

Quantitative Reasoning

Description and Learning Outcomes

1. Students are able to interpret quantitative information (i.e., formulas, graphs, tables, models, and schematics) and draw inferences from them.
2. Given a quantitative problem, students are able to formulate the problem quantitatively and use appropriate arithmetical, algebraic, and/or statistical methods to solve the problem.
3. Students are able to evaluate logical arguments using quantitative reasoning.
4. Students are able to communicate and present quantitative results effectively.

Approved Courses and Enrollment

Students are required to pass one course approved for Quantitative Reasoning or transfer in an appropriate course. During the assessment period, eleven courses were approved to meet the Quantitative Reasoning requirement:

HNRT 125 A Liberal Arts Approach to Calculus
MATH 106 Quantitative Reasoning
MATH 108 Introductory Calculus with Business Applications
MATH 110 Introductory Probability
MATH 111 Linear Mathematical Modeling
MATH 113 Analytic Geometry and Calculus I
MATH 115 Analytic Geometry and Calculus I(Honors)
MATH 124 Calculus with Algebra/Trigonometry, Part B
MATH 125 Discrete Mathematics I
SOCL 313 Statistics for the Behavioral Sciences
STAT 250 Introductory Statistics I

Quantitative Reasoning courses enroll over 6,500 students each year with an average lecture class size of between 18 and 80 students; recitations maintain smaller class sizes for focused, practical instruction. Table 23 and figure 52 show enrollment trends over the past five years. STAT 250 is the highest enrolled course (28.4% of AY19 enrollment), followed by MATH 113 (19.3%), and MATH 106 (16.2%).

Courses Included in Assessment

The assessment period included 42 sections of HNRT 125; MATH 106, 108, 110, 111; SOCL 313; and STAT 250 courses taught on all of Mason's campuses and via distance learning in spring 2019. Of the 42 course sections included in the assessment period, 88% submitted materials.

Calculus courses (MATH 113, 108, 124) were part of a long-term pedagogical change project for which learning outcomes and an assessment strategy are in the process of being defined.

Enrollment and Grades Distribution

A total of 3,121 students enrolled in Quantitative Reasoning courses in the assessment period. Of these students, 72.3% passed their courses with a C or above (see Figure 41). Figure 42 displays average final grades by course. Note that the DFW rate approaches 25%.

Figure 41. Grades Distribution for Mason Core Quantitative Reasoning Courses, Spring 2019

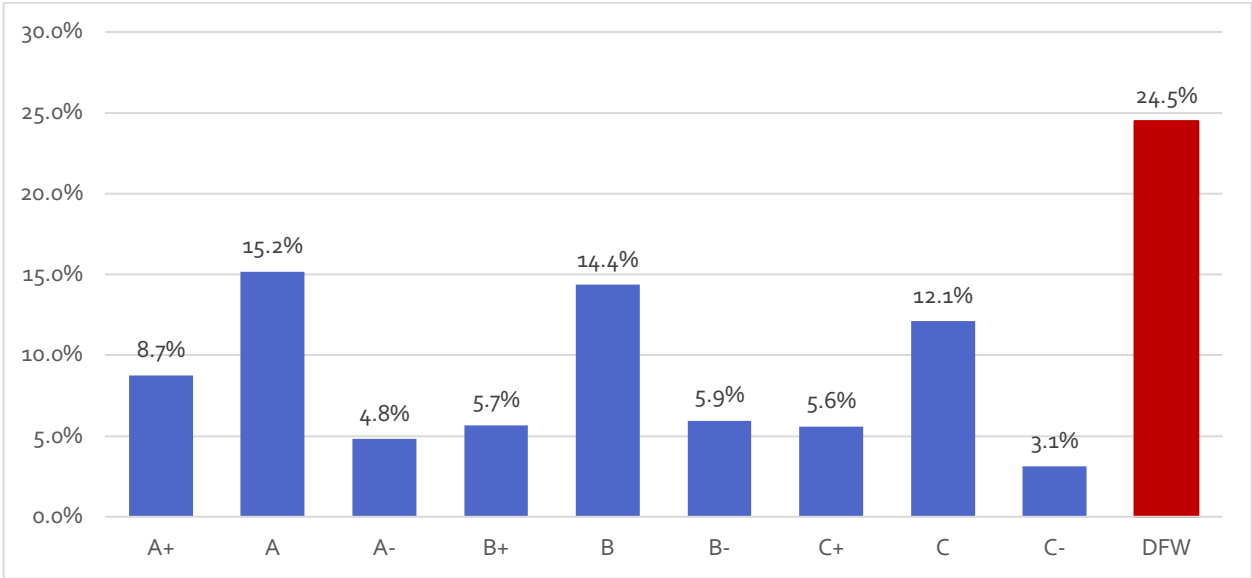
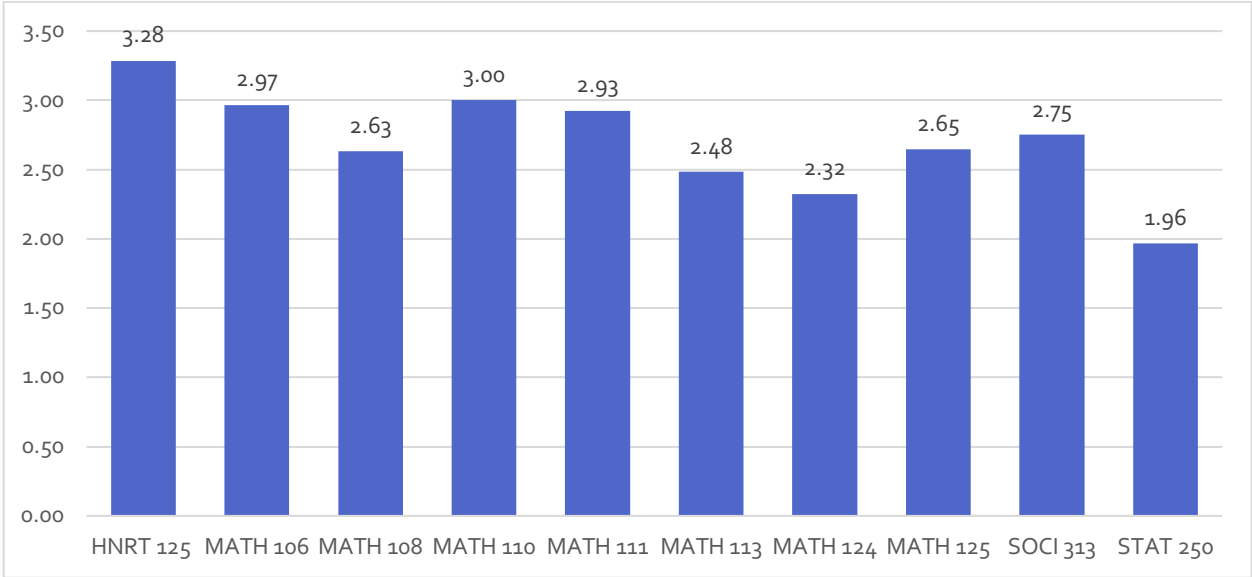


