Analysis of Artifacts aligned to Marshall's Baccalaureate Degree Profile (BDP)
Academic Year 2019 – 2020

We dedicate this report to the memory of Professor Joan St. Germain, who was a dedicated member of this Team for seven years (from 2013-2019). We miss her!!

Summer Assessment Team Members: Marie Archambault, Cam Brammer, Kim DeTardo-Bora, Robert Ellison, Victor Fet, Marty Laubach, Anita Walz, and Mary Welch

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

Executive Summary

Background

Recommendations from the 2019 Assessment Workgroup (Updates are in red).

1. Interrater reliability analyses this year showed that the need for reviewers to make judgments regarding which outcome traits instructors had intended to align their assignments resulted in much weaker score agreement than was the case when raters were certain of the instructors' intended trait alignments. For *Intercultural Thinking* and for *Ethical and Civic Thinking* the percentage of rater dyads where one rater gave the artifact a score and the other said it was not aligned to the trait ranged between 19% and 35%. The best way to address this problem is to reiterate to instructors that it is important that they provide assignment instructions that clearly state to which traits of the

BDP outcome their assignments align. Currently, directions for doing this are included in the instructor tab within Blackboard. The Summer Assessment Team suggested that the Office of Assessment and Quality Initiatives create an organization within Blackboard to make these instructions more prominent, as well as to feature other important information about aligning assignments to Marshall's BDP and other important outcomes (e.g. capstone, FYS, and baseline). Instructors can also receive emails through the Blackboard organization drawing their attention to these instructions. After discussion in our final meeting we decided that the Online Design Center will work with the Office of Assessment and Quality Initiatives to design a more explicit assignment to outcome trait alignment system that will make more explicit how to align to various traits within an outcome. Specific outcome statements for each trait will be provided for this purpose. A subcommittee of the University's General Education Committee was formed in fall 2019. Among their recommendations was that we create organizations within Blackboard to provide shorter general education results summaries and recommendations to improve pedagogical practices. Chris Sochor has created Blackboard organizations for FYS, Multicultural and International Courses, CT courses, and Writing Intensive Courses. Following discussion among Assessment Team members this summer, we decided on some modifications to last year's recommendation (please refer to recommendations 1 and 3 at the end of this report). The Office of Assessment and Quality Initiatives will create short reports to be disseminated to faculty via the methods discussed in recommendation 3 (end of report) following this summer's assessment.

2. An alternative to recommendation #1 is to allow the Summer Assessment Team to evaluate each assignment and arrive at consensus regarding the traits to which it aligns <u>before</u> beginning the assessment process. This recommendation also was made at the end of the summer 2018 assessment process. However, the team believes that recommendation #1 would be preferable, primarily due to the large number of potential assignments to be reviewed and the best use of time during the assessment project. We have not yet implemented this suggestion, but it is still under consideration.

3. The Summer Assessment Team recommends that results of these assessments be reported to Faculty Senate and to the General Education Council. In reviewing artifacts aligned to *Intercultural Thinking* this year, it became apparent that very few courses with **International** designations aligned assignments to the *Intercultural Thinking* outcome in Blackboard. The Office of Assessment and Quality Initiatives will provide the General Education Council with an inventory of *Multicultural* and *International* courses that are providing uploads aligned to *Intercultural Thinking*. The Summer Assessment Team recommends that the General Education Council follow-up with Department Chairs regarding this requirement and that, if courses fail to comply, their **Multicultural** or **International** designations be removed. This has not been accomplished (please refer to recommendation 5 at the end of this report.

Procedures for 2020 Assessment

General Procedures

In May 2020 we evaluated student artifacts produced in response to course assignments aligned to *Information Literacy*, *Integrative Thinking*, and *Metacognitive Thinking*.   A group of eight faculty representing several academic colleges from across the university evaluated a sample of these artifacts using Marshall's outcome specific rubrics. These rubrics are included in the supporting documentation. Our sample initially consisted of 300 artifacts, 100 per outcome. However, during scoring raters agreed that 41 artifacts (13 aligned to *Information Literacy*, nine to *Integrative Thinking,* and 19 to *Metacognitive Thinking*) were not aligned to the outcomes to which they had been tagged. This reduced the number of artifacts to 259 (87 *Information Literacy*, 91 *Integrative Thinking*, and 81 *Metacognitive Thinking*). One additional artifact aligned to *Integrative Thinking* and another aligned to *Metacognitive Thinking* could not be accessed, reducing the number of scorable artifacts to 257 (87 for *Information Literacy*, 90 for *Integrative Thinking*, and 80 for *Metacognitive 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 and Quality Initiatives.

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, or the evaluator saw no evidence of the trait in the student's work.
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.	
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 forty-one (241; 80%) of the artifacts in our sample were drawn from courses at the 100/200 level, with the remaining 59 (20%) drawn from courses at the 300/400 level.

Results and Analysis

One challenge in reporting results of Baccalaureate Degree Profile (BDP) assessment is that, although we began with a sample of 300 artifacts (each of which was designated by instructors to align to one of the BDP outcomes assessed this year), each artifact was analyzed by outcome trait. The total number of traits across the three outcomes was 10 (four traits each for *Information Literacy* and for *Integrative Thinking* and two traits for *Metacognitive Thinking*). As mentioned previously, 41 artifacts were judged to not be aligned to any trait of the outcome to which they had been tagged (or there was no evidence in the student work that they had been). An additional two artifacts could not be scored, one because it was blank and the second because it was saved in a format that assessors could not open. This left 257 scorable artifacts. However, not all 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 Team aligned to a total of 767 traits using the MU rubrics (312 for *Information Literacy*, 314 for *Integrative Thinking*, and 141 for *Metacognitive Thinking*).

Outcome	Trait (MU rubric)	Total Traits Aligned
Information Literacy	Relevance of Sources	82
	Integration of Information	77
	Assumptions and Biases	67
	Citation	86
Total for Information Literacy		**312**
Integrative Thinking	Connections among Disciplines	78
	Relation among Domains of Thinking	79
	Transfer	71
	Connections to Experience	86
Total for Integrative Thinking		**314**
Metacognitive Thinking	Project Management	63
	Self-Evaluation	78
Total for Metacognitive Thinking		**141**
Totals		**767**

Results based on course level were as follows:

Information Literacy				Integrative Thinking				Metacognitive Thinking			
Trait	Course Level	Number	Mean (SD)	Trait	Course Level	Number	Mean (SD)	Trait	Course Level	Number	Mean (SD)
Relevance of Sources	100/200	67	2.77 (0.90)	Connections among Disciplines	100/200	64	1.79 (0.65)	Project Management	100/200	34	1.62 (0.70)
	300/400	15	3.43 (0.46)		300/400	14	2.12 (0.82)		300/400	29	2.12 (0.78)
Integration of Information	100/200	64	2.57 (0.83)	Relation among Domains of Thinking	100/200	65	1.81 (0.54)	Self-Evaluation	100/200	49	1.95 (0.82)
	300/400	13	3.04 (0.63)		300/400	14	2.07 (0.65)		300/400	29	2.33 (0.77)
Assumptions and Biases	100/200	57	1.41 (0.68)	Transfer	100/200	57	1.99 (0.67)				
	300/400	10	1.45 (0.64)		300/400	14	2.46 (0.91)				
Citation	100/200	71	2.25 (0.89)	Connections to Experience	100/200	72	1.96 (0.69)				
	300/400	15	2.70 (0.73)		300/400	14	2.25 (0.70)				

For all traits of each outcome, students enrolled in courses at the 300/400 level had higher mean scores than did students enrolled in courses at the 100/200 level. However, the only mean differences that were statistically significant were those for *Information Literacy*: relevance of sources and for *Metacognitive Thinking*: project management.   We note that 80% of the artifacts assessed were from students enrolled in courses at the 100/200 level.

A perusal of the chart above shows mean performance for artifacts uploaded from 100/200 level courses ranged from 1.41 for *Information Literacy*: assumptions and biases to 2.77 for *Information Literacy*: relevance of sources. Means for 300/400 level courses ranged from 1.45 for I*nformation Literacy*: assumptions and biases to 3.43 for *Information Literacy*: relevance of sources. With the exception of the trait assumptions and biases, student performance on artifacts aligned to *Information Literacy* was stronger than their performance on *Integrative* and *Metacognitive Thinking*.

Frequency Analysis

Information Literacy				Integrative Thinking				Metacognitive Thinking			
Trait	Course Level	Percent Scoring 3.5 or 4.0	Percent Scoring 2.5 to 4.0	Trait	Course Level	Percent Scoring 3.5 or 4.0	Percent Scoring 2.5 to 4.0	Trait	Course Level	Percent Scoring 3.5 or 4.0	Percent Scoring 2.5 to 4.0
Relevance of Sources	100/200	36%	73%	Connections among Disciplines	100/200	3%	23%	Project Management	100/200	3%	15%
	300/400	67%	100%		300/400	7%	43%		300/400	7%	52%
Integration of Information	100/200	22%	69%	Relation among Domains of Thinking	100/200	0	18%	Self-Evaluation	100/200	4%	37%
	300/400	46%	85%		300/400	0	36%		300/400	14%	55%
Assumptions and Biases	100/200	2%	16%	Transfer	100/200	4%	28%				
	300/400	0	10%		300/400	29%	57%				
Citation	100/200	13%	51%	Connections to Experience	100/200	6%	33%				
	300/400	27%	67%		300/400	7%	43%				
Overall	100/200	19%	53%	Overall	100/200	3%	26%	Overall		4%	28%
	300/400	38%	70%		300/400	11%	45%			10%	53%

Frequency analysis showed that, for *Information Literacy,* 38% of students enrolled in courses at the 300/400 level received final scores of 3.5 or 4.0, while 70% of students in these courses received final scores of at least 2.5. However, only 11% and 10% of students enrolled in 300/400 level courses received scores of 3.5 or 4.0 in *Integrative Thinking* and in *Metacognitive Thinking*, respectively. We must keep in mind, however, that there were very few artifacts drawn from students in 300/400 level courses for both *Information Literacy* and *Integrative Thinking*. Even for *Metacognitive Thinking*, the majority of the assignments we scored this year were from 100/200 level courses.

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 analyzed this year (Critical Thinking [CT], Writing Intensive [WI], Core II, and Capstone) 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.

Critical Thinking (CT) Courses

CT courses in the assessment sample included those that aligned to each of the outcomes assessed. All CT courses are at the 100/200 level. Results are below:

Information Literacy			Integrative Thinking			Metacognitive Thinking		
Trait	Number	Mean Score	Trait	Number	Mean Score	Trait	Number	Mean Score
Relevance of Sources	40	2.79	Connections among Disciplines	54	1.84	Project Management	7	1.64
Integration of Information	41	2.59	Relation among Domains of Thinking	54	1.80			
Assumptions and Biases	36	1.46	Transfer	51	2.0	Self-Evaluation	18	2.06
Citation	41	2.24	Connections to Experience	62	2.0			

While, due to the relatively small /n/s, the results should be interpreted with caution, mean scores for students in Marshall's CT courses (which are at the 100 and 200 level) suggest performance at level 2 or higher on three of the four traits of *Information Literacy*, on two of the four traits of *Integrative Thinking*, and on one of the two traits of *Metacognitive Thinking*.

Core II Courses

Core II courses in the assessment sample included those that aligned to each of the outcomes assessed: *Information Literacy, Integrative Thinking,* and *Metacognitive Thinking.* All Core II courses are at the 100/200 level. Results are below:

Information Literacy			Integrative Thinking			Metacognitive Thinking		
Trait	Number	Mean Score	Trait	Number	Mean Score	Trait	Number	Mean Score
Relevance of Sources	32	2.73	**Connections among Disciplines**	48	1.84	**Project Management**	7	1.64
Integration of Information	31	2.40	**Relation among Domains of Thinking**	48	1.82			
Assumptions and Biases	29	1.48	**Transfer**	45	1.93	**Self-Evaluation**	18	2.06
Citation	31	2.36	**Connections to Experience**	55	1.96			

While, due to the relatively small /*n*/s, the results should be interpreted with caution, mean scores for students in Marshall's Core II courses (which are all at the 100 and 200 level) suggest performance at level 2 or higher for three of the four traits of *Information Literacy* and for one of the two traits of *Metacognitive Thinking*. Mean scores approached, but did not reach, Level 2 for all traits of *Integrative Thinking*.

Writing Intensive (WI) Courses

WI courses in the assessment sample aligned to all outcomes assessed: *Information Literacy, Integrative Thinking,* and *Metacognitive Thinking.* Results are given below:

Information Literacy			Integrative Thinking			Metacognitive Thinking		
Trait	Number	Mean Score	Trait	Number	Mean Score	Trait	Number	Mean Score
Relevance of Sources	32	3.06	**Connections among Disciplines**	42	1.92	**Project Management**	22	1.77
Integration of Information	28	2.59	**Relation among Domains of Thinking**	42	1.93			
Assumptions and Biases	27	1.33	**Transfer**	40	2.06	**Self-Evaluation**	32	2.31
Citation	32	2.52	**Connections to Experience**	45	2.02			

While, due to the relatively small /*n*/s, the results should be interpreted with caution, mean scores for students in Marshall's WI courses suggest performance at level 2 or higher in three of the four traits of *Information Literacy*, in two of the four traits of *Integrative Thinking*, and on one of the two traits of *Metacognitive Thinking*. Although WI courses were at both 100/200 and 300/400 levels, the proportion of artifacts from 300/400 levels included in this analysis was small. Please refer to the supporting documentation for a specific breakdown.

Capstone Courses

Capstone courses in this assessment sample primarily aligned to *Metacognitive Thinking* and all courses were at the 400 level. Please see the supporting documentation that follows this summary for additional detail.

Trait	Number	Mean Score
Project Management	22	2.20
Self-Evaluation	22	2.11

While, due to the relatively small /*n*/s, the results should be interpreted with caution, mean scores for students in the capstone sample suggest performance between Levels 2 and 3.

FYS Courses

FYS courses in the assessment sample aligned to all outcomes assessed: *Information Literacy, Integrative Thinking,* and *Metacognitive Thinking.* Results are given below:

Information Literacy			Integrative Thinking			Metacognitive Thinking		
Trait	Number	Mean Score	Trait	Number	Mean Score	Trait	Number	Mean Score
Relevance of Sources	24	2.77	**Connections among Disciplines**	7	1.50	**Project Management**	26	1.60
Integration of Information	21	2.55	**Relation among Domains of Thinking**	8	1.88			
Assumptions and Biases	20	1.35	**Transfer**	5	1.70	**Self-Evaluation**	30	1.88
Citation	27	2.30	**Connections to Experience**	8	1.69			

While, due to the relatively small /*n*/s, the results should be interpreted with caution, mean scores for students in Marshall's FYS course, the majority of whom are freshmen, suggest performance between Levels 2 or higher on three of the four traits of Information Literacy, at Levels 1 or 2 on all traits of *Integrative Thinking* and *Metacognitive Thinking*.

Conclusion

We used rubrics this year that measured student performance according to the level of sophistication they demonstrated in achievement of each trait of the three Baccalaureate Degree Profile (BDP) outcomes we assessed. BDP outcomes specify what students are expected to achieve at the time they receive their baccalaureate degrees. Admittedly, the proportion artifacts from 300/400 level courses in our sample was small this year, with only 15 artifacts aligning to *Information Literacy*, 14 to *Integrative Thinking*, and 30 to *Metacognitive Thinking*. However, we were pleased that 70% of students who submitted artifacts from 300/400 level courses received overall scores of 2.5 or higher in *Information Literacy*, with 38% receiving scores of 3.5 or 4.0. Although not as high, 53% and 45% of students who submitted artifacts from 300/400 level courses received overall scores of 2.5 or higher on *Metacognitive* and *Integrative Thinking*, respectively.   A score of 2.5 indicates that at least one rater assigned a score of Level 3 to the artifact, a score of 3 indicates that both raters assigned a score of Level 3.0, a score of 3.5 indicates that at least one rater assigned a score of Level 4, and a score of 4.0 indicates that both raters assigned a score of Level 4.

When examining mean performance across all artifacts, we noted that, for *Information Literacy,* assumptions and biases emerged as a relative weakness (*mean* = 1.42; *n* = 67) among the traits of this outcome. Only 14% of the 67 artifacts received scores between 2.5 and 4.0 (as compared to 78% for relevance of sources, 71% for integration of information, and 53% for citation).   For *Integrative Thinking*, we noted little variation among means scores, which ranged from 1.85 for relation among domains of thinking to 2.09 for transfer. Likewise, for *Metacognitive Thinking* (which had only two traits), mean scores were 1.85 for project management and 2.09 for self-evaluation.

Although the majority of *Integrative Thinking* artifacts were drawn from 100/200 level courses, we argue that there is room to improve performance, especially given the emphasis placed on this outcomes in Marshall's CT courses. We also note that the rubric level descriptions we used this assessment cycle may have been too general and suggest studying the AAC&U's rubric for *Integrative Learning* to see if it might better capture student performance on this outcome.

For *Metacognitive Thinking*, two large categories of artifacts were included in this year's analysis. The first was a group of 40 artifacts from FYS that aligned to at least one trait (26 aligned to project management and 30 aligned to self-evaluation) and the second was a group of 22 artifacts from capstone courses aligned to both traits. While mean scores for capstone artifacts were significantly higher than those of FYS for project management (2.20 for capstone as compared to 1.60 for FYS), the mean differences for self-evaluation (2.11 for capstone and 1.88 for FYS) did not differ significantly.   We suggest there is room for improvement on this outcome for students about to graduate from Marshall University. We recommend that more emphasis be placed on these skills throughout the Marshall curriculum.

The Office of Assessment and Quality Initiatives is currently examining the mapping of degree program outcomes to traits of the BDP. We will continue to work with programs that have not yet completed this analysis.

Recommendations from the 2020 Assessment Team

The Summer Assessment Team made the following recommendations:

1. That the Baccalaureate Degree Profile outcomes be reconfigured in Blackboard to allow instructors to align each assignment to individual outcome traits. Chris Sochor, Instructional Designer in Online Learning, said that it is possible to do this. The rationale for this recommendation is that it will focus instructors more closely on the specific elements (and definitions) of each outcome. Aligning to traits rather than holistically to an outcome should reduce the number of artifacts that assessors judge to lack specific outcome trait alignment.
2. That faculty be reminded to have students upload *final* versions of *summative* assignments for assessment.
3. That we use Microsoft Teams (rather than Blackboard Organizations) to communicate general education assessment information to specific constituencies. A special emphasis during 2020-2021 will be to communicate with faculty the need to align assignments to outcome traits. We will include the definitions of all outcome traits.

