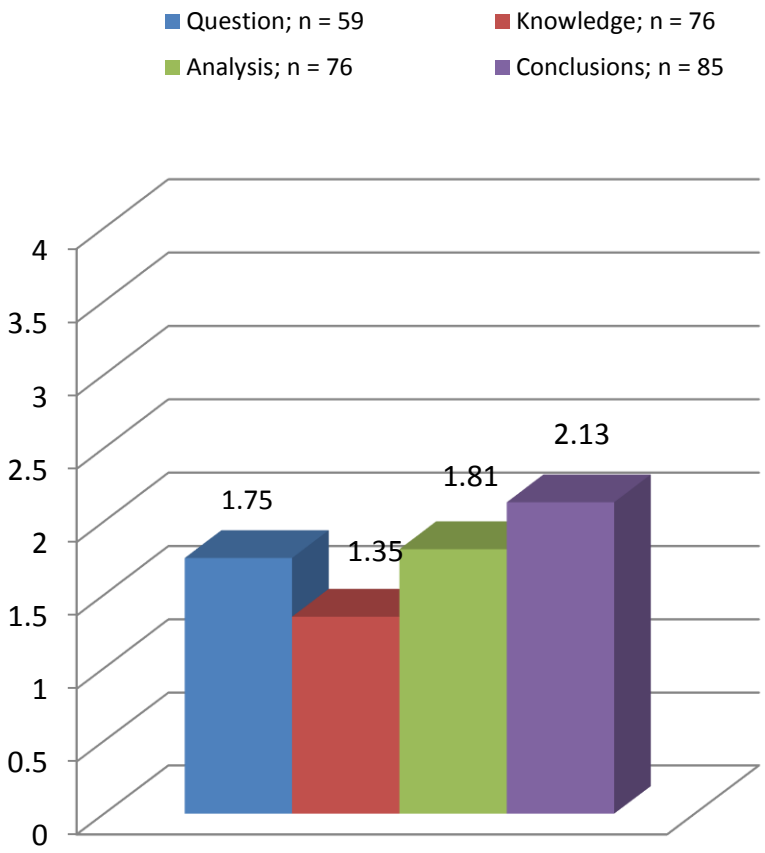
Trait/ Performance Level	Competencies; Kappa = -.012 (All scores); Kappa = -.027 (Not aligned , unable to score, and one rater score missing excluded) Kappa Liberal = .436 (All Scores); Kappa Liberal = .727 (Exclusions Noted Above)	Risks; Kappa = .165 (All scores); Kappa = .206 (Not aligned , unable to score, and one rater score missing excluded) Kappa Liberal = .511 (All Scores); Kappa Liberal = .806 (Exclusions Noted Above)	Solving Problems; .219 (All scores); Kappa = .229 (Not aligned , unable to score, and one rater score missing excluded) Kappa Liberal = .508 (All Scores); Kappa Liberal = .896 (Exclusions Noted Above)	Contradictions; Kappa = .217 (All scores); Kappa = .214 (Not aligned , unable to score, and one rater score missing excluded) Kappa Liberal = .475 (All Scores); Kappa Liberal = . 742 (Exclusions Noted Above)	Innovation; Kappa = .121 (All scores); Kappa = .039 (Not aligned , unable to score, and one rater score missing excluded) Kappa Liberal = .609 (All Scores); Kappa Liberal = . 797 (Exclusions Noted Above)	Connecting; Kappa = .132 (All scores); Kappa = .074 (Not aligned , unable to score, and one rater score missing excluded) Kappa Liberal = .627 (All Scores); Kappa Liberal = . 908 (Exclusions Noted Above)
Agree on score	15 (14%)	31 (29%)	25 (23%)	23 (21%)	27 (25%)	28 (26%)
Difference = 1 point or less	35 (32%)	24 (22%)	21 (19%)	20 (19%)	36 (33%)	37 (34%)
Difference = 1.5 to 2 points	10 (9%)	9 (8%)	4 (4%)	9 (8%)	10 (9%)	5 (5%)
Difference = > 2 points	5 (5%)	1 (1%)	0	3 (3%)	2 (2%)	0
Agree on Not Aligned	4 (4%)	9 (8%)	19 (18%)	16 (15%)	9 (8%)	8 (7%)
Agree on Unable to Score due to error	1 (1%)	1 (1%)	1 (1%)	1 (1%)	1 (1%)	1 (1%)
Score + Not Aligned	37 (34%)	33 (31%)	35 (32%)	34 (31%)	21 (19%)	27 (25%)
Score + Missing Second Rater Score	1 (1%)	0	3 (3%)	2 (2%)	2 (2%)	2 (2%)
Total	108 (100 %)	108 (100%)	108 (100%)	108 (100%)	108 (100%)	108 (100%)

Inquiry-Based Thinking: Overall Analysis

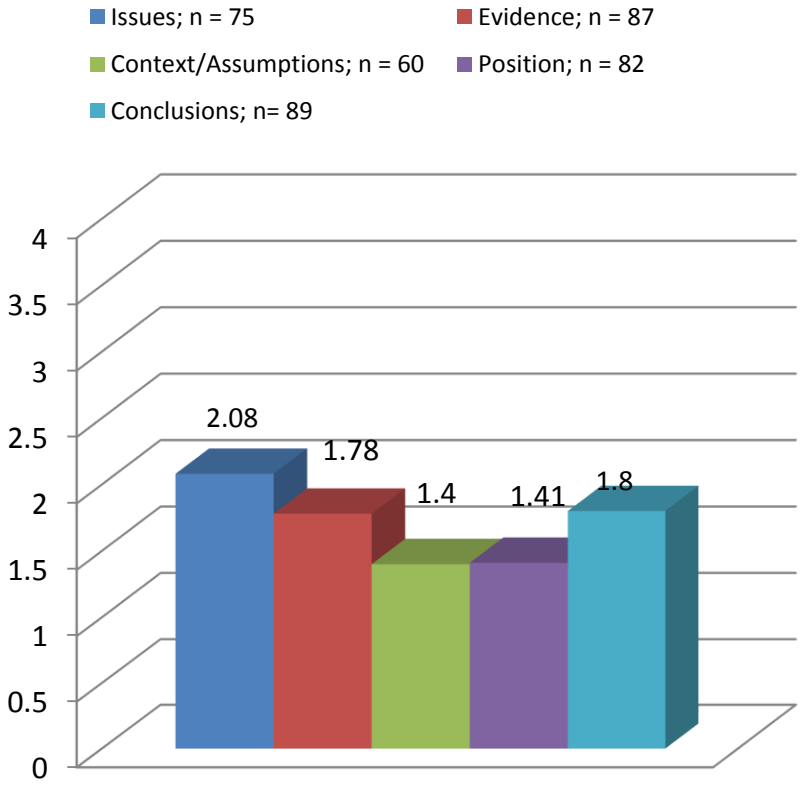
Mean Scores on a scale of 0 – 4, with 4 being the highest possible score.

(Although there were 93 IBT artifacts in the MU rubric analysis and 95 in the AAC&U rubric analysis, not all artifacts aligned to every trait)

MU Rubric



AAC&U Rubric



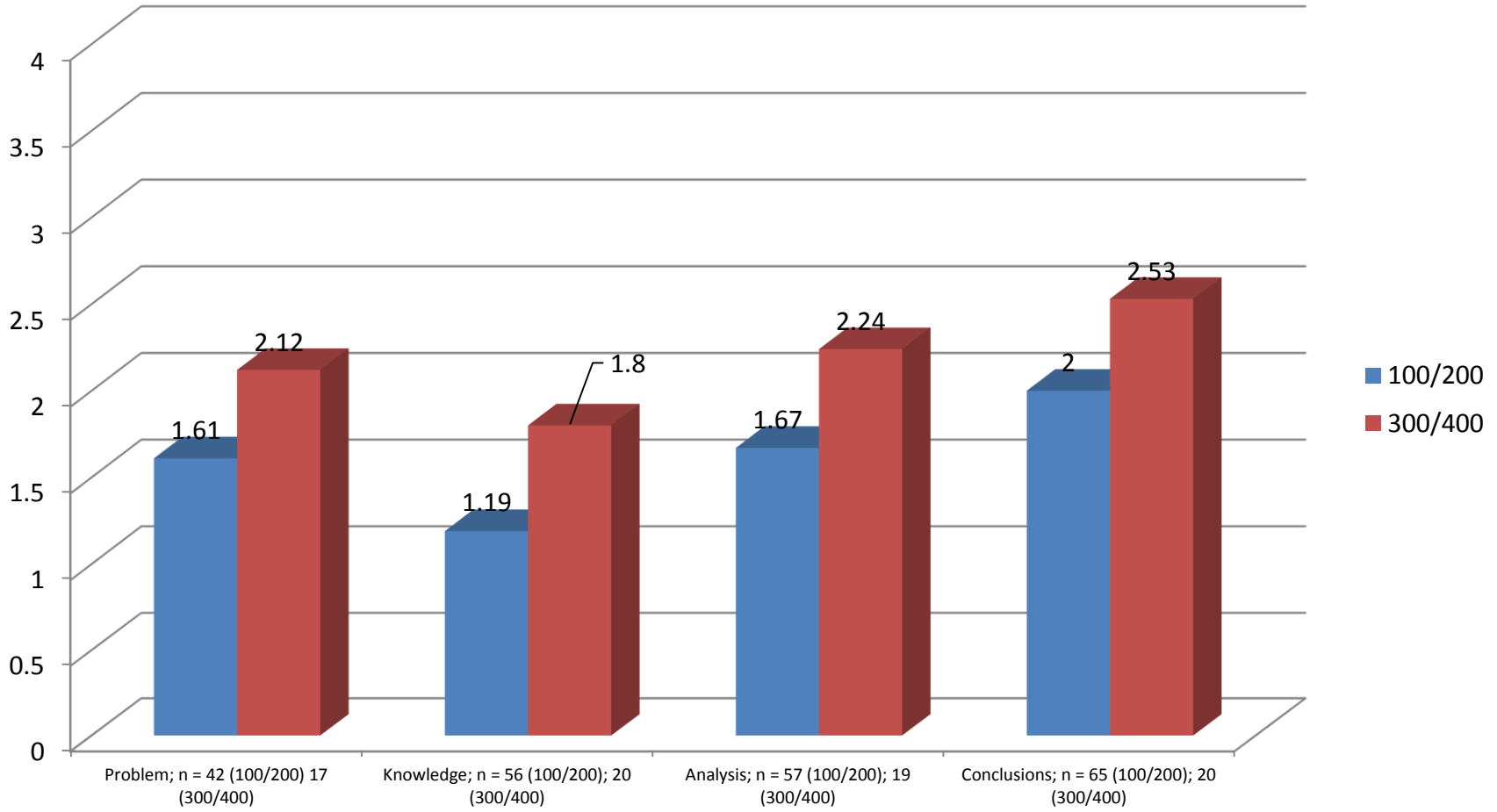
Inquiry-Based Thinking: Course Level Analysis

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score

Mean for all traits were significantly higher for 300/400 level courses than for 100/200 level courses.

Overall means for Inquiry-Based Thinking were 1.63 for 100/200 level courses and 2.17 for 300/400 level courses.

MU Rubric



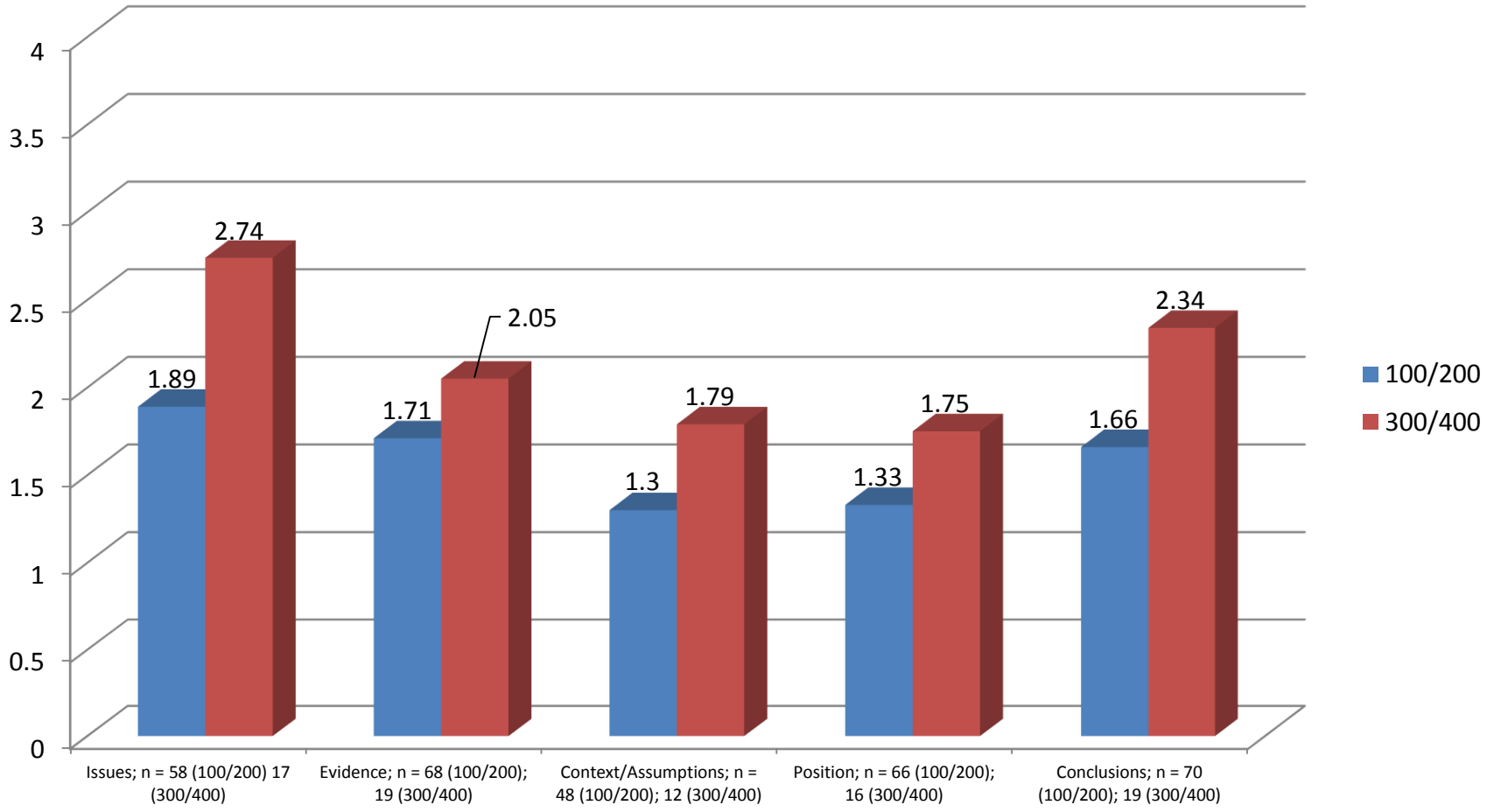
Inquiry-Based Thinking: Course Level Analysis

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score

Means for all traits *except evidence* and *context/assumptions* were significantly higher for 300/400 level than for 100/200 level courses.

Overall means using the AAC&U *Critical Thinking* rubric were 1.59 for 100/200 level courses and 2.16 for 300/400 level courses.