Figure 42. Average Final Grade by Course, Spring 2019



Assessment Methods

Three kinds of work samples were collected for this assessment:

1. Samples from Quantitative Reasoning courses taught in spring 2019
 - a. Project-based samples in which students were expected to analyze data
 - b. Calculation-based samples in which students solved equations or derived solutions (e.g. exams)
2. Samples from Capstone or Writing Intensive (WI) courses taught in spring 2018—a convenience sample of senior-level work collected across majors from the Critical Thinking and Written Communication assessments conducted in this period. Although the samples were identified as being appropriate for the Quantitative Reasoning assessment, it was not an intentional sampling for quantitative projects. Additionally, the samples represent individual work and not team-based projects that are typical in many fields that emphasize quantitative methods (i.e. engineering, business).

Faculty were asked to submit samples completed in the final third part of the semester and that allowed students to demonstrate their learning on one or more of the Quantitative Reasoning learning outcomes. Samples were selected using randomized course enrollment lists to insure the best possible representative sample.

The **AAC&U Quantitative Literacy VALUE Rubric** was used for this assessment. The rubric was selected by Mason faculty as a tool to assess individual student work on six learning tasks or outcomes that align well to the Mason Core Quantitative Reasoning learning outcomes. The rubric uses four performance descriptors: Benchmark, Emerging Milestone, Advanced Milestone, and Capstone, as well as an option for "no evidence." The performance descriptors are developmental, identifying student performance levels in a context of learning and growth. The rubric is intended to be used across all of the years of a student's college experience, and is not limited to a single course, assignment, or student class level.

Using a process modeled after the VALUE Institute reviewer calibration, faculty reviewers were trained to use the rubric to assess student work. Reviews were normed to produce consistent ratings across reviewers. Reviewers met for an in-person, one-day training and review session and completed the reviews of student work by the end of the day. Reviewers were faculty members who have taught Mason Core Quantitative Reasoning courses and related courses. Reviewers earned a small stipend for their efforts.

Each student work sample was assessed twice. Project-based samples were rated using the rubric. Calculation-based samples were rated a bit differently; the assignment itself was rated on the rubric for expected performance.

Learning Outcomes Assessment Results

Project-Based and Capstone Samples

Figures 43-48 display results from 224 randomly selected student work samples rated on the rubric, disaggregated by level: "QR" (n = 123) represents samples from the courses approved for Quantitative Reasoning and "In the Major" (n = 101) represents samples from the Capstone courses. The figures include "no evidence" ratings. A rating of "no evidence" was used when the learning outcome could not be detected in the sample; this could mean that either the assignment did not require application of the outcome, or that the student did not demonstrate it. A "no evidence" rating provides important information in aggregate but is given no value for an individual sample.