4. If recommendations 1 and 3 are not accomplished, consider having the Summer Assessment Team review all assignments that will be part of its reviews <u>before</u> beginning to score artifacts to determine the appropriateness of these assignments to each outcome trait.
5. That we communicate the results of general education assessment <u>and</u> recommendations of the Summer Assessment Team to the General Education Council.
6. That we examine the mapping of degree program outcomes to those of Marshall's Baccalaureate Degree Profile (BDP). These data are currently being collected in Taskstream and the Office of Assessment and Quality Initiatives will analyze the mappings completed to date and continue to work with degree programs that have not completed the mappings.
7. Use the analysis from recommendation 6 to recommend possible modifications to BDP traits.

Supporting Documentation

Baccalaureate Degree Profile Artifact Assessment

Academic Year 2019 – 2020

Outcomes Assessed: MU Rubrics

Outcome	Abbreviation	Traits	Abbreviations
Information Literacy	**IL**	Relevance of Sources	Relevance
		Integration of Information	Integration
		Assumptions and Biases	A & B
		Citation	Citation
Integrative Thinking	**IT**	Connections among Disciplines	Discipline
		Relation among Domains of Thinking	Domain
		Transfer	Transfer
		Connections to Experience	Experience
Metacognitive Thinking	**MT**	Project Management	Project
		Self-Evaluation	Self

Course Types

Course Type	Abbreviation
Critical Thinking	CT
Core II	Core II
Writing Intensive	WI
Senior Capstone	Capstone
First Year Seminar in Critical Thinking	FYS

Course Types in ICT, ECT, and CF Outcome Sample
Each Course Counted Separately for Each Category
(i.e. sample *n* does not add to 300)

Course Type	Course Level	Sample *n*	Total Sample *n*
CT	100-200	147	147
	300-400	N/A	
Core II	100-200	126	126
	300-400	N/A	
WI	100-200	86	118
	300-400	32	
Senior Capstone	100-200	N/A	25
	300-400	25	
FYS	100-200	87	87
	300-400	N/A	
Total	100-200	384	503
	300-400	57	

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
Information Literacy	838	85	10%	71	15	21%
Integrative Thinking	1,289	86	7%	242	14	6%
Metacognitive Thinking	406	70	17%	78	30	38%
Total	**2,533**	**241**	**10%**	**391**	**59**	**15%**

Sample Frequencies
Total # of artifacts assessed = 100 per outcome

Course Level Frequencies: Information Literacy
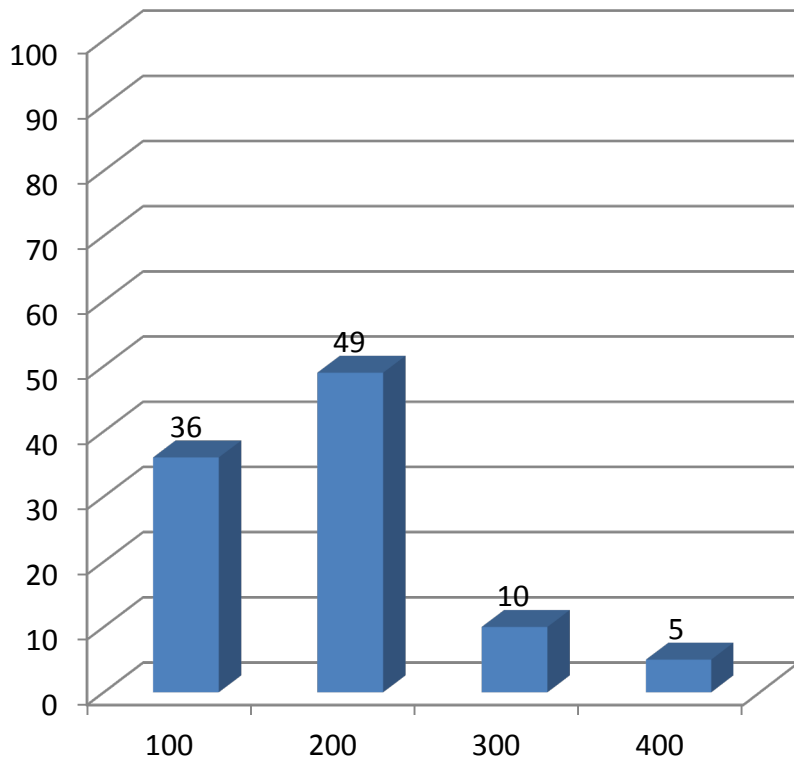
Course Level Frequencies: Integrative Thinking
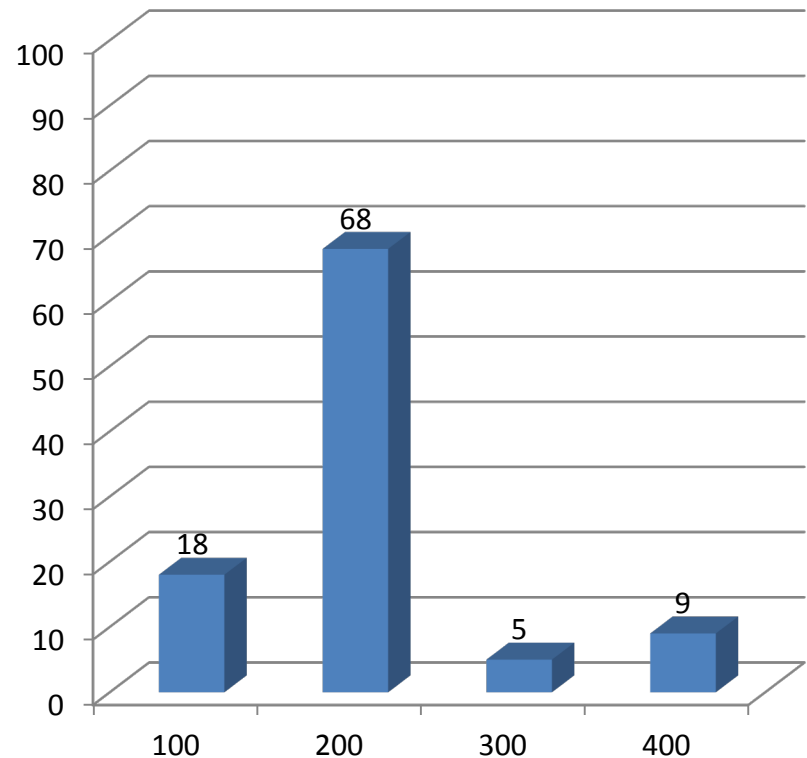
Sample Frequencies

Total # of artifacts assessed = 100 per outcome
Total = 300

Course Level Frequencies: Metacognitive Thinking
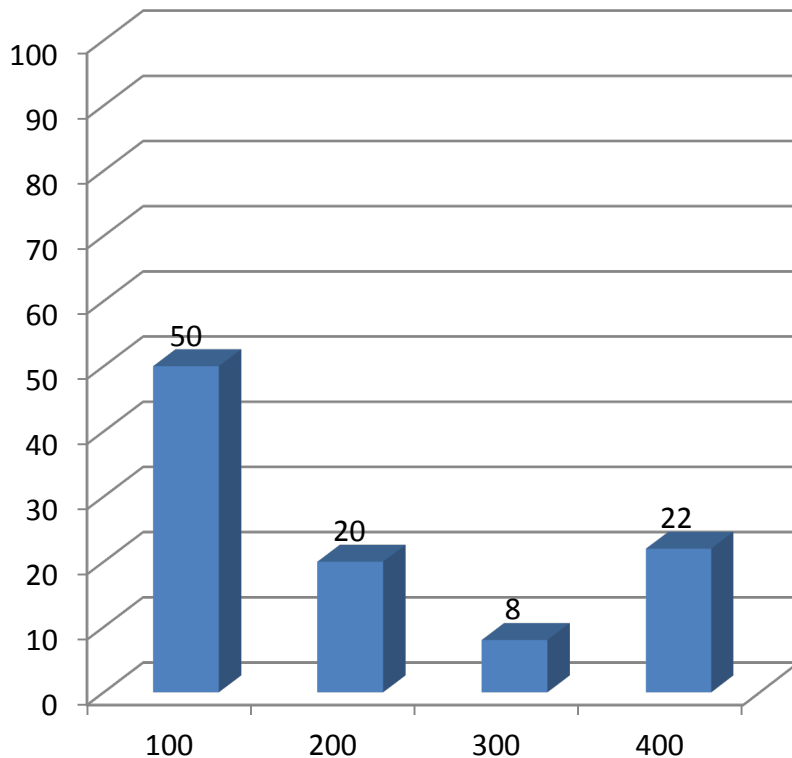
Course Level Frequencies: Total across the three outcomes
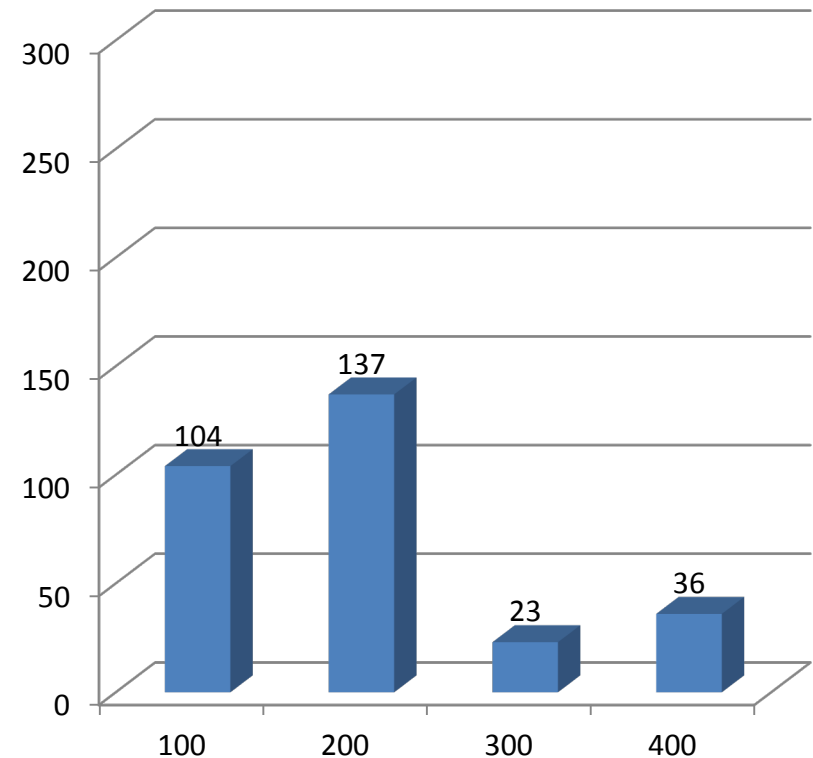
Review Procedures

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

Review Procedures

- We also allowed reviewers to assign a score of *N/A (not applicable)* when they did not see evidence of the trait in the artifact. When one rater assigned a score of *N/A* and the second rater assigned a score of 1 – 4, they also met to discuss the rationale for their scores to see if they could agree on the presence (or not) of the trait in the assignment or artifact. If they could not agree, a third reader was assigned.

Third Readers for this Year's Review

- We had four artifacts that required a third review. For two of the artifacts, reviewers could not agree between a score of N/A and a numerical score for one trait, while they could not agree between scores of N/A and numerical scores for four traits of the third artifact. In all three cases, a third reviewer, who was unaware of the scores of the first two reviewers for each of the three artifacts, provided either numerical or N/A scores that matched the scores of one of the original reviewers or differed by one point.

- The original reviewers for a fourth artifact settled on numerical scores that were 2 points apart. The third reviewer was able to resolve this disagreement, which was one point apart from one of the original reviewers.

Interrater Reliability

- We conducted interrater reliability analyses using the Cohen's Kappa statistical procedure. In so doing, we used the following rules, similar to those suggested Stellmack, Kohneim-Kalkstein, Manor, Massey, & Schmitz (2009):
  - Since our scoring procedure was to average final scores between two raters when scores differed by only one point, we used that averaged score (e.g. 1.5) as the score for both raters, counting it as an agreement in the interrater reliability analysis.
  - When each evaluator rated an artifact trait as *N/A* (i.e. not aligned to the rubric trait), these ratings were counted as agreements in the interrater reliability analysis.
  - For scores that were two or more points apart, the original score of each reviewer was used in the analysis. Therefore, these scores were counted as disagreements.
  - Any time one rater scored the artifact as *N/A* and another provided a score, the scores were counted as disagreements in the analysis.

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

Outcome	Total Artifacts	Total Artifacts Not Able to be Scored	Total Used for Analysis	Notes:
Information Literacy	100	13	87	All 13 received scores of N/A
Integrative Thinking	100	10	90	9 received all scores of N/A and one was not uploaded correctly.
Metacognitive Thinking	100	20	80	19 received all scores of N/A and one was not able to be opened.
Total	**300**	**43**	**257**	

Revised Information Literacy MU Rubric

Information Literacy: Students will **employ** appropriate research tools, **integrate** relevant information from reliable sources, **question and evaluate** information and its sources, and **cite sources** in an academic manner.

Traits: Performance Indicators/Performance Levels	Level 0 (N/A)	Level 1	Level 2	Level 3	Level 4
Relevance of Sources: Employs appropriate research tools (i.e. sources).	Cannot assess due to issues with the assignment and/or the artifact.	**Uses** questionable sources or no sources at all.	**Uses** a combination of reliable and questionable sources.	**Uses** mostly reliable sources or does not reflect the depth required by the artifact.	**Uses** reliable sources that are adequate to the depth required by the artifact.
Integration of Information: Integrates relevant information from reliable sources.	Cannot assess due to issues with the assignment and/or the artifact.	**Fails to integrate** information (disorganized presentation).	**Inappropriate integration** of information (inappropriate sequential presentation, e.g. one source at a time).	**Some integration** of information relevant to the type of artifact (mix of appropriate and inappropriate presentation).	**Appropriately integrates** information relevant to the type of artifact.
Assumptions and Biases: Questions and **evaluates** information and its sources.	Cannot assess due to issues with the assignment and/or the artifact.	**Reports** information at face value.	**Shows** some evidence of appropriately questioning assumptions and biases of sources.	**Shows** evidence of appropriately questioning multiple assumptions and biases of sources.	**Analyzes** assumptions and biases and **evaluates** the relevance of contexts as described in sources.
Citation: Cites sources in an academic manner.	Cannot assess due to issues with the assignment and/or the artifact.	**Inappropriate** use of paraphrasing, quotes, and/or citations.	**Inconsistent** use of paraphrasing, quotes, and/or citations.	**Acceptable** use of paraphrasing, quotes, and/or citations.	**Correct** use of paraphrasing, quotes, and citations using the style manual of the discipline (as appropriate).

Revised Integrative Thinking MU Rubric

Integrative Thinking: Students will **make connections** and **transfer** skills and learning among varied disciplines, domains of thinking, experiences, and situations.

| Traits: Performance Indicators/Performance Levels	Level 0 (N/A)	Level 1	Level 2	Level 3	Level 4
Connections among Disciplines: Connects examples, facts, or theories from more than one discipline.	Cannot assess due to issues with the assignment and/or the artifact.	**Connects** in a rudimentary manner.	**Connects** in an emerging manner.	**Connects** in a thorough manner.	**Creates** wholes out of multiple parts (synthesizes) or draws conclusions.
Relation among Domains of Thinking: Connects examples, facts, or theories from more than one of Marshall's Domains of Thinking.	Cannot assess due to issues with the assignment and/or the artifact.	**Connects** in a rudimentary manner.	**Connects** in an emerging manner.	**Connects** in a thorough manner.	**Creates** wholes out of multiple parts (synthesizes) or draws conclusions.
Transfer: Adapts and applies skills, abilities, theories, or methodologies gained in one situation and/or discipline to other situations and/or other disciplines.	Cannot assess due to issues with the assignment and/or the artifact.	**Adapts or applies** in a rudimentary manner.	**Adapts or applies** in an emerging manner.	**Adapts and applies** in a thorough manner.	**Adapts and applies** in an original or complex manner.
Connections to Experience: Connects relevant experience and academic knowledge.	Cannot assess due to issues with the assignment and/or the artifact.	**Connects** in a rudimentary manner.	**Connects** in an emerging manner.	**Connects** in a thorough manner.	**Creates** wholes out of multiple parts (synthesizes) or draws conclusions.

Revised Metacognitive MU Rubric

Metacognitive Thinking: Students will **evaluate** the effectiveness of a project plan or strategy to **determine** their improvement in knowledge and skills.

Traits: Performance Indicators/Performance Levels	Level 0 (N/A)	Level 1	Level 2	Level 3	Level 4
Project Management: Evaluates the effectiveness of a project plan or strategy.	Cannot assess due to issues with the assignment and/or the artifact.	Evaluation: Superficial (If merely reporting)	Evaluation: Clearly identifies strengths and weaknesses.	Evaluation: • Clearly identifies strengths and weaknesses. • Evidence of continual reflection or improvement. • Proposes an improvement plan that is general in nature.	Evaluation: • Clearly identifies strengths and weaknesses. • Evidence of continual reflection or improvement. • Proposes an improvement plan that is detailed.
Self-evaluation: Evaluates improvement in knowledge and skills.	Cannot assess due to issues with the assignment and/or the artifact.	Reflects: In a superficial manner (If merely reporting)	Reflects: • With some depth • Without evidence of continual reflection or improvement.	Reflects: • In depth • Evidence of continual reflection or improvement. • Acknowledges general changes in perspectives regarding his/her own learning.	Reflects: • In depth • Evidence of continual reflection or improvement. • Evaluates specific changes in perspectives regarding his/her own learning.