AAC&U Rubric

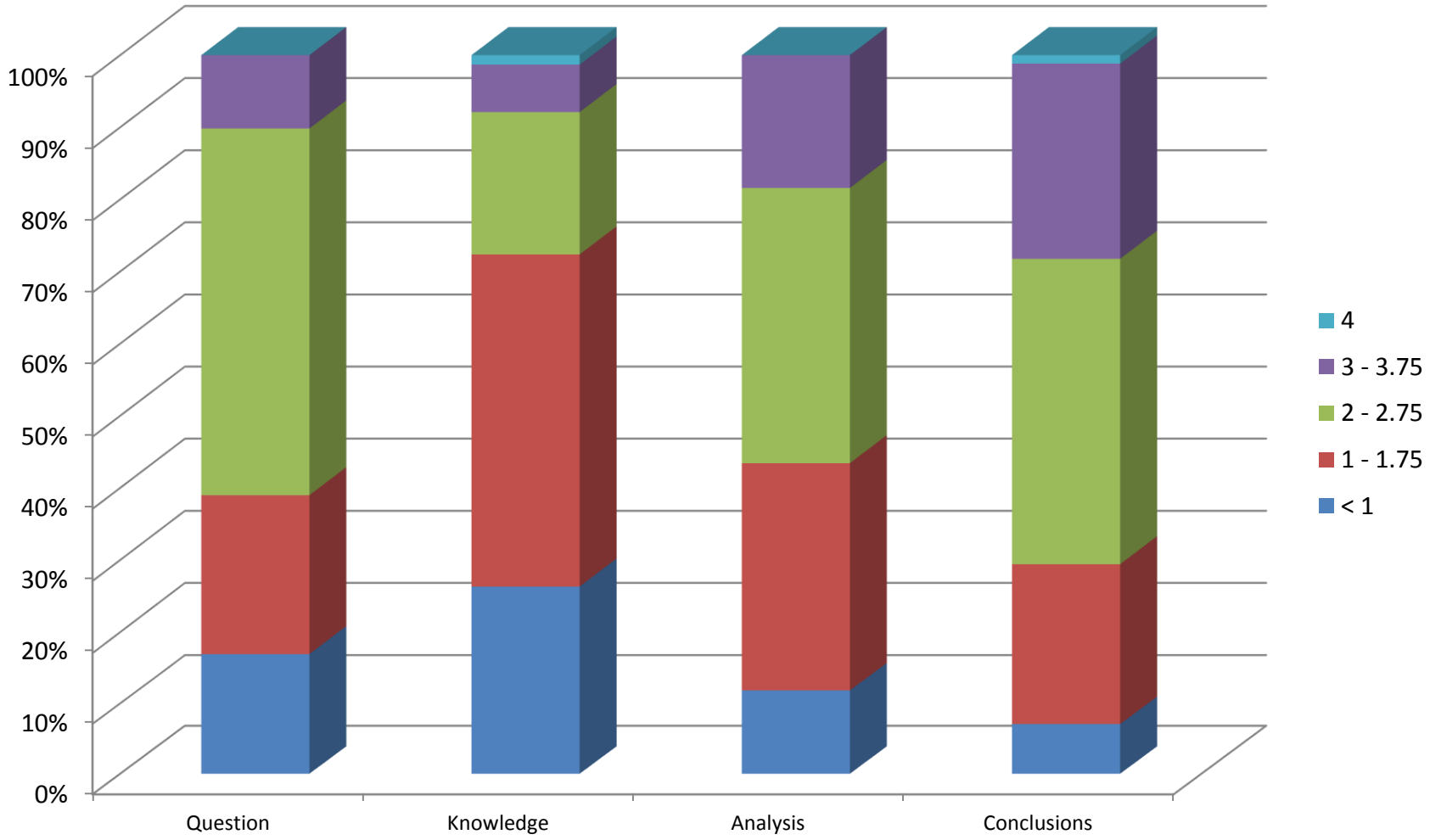


Inquiry-Based Thinking (MU Rubric)

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

Trait/ Performance Level	Question	Knowledge	Analysis	Conclusions	To91tal
< 1	10 (17%)	20 (26%)	9 (12%)	6 (7%)	45 (15%)
1 – 1.75	13 (22%)	35 (46%)	24 (32%)	19 (22%)	91 (31%)
2 – 2.75	30 (51%)	15 (20%)	29 (38%)	36 (42%)	110 (37%)
3 – 3.75	6 (10%)	5 (7%)	14 (18%)	23 (27%)	48 (16%)
4	0	1 (1%)	0	1 (1%)	2 (1%)
Totals	59 (100)	76 (100%)	76 (100%)	85 (100%)	296 (100%)

Inquiry-Based Thinking (MU Rubric)

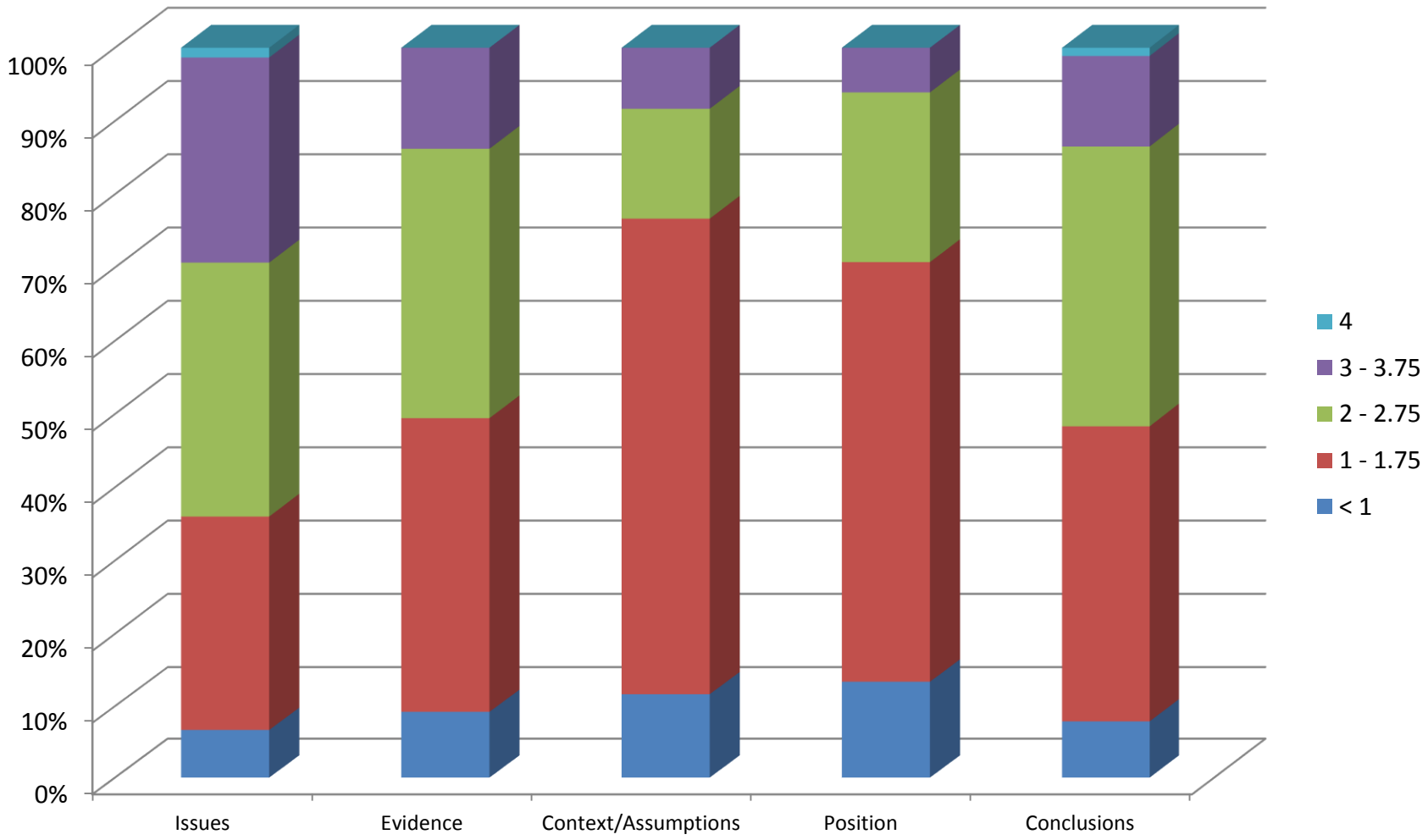


Inquiry-Based Thinking (AAC&U Rubric)

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

Trait/ Performance Level	Issues	Evidence	Context/ Assumptions	Position	Conclusions	Total
< 1	5 (7%)	8 (9%)	7 (12%)	11 (13%)	7 (8%)	38 (10%)
1 – 1.75	22 (29%)	35 (40%)	39 (65%)	47 (57%)	36 (40%)	179 (46%)
2 – 2.75	26 (35%)	32 (37%)	9 (15%)	19 (23%)	34 (38%)	120 (31%)
3 – 3.75	21 (28%)	12 (14%)	5 (8%)	5 (6%)	11 (12%)	54 (14%)
4	1 (1%)	0	0	0	1 (1%)	2 (1%)
Totals	75 (100%)	87 (100%)	60 (100%)	82 (100%)	89 (100%)	393 (100%)

Inquiry-Based Thinking (AAC&U Rubric)



Inquiry-Based Thinking (MU)

Inter-Rater Agreement Results

Trait/ Performance Level	Question; Kappa = .339 (All scores); Kappa = .268 (Not aligned , unable to score, and one rater score missing excluded) Kappa Liberal = .622 (All Scores); Kappa Liberal = .863 (Exclusions Noted Above)	Knowledge; Kappa = .224 (All scores); Kappa = .033 (Not aligned , unable to score, and one rater score missing excluded); Kappa Liberal = .663 (All Scores); Kappa Liberal = .802 (Exclusions Noted Above)	Analysis; Kappa = .355 (All scores); Kappa = .331 (Not aligned and unable to score excluded); Kappa Liberal = .683 (All Scores); Kappa Liberal = .924 (Exclusions Noted Above)	Conclusions; Kappa = .359 (All scores); Kappa = .297 (Not aligned , unable to score, and one rater score missing excluded); Kappa Liberal = .731 (All Scores); Kappa Liberal = .836 (Exclusions Noted Above)
Agree on score	25 (23%)	20 (19%)	33 (31%)	41 (38%)
Difference = 1 point or less	22 (20%)	36 (33%)	26 (24%)	29 (27%)
Difference = 1.5 to 2 points	4 (4%)	11 (10%)	3 (3%)	9 (8%)
Difference > 2 points	2 (2%)	0	1 (1%)	2 (2%)
Agree on Not Aligned	25 (23%)	18 (17%)	17 (16%)	10 (9%)
Agree on Unable to Score due to error	3 (3%)	3 (3%)	3 (3%)	3 (3%)
Score + Not Aligned	26 (24%)	20 (19%)	24 (22%)	14 (13%)
Score + Missing Second Rater Score	1 (1%)	0	1 (1%)	0
Total	108 (100%)	108 (100%)	108 (100%)	108 (100%)

Inquiry-Based Thinking (AAC&U)

Inter-Rater Agreement Results

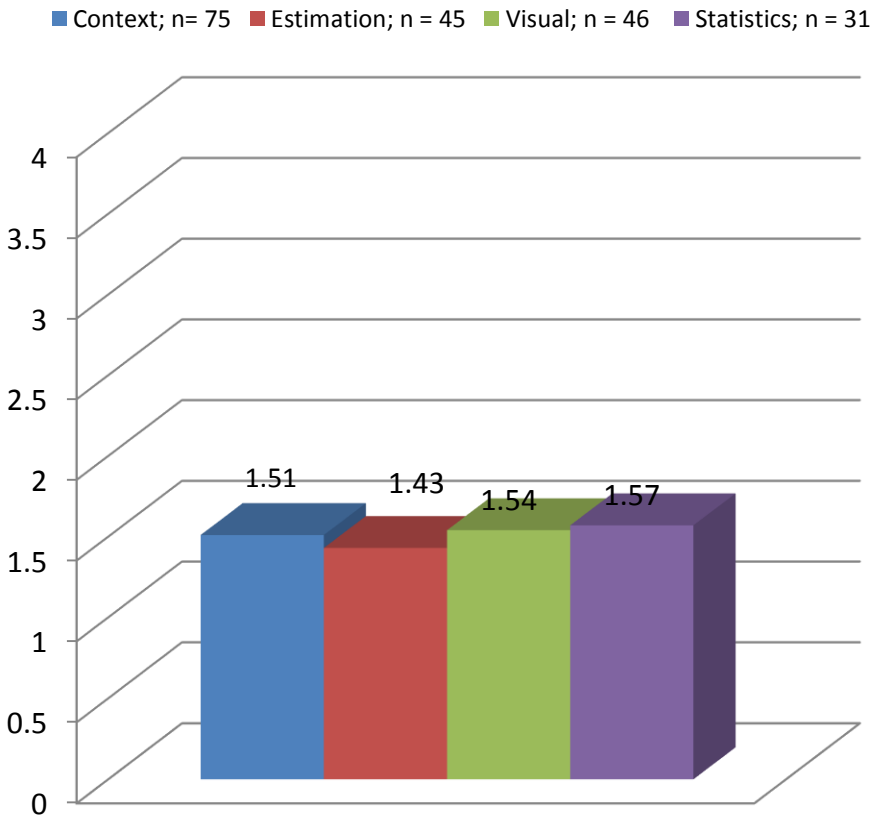
Trait/ Performance Level	Issues; Kappa = .271 (All scores); Kappa = .231 (Not aligned , unable to score, and one rater score missing excluded) Kappa Liberal = .583 (All Scores); Kappa Liberal = .793 (Exclusions Noted Above)	Evidence; Kappa = .285 (All scores); Kappa = .144 (Not aligned , unable to score, and one rater score missing excluded); Kappa Liberal = .781 (All Scores); Kappa Liberal = .919 (Exclusions Noted Above)	Context/Assumptions; Kappa = .148 (All scores); Kappa = -.009 (Not aligned and unable to score excluded); Kappa Liberal = .455 (All Scores); Kappa Liberal = .814 (Exclusions Noted Above)	Position; Kappa = .186 (All scores); Kappa = .166 (Not aligned , unable to score, and one rater score missing excluded); Kappa Liberal = .528 (All Scores); Kappa Liberal = .748 (Exclusions Noted Above)	Conclusions; Kappa = .295 (All scores); Kappa = .227 (Not aligned , unable to score, and one rater score missing excluded); Kappa Liberal = .737 (All Scores); Kappa Liberal = .863 (Exclusions Noted Above)
Agree on score	29 (27%)	33 (31%)	13 (12%)	29 (27%)	39 (36%)
Difference = 1 point or less	24 (22%)	39 (36%)	22 (20%)	25 (23%)	34 (31%)
Difference = 1.5 to 2 points	9 (8%)	5 (5%)	4 (4%)	13 (12%)	8 (7%)
Difference > 2 points	2 (2%)	0	2 (2%)	1 (1%)	1 (1%)
Agree on Not Aligned	14 (13%)	13 (12%)	24 (22%)	8 (7%)	8 (7%)
Agree on Unable to Score due to error	3 (3%)	3 (3%)	3 (3%)	3 (3%)	3 (3%)
Score + Not Aligned	27 (25%)	15 (14%)	38 (35%)	28 (26%)	13 (12%)
Score + Missing Second Rater Score	0	0	2 (2%)	1 (1%)	2 (2%)
Total	108 (100%)	108 (100%)	108 (100%)	108 (100%)	108 (100%)