Figure 43. Assessment Results for Outcome 1, Interpretation

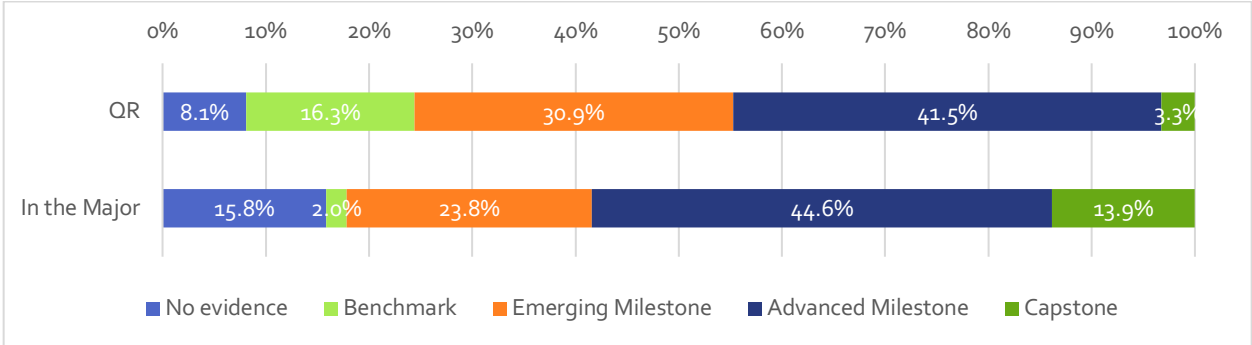


Figure 44. Assessment Results for Outcome 2, Representation

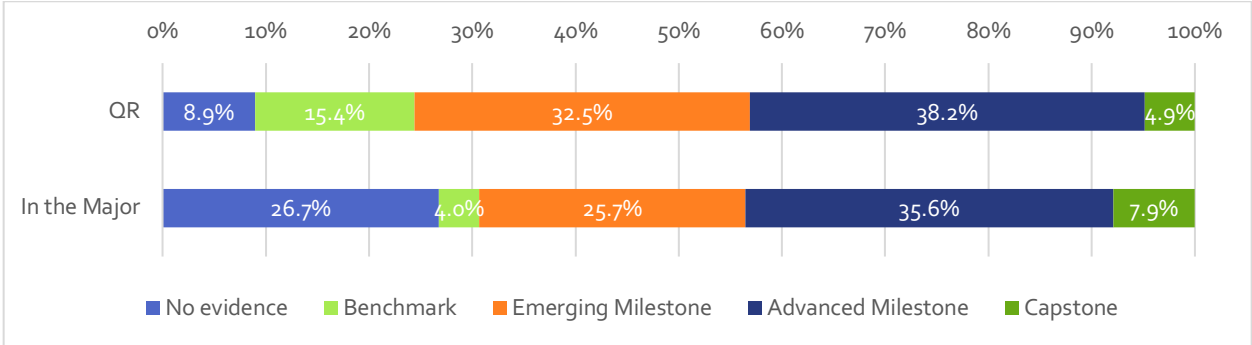


Figure 45. Assessment Results for Outcome 3, Calculation

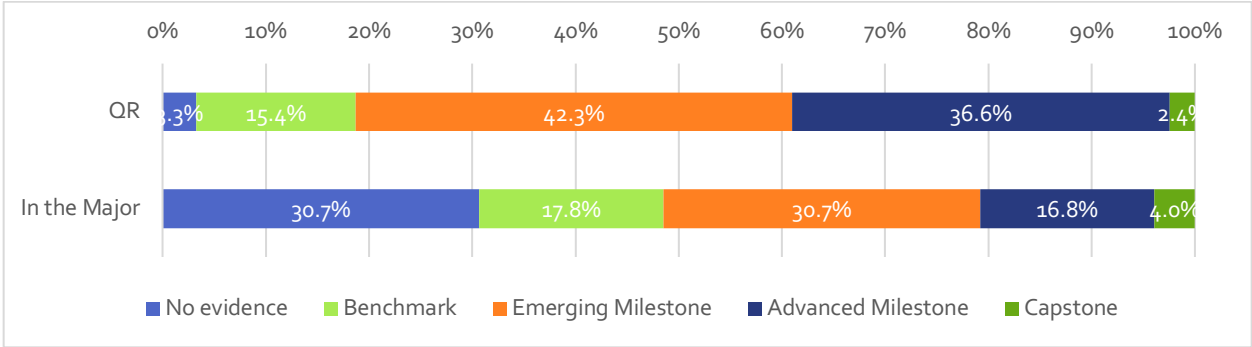


Figure 46. Assessment Results for Outcome 4, Application/Analysis

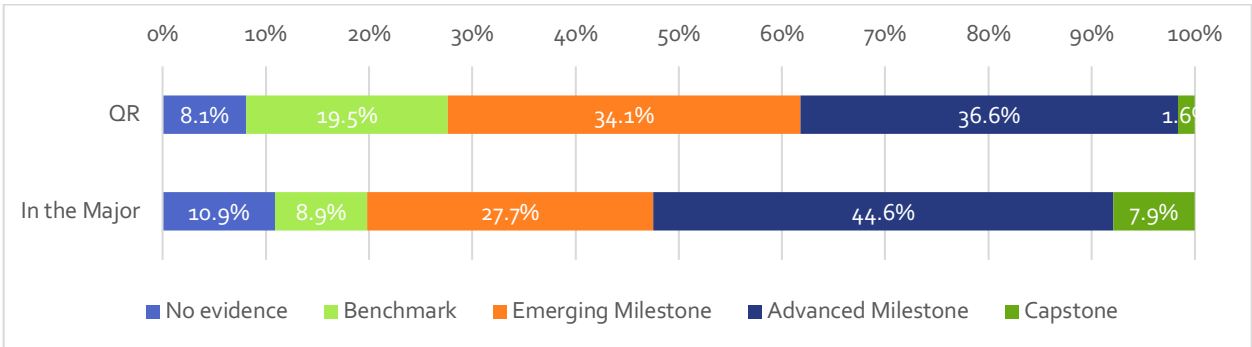


Figure 47. Assessment Results for Outcome 5, Assumptions

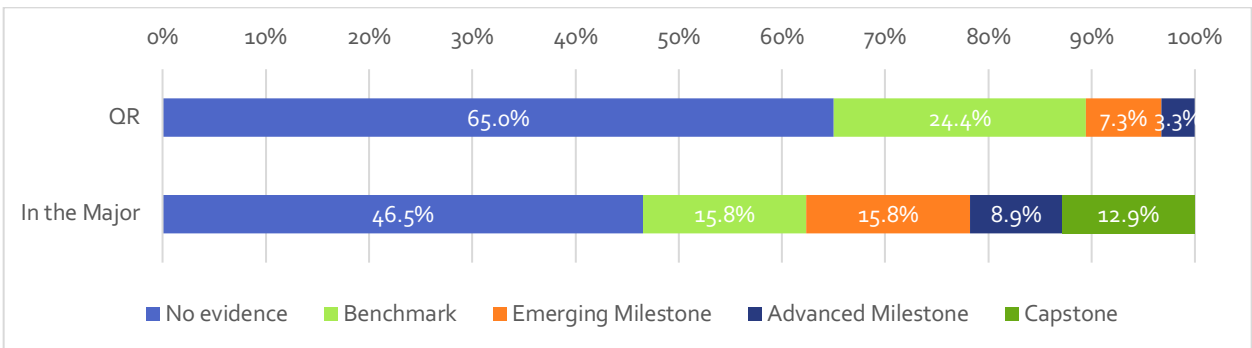
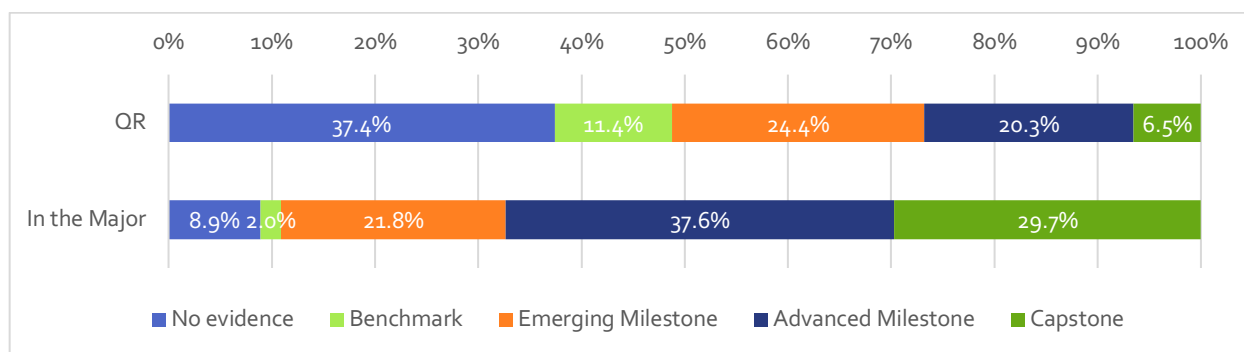