Information Literacy

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

(Although there were 87 artifacts in the analysis, not all artifacts aligned to every trait)

Overall Analysis

■ Relevance; n = 82   ■ Integration; n = 77   ■ A&B; n = 67   ■ Citation; n = 86
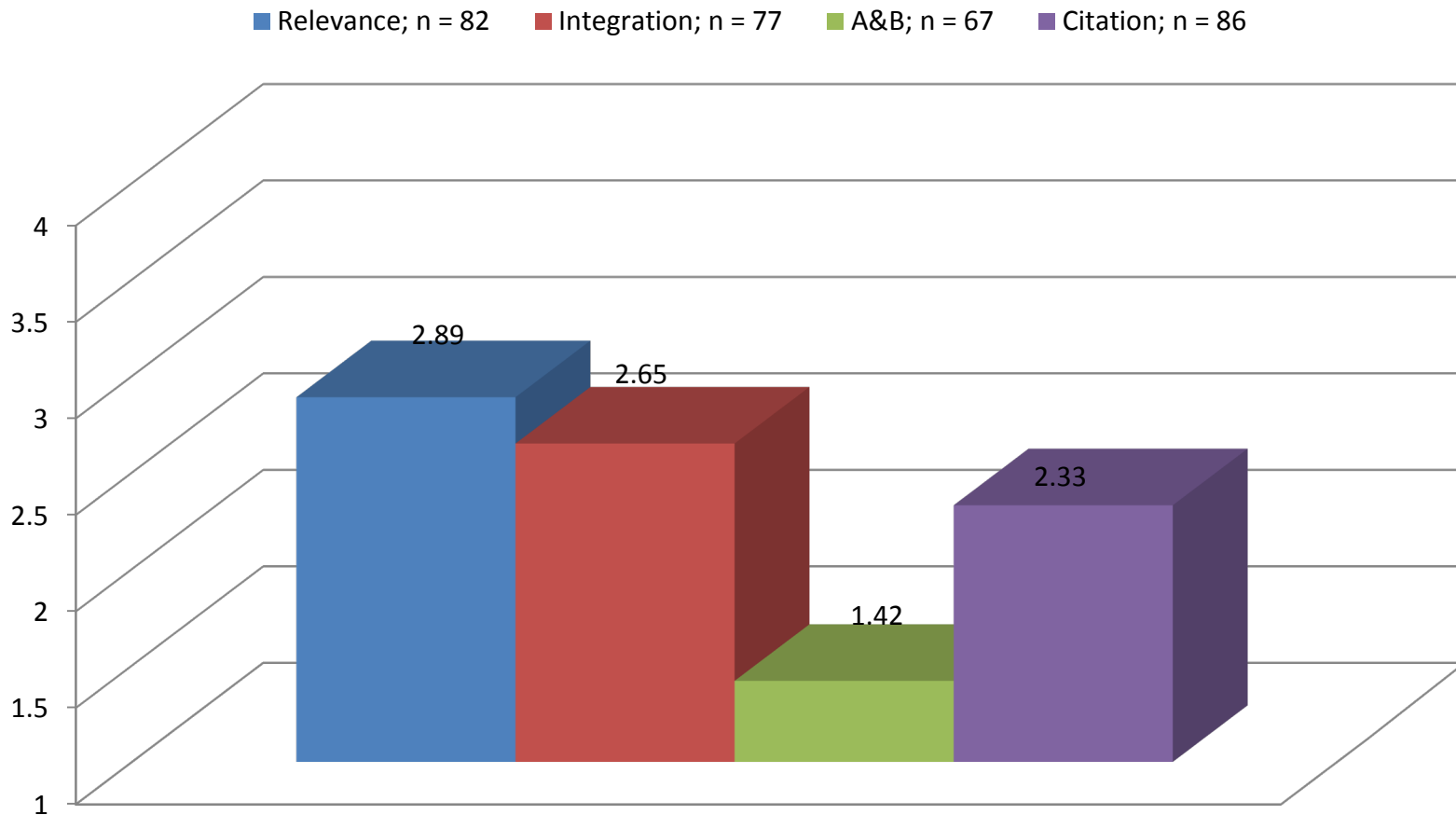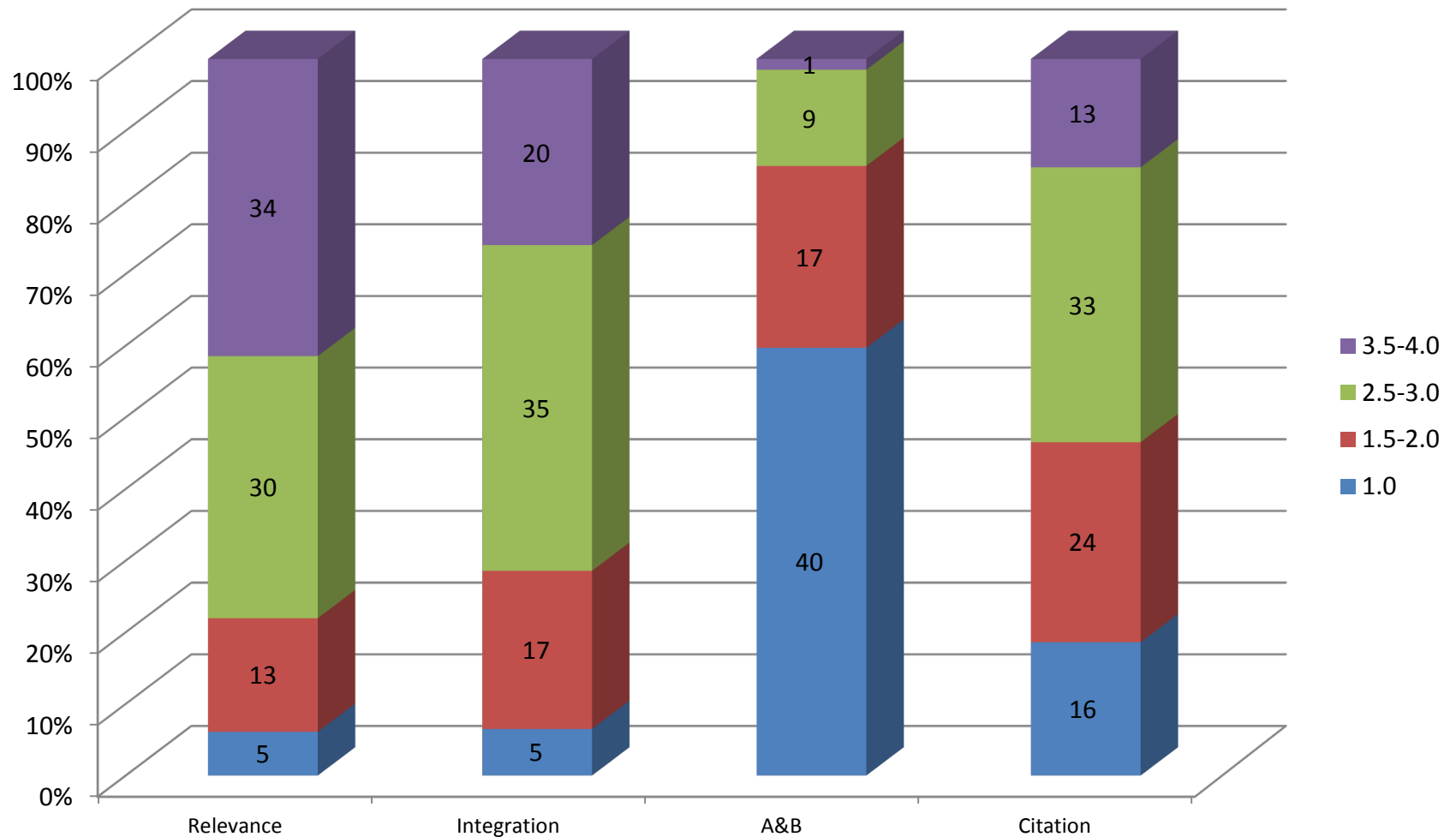
Information Literacy

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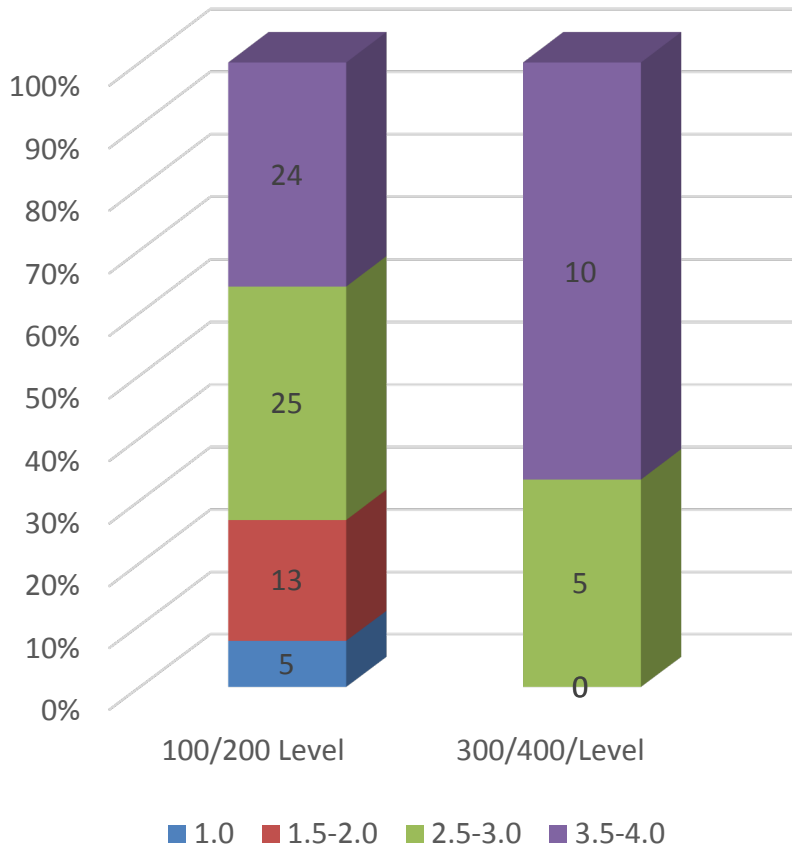
Number of artifacts (with usable scores) scoring at each performance level

Trait/ Performance Level	Relevance of Sources	Integration of Information	Assumptions and Biases	Citation	Total
1.0	5 (6%)	5 (6%)	40 (60%)	16 (19%)	66 (21%)
1.5 – 2.0	13 (16%)	17 (22%)	17 (25%)	24 (28%)	71 (23%)
2.5 – 3.0	30 (37%)	35 (45%)	9 (13%)	33 (38%)	107 (34%)
3.5 – 4.0	34 (41%)	20 (26%)	1 (1%)	13 (15%)	68 (22%)
Total Traits with Usable Scores	**82 (100%)**	**77 (100%)**	**67 (100%)**	**86 (100%)**	**312 (100%)**

Information Literacy

Frequency Analysis

Information Literacy

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score
The course mean difference for *relevance* was significant. All others were not.

Course Level Analysis
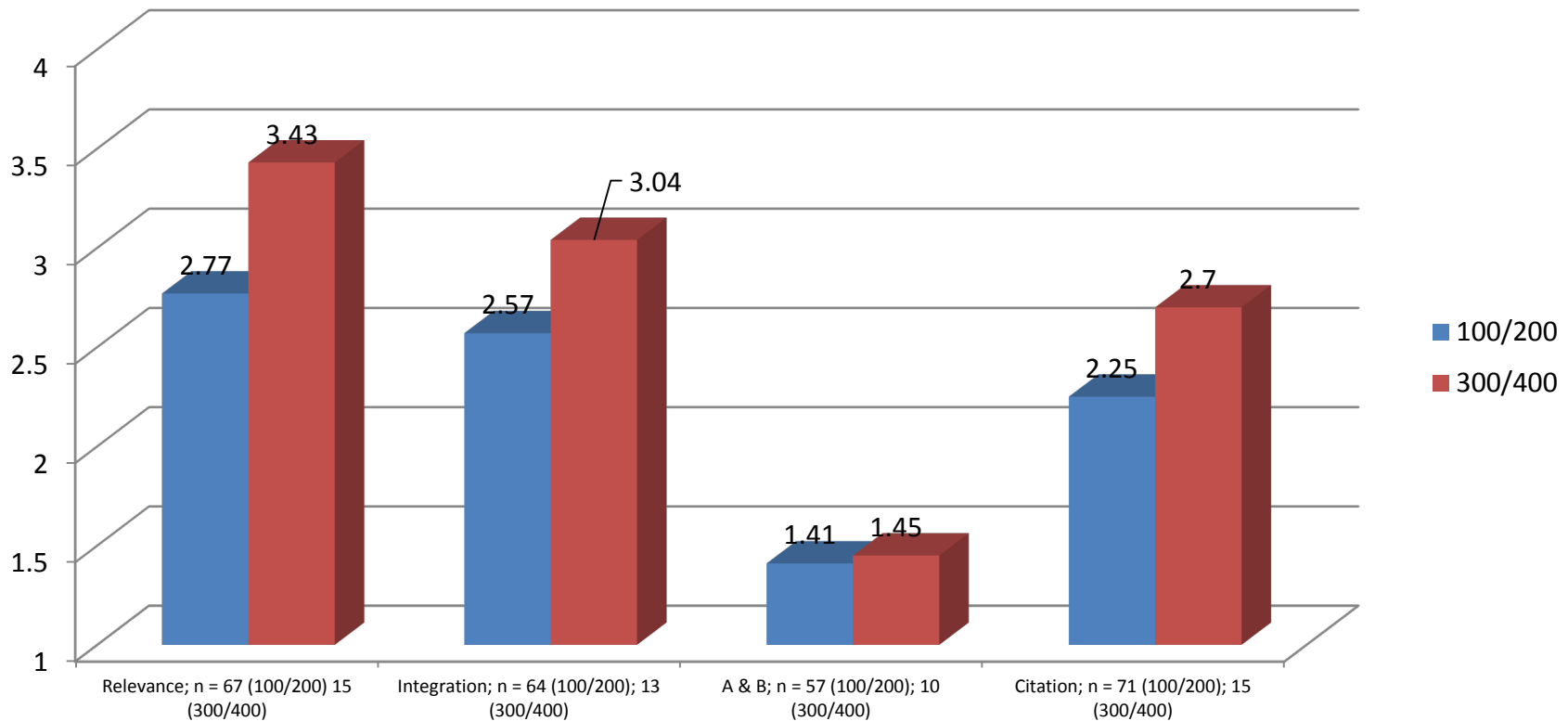
Legend:
- 100/200
- 300/400

Chart data:

Category	100/200	300/400
Relevance; n = 67 (100/200) 15 (300/400)	2.77	3.43
Integration; n = 64 (100/200); 13 (300/400)	2.57	3.04
A & B; n = 57 (100/200); 10 (300/400)	1.41	1.45
Citation; n = 71 (100/200); 15 (300/400)	2.25	2.7

Information Literacy

Frequency Analysis by Course Level

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

Course Level	Trait/ Performance Level	Relevance of Sources	Integration of Information	Assumptions and Biases	Citation	Total
100/200	1.0	5 (7%)	5 (8%)	35 (61%)	16 (23%)	61 (24%)
300/400		0	0	5 (50%)	0	5 (9%)
100/200	1.5 – 2.0	13 (19%)	15 (23%)	13 (23%)	19 (27%)	60 (23%)
300/400		0	2 (15%)	4 (40%)	5 (33%)	11 (21%)
100/200	2.5 – 3.0	25 (37%)	30 (47%)	8 (14%)	27 (38%)	90 (35%)
300/400		5 (33%)	5 (38%)	1 (10%)	6 (40%)	17 (32%)
100/200	3.5 – 4.0	24 (36%)	14 (22%)	1 (2%)	9 (13%)	48 (19%)
300/400		10 (67%)	6 (46%)	0	4 (27%)	20 (38%)
100/200	**Total Tags with Usable Scores**	**67 (100%)**	**64 (100%)**	**57 (100%)**	**71 (100%)**	**259 (100%)**
300/400		**15 (100%)**	**13 (100%)**	**10 (100%)**	**15 (100%)**	**53 (100%)**
All Course Levels	**Grand Totals**	**82**	**77**	**67**	**86**	**312**

Information Literacy
Frequency Analysis by Course Level

Relevance of Sources

100/200 Level	300/400/Level

100/200 Level: 5, 13, 25, 24
300/400/Level: 0, 5, 10

Integration of Information

100/200 Level	300/400 Level

100/200 Level: 5, 15, 30, 14
300/400 Level: 0, 2, 5, 6

Legend: ■ 1.0 ■ 1.5-2.0 ■ 2.5-3.0 ■ 3.5-4.0
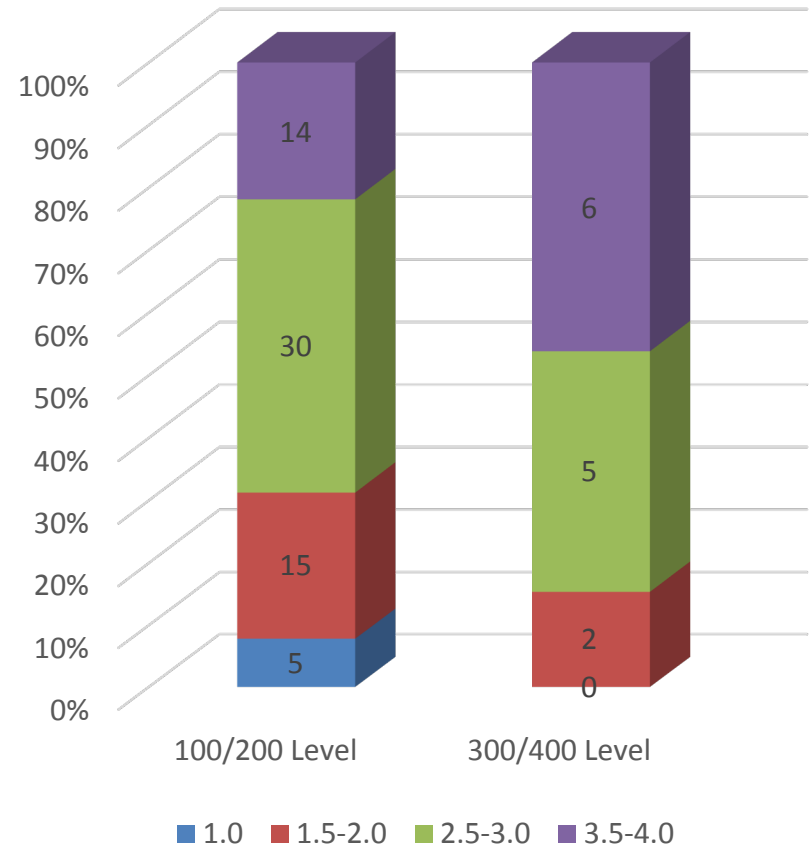
Information Literacy
Frequency Analysis by Course Level