Quantitative Thinking: Overall Analysis

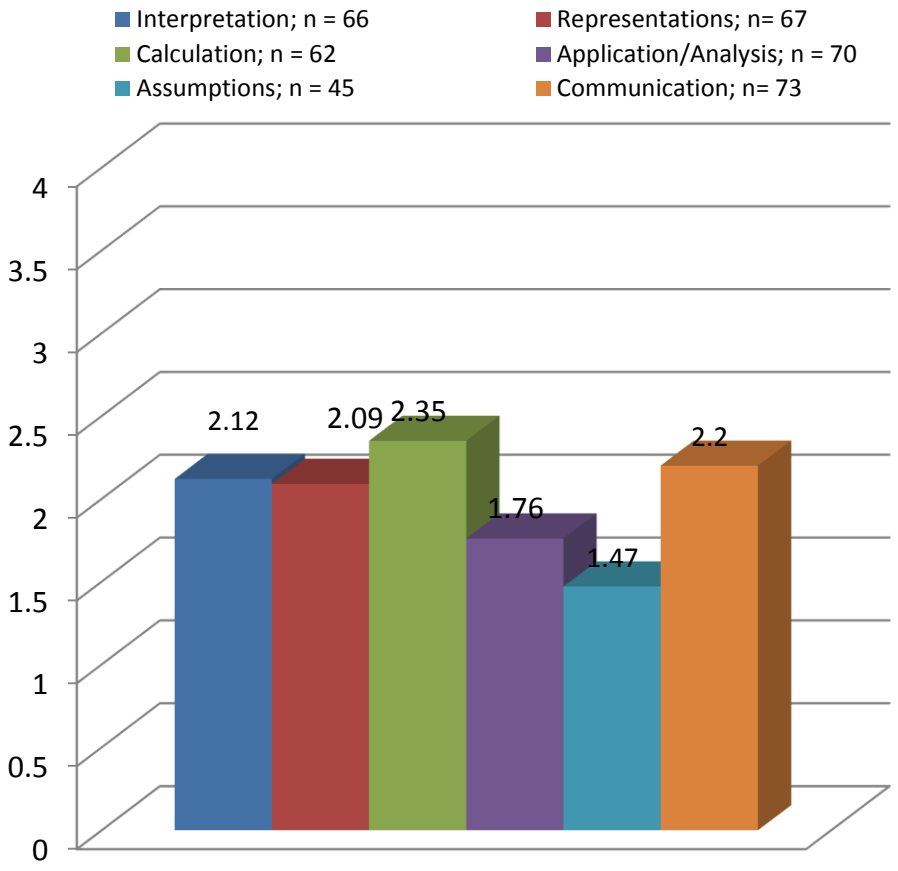
Mean Scores on a scale of 0 – 4, with 4 being the highest possible score.

(Although there were 77 QT artifacts in the MU rubric analysis and 80 in the AAC&U rubric analysis, not all artifacts aligned to every trait)

MU Rubric



AAC&U Rubric



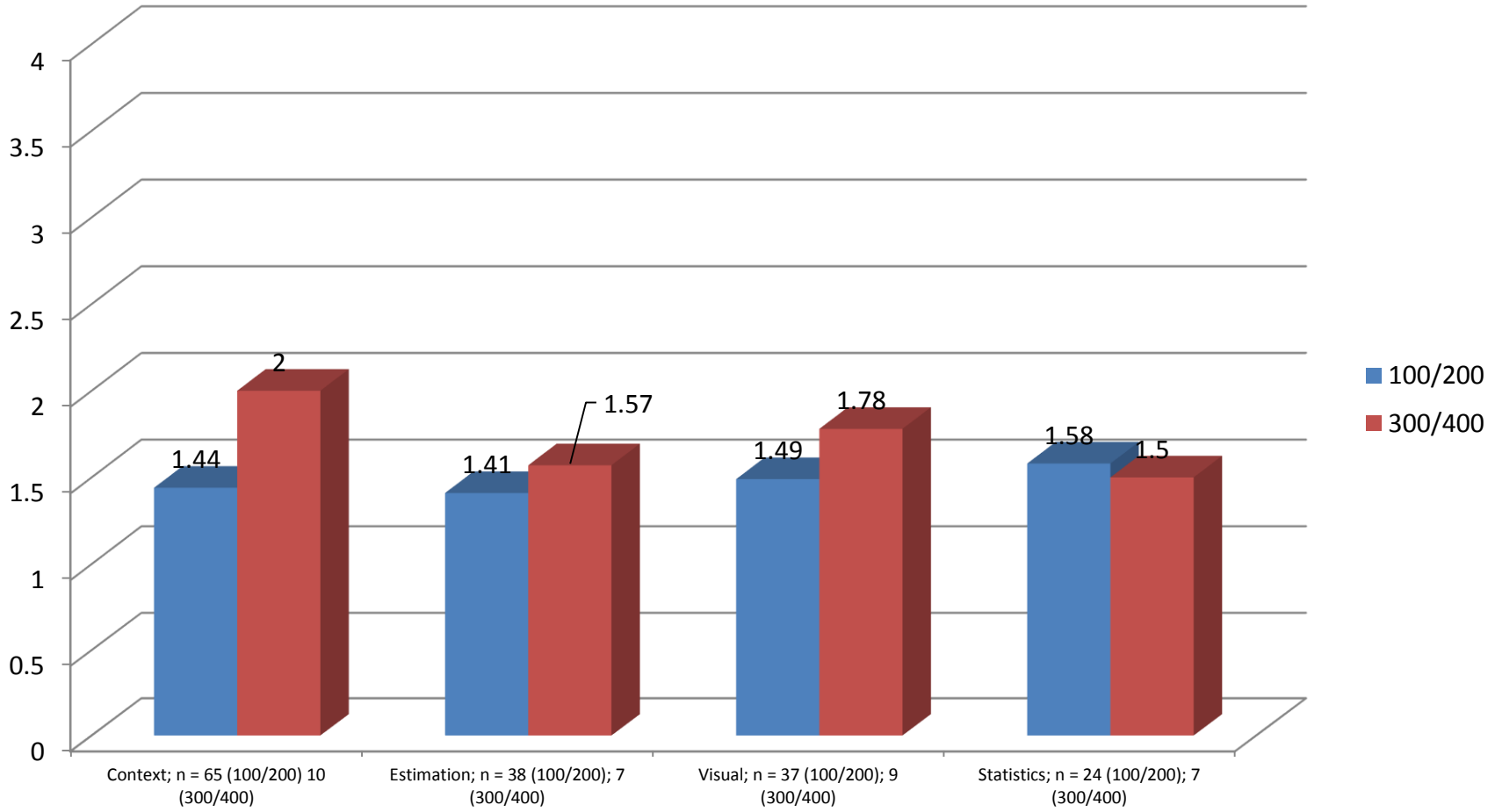
Quantitative Thinking: Course Level Analysis

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score

Course level differences were not significant for any trait.

Overall means for Quantitative Thinking were 1.46 for 100/200 level courses and 1.74 for 300/400 level courses.

MU Rubric



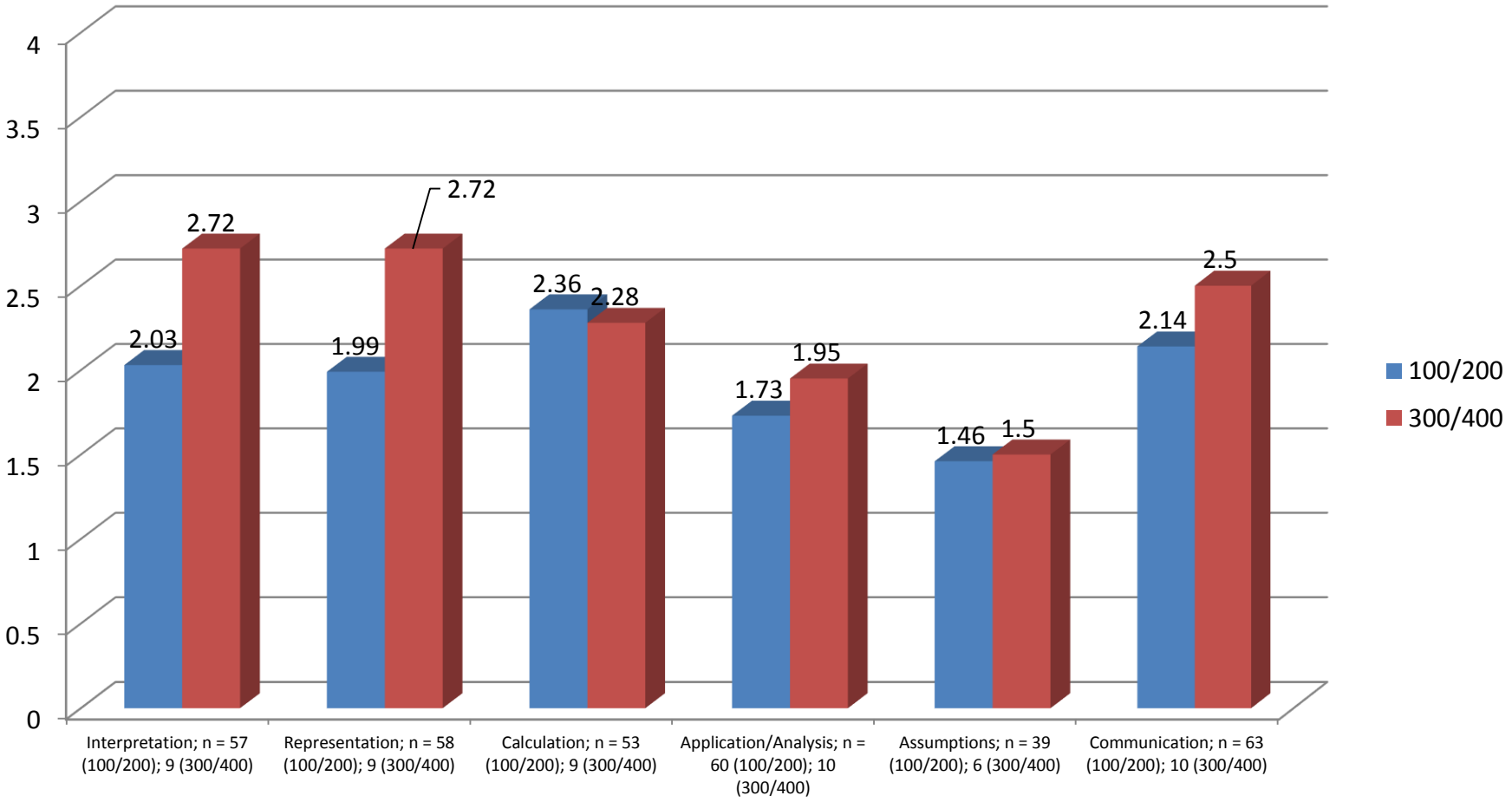
Quantitative Literacy: Course Level Analysis

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score

Course level differences were not significant for any trait.

Overall means for Quantitative Literacy were 1.98 for 100/200 level courses and 2.32 for 300/400 level courses.

AAC&U Rubric

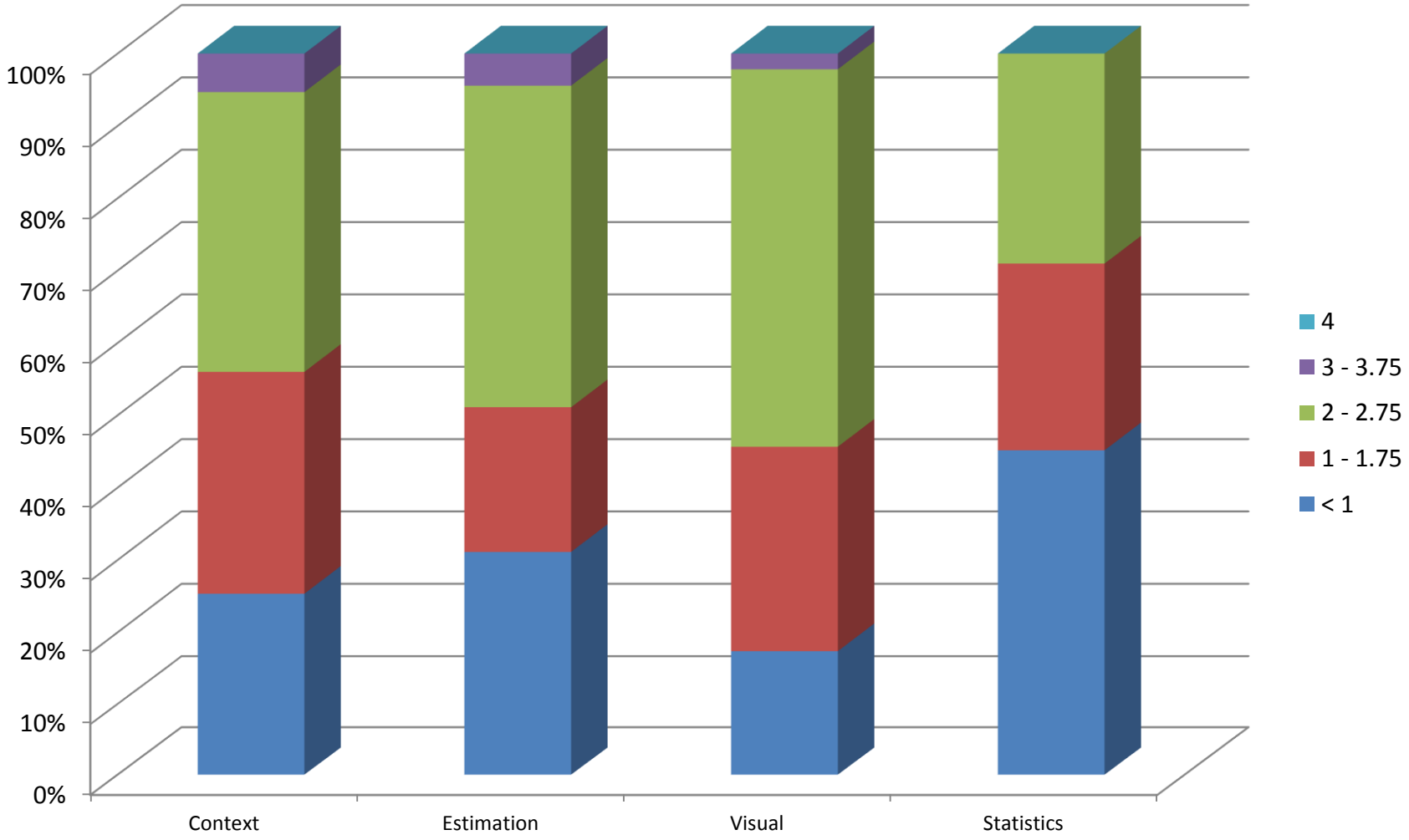


Quantitative Thinking (MU Rubric)