Figure 48. Assessment Results for Outcome 6, Communication

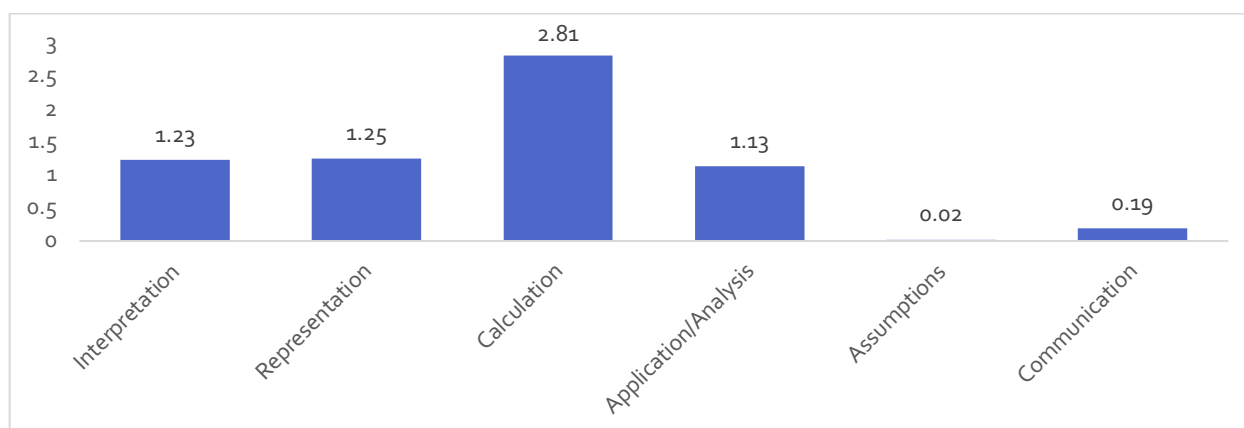


Calculation-Based Samples

Several lower-division QR courses (MATH 106, 108, 111; HNRT 125) submitted exams or quizzes as samples. These calculation-based assignments expected right or wrong answers and, for the most part, did not allow students to show their thinking. Because this presented a challenge to scoring samples on the rubric, it was determined that the *assignment* be rated on the rubric, to understand the level at which QR courses are *expecting* students to perform. This provides some information about course emphases for assessing learning outcomes.

Figure 49 displays mean ratings for the assignments across the four courses; note that because of multiple sections and instructors, the expectations across sections of the same course varied slightly. Overall, **Calculation** was the most emphasized outcome. MATH 108, 111 and HNRT 125 all emphasized **Interpretation** and **Representation** in addition to **Calculation**, but at lower levels. MATH 108 and 111 also emphasized **Application/Analysis**. Very few of these samples expected students to make **Assumptions** or **Communicate** quantitative evidence in support of an argument.

Figure 49. Mean Ratings of QR Calculation-Based Assignments



N = 123; Based on a scale of 0-4 where 0=No Evidence, 1=Benchmark, 2=Emerging, 3=Advanced, and 4=Capstone.

Highlights from Analysis of Results

Data were analyzed to ascertain differences among students in achieving the six learning outcomes. Comparison tests were conducted using nonparametric statistics because rubric data are ordinal; Independent-Samples Mann-Whitney U , was used when analyzing differences between two groups, and Independent-Samples Kruskal–Wallis H test was used to analyze differences across three or more groups or courses. “No evidence” was treated as missing.

An Independent-Samples Mann-Whitney U test found differences between the lower-division QR samples and the upper-division In the Major samples on all outcomes except Calculation, with In the Major samples rated significantly higher ($p < .05$) on all five outcomes (see Table 24).

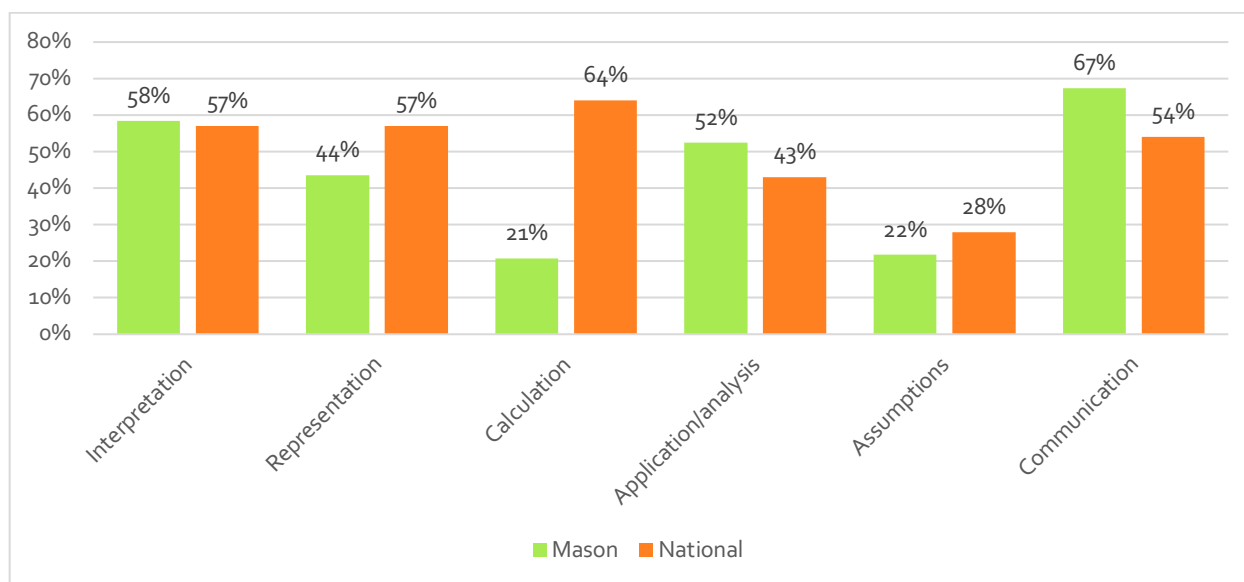
It was determined that the variations in courses and subsequent sample sizes were insufficient to do adequate comparisons by student demographics. Analyses comparing In the Major samples in aggregate did not reveal differences by gender, race, nor transfer status.