Assumptions and Biases

	100/200 Level	300/400 Level
3.5-4.0	1	0
2.5-3.0	8	1
1.5-2.0	13	4
1.0	35	5

Legend: ■ 1.0 ■ 1.5-2.0 ■ 2.5-3.0 ■ 3.5-4.0

Citation

	100/200 Level	300/400 Level
3.5-4.0	9	4
2.5-3.0	27	6
1.5-2.0	19	5
1.0	16	0

Legend: ■ 1.0 ■ 1.5-2.0 ■ 2.5-3.0 ■ 3.5-4.0
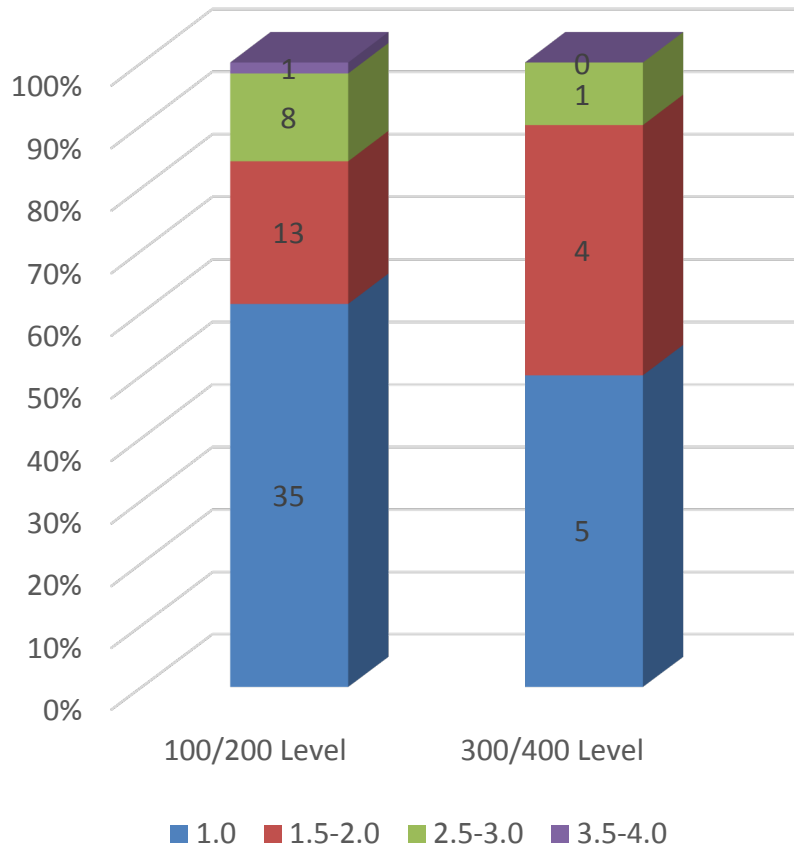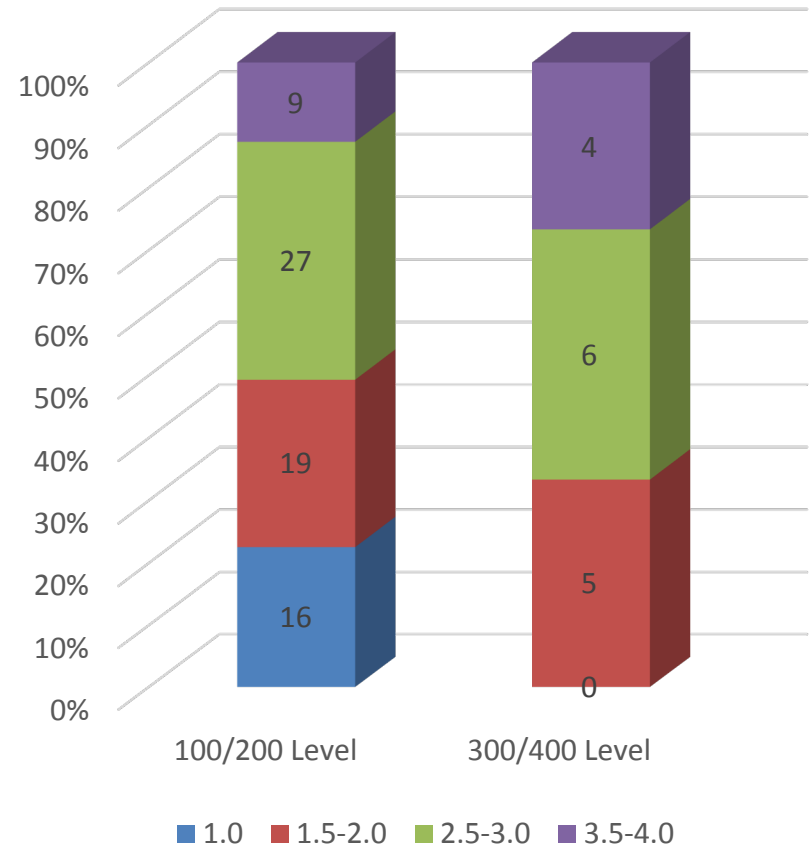
Information Literacy

Inter-Rater Agreement Results

Trait/ Performance Level	Relevance of Sources Kappa Liberal = .627	Integration of Information Kappa Liberal = .578	Assumptions and Biases Kappa Liberal = .574	Citation Kappa Liberal = .582
Agree on score	31	25	38	30
Difference = 1 point	30	32	11	26
Difference = 2 points	14	6	9	17
Difference = 3 points	1	3	2	2
Agree on Not Aligned	7	7	22	8
Score + Not Aligned	17	27	18	17
Total	100	100	100	100

Integrative Thinking

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

(Although there were 90 artifacts in the analysis, not all artifacts aligned to every trait)

Overall Analysis

Discipline; n = 78   Domain; n = 79   Transfer; n = 71   Experience; n = 86

1.86   1.85   2.09   2.01
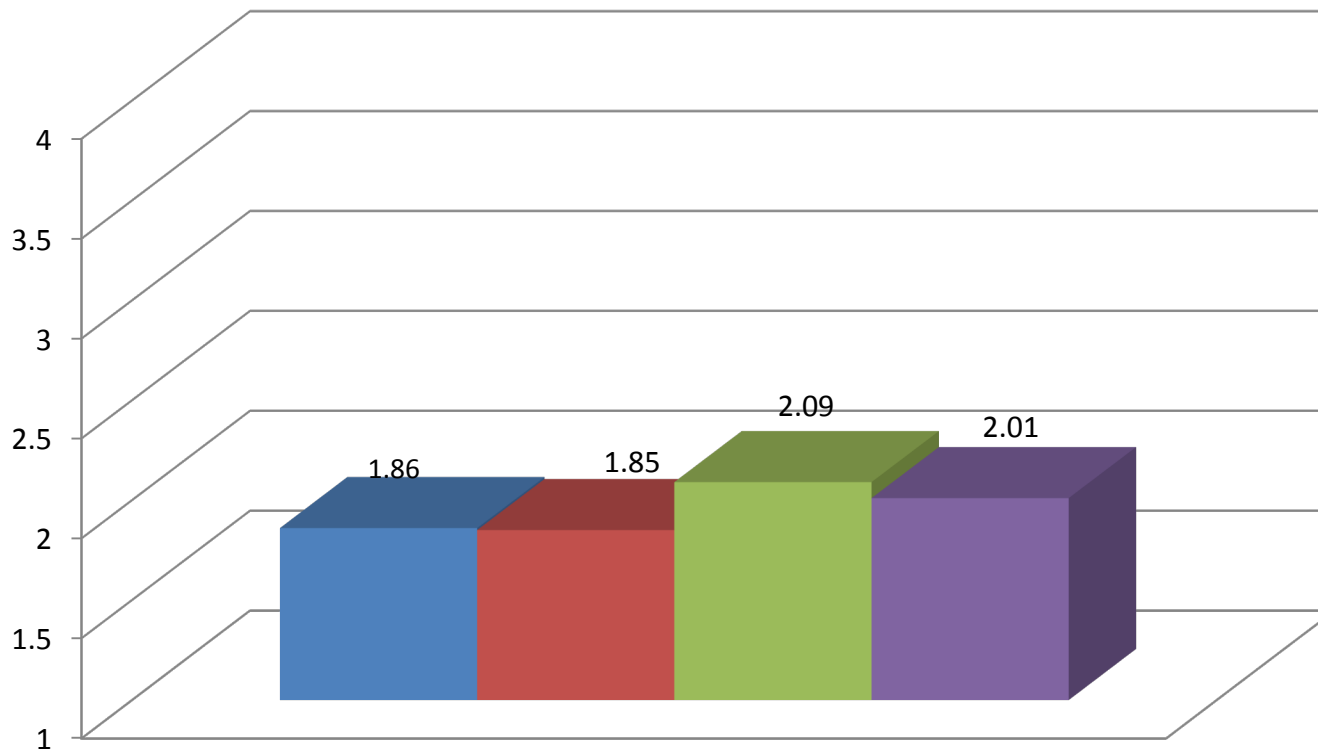
Integrative Thinking

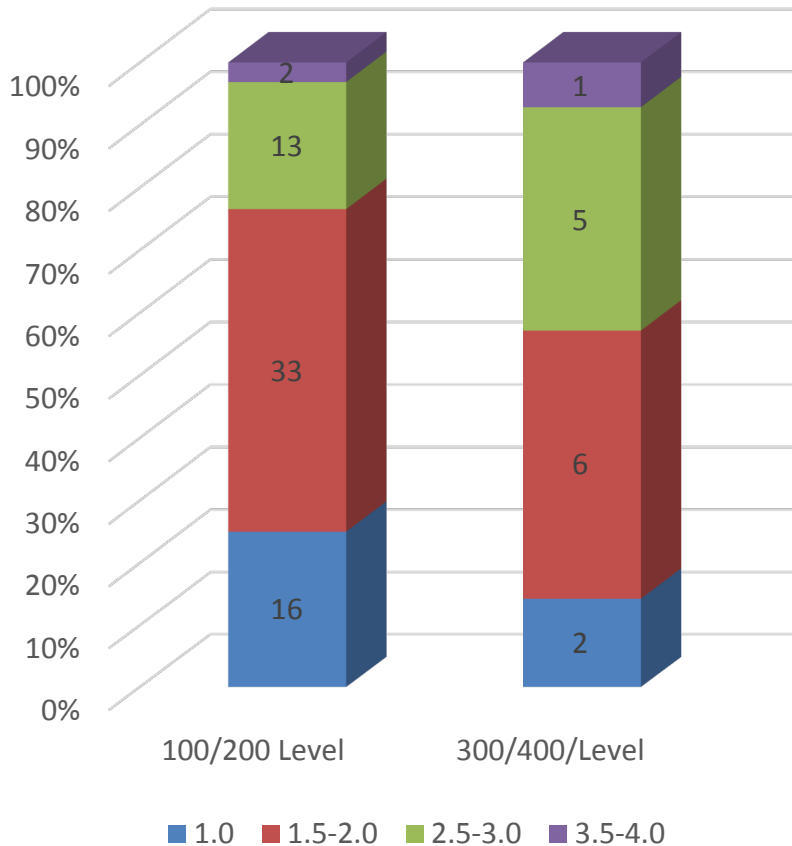Number of artifacts (with usable scores) scoring at each performance level

Trait/ Performance Level	Connections among Disciplines	Relation among Domains of Thinking	Transfer	Connections to Experience	Total
1.0	18 (23%)	13 (16%)	9 (13%)	12 (14%)	**52 (17%)**
1.5 – 2.0	39 (50%)	49 (62%)	38 (54%)	44 (51%)	**170 (54%)**
2.5 – 3.0	18 (23%)	17 (22%)	18 (25%)	25 (29%)	**78 (25%)**
3.5 – 4.0	3 (4%)	0	6 (8%)	5 (6%)	**14 (4%)**
Totals	**78 (100%)**	**79 (100%)**	**71 (100%)**	**86 (100%)**	**314 (100%)**

Integrative Thinking
Frequency Analysis

	Discipline	Domain	Transfer	Experience
3.5-4.0	3	0	6	5
2.5-3.0	18	17	18	25
1.5-2.0	39	49	38	44
1.0	18	13	9	12
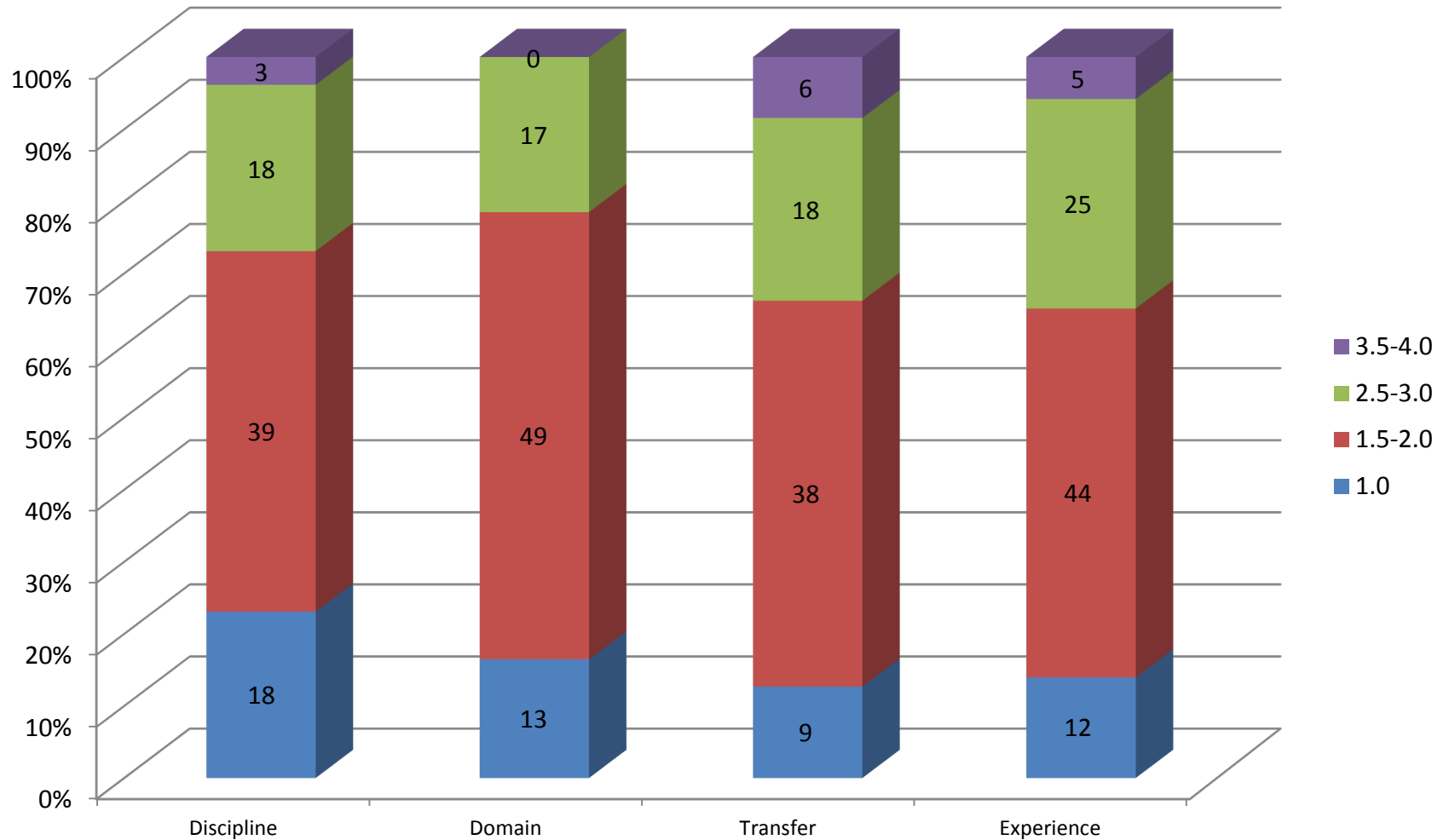
Integrative Thinking: Course Level Analysis

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

There were no significant mean differences based on course level for any trait.

Course Level Analysis
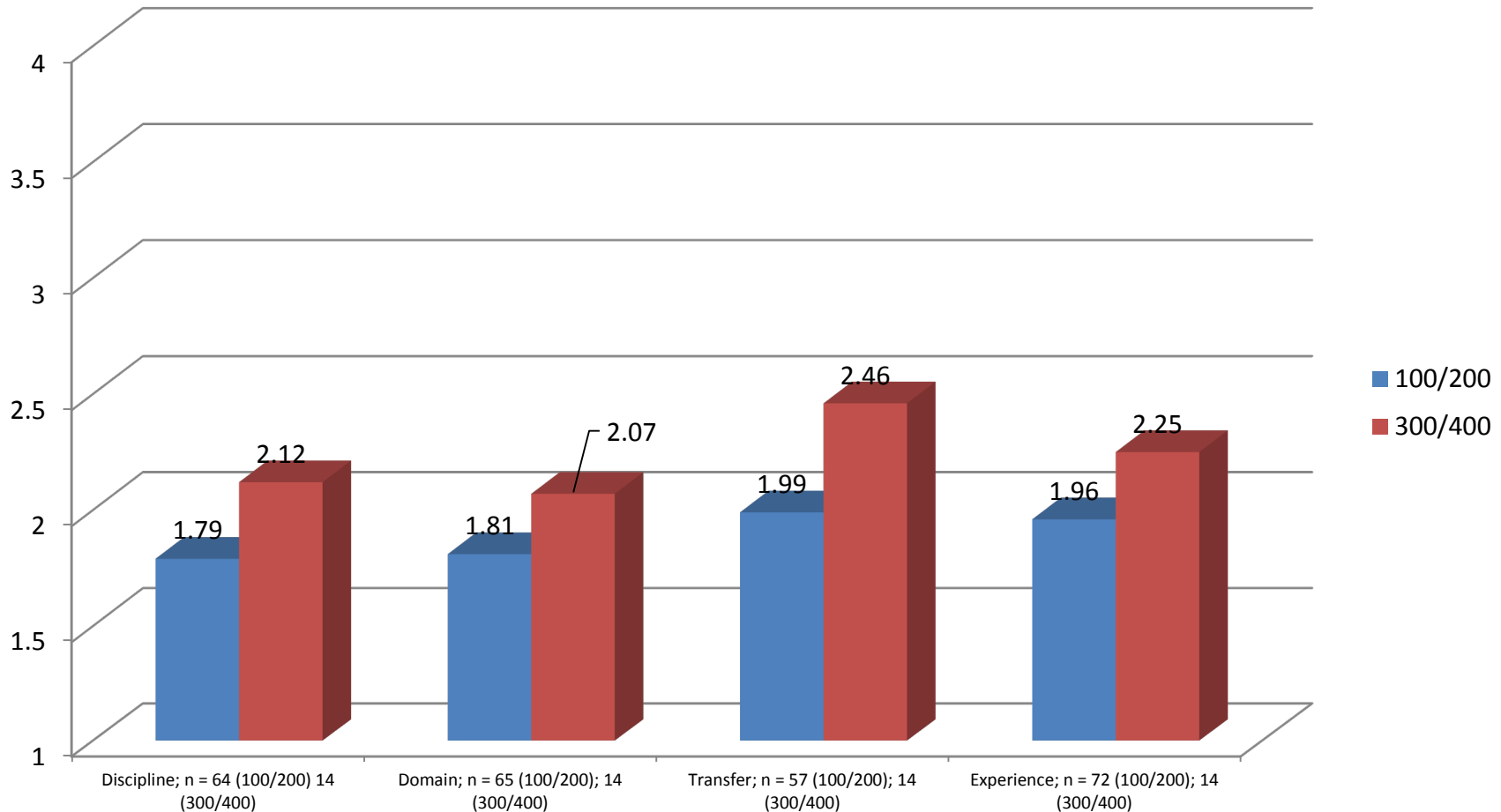
Legend:
- 100/200
- 300/400