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

Trait/ Performance Level	Context	Estimation	Visual	Statistics	Total
< 1	19 (25%)	14 (31%)	8 (17%)	14 (45%)	55 (28%)
1 – 1.75	23 (31%)	9 (20%)	13 (28%)	8 (26%)	53 (27%)
2 – 2.75	29 (39%)	20 (44%)	24 (52%)	9 (29%)	82 (42%)
3 – 3.75	4 (5%)	2 (4%)	1 (2%)	0	7 (4%)
4	0	0	0	0	0
Totals	75 (100%)	45 (100%)	46 (100%)	31 (100%)	197 (100%)

Quantitative Thinking (MU Rubric)

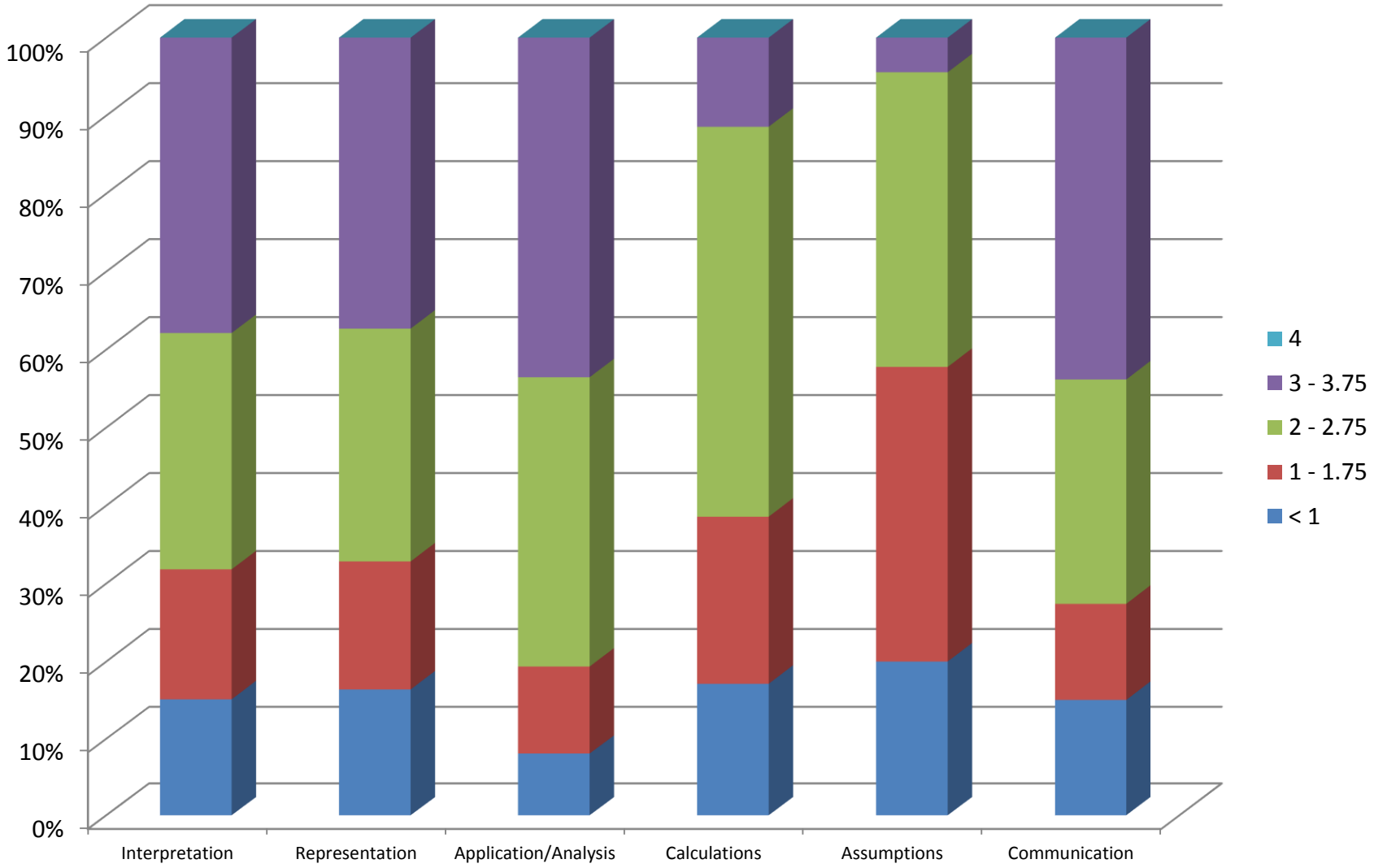


Quantitative Thinking (AAC&U Rubric)

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

Trait/ Performance Level	Interpretation	Representation	Application/ Analysis	Calculations	Assumptions	Communication	Total
< 1	10 (15%)	11 (16%)	5 (8%)	12 (17%)	9 (20%)	11 (15%)	58 (15%)
1 – 1.75	11 (17%)	11 (16%)	7 (11%)	15 (21%)	17 (38%)	9 (12%)	70 (18%)
2 – 2.75	20 (30%)	20 (30%)	23 (37%)	35 (50%)	17 (38%)	21 (29%)	136 (36%)
3 – 3.75	25 (38%)	25 37%)	27 (44%)	8 (11%)	2 (4%)	32 (44%)	119 (31%)
4	0	0	0	0	0	0	0
Totals	66 (100%)	67 (100%)	62 (100%)	70 (100%)	45 (100%)	73 (100%)	383 (100%)

Quantitative Thinking (AAC&U Rubric)



Quantitative Thinking (MU)

Inter-Rater Agreement Results

Trait/ Performance Level	Context; Kappa = .195 (All scores); Kappa = -.018 (Not aligned and Unable to Score Excluded) Kappa Liberal = .677 (All Scores); Kappa Liberal = .800 (Exclusions Noted Above)	Estimation; Kappa = .040 (All Scores); Kappa = -.044 (Not Aligned, Unable to Score, and One Rater Score Missing Excluded); Kappa Liberal = .270 (All Scores); Kappa Liberal = .734 (Exclusions Noted Above)	Visual; Kappa = .284 (All scores); Kappa = .137 (Not Aligned, Unable to Score, and One Rater Score Missing Excluded); Kappa Liberal = .460 (All Scores); Kappa Liberal = .786 (Exclusions Noted Above)	Statistics; Kappa = .245 (All scores); Kappa = .439 (Not Aligned and Unable to Score Excluded); Kappa Liberal = .354 (All Scores); Kappa Liberal = .874 (Exclusions Noted Above)
Agree on score	16 (15%)	6 (6%)	19 (18%)	11 (10%)
Difference = 1 point or less	39 (36%)	15 (14%)	11 (10$)	6 (6%)
Difference = 1.5 to 2 points	9 (8%)	5 (5%)	5 (5%)	1 (1%)
Difference > 2 points	2 (2%)	1 (1%)	0	1 (1%)
Agree on Not Aligned	24 (22%)	32 (30%)	39 (36%)	52 (48%)
Agree on Unable to Score due to error	0	0	0	0
Score + Not Aligned	18 (17%)	49 (45%)	34 (31%)	35 (32%)
Score or Misaligned + Missing Second Rater Score	0	0	0	2 (2%)
Total	108 (100%)	108 (100%)	108 (100%)	108 (100%)

Quantitative Thinking (AAC&U)

Inter-Rater Agreement Results

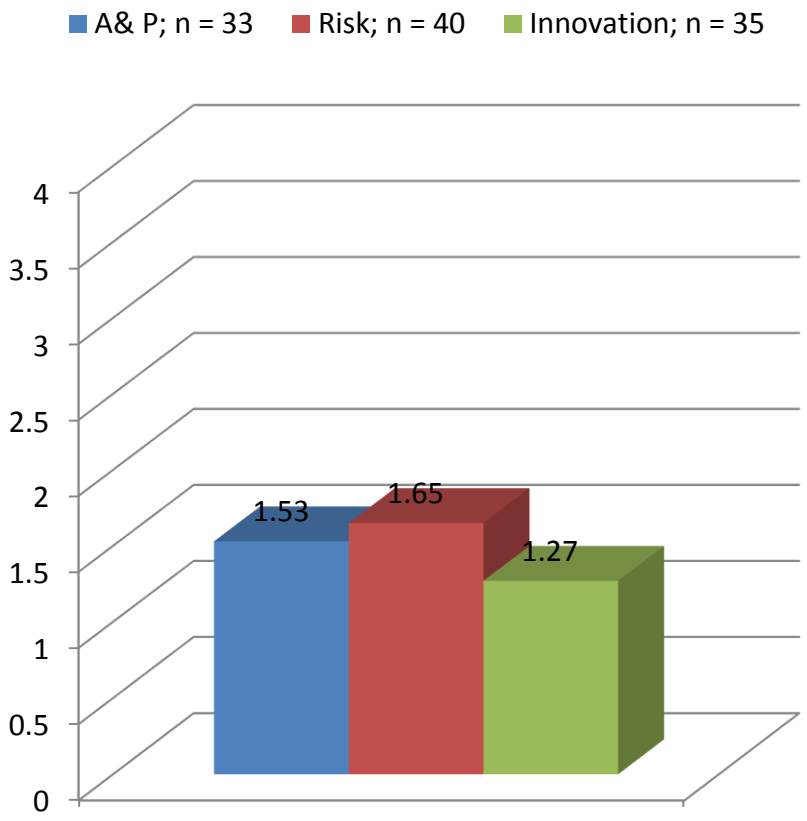
Trait/ Performance Level	Interpretation; Kappa = .306 (All scores); Kappa = .152 (Not aligned and Unable to Score Excluded) Kappa Liberal = .573 (All Scores); Kappa Liberal = .688 (Exclusions Noted Above)	Representation; Kappa = .263 (All Scores); Kappa = .248 (Not Aligned, Unable to Score, and One Rater Score Missing Excluded); Kappa Liberal = .528 (All Scores); Kappa Liberal = .909 (Exclusions Noted Above)	Calculation; Kappa = .338 (All scores); Kappa = .198 (Not Aligned, Unable to Score, and One Rater Score Missing Excluded); Kappa Liberal = .606 (All Scores); Kappa Liberal = .863 (Exclusions Noted Above)	Application/Analysis ; Kappa = .231 (All scores); Kappa = .093 (Not Aligned and Unable to Score Excluded); Kappa Liberal = .529 (All Scores); Kappa Liberal = .681 (Exclusions Noted Above)	Assumptions; Kappa = .173 (All scores); Kappa = .013 (Not Aligned and Unable to Score Excluded); Kappa Liberal = .375 (All Scores); Kappa Liberal = .728 (Exclusions Noted Above)	Communication; Kappa = .266 (All scores); Kappa = .145 (Not Aligned and Unable to Score Excluded); Kappa Liberal = .555 (All Scores); Kappa Liberal = .777 (Exclusions Noted Above)
Agree on score	22 (20%)	22 (20%)	27 (25%)	21 (19%)	9 (8%)	26 (24%)
Difference = 1 point or less	20 (19%)	19 (18%)	19 (18%)	22 (20%)	13 (12%)	21 (19%)
Difference = 1.5 to 2 points	8 (7%)	2 (2%)	3 (3%)	13 (12%)	6 (6%)	9 (8%)
Difference > 2 points	6 (6%)	1 (1%)	2 (2%)	2 (2%)	0	0
Agree on Not Aligned	30 (28%)	29 (27%)	30 (28%)	24 (22%)	41 (38%)	24 (22%)
Agree on Unable to Score due to error	0	0	0	0	0	0
Score + Not Aligned	22 (20%)	34 (31%)	25 (23%)	25 (23%)	36 (33%)	25 (23%)
Score or Misaligned + Missing Second Rater Score	0	1 (1%)	2 (2%)	1 (1%)	3 (3%)	3 (3%)
Total	108 (100%)	108 (100%)	108 (100%)	108 (100%)	108 (100%)	108 (100%)

Course Type Analysis

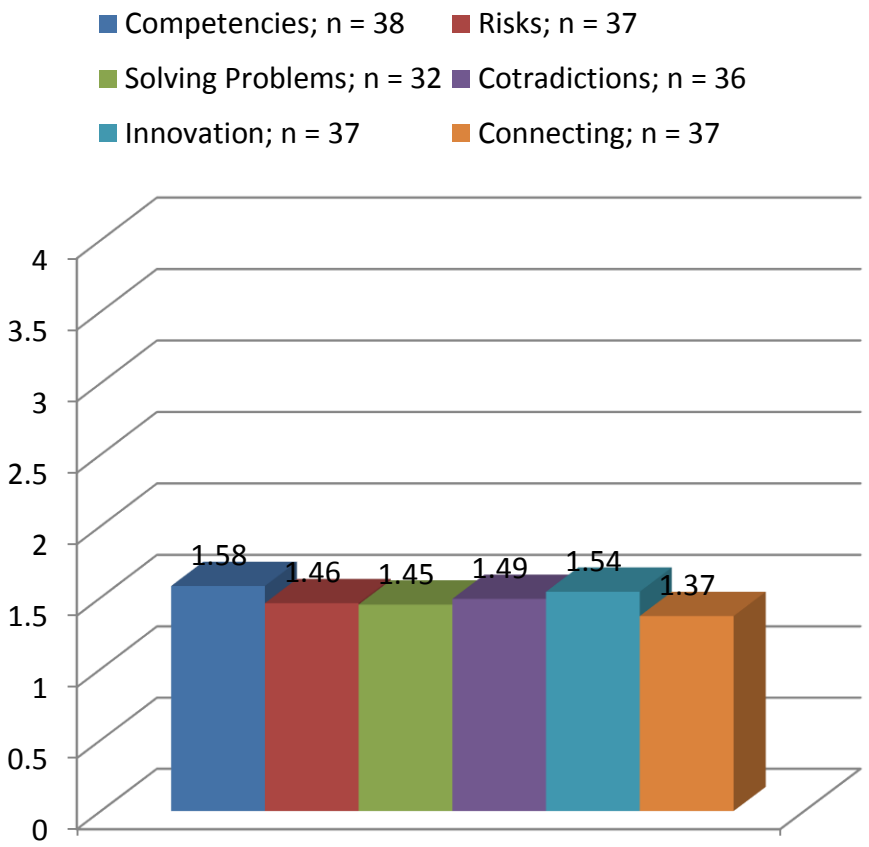
CT Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. All CT courses are 100/200 Level. Some artifacts were from courses that, in addition to being CT, also were Core II, multicultural, international, and/or writing intensive.