How do Mason Students Compare?

In comparing results from a 2017 national study (McConnell & Rhodes, 2017) using samples of student work from seniors at 4-year institutions and Mason’s Capstone samples, this assessment provides some information about how Mason students perform compared to their peers on combined rubric ratings of Advanced + Capstone. Similar to Mason, national data revealed that students were *as likely* to show strength in **Interpretation**. Mason students seemed to perform *better* than their national peers in **Application/Analysis** and **Communication**. Note that this is an observational comparison; the raw data from the national study was not available to perform a statistical comparison. See Figure 50.

It is important to note that the samples for Mason’s assessment were drawn from a convenience sample from the Capstone assessment efforts, and were not specifically requested for the QR assessment. This likely has bearing on being able to accurately compare student performance to the national samples. In addition, we did not receive individual samples from disciplines for which quantitative reasoning in the senior year is paramount, such as Business and Engineering.

Figure 50. Comparison of Mason Capstone Samples to National Results



Calculus Assessment

During the assessment period, the Calculus series MATH 113, *Analytic Geometry and Calculus I*, and MATH 114, *Analytic Geometry and Calculus II*, was involved in an NSF-funded initiative⁷ to increase the use of active learning instructional techniques. In the initial phase of this project, enrollment data were analyzed to determine possible inequities in student performance. Also, project faculty in the Math department sought to understand how a placement test taken at the beginning of MATH 113 could help students select into the appropriate level of Calculus—MATH 113 (standard), MATH 123/124 (two semester sequence for students with limited math background), or MATH 105 (Pre-Calculus). The following sections offers a summary of these analyses in lieu of learning outcomes assessment results.

Calculus Enrollment and Grade Performance, AY17-19

Student enrollment data from AY17-19 were analyzed to understand the enrollment profile and grade performance by gender and race.

- Although institutional undergraduate enrollment was 50/50 female/male during this time period, the Calculus series was more heavily male, with 68% of MATH 113 and 74% of MATH 114 (see Figure 53).

⁷ Nelson, J. (PI), Lester, J., Sachs, R., Rosenberg, J., & Foster, S. (Co-PIs). (Funded 2018-2023). *Collaborative research: Fostering a culture of inquiry-based learning by building course-based communities of transformation*. National Science Foundation, Improving Undergraduate STEM Education (IUSE) Program. \$1.745 million. Award # NSF 1821589

- Female students earned statistically significantly higher mean grades than male students in both courses (see Figure 54).
- Enrolled students in MATH 113/114 were more likely to be Asian and less likely to be Black/African American or Hispanic/Latino than the institutional undergraduate population (see Figure 55).
- Asian and White students had the highest grades in Calculus courses, while African American students had the lowest grades; differences were statistically significant (see Figure 56).

Analysis of Pilot Placement Test, Standardized Test Scores, and Grade Performance

The Math faculty piloted a math placement test to replace the existing placement test. The test was developed by Math faculty at the University of Colorado Boulder, and comprises 30 multiple-choice items with content in algebra and trigonometry, with only one correct choice per item. The placement test was administered in paper form in MATH 113 recitation sections in the first week of classes of the fall 2019 semester. Results were merged with enrollment data to analyze student performance.

Using Pearson bivariate correlation, placement scores were analyzed against ACT Math scores, SAT Math scores⁸, and final course grade for MATH 113 completers. Placement scores were significantly correlated with SAT Math scores ($r^2 = 0.355$, $p < .01$) and course grade ($r^2 = 0.437$, $p < .01$), but not ACT Math scores (see Table 6). SAT Math scores showed a significant but weak positive correlation with course grade ($r^2 = 0.140$, $p < .01$). See Table 22.

Based on the analysis of Mason student scores on the Placement Exam, a score range was proposed to provide guidance to students for enrolling in the course most appropriate to their needs. The Math department is considering adopting the new placement test for an upcoming academic year. The full report is available upon request.⁹

⁸ Note that students typically report either SAT or ACT scores to the university, but not both. Transfer students are not required to report standardized test scores, though some do and those are included.

⁹ Foster, S. L. (2020). Results of analysis of CU math placement test pilot, fall 2019. George Mason University Office of the Provost.

Table 22. Placement Exam Scores Correlation with ACT Scores, SAT Scores, and MATH 113 Course Grade