Trait	100/200	300/400
Discipline; n = 64 (100/200) 14 (300/400)	1.79	2.12
Domain; n = 65 (100/200); 14 (300/400)	1.81	2.07
Transfer; n = 57 (100/200); 14 (300/400)	1.99	2.46
Experience; n = 72 (100/200); 14 (300/400)	1.96	2.25

Integrative Thinking

Frequency Analysis by Course Level

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

Course Level	Trait/ Performance Level	Connections among Disciplines	Relation among Domains of Thinking	Transfer	Connections to Experience	Total
100/200	1.0	16 (25%)	12 (18%)	7 (12%)	11 (15%)	46 (18%)
300/400		2 (14%)	1 (7%)	2 (14%)	1 (7%)	6 (11%)
100/200	1.5 – 2.0	33 (52%)	41 (63%)	34 (60%)	37 (51%)	145 (56%)
300/400		6 (43%)	8 (57%)	4 (29%)	7 (50%)	25 (45%)
100/200	2.5 – 3.0	13 (20%)	12 (18%)	14 (25%)	20 (28%)	59 (23%)
300/400		5 (36%)	5 (36%)	4 (29%)	5 (36%)	19 (34%)
100/200	3.5 – 4.0	2 (3%)	0	2 (4%)	4 (6%)	8 (3%)
300/400		1 (7%)	0	4 (29%)	1 (7%)	6 (11%)
100/200	**Total Traits with Usable Scores**	**64 (100%)**	**65 (100%)**	**57 (100%)**	**72 (100%)**	**258 (100%)**
300/400		**14 (100%)**	**14 (100%)**	**14 (100%)**	**14 (100%)**	**56 (100%)**
All Course Levels	**Grand Totals**	**78**	**79**	**71**	**86**	**314**

Integrative Thinking
Frequency Analysis by Course Level
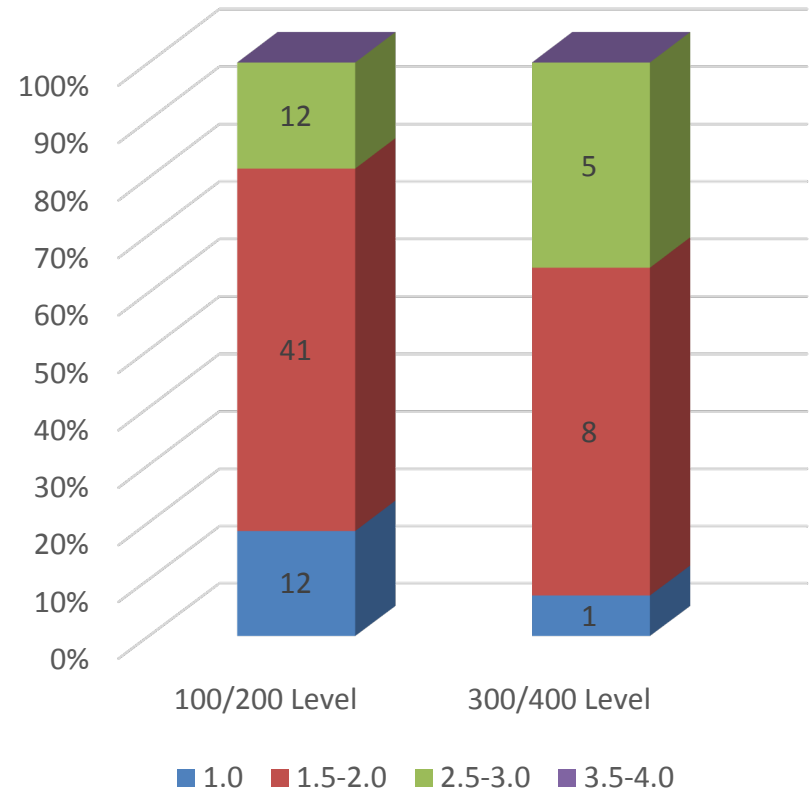
Integrative Thinking
Frequency Analysis by Course Level

Transfer

	100/200 Level	300/400/Level
3.5-4.0	2	4
2.5-3.0	14	4
1.5-2.0	34	4
1.0	7	2

Legend: ■ 1.0 ■ 1.5-2.0 ■ 2.5-3.0 ■ 3.5-4.0

Connections to Experience

	100/200 Level	300/400 Level
3.5-4.0	4	1
2.5-3.0	20	5
1.5-2.0	37	7
1.0	11	1

Legend: ■ 1.0 ■ 1.5-2.0 ■ 2.5-3.0 ■ 3.5-4.0
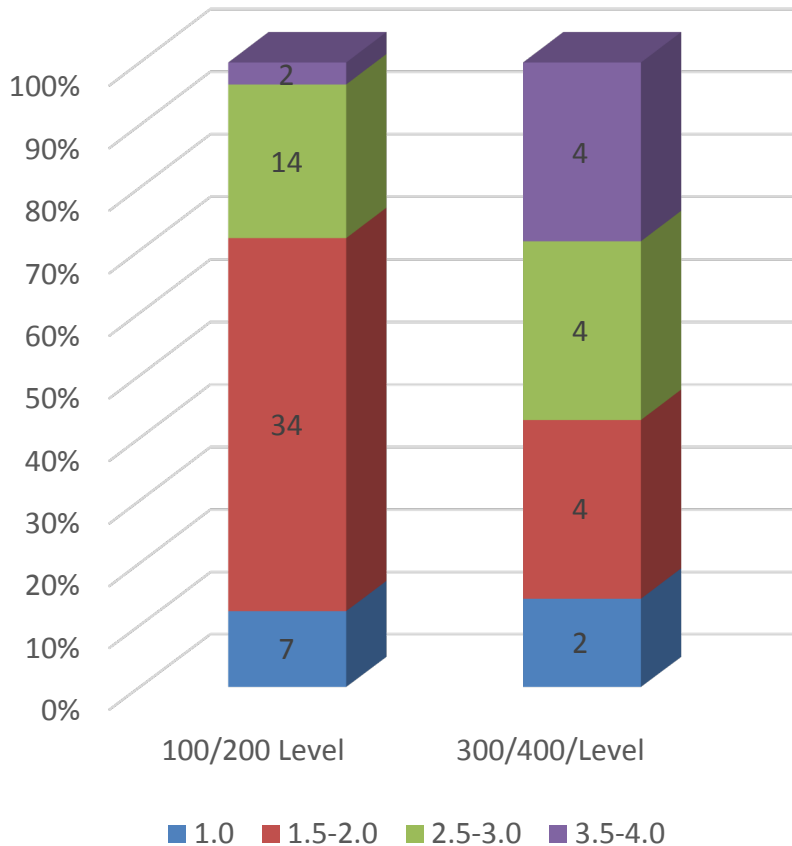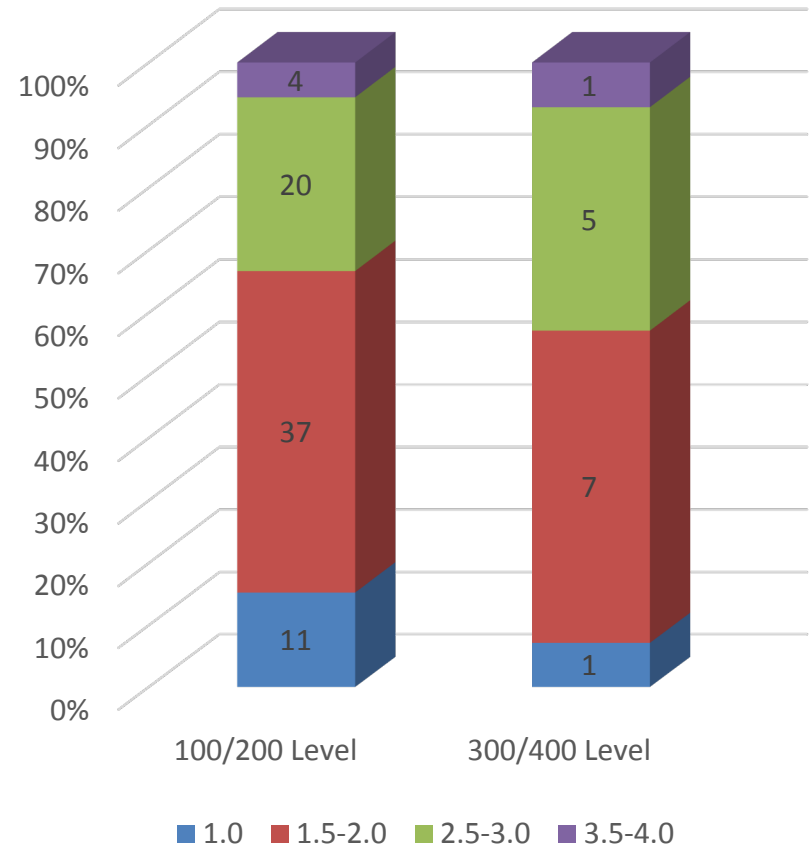
Integrative Thinking
Inter-Rater Agreement Results

Trait/ Performance Level	Connections among Disciplines Kappa Liberal = .535	Relation among Domains of Thinking Kappa Liberal = .425	Transfer Kappa Liberal = .430	Connections to Experience Kappa Liberal = . 544
Agree on score	20	16	13	20
Difference = 1 point	24	26	25	32
Difference = 2 points	7	13	14	13
Difference = 3 points	2	1	0	2
Agree on Not Aligned	14	10	14	9
Score + Not Aligned	29	33	33	23
Total	99	99	99	99

Metacognitive Thinking

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

(Although there were 80 artifacts in the analysis, not all artifacts aligned to every trait)

Overall Analysis

Legend: ■ Project; n = 63   ■ Self; n = 78
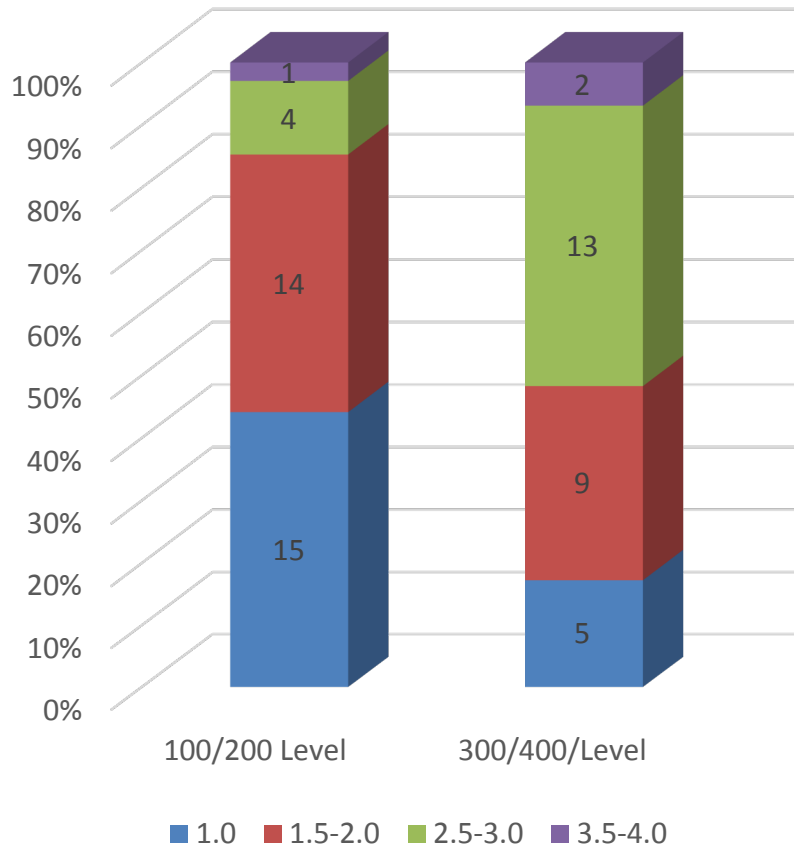
Project: 1.85
Self: 2.09
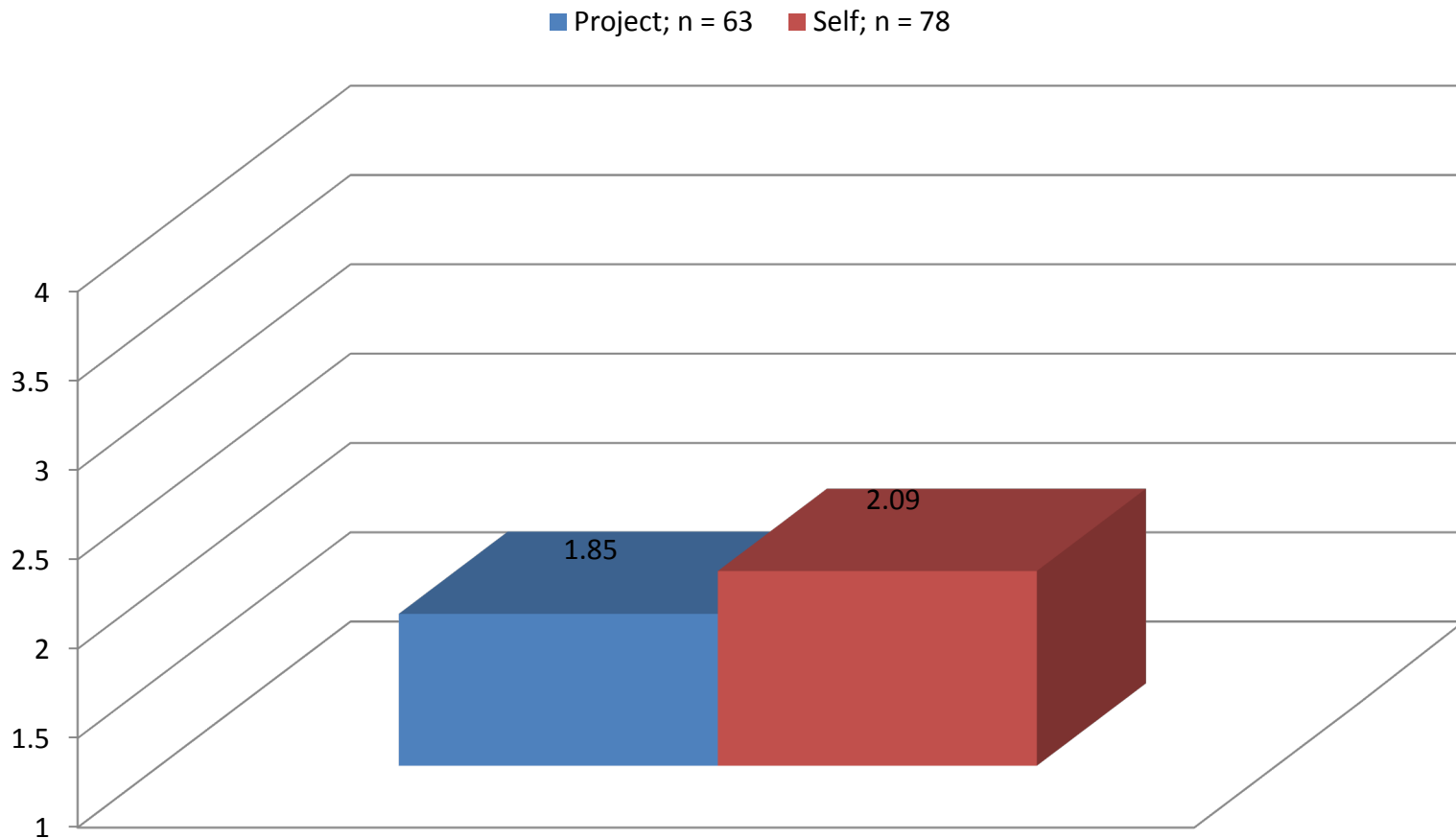
Metacognitive Thinking

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

Trait/ Performance Level	Project Management	Self-Evaluation	Total
1.0	20 (32%)	17 (22%)	**37 (26%)**
1.5 – 2.0	23 (37%)	27 (35%)	**50 (36%)**
2.5 – 3.0	17 (27%)	28 (36%)	**45 (32%)**
3.5 – 4.0	3 (5%)	6 (8%)	**9 (6%)**
Totals	**63 (100%)**	**78 (100%)**	**141 (100%)**