Creative Thinking (MU)



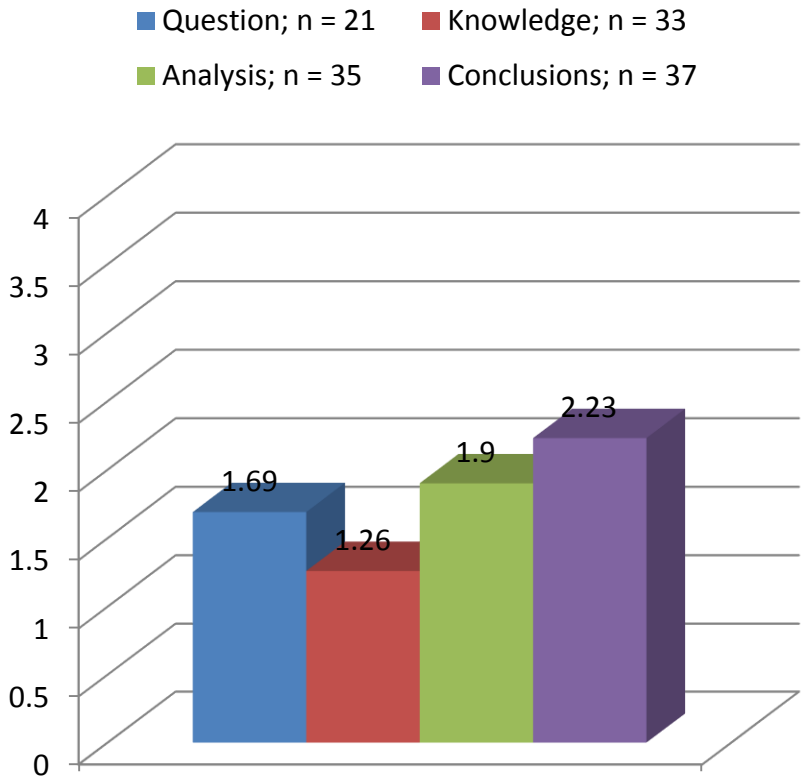
Creative Thinking (AAC&U)



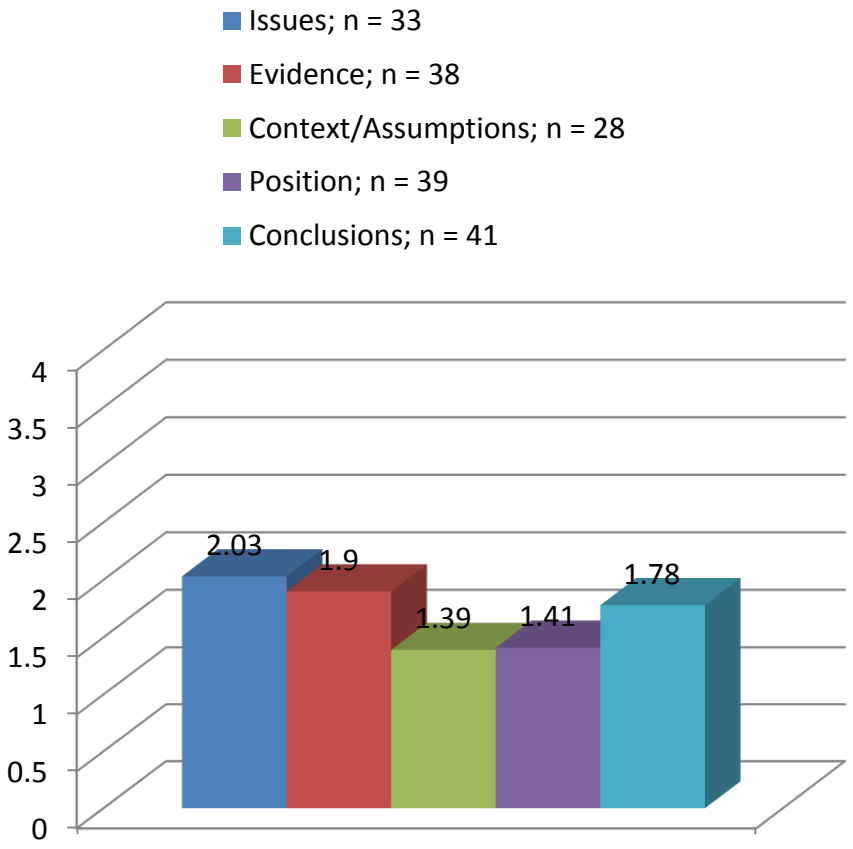
CT Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. All CT courses are 100/200 Level. Some artifacts were from courses that, in addition to being CT, also were Core II, multicultural, international, and/or writing intensive.

Inquiry-Based Thinking (MU)



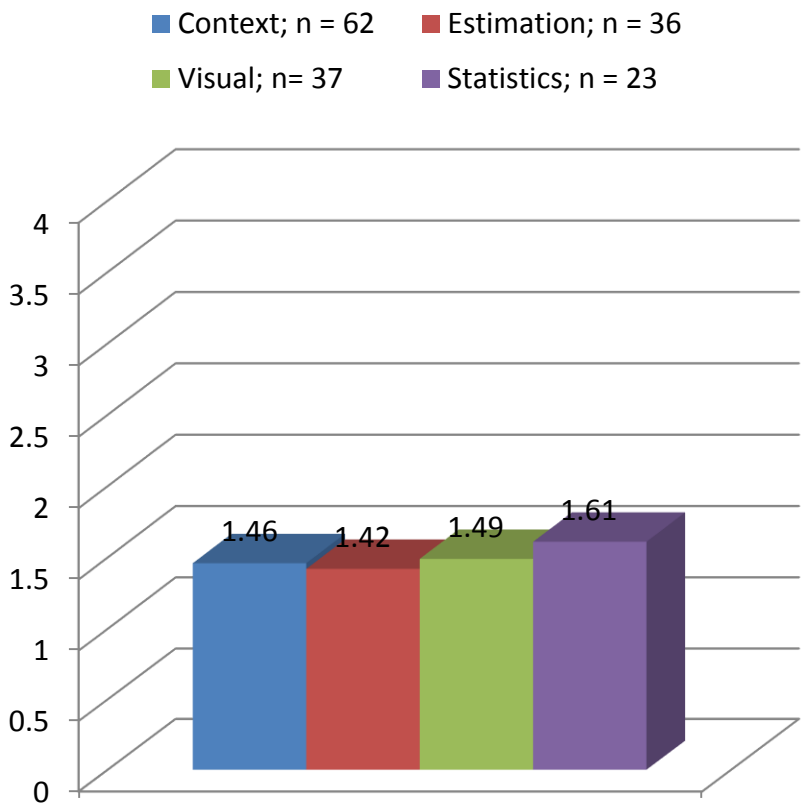
Inquiry-Based Thinking (AAC&U)



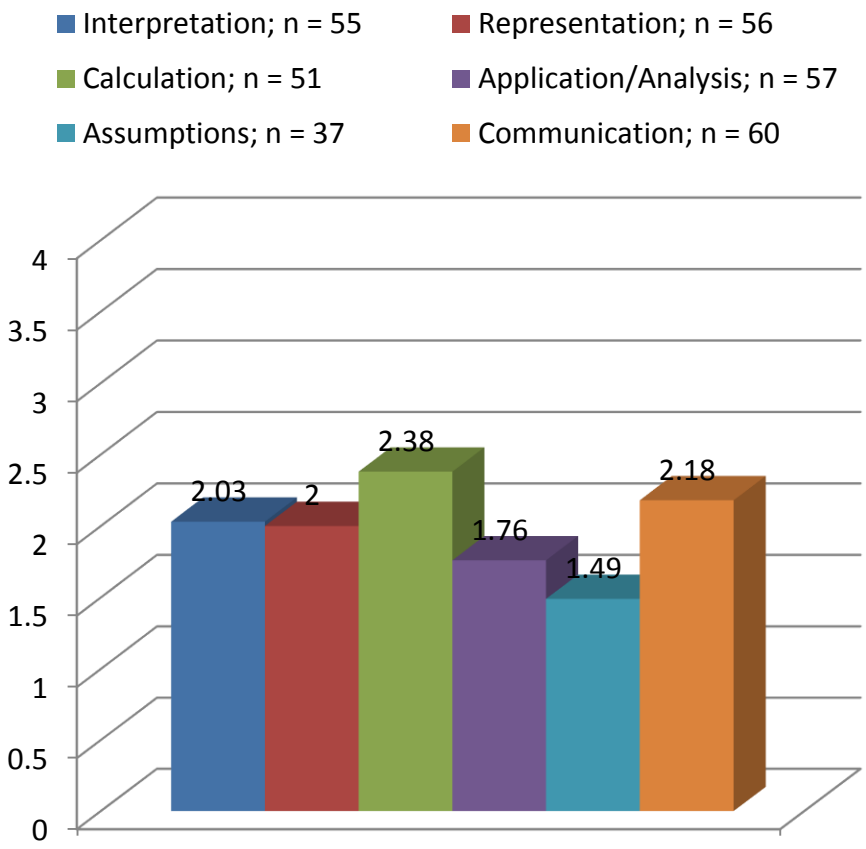
CT Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. All CT courses are 100/200 Level. Some artifacts were from courses that, in addition to being CT, also were Core II, multicultural, international, and/or writing intensive.

Quantitative Thinking (MU)



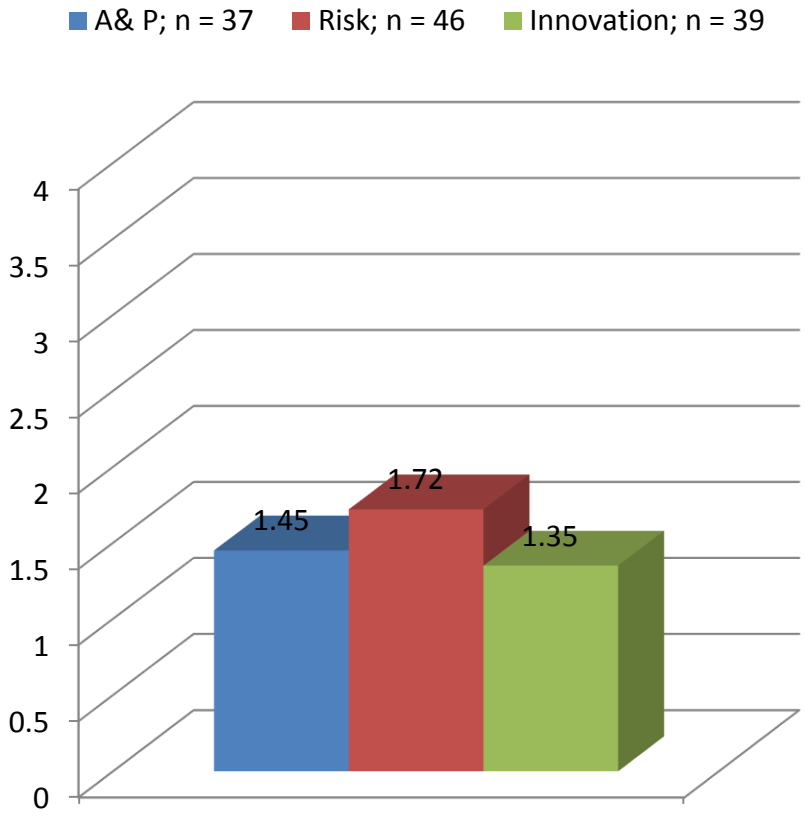
Quantitative Thinking (AAC&U)



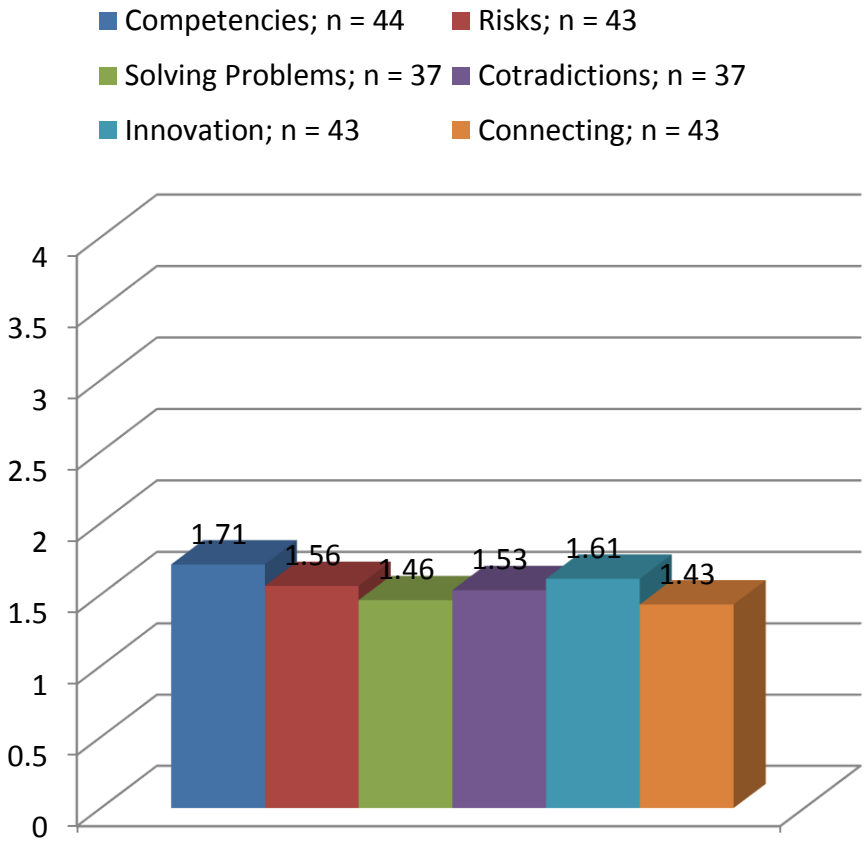
Core II Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. All Core II courses are 100/200 Level. Some artifacts were from courses that, in addition to being Core II, also were CT, multicultural, international, and/or writing intensive.

Creative Thinking (MU)



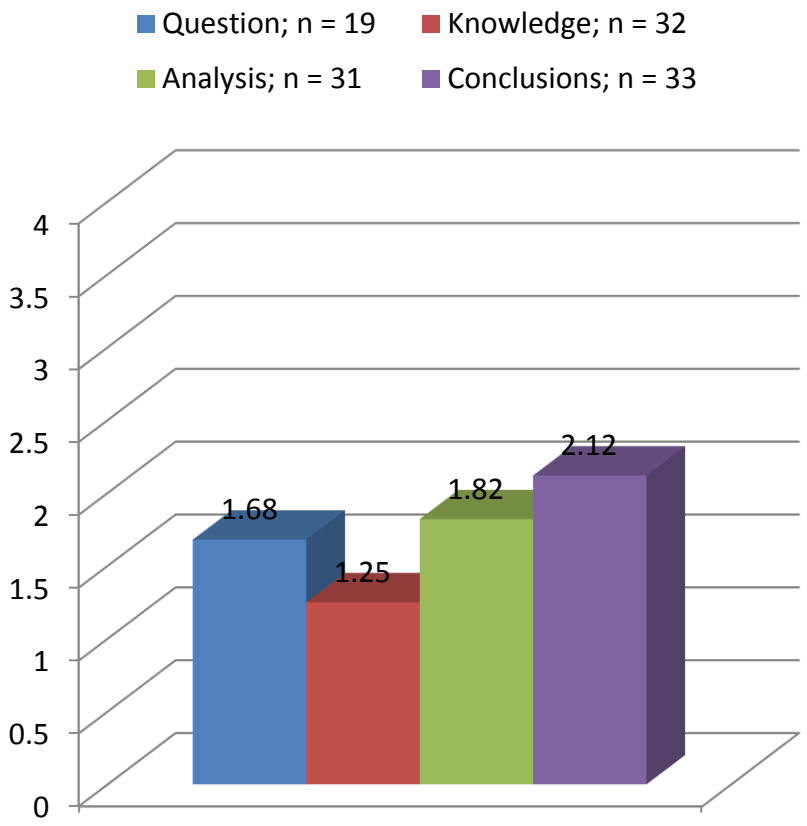
Creative Thinking (AAC&U)