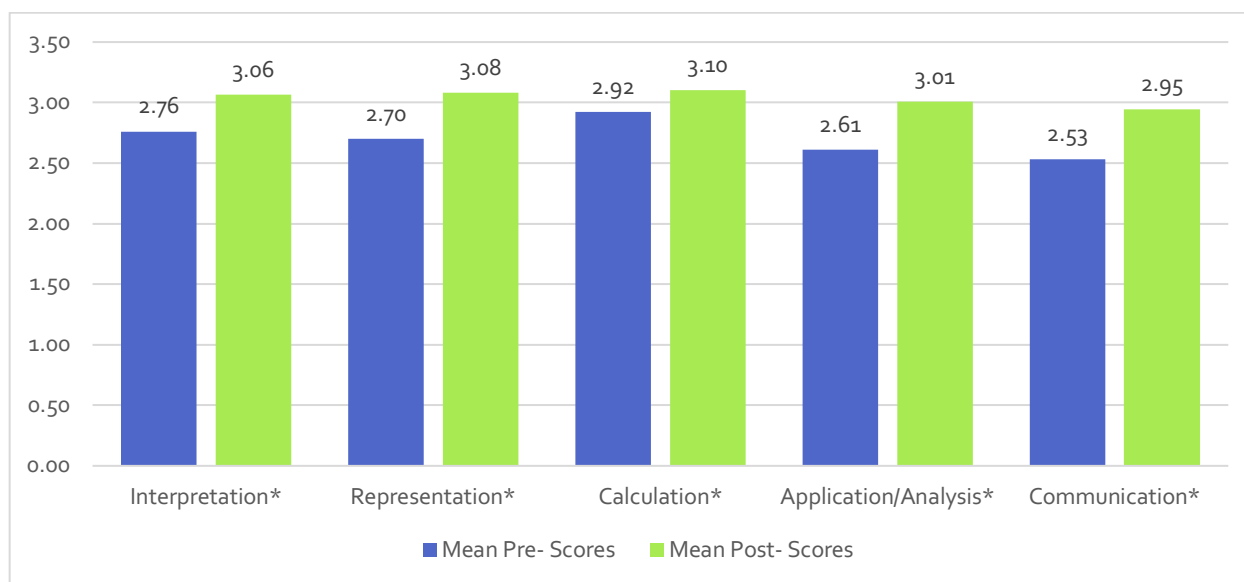
Descriptive Statistics					
	Mean	Std. Deviation	N		
Total Raw Score	20.73	4.875	617		
ACTMathHigh (Official)	3.38	8.886	703		
SATMathHigh (Official)	486.64	280.202	703		
Grade Points	2.5743	1.30853	674		
Correlations					
		Total Raw Score	ACT Math	SAT Math	Course Grade
Total Raw Score	Pearson Correlation	1	0.073	.355**	.437**
	Sig. (2-tailed)		0.068	0.000	0.000
	N	617	617	617	595
ACTMathHigh (Official)	Pearson Correlation	0.073	1	-0.040	-0.026
	Sig. (2-tailed)	0.068		0.288	0.503
	N	617	703	703	674
SATMathHigh (Official)	Pearson Correlation	.355**	-0.040	1	.140**
	Sig. (2-tailed)	0.000	0.288		0.000
	N	617	703	703	674
Grade Points	Pearson Correlation	.437**	-0.026	.140**	1
	Sig. (2-tailed)	0.000	0.503	0.000	
	N	595	674	674	674

** . Correlation is significant at the 0.01 level (2-tailed).

Student Self-Assessment

All students who were enrolled in a Mason Core Quantitative Reasoning course during the assessment period received an online self-assessment survey at the end of the semester. The retrospective pre-post self-assessment asked students to rate their knowledge and skills on five learning outcomes at the beginning of the semester (pre), and then again at the end of the semester (post). In total, 327 students completed both the pre and post items, resulting in a 10.4% response rate. A t-test pairwise comparison showed significant perceived learning gains on all five outcomes (see Figure 51).

Figure 51. Mean Scores on Student Learning Self-Assessment



Mean scores, self-reported on a scale of 1-4, n=327, * $p < .05$

How do the Results Meet Expectations?

Because this was the first time that Mason used the **Quantitative Literacy rubric** to assess student work, these data provide baseline information only. Math faculty and faculty in majors for which quantitative literacy is emphasized should consider these results in terms of the learning outcomes identified for their academic programs. Results for the In the Major samples are inconclusive; to test the efficacy of the rubric for use in the Capstone, the assessment should be repeated with an intentional sample across the majors.

How are Results Being Used to Improve Students' Educational Experience?

Results have been shared with the Mason Core Committee and the Math department. The NSF IUSE project is currently focused on changing the culture of instruction in Calculus courses, and in the near term, efforts will be expanded to Computer Science faculty. As the rubric was well-received by faculty on the QR working group, it is recommended that the rubric be used as one tool to guide course and assignment design for the development of quantitative literacy.

Limitations of this Assessment

As this was the first time that quantitative literacy was assessed using this method, caution should be taken in interpreting the results. The rubric shows promise as a tool for guiding the language and expectations for quantitative literacy across the Mason undergraduate experience, allowing faculty to plan learning experience that support development of these skills from first through senior years. Overall, this assessment was well-designed for project-

based work because it allows students to demonstrate their reasoning ability. However, the rubric is limited for use with calculation-based assignments for which there is a right answer and students are not asked to document their mathematical thinking. This finding is consistent with the experience at Fitchburg State University (Berg et. al., 2014), which concluded that more carefully constructed assignment prompts were needed to elicit higher-order thinking. Additionally, the sample sizes for many of the courses were insufficient to perform a robust analysis of results by student demographics; in future assessments, efforts should be made to collect bigger samples of student work that best align with the rubric method.

Assessment Rubric(s)

The **Quantitative Literacy VALUE Rubric** was selected by a team of Mason Quantitative Reasoning faculty to evaluate student work for the Mason Core learning outcomes in Quantitative Reasoning. The team agreed that the outcomes and performance descriptors were appropriate for the courses they teach, as well as for desired outcomes for undergraduates completing non-mathematics majors.

AAC&U Quantitative Literacy VALUE Rubric is reprinted with permission from "*VALUE: Valid Assessment of Learning in Undergraduate Education*." Copyright 2019 by the Association of American Colleges and Universities. <http://www.aacu.org/value/index.cfm>.