Metacognitive Thinking
Frequency Analysis
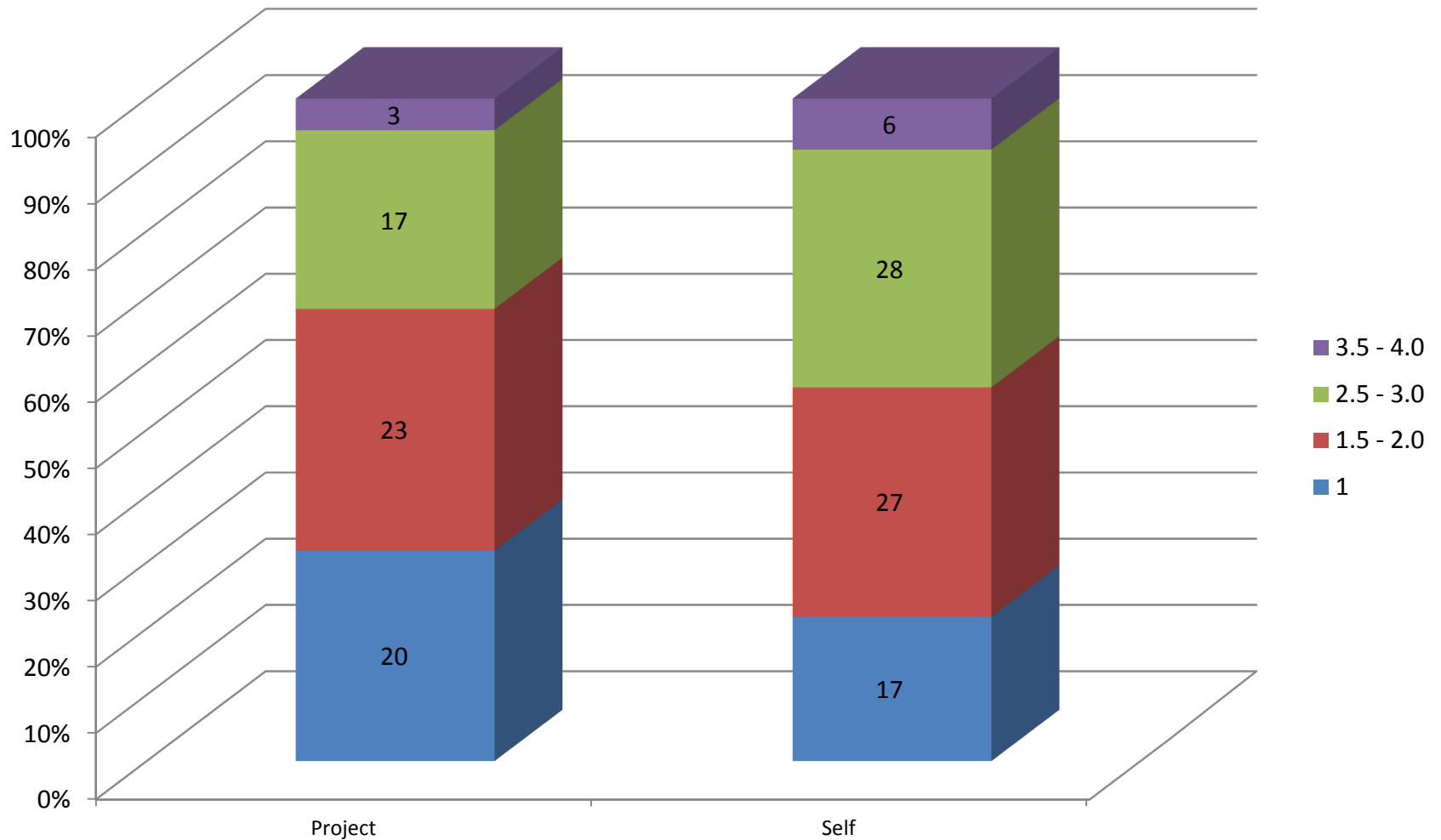
	Project	Self
3.5 - 4.0	3	6
2.5 - 3.0	17	28
1.5 - 2.0	23	27
1	20	17

Metacognitive Thinking

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score
Respondents from 300/400 level courses scored significantly higher than those from 100/200 level courses on *project management.*

Course Level Analysis
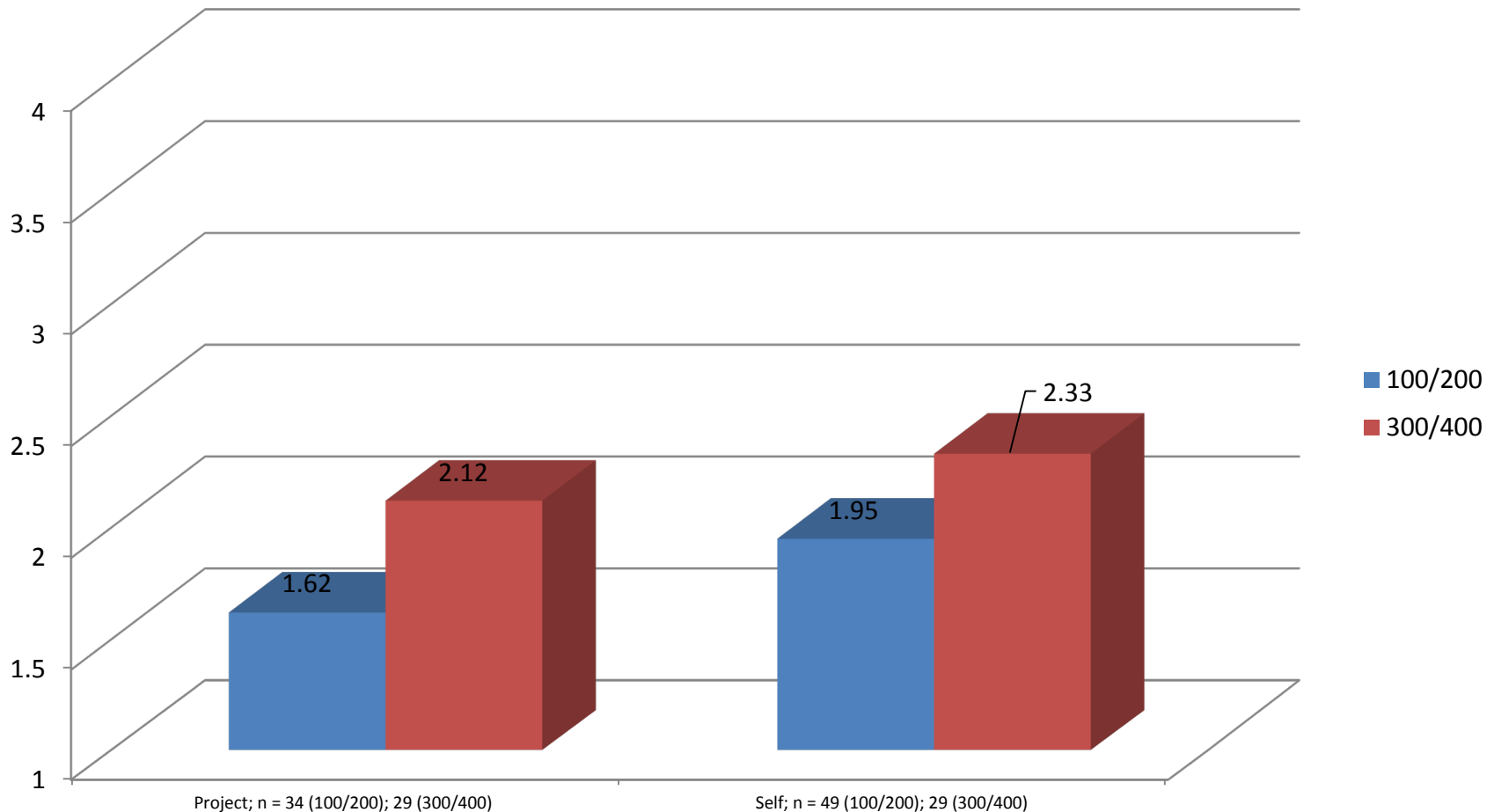
Legend:
- 100/200
- 300/400

Data labels:
- Project: 1.62 (100/200), 2.12 (300/400)
- Self: 1.95 (100/200), 2.33 (300/400)

X-axis: Project; n = 34 (100/200); 29 (300/400) | Self; n = 49 (100/200); 29 (300/400)

Metacognitive Thinking

Frequency Analysis by Course Level

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

Course Level	Trait/ Performance Level	Project Management	Self-Evaluation	Total
100/200	1.0	15 (44%)	14 (29%)	29 (35%)
300/400		5 (17%)	3 (10%)	8 (14%)
100/200	1.5 – 2.0	14 (41%)	17 (35%)	31 (37%)
300/400		9 (31%)	10 (35%)	19 (33%)
100/200	2.5 – 3.0	4 (12%)	16 (33%)	20 (24%)
300/400		13 (45%)	12 (41%)	25 (43%)
100/200	3.5 – 4.0	1 (3%)	2 (4%)	3 (4%)
300/400		2 (7%)	4 (14%)	6 (10%)
100/200	Total Traits with Usable Scores	34 (100%)	49 (100%)	83 (100%)
300/400		29 (100%)	29 (100%)	58 (100%)
All Course Levels	Grand Totals	63	78	141

Metacognitive Thinking
Frequency Analysis by Course Level

Project Management

1	2
4	13
14	9
15	5

100/200 Level — 300/400/Level

■ 1.0 ■ 1.5-2.0 ■ 2.5-3.0 ■ 3.5-4.0

Self-Evaluation

2	4
16	12
17	10
14	3

100/200 Level — 300/400 Level

■ 1.0 ■ 1.5-2.0 ■ 2.5-3.0 ■ 3.5-4.0
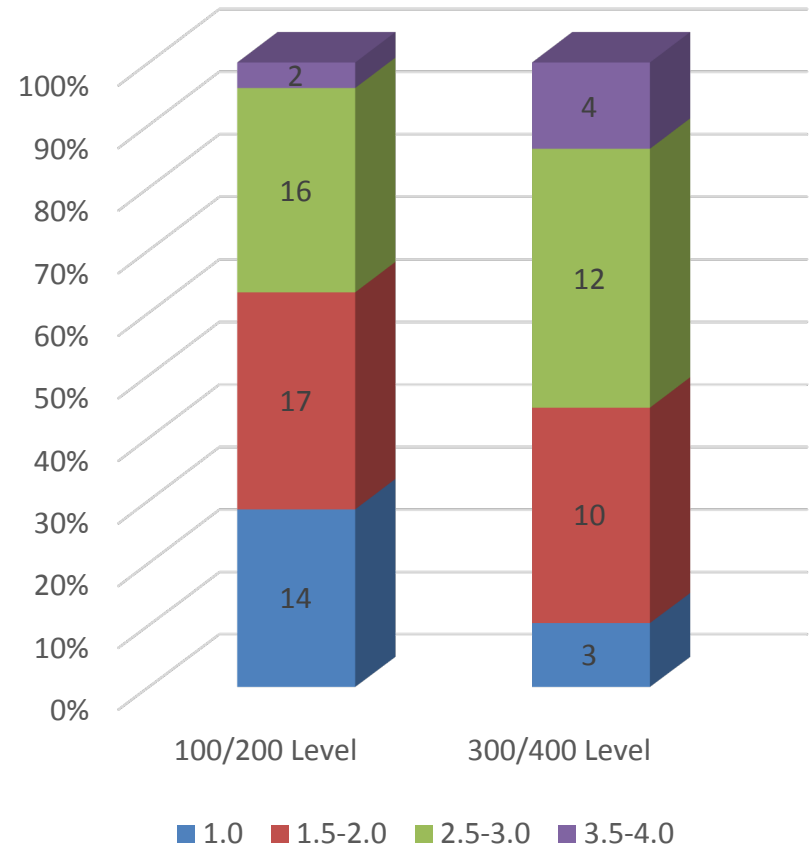
Metacognitive Thinking
Inter-Rater Agreement Results

Trait/ Performance Level	Project Management Kappa Liberal = .575	Self-Evaluation Kappa Liberal = .734
Agree on score	23	34
Difference = 1 point	17	26
Difference = 2 points	7	10
Difference = 3 points	1	0
Agree on Not Aligned	27	17
Score + Not Aligned	24	12
Total	99	99

Course Type Analysis

CT Courses

Mean Scores on a scale of 1 – 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 and Writing Intensive.

Information Literacy

- Relevance; n = 40
- Integration; n = 41
- A & B; n = 36
- Citation; n = 41

Relevance	2.79
Integration	2.59
A & B	1.46
Citation	2.24
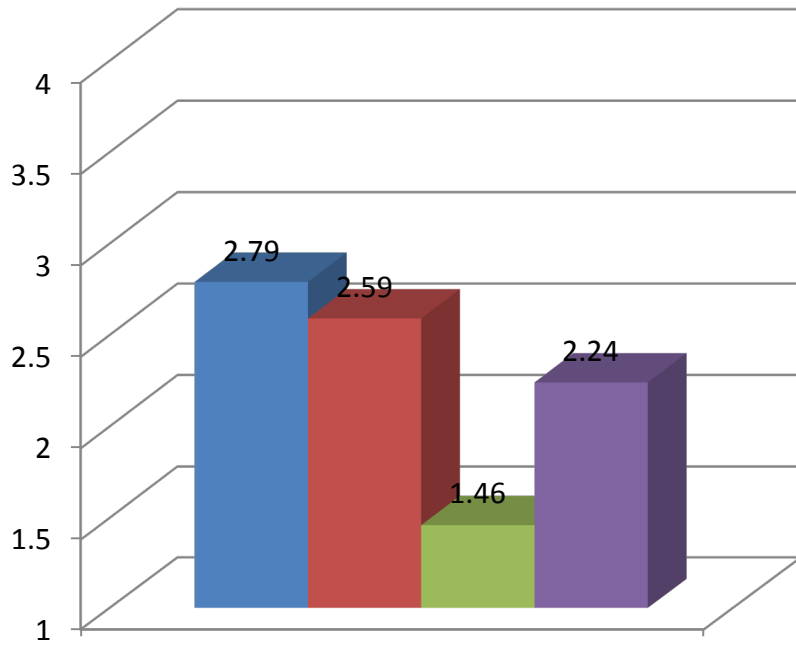
Integrative Thinking

- Discipline; n = 54
- Domain; n = 54
- Transfer; n = 51
- Experience; n = 62
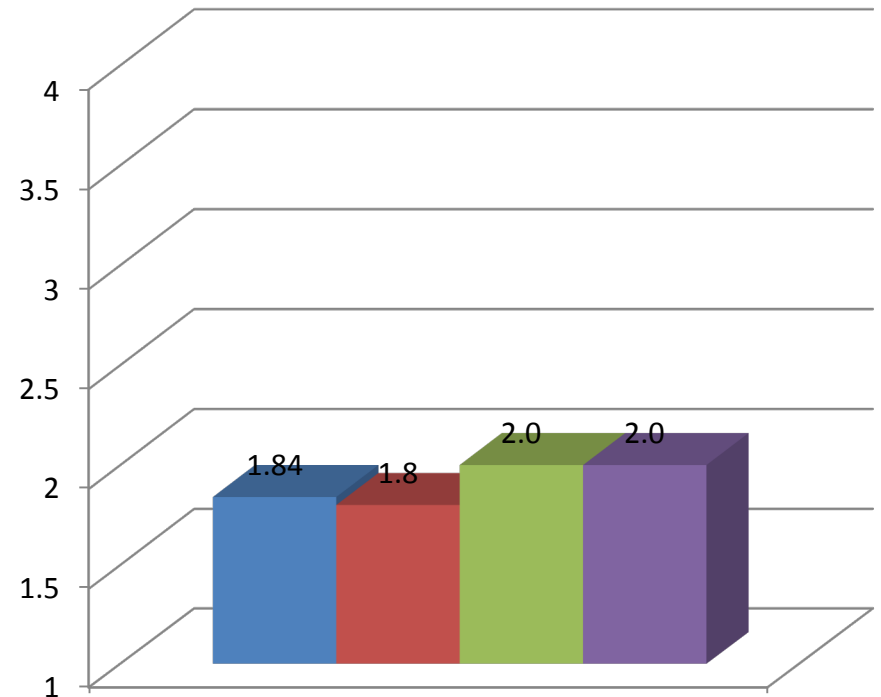
Discipline	1.84
Domain	1.8
Transfer	2.0
Experience	2.0

CT Courses

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score. Mean Scores on a scale of 1 – 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, and/or writing intensive.
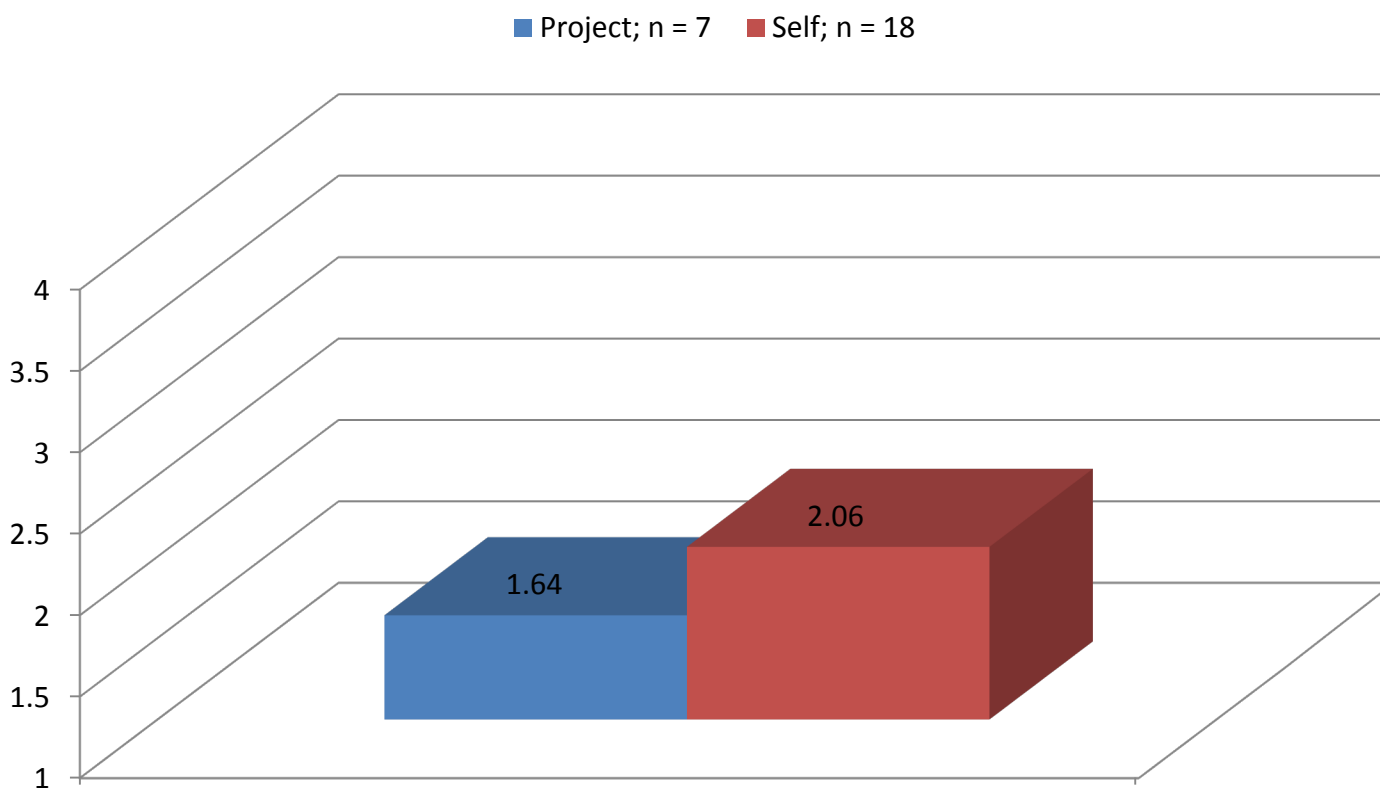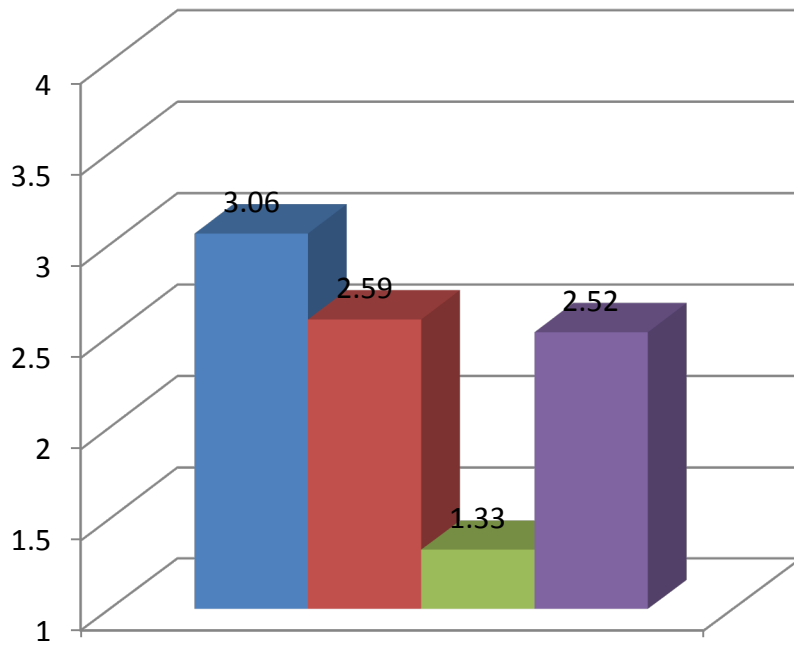
Metacognitive Thinking

■ Project; n = 7   ■ Self; n = 18

1.64

2.06

Core II Courses

Mean Scores on a scale of 1 – 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 and/or Writing Intensive.