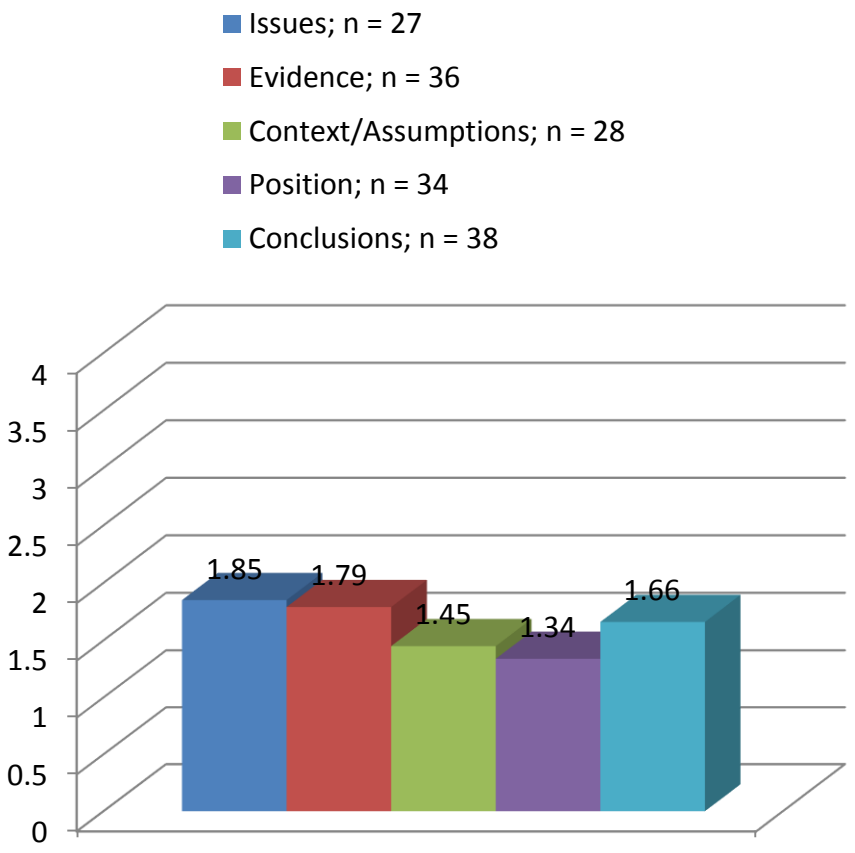
Core II Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. All Core II courses are 100/200 Level. Some artifacts were from courses that, in addition to being Core II, also were CT, multicultural, international, and/or writing intensive.

Inquiry-Based Thinking (MU)



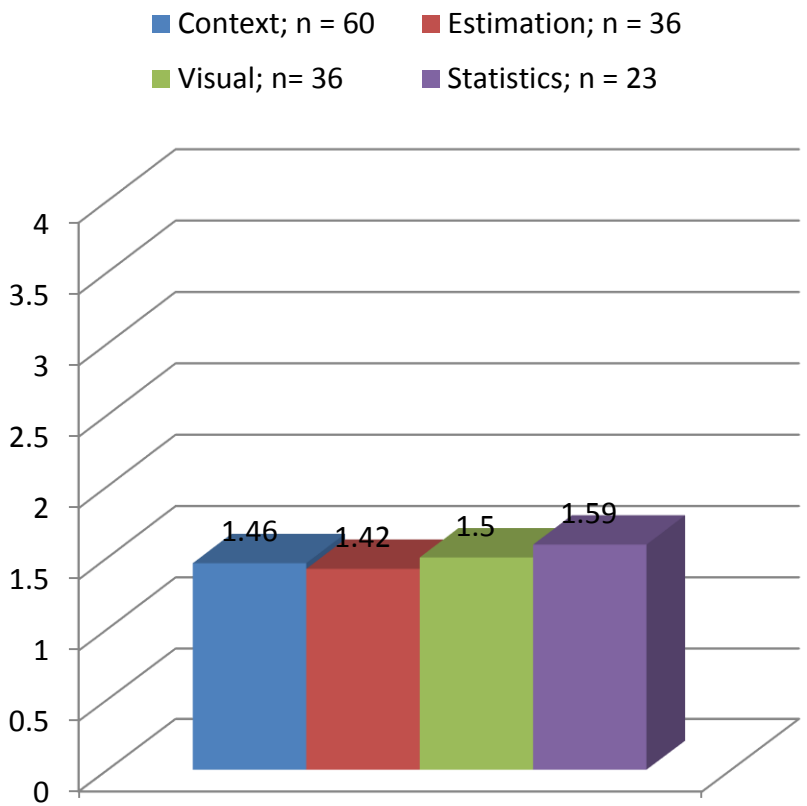
Inquiry-Based Thinking (AAC&U)



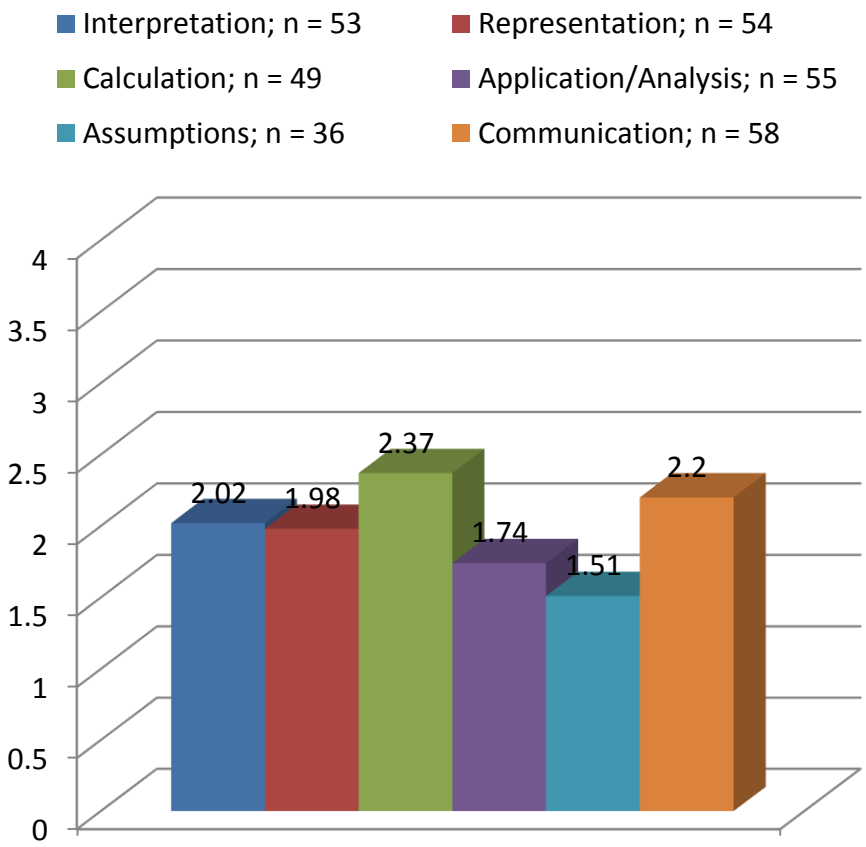
Core II Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. All Core II courses are 100/200 Level. Some artifacts were from courses that, in addition to being Core II, also were CT, multicultural, international, and/or writing intensive.

Quantitative Thinking (MU)



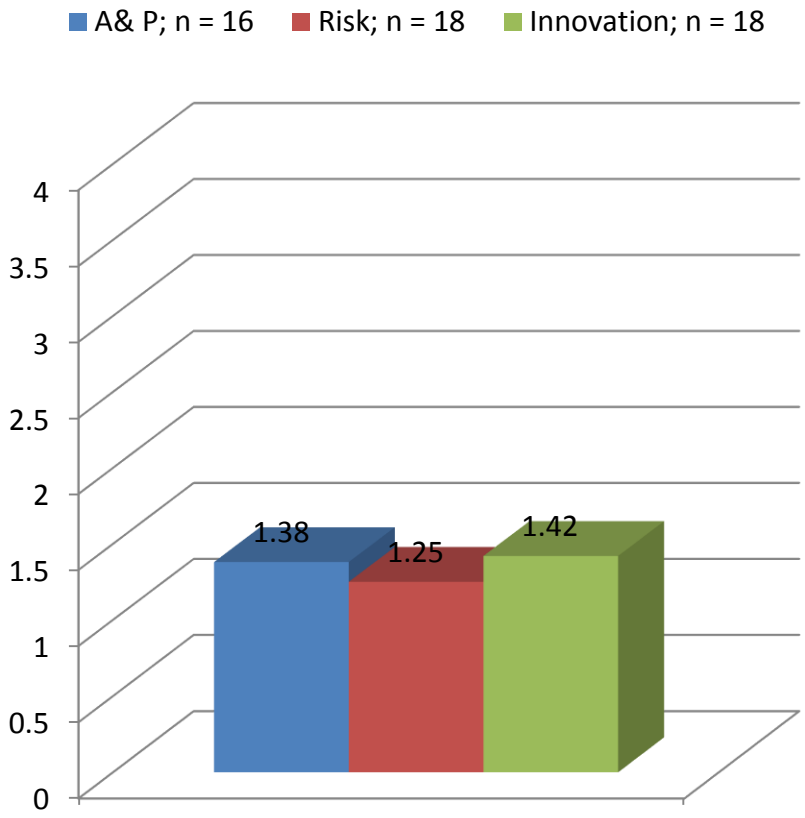
Quantitative Thinking (AAC&U)



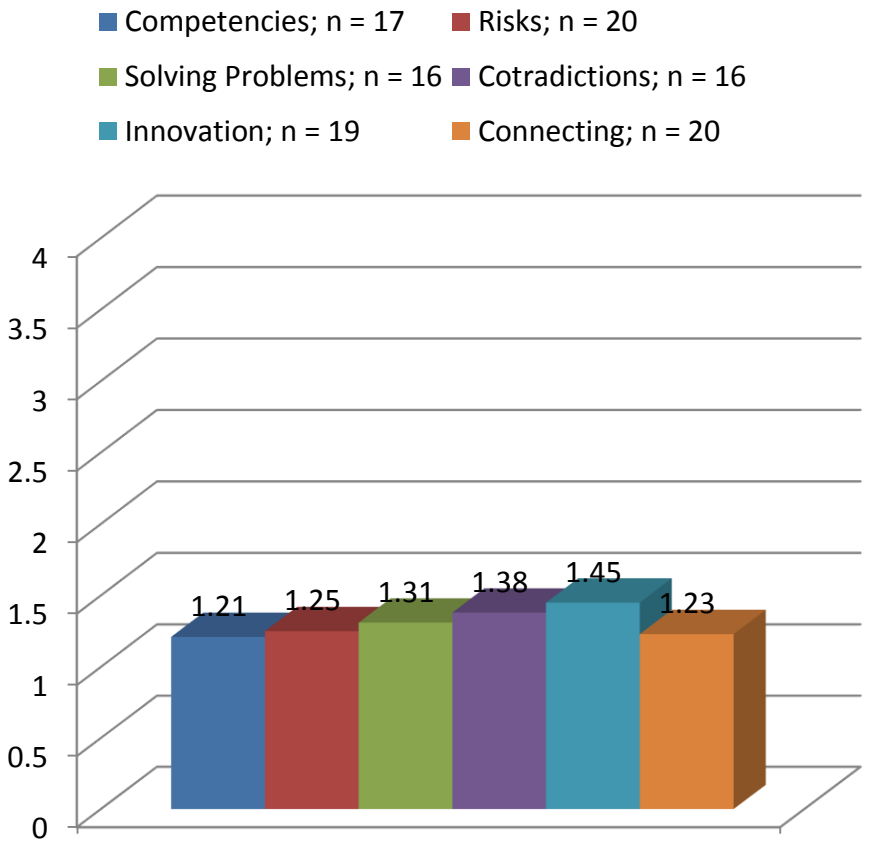
FYS Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. All FYS courses are 100-level.

Creative Thinking (MU)



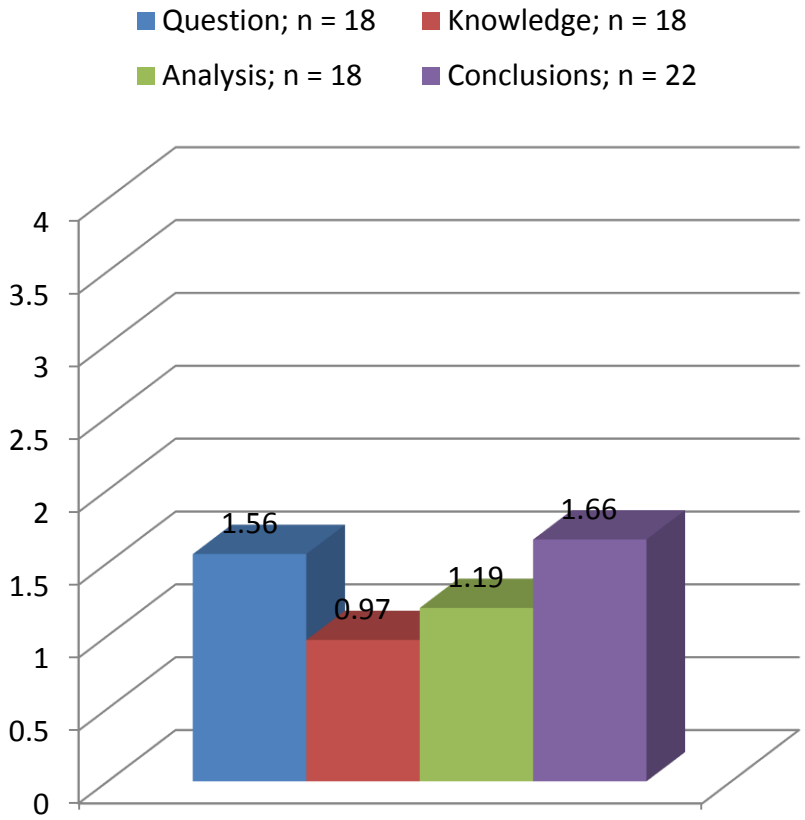
Creative Thinking (AAC&U)



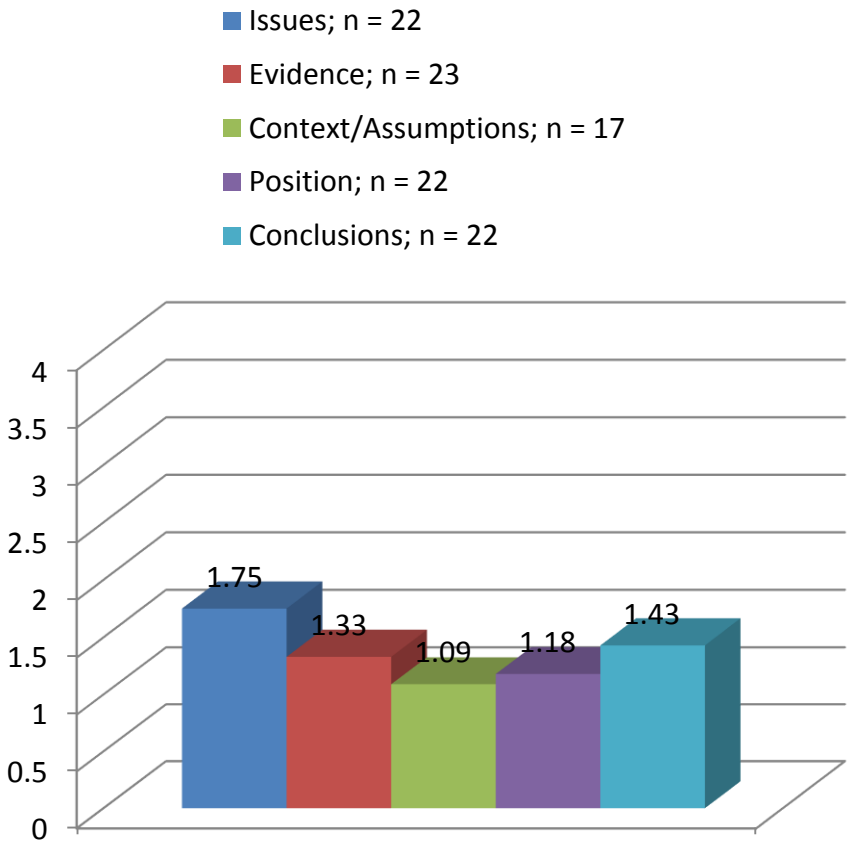
FYS Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. All FYS courses are 100-level.