Table 23. Enrollment in Mason Core Quantitative Reasoning Courses, AY2015-19

	AY2015		AY2016		AY2017		AY2018		AY2019	
	#Course Sections	Enroll	#Course Sections	Enroll	#Course Sections	Enroll	#Course Sections	Enroll	#Course Sections	Enroll
HNRT 125	3	83	3	89	3	88	3	87	4	107
MATH 106	39	1,184	40	1,105	36	1,032	34	1,059	29	1,072
MATH 108	17	880	17	931	18	970	15	947	15	817
MATH 110	3	83	2	27	5	151	5	152	6	157
MATH 111	5	168	7	278	7	217	7	251	7	227
MATH 113	22	1,210	21	1,185	20	1,181	21	1,249	21	1,282
MATH 124			2	51	2	60	4	105	5	141
MATH 125	9	437	12	530	12	626	13	701	13	798
SOCI 313					4	105	4	108	4	112
STAT 250	24	1,768	24	1,731	20	2,028	23	1,881	23	1,886
TOTAL	122	5,813	128	5,927	127	6,458	129	6,540	127	6,599

Figure 52. Five-Year Enrollment Trends in Mason Core Quantitative Reasoning Courses, AY2015-19

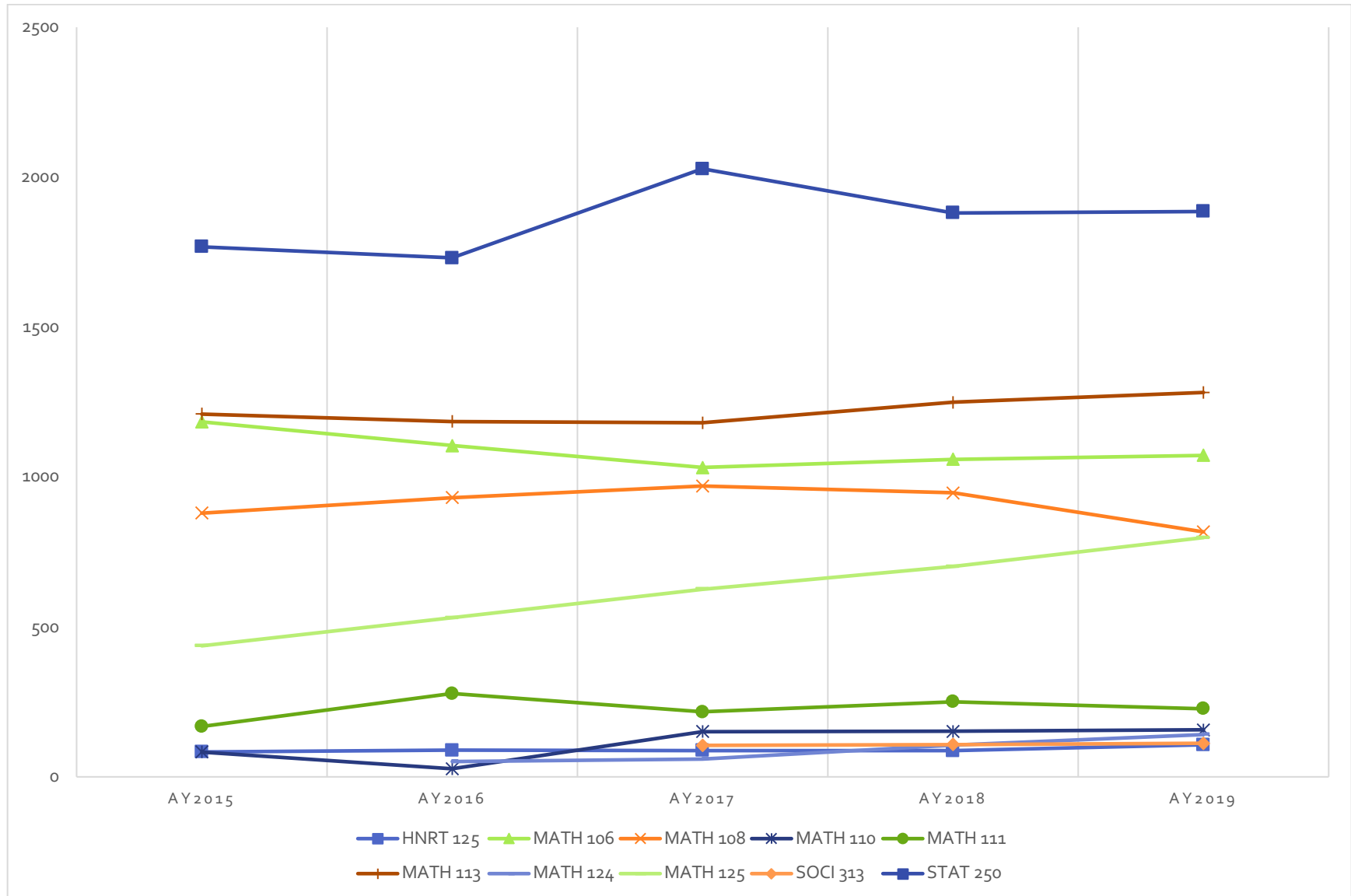


Table 24. Mann-Whitney U Test: Comparison of Rubric Ratings, Lower-Division QR vs. Upper-Division In the Major

	Mean Rank (n)		U	Z	p	Sig.
	Lower	Upper				
Interpretation	86.54 (113)	116.74 (85)	3337.500	-3.974	0.000	*
Representation	86.80 (112)	103.64 (74)	3394.000	-2.245	0.025	*
Calculation	99.53 (119)	87.29 (70)	3625.500	-1.590	0.112	
Application/analysis	91.72 (113)	114.91 (90)	3923.500	-2.999	0.003	*
Assumptions	36.00 (43)	59.35 (54)	602.000	-4.348	0.000	*
Communication	66.21 (77)	100.73 (92)	2095.000	-4.796	0.000	*

Figure 53. Enrollment by Course and Gender, AY17-19

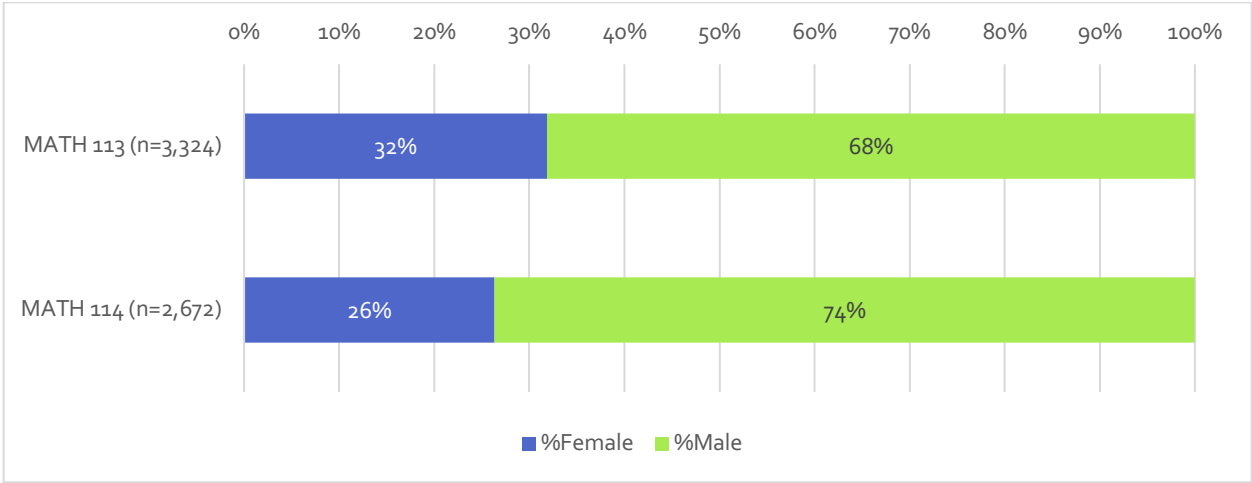
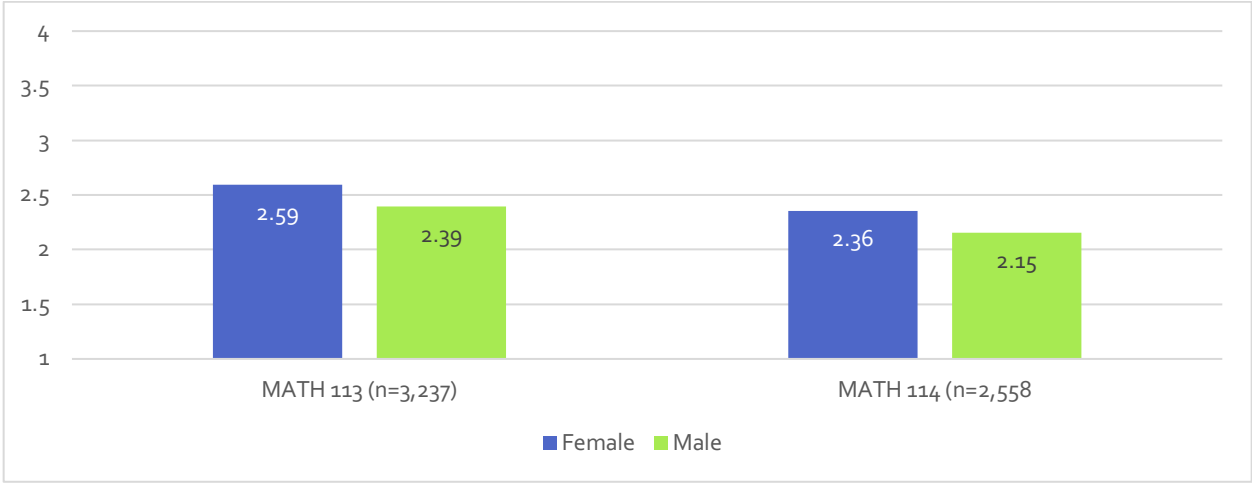


Figure 54. Mean Grades by Course and Gender, AY17-19



On a 4.0 scale; calculations do not include withdrawals

Figure 55. MATH 113/114 Enrollment by Race, Compared to Fall 2019 Mason Undergraduate Enrollment

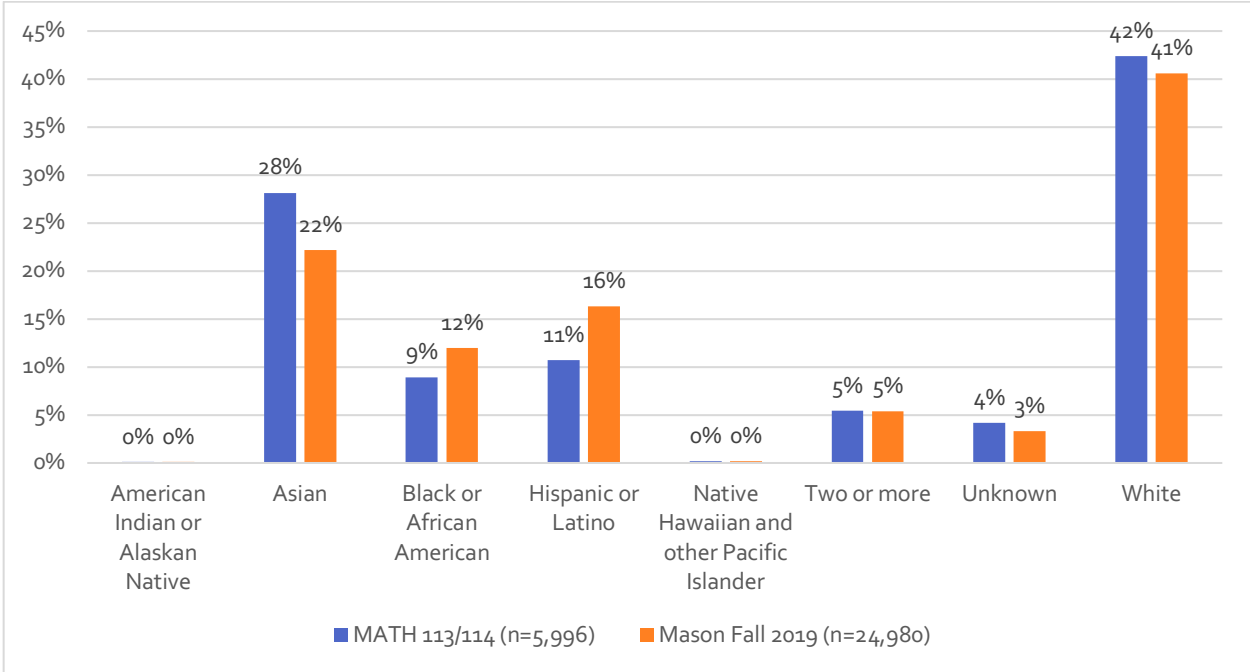