Information Literacy

- ■ Relevance; n = 32
- ■ Integration; n = 31
- ■ A & B; n = 29
- ■ Citation; n = 31
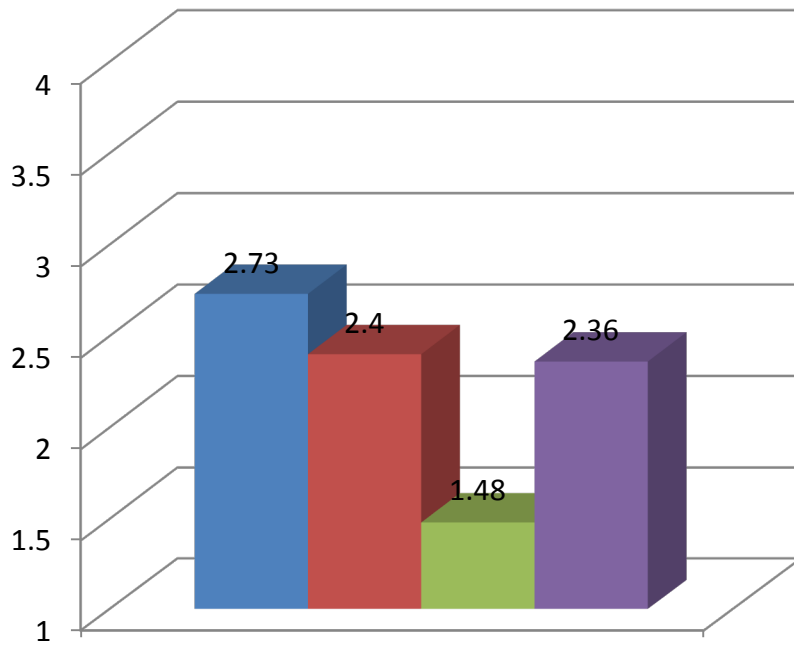
Integrative Thinking

- ■ Discipline; n = 48
- ■ Domain; n = 48
- ■ Transfer; n = 45
- ■ Experience; n = 55
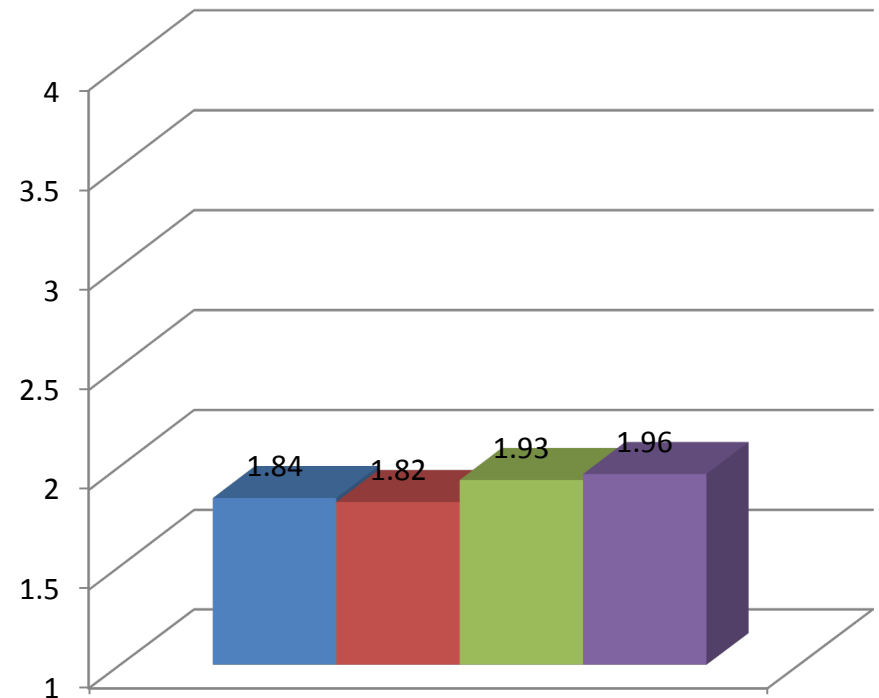
Core II Courses

Mean Scores on a scale of 1 – 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, and/or writing intensive.

Metacognitive Thinking

■ Project; n = 7   ■ Self; n = 18

Project: 1.64
Self: 2.06
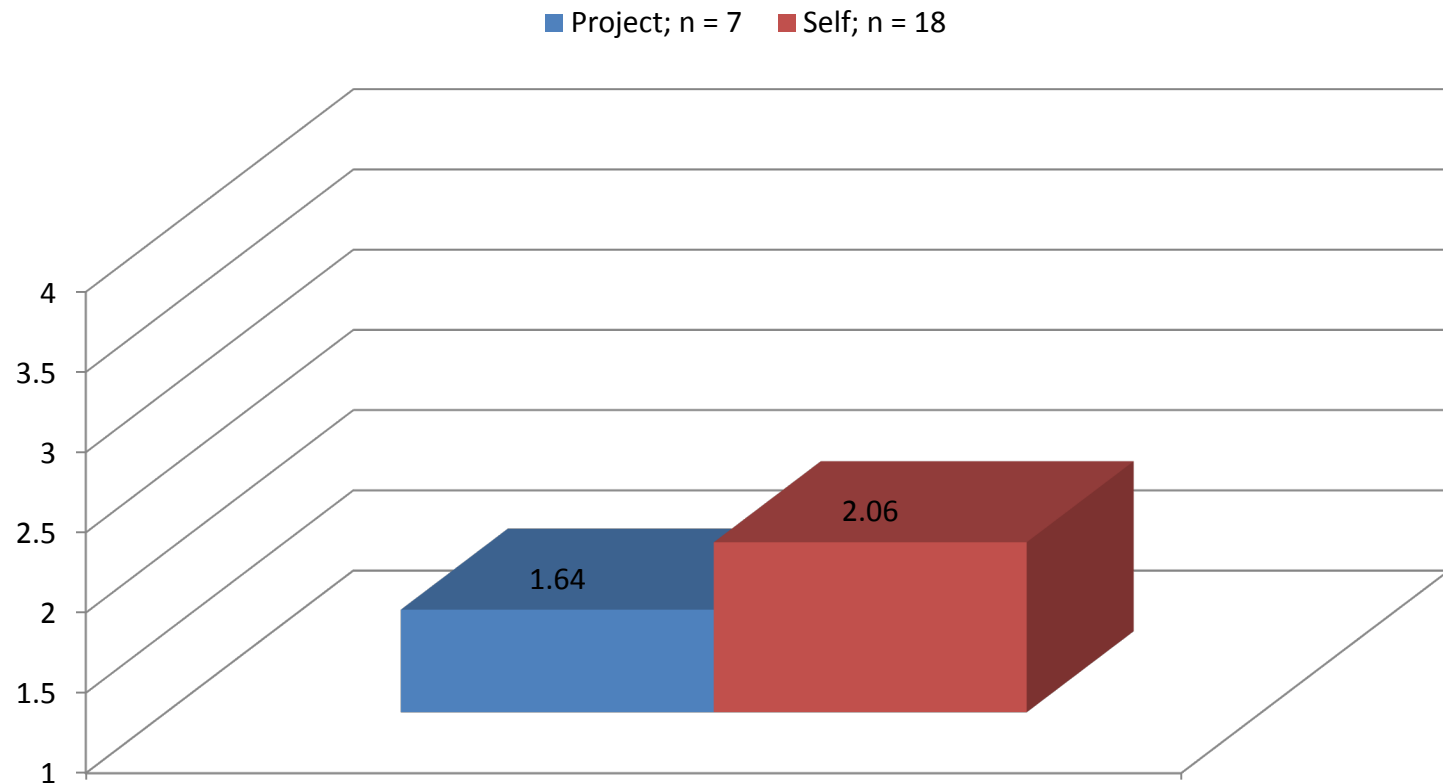
Writing Intensive Courses

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score. Some artifacts were from courses that, in addition to being WI, also were CT, Core II, and/or capstone.
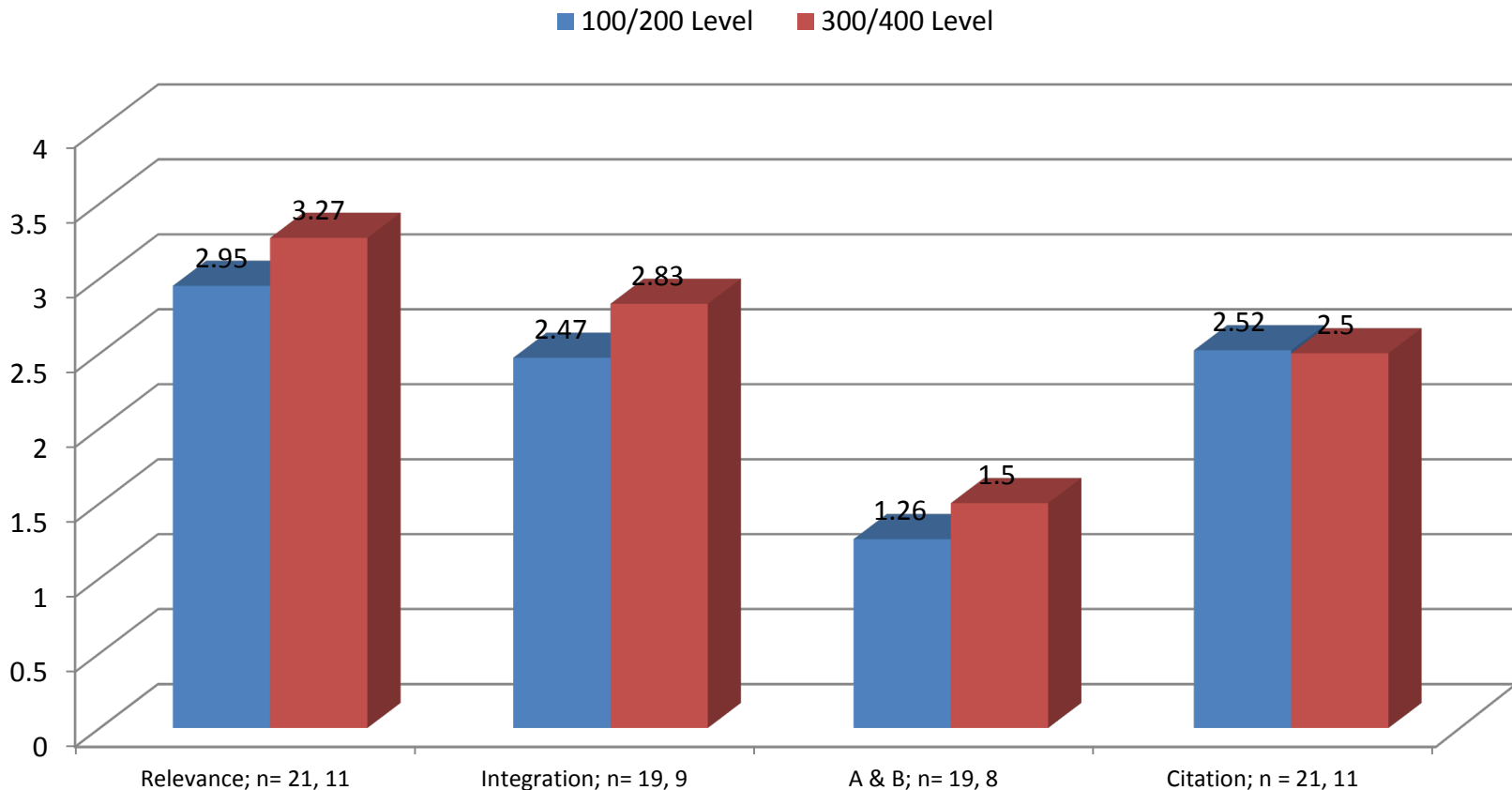
Information Literacy

- Relevance; n = 32
- Integration; n = 28
- A & B; n = 27
- Citation; n = 32

Integrative Thinking

- Discipline; n = 42
- Domain; n = 42
- Transfer; n = 40
- Experience; n = 45
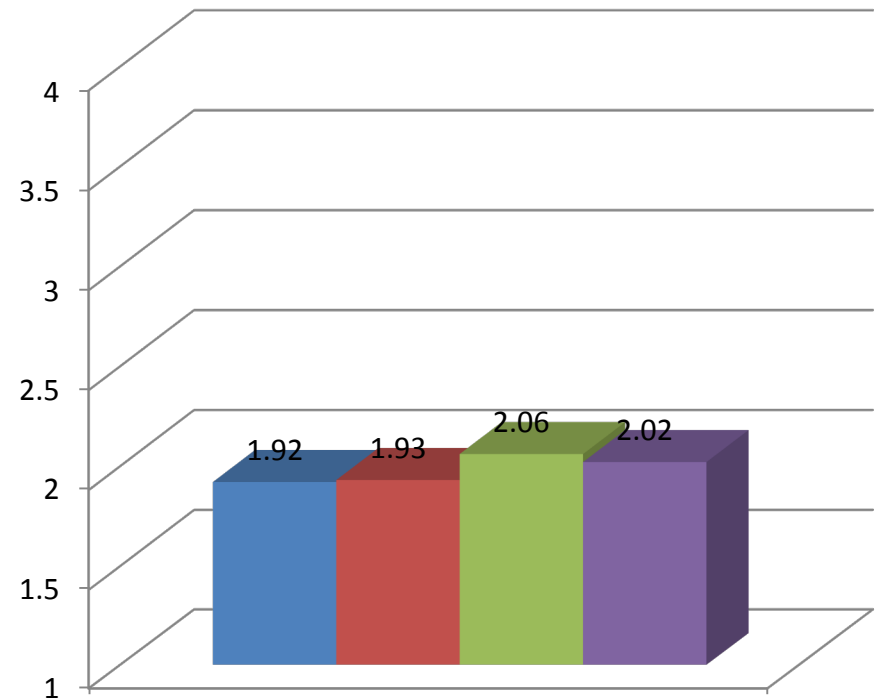
Writing Intensive Courses

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score. Mean Scores on a scale of 1 – 4, with 4 being the highest possible score. Some artifacts were from courses that, in addition to being WI, also were, CT, Core II, and/or capstone.

Metacognitive Thinking
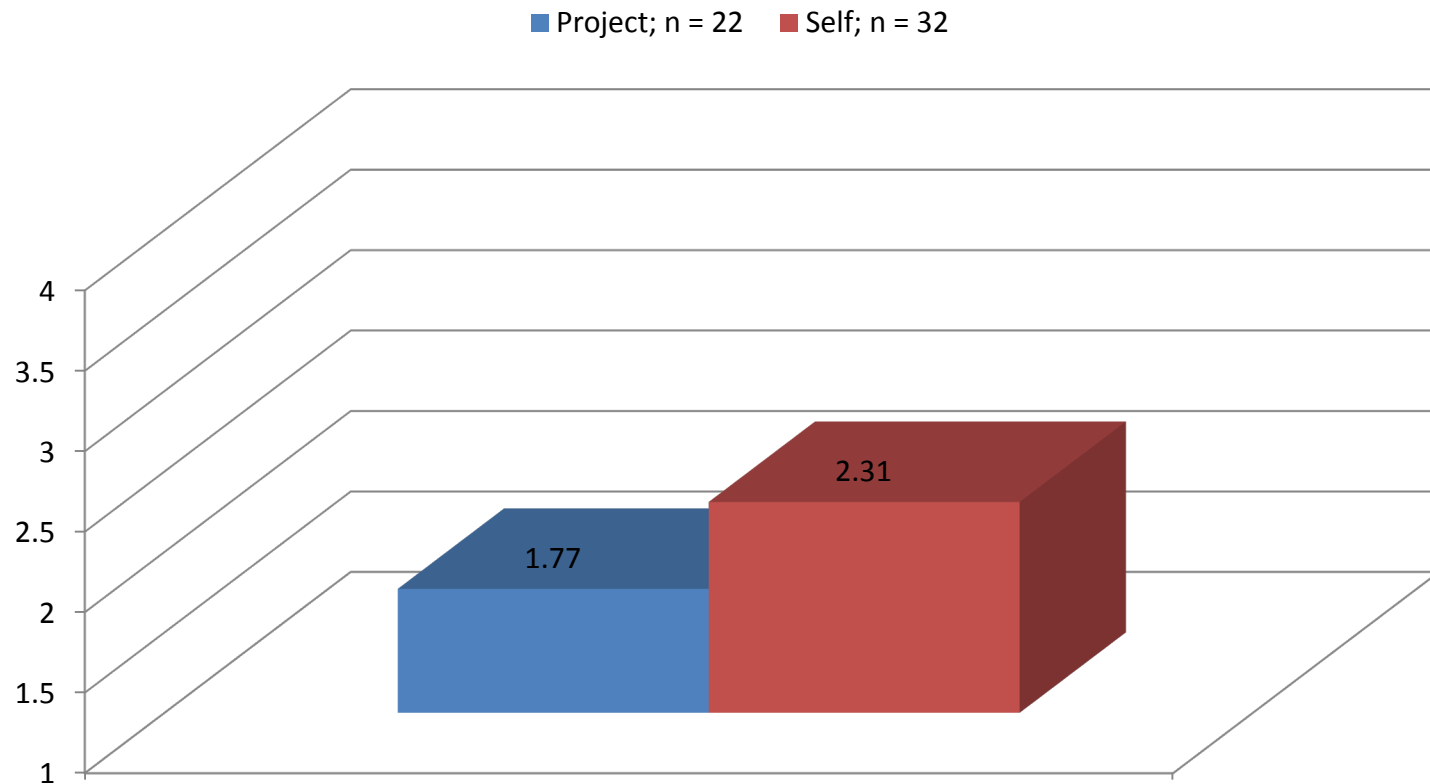
Legend: ■ Project; n = 22   ■ Self; n = 32

Project: 1.77
Self: 2.31

Writing Intensive Courses: Course Level Comparisons

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score. Mean Scores on a scale of 1 – 4, with 4 being the highest possible score. Some artifacts were from courses that, in addition to being WI, also were CT, Core II, and/or capstone. Note: There were only two artifacts aligned to any of the traits of *Integrative Thinking*. Therefore, we did not conduct a course-level analysis for WI courses aligned to this outcome.