Inquiry-Based Thinking (MU)



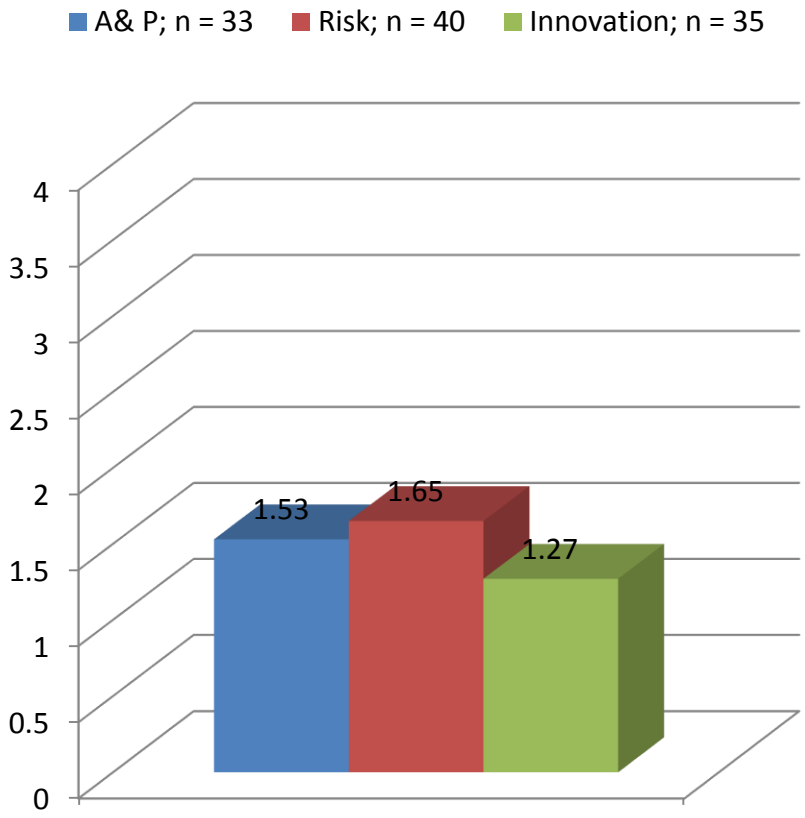
Inquiry-Based Thinking (AAC&U)



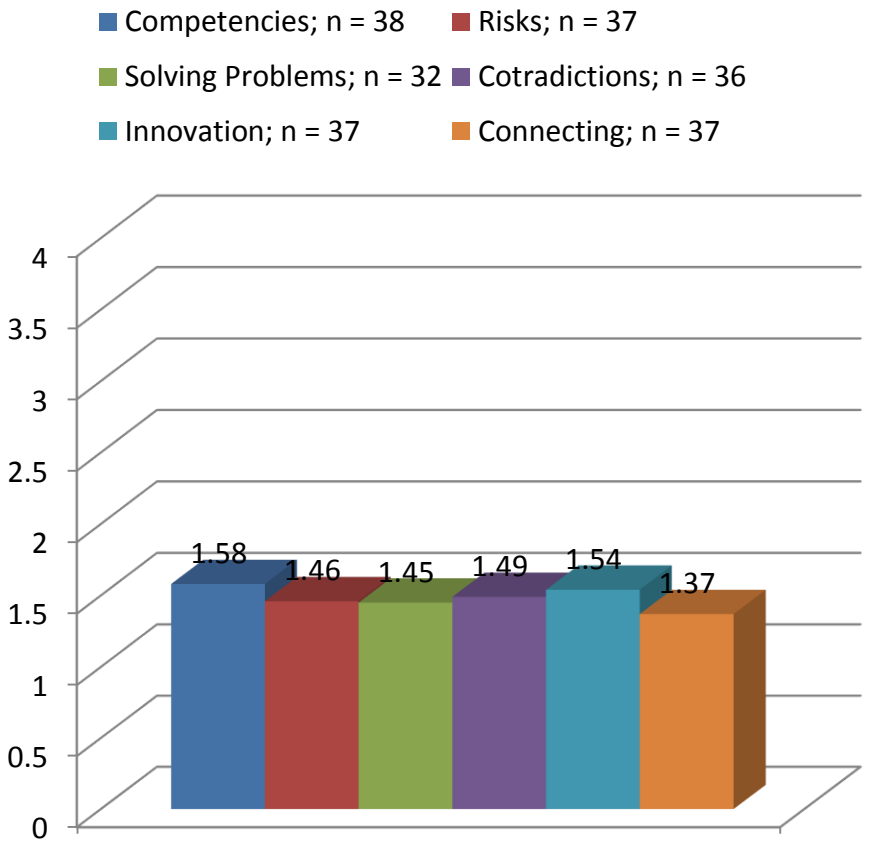
Writing Intensive Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. Some artifacts were from courses that, in addition to being WI, also were multicultural, international, CT, or Core II.

Creative Thinking (MU)



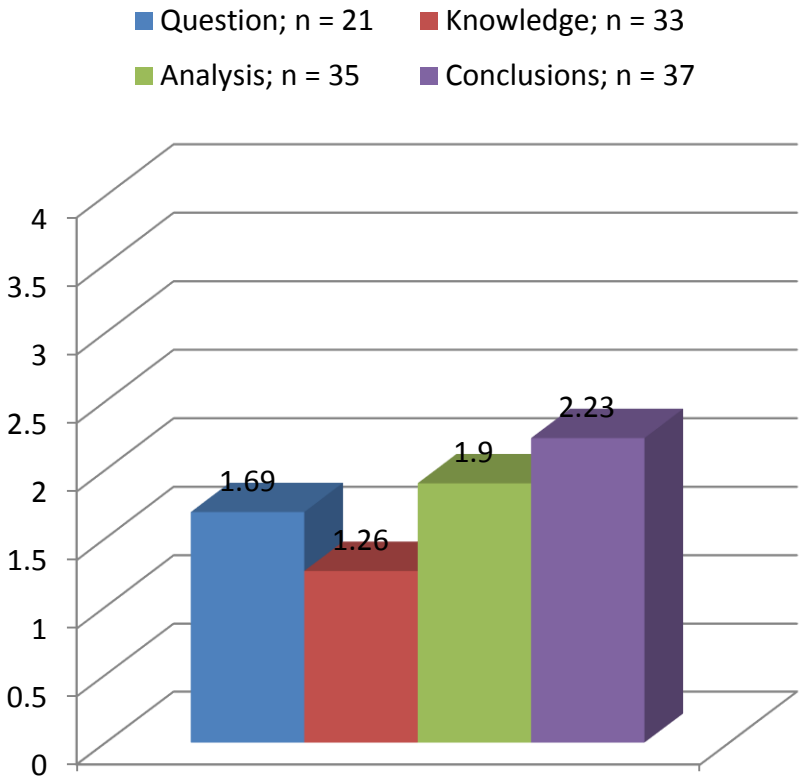
Creative Thinking (AAC&U)



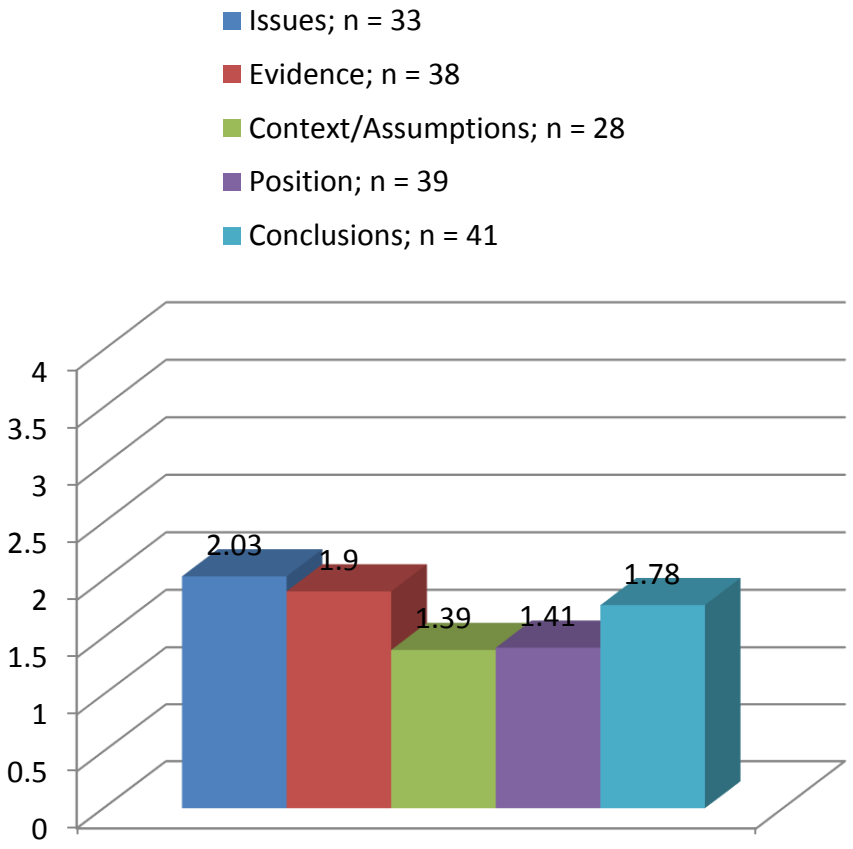
Writing Intensive Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. Some artifacts were from courses that, in addition to being WI, also were multicultural, international, CT, or Core II.

Inquiry-Based Thinking (MU)



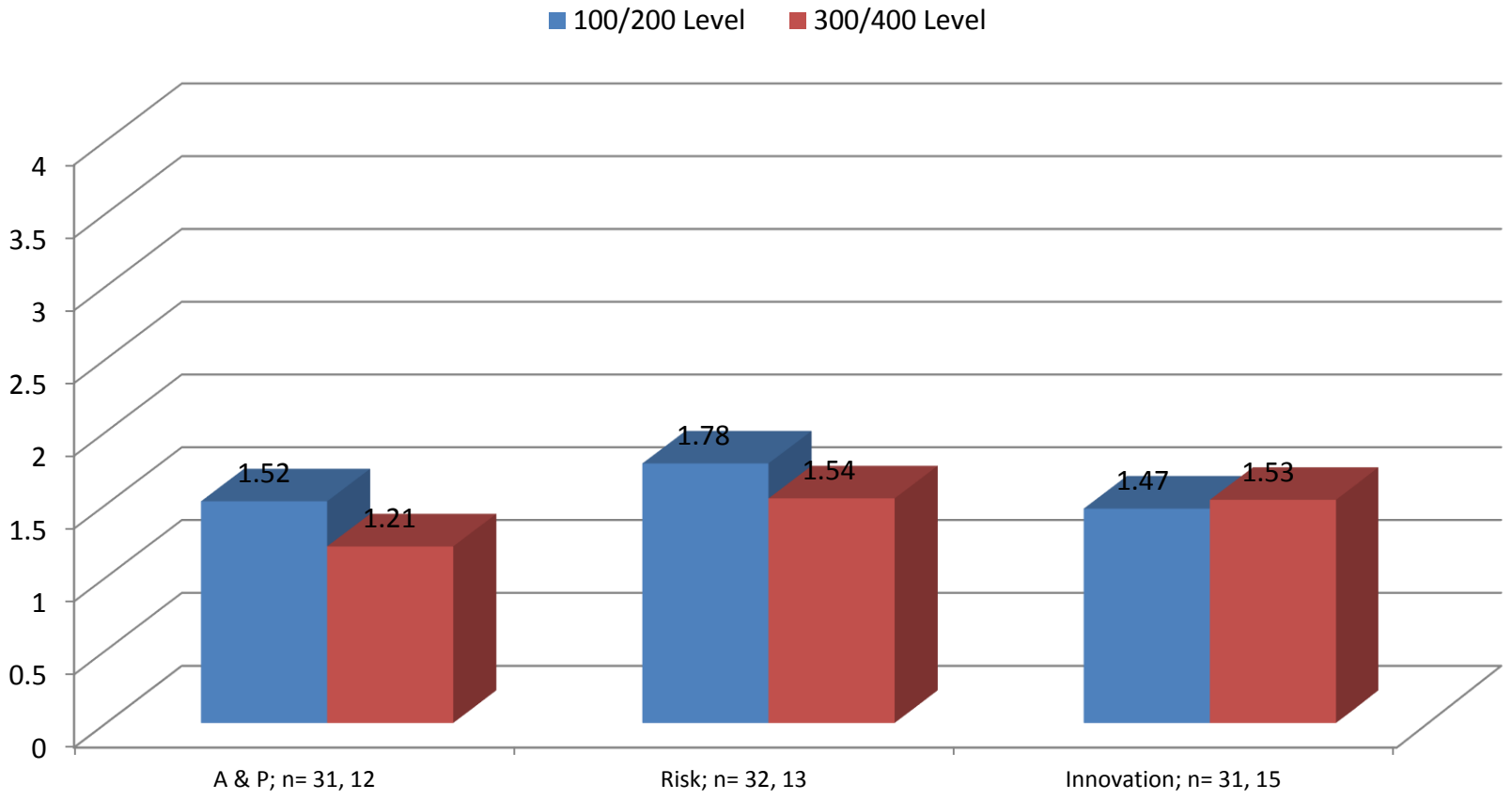
Inquiry-Based Thinking (AAC&U)



Writing Intensive Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. Some artifacts were from courses that, in addition to being WI, also were multicultural, international, CT, or Core II.