Information Literacy

■ 100/200 Level   ■ 300/400 Level

	100/200 Level	300/400 Level
Relevance; n= 21, 11	2.95	3.27
Integration; n= 19, 9	2.47	2.83
A & B; n= 19, 8	1.26	1.5
Citation; n = 21, 11	2.52	2.5

Writing Intensive Courses: Course Level Comparisons
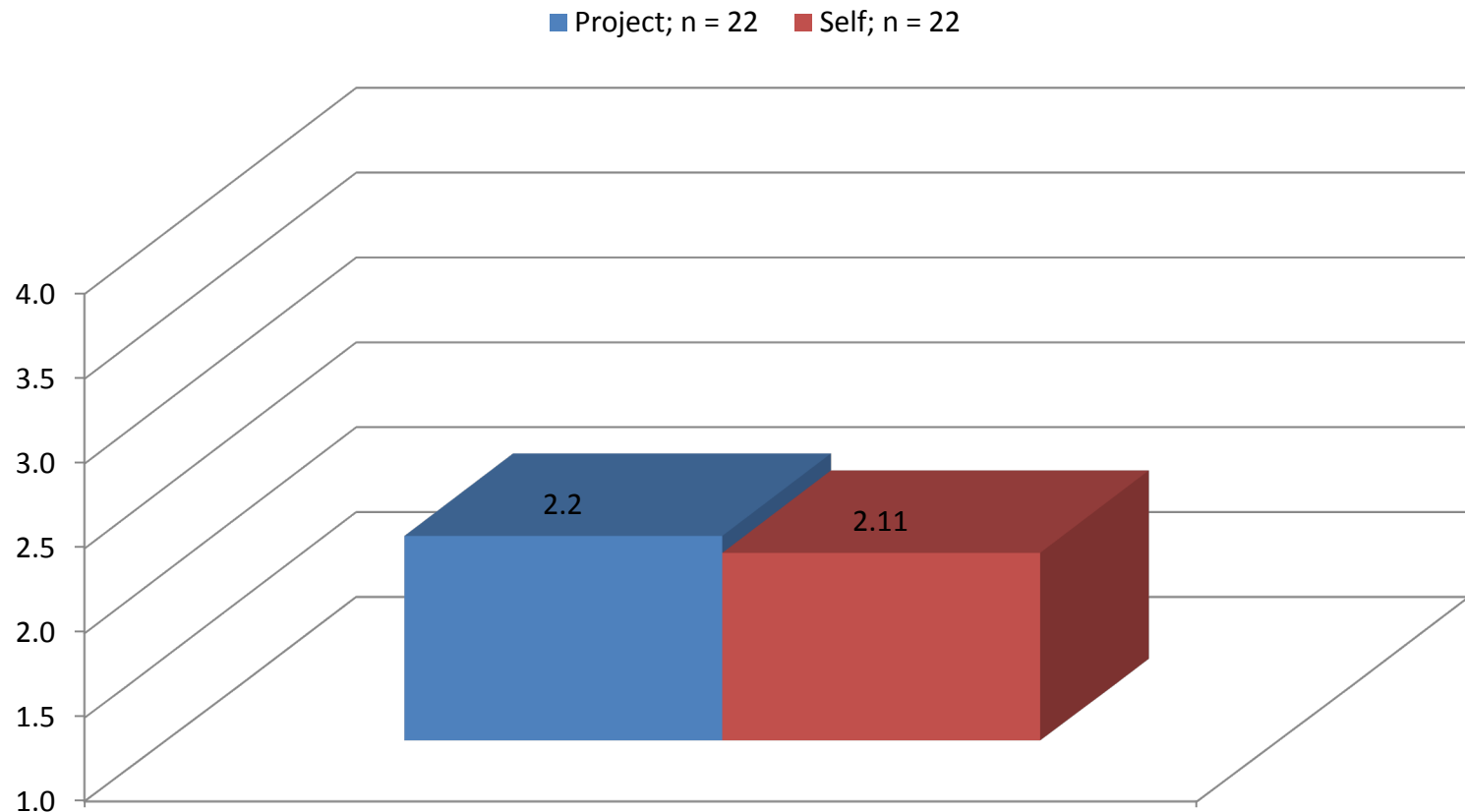
Mean Scores on a scale of 1 – 4, with 4 being the highest possible score Mean Scores on a scale of 1 – 4, with 4 being the highest possible score. Some artifacts were from courses that, in addition to being WI, CT, Core II, and/or capstone. Note: There were only four artifacts aligned to Project Management at the 100/200 level; therefore we did not include this trait in the course level comparisons.

Metacognitive Thinking

■ 100/200 Level   ■ 300/400 Level

2.25   2.36

Self; n = 14, 18
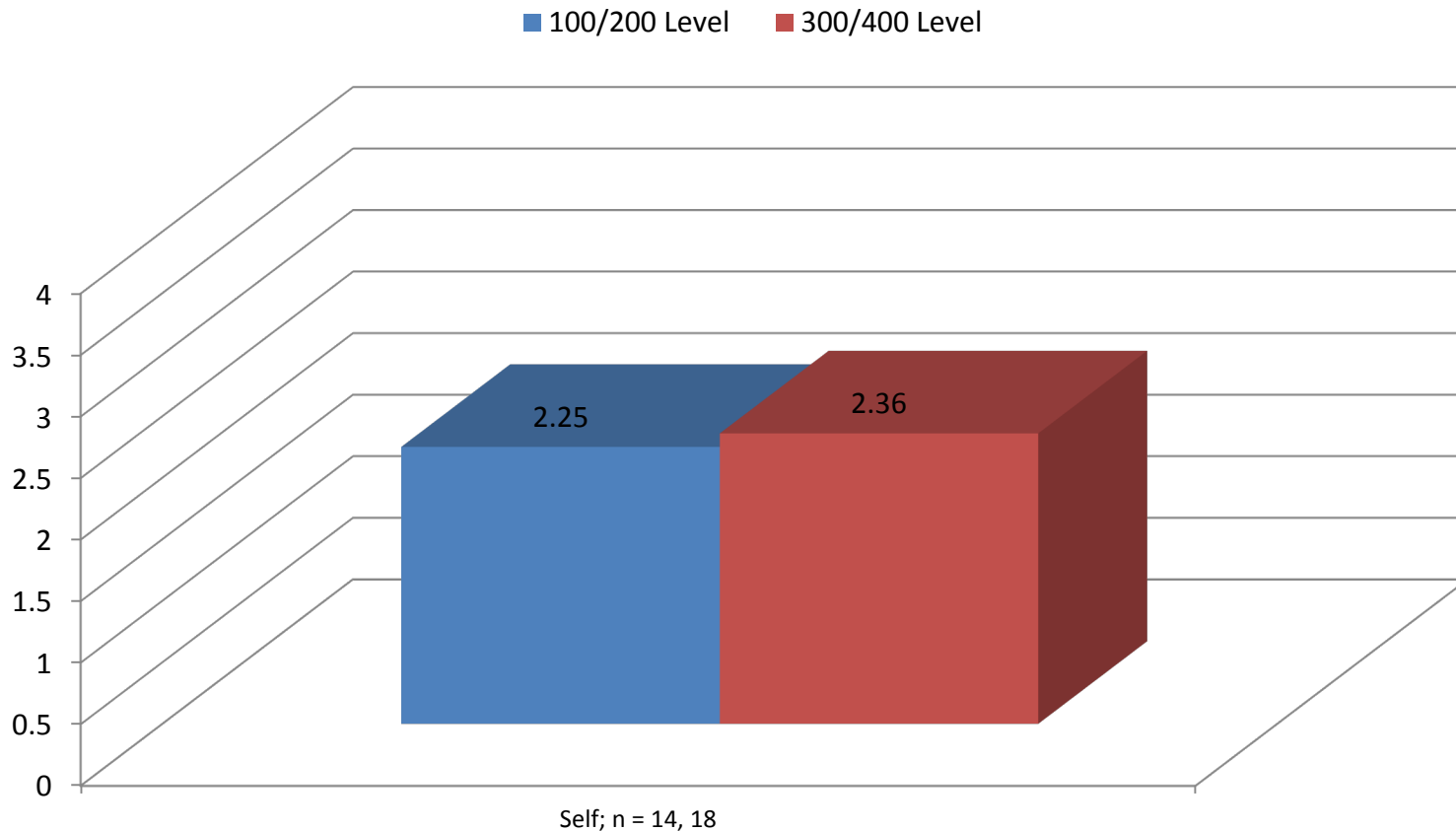
Capstone Courses

Mean Scores on a scale of 1 – 4, with 4 being the highest possible score. Some artifacts were from courses that, in addition to being Capstone, also were Writing Intensive
Note: There were no capstone courses in our sample that aligned to Integrative Thinking and only three that aligned to Information Literacy.
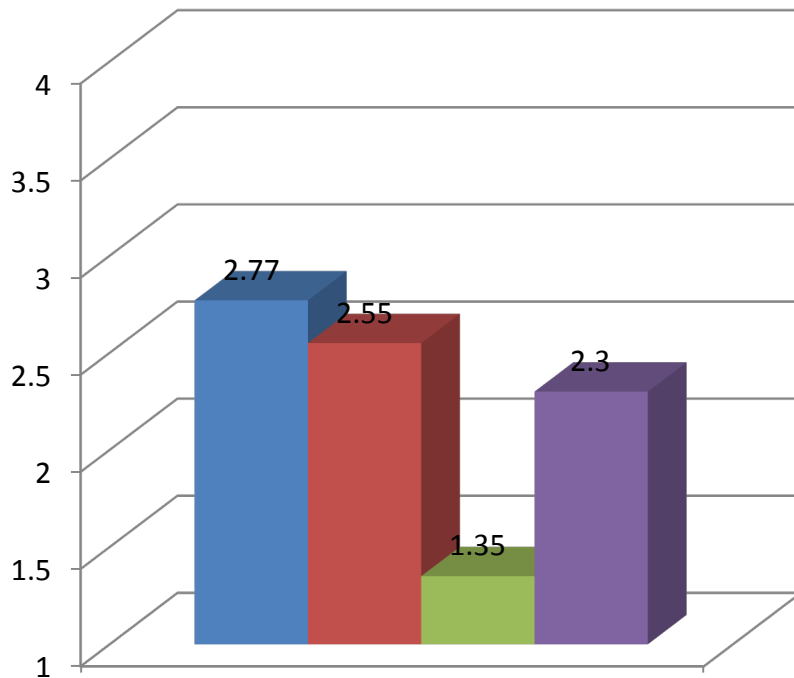
Metacognitive Thinking

■ Project; n = 22   ■ Self; n = 22

First Year Seminar (FYS) in Critical Thinking

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

Information Literacy
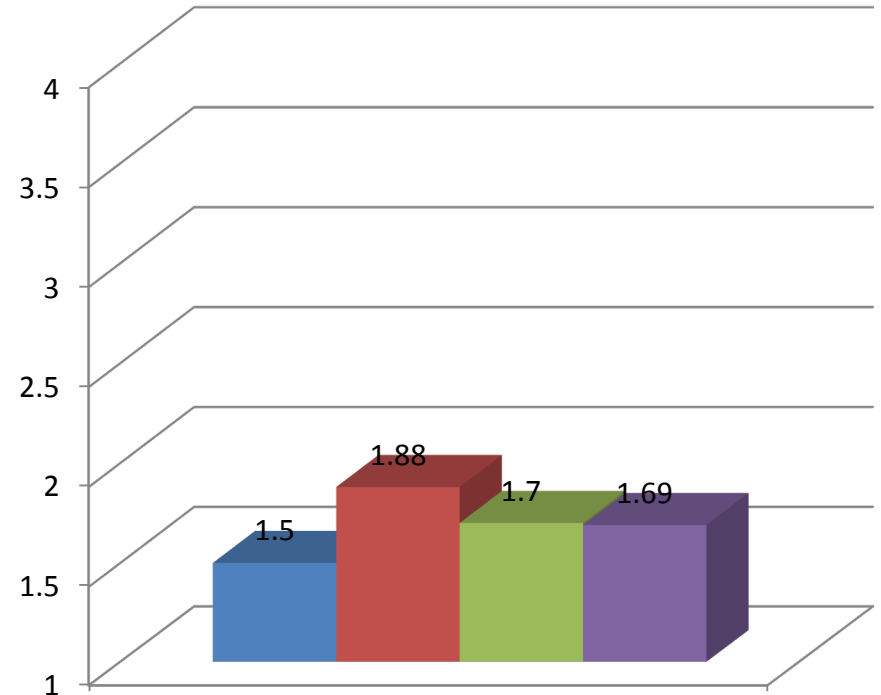
- Relevance; n = 24
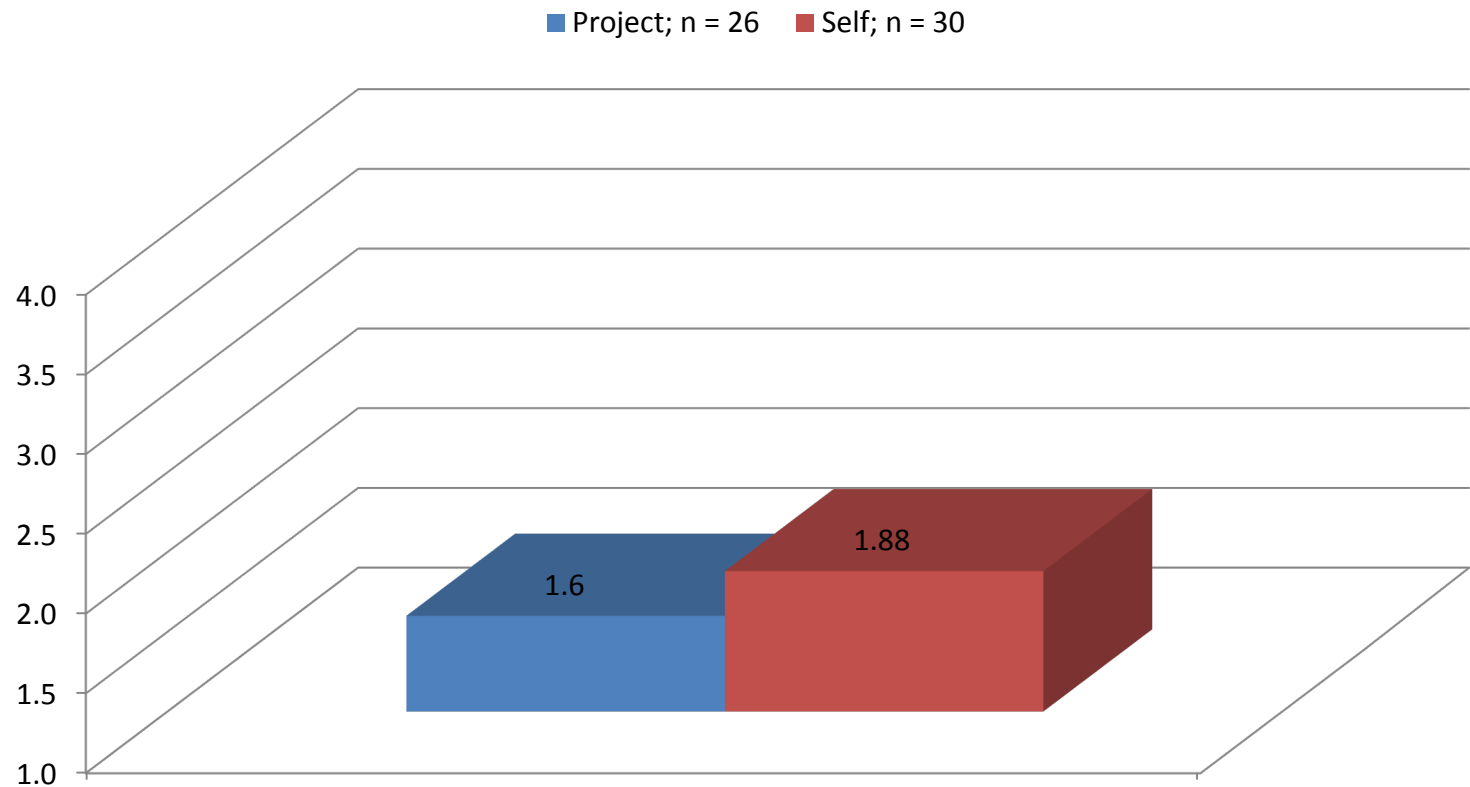- Integration; n = 21
- A & B; n = 20
- Citation; n = 27

2.77
2.55
1.35
2.3

Integrative Thinking

- Discipline; n = 7
- Domain; n = 8
- Transfer; n = 5
- Experience; n = 8

1.5
1.88
1.7
1.69

Reference

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