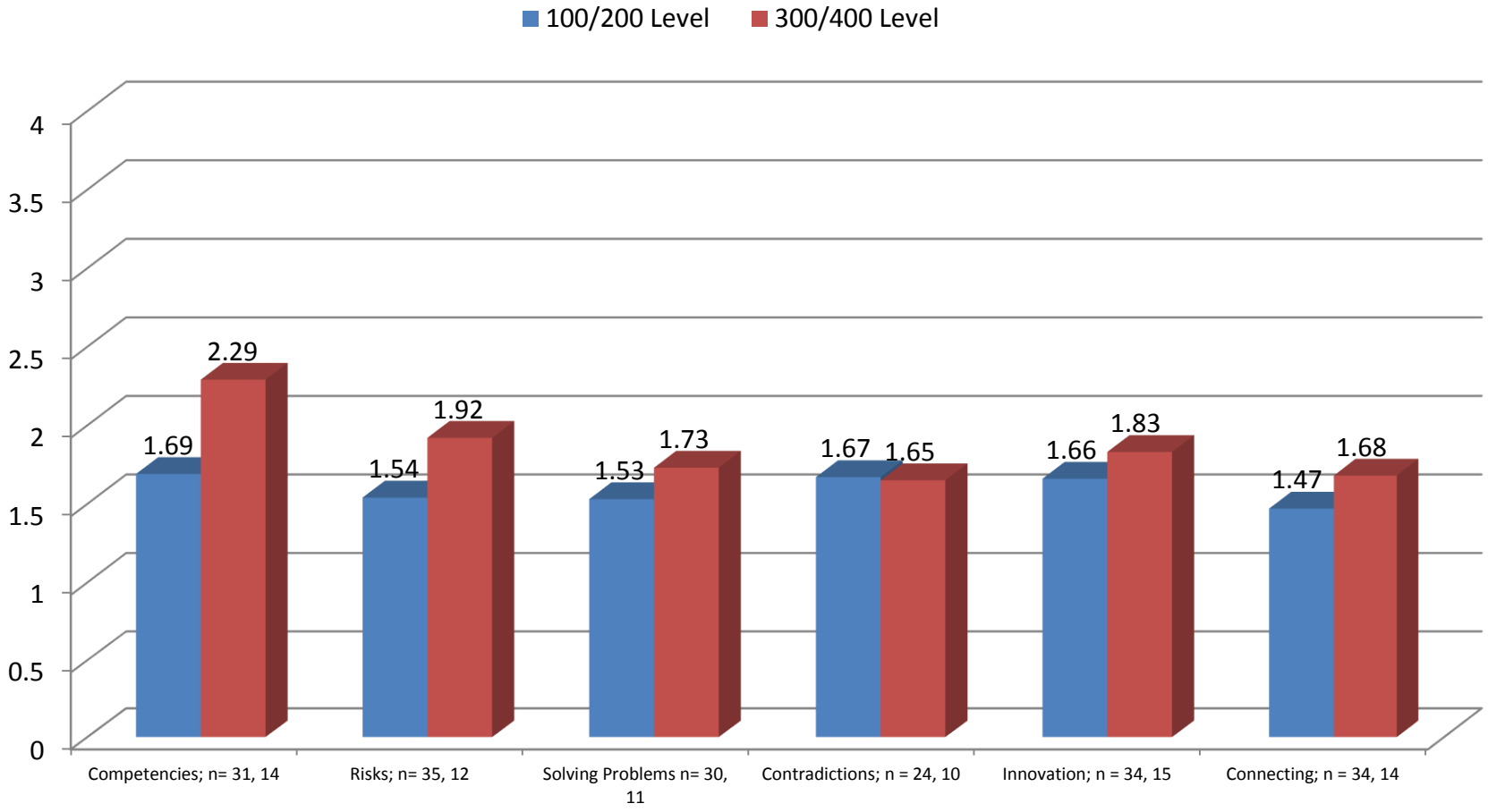
Creative Thinking (MU)



Writing Intensive Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. Some artifacts were from courses that, in addition to being WI, also were multicultural, international, CT, or Core II.

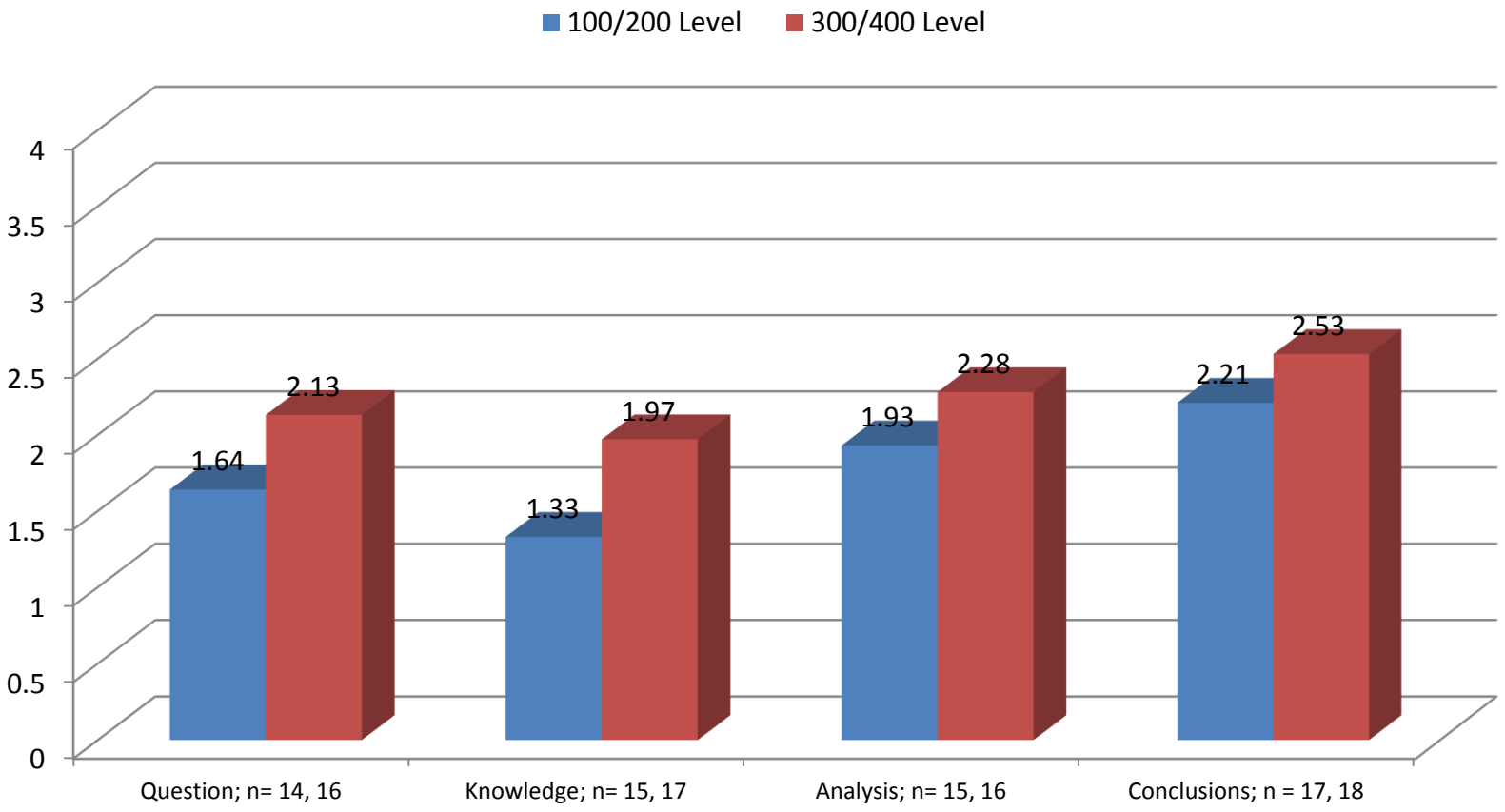
Creative Thinking (AAC&U)



Writing Intensive Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. Some artifacts were from courses that, in addition to being WI, also were multicultural, international, CT, or Core II.

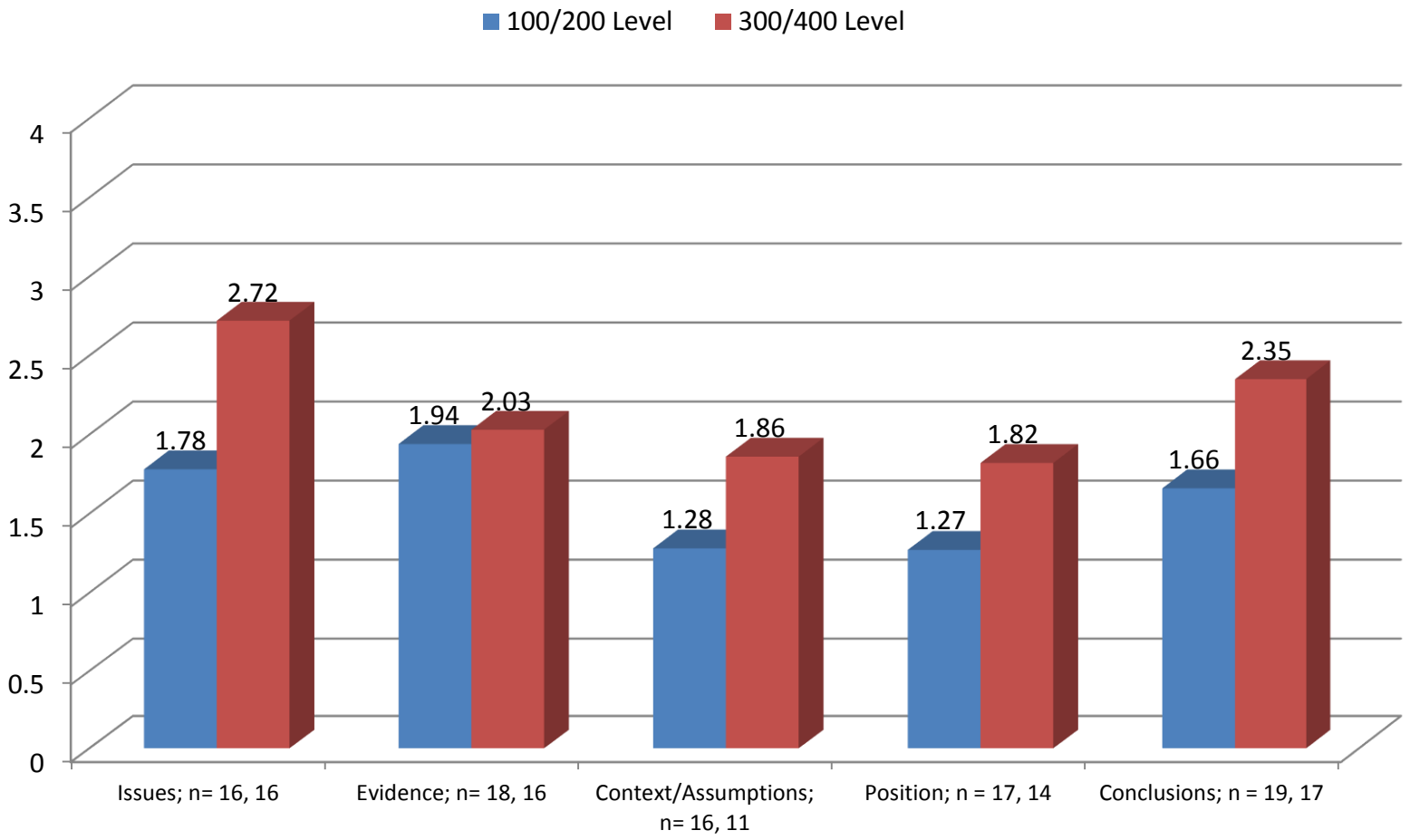
Inquiry-Based Thinking (MU)



Writing Intensive Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. Some artifacts were from courses that, in addition to being WI, also were multicultural, international, CT, or Core II.

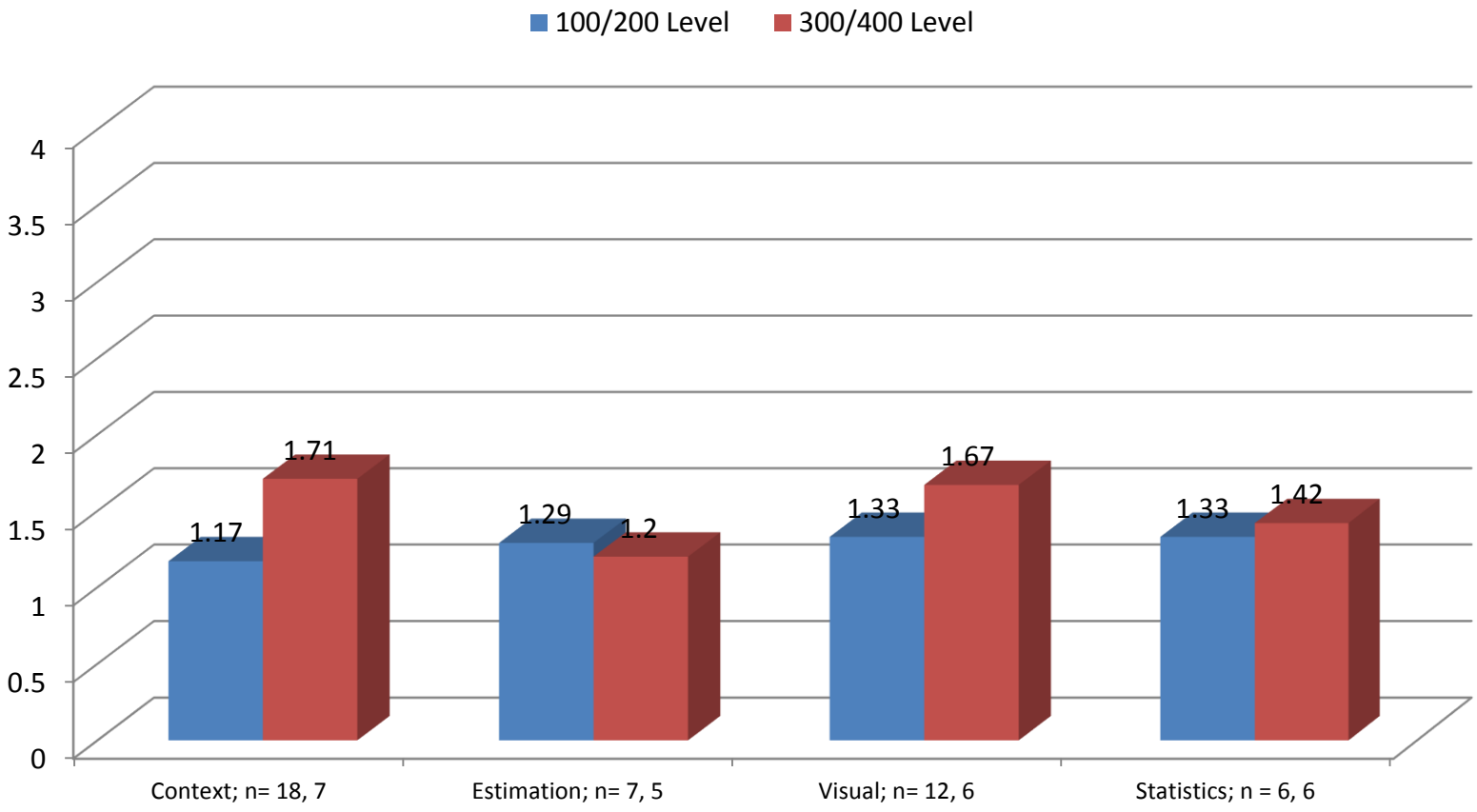
Inquiry-Based Thinking (AAC&U)



Writing Intensive Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. Some artifacts were from courses that, in addition to being WI, also were multicultural, international, CT, or Core II.

Quantitative Thinking (MU)



Writing Intensive Courses

Mean Scores on a scale of 0 – 4, with 4 being the highest possible score. Some artifacts were from courses that, in addition to being WI, also were multicultural, international, or CT.

Quantitative Thinking (AAC&U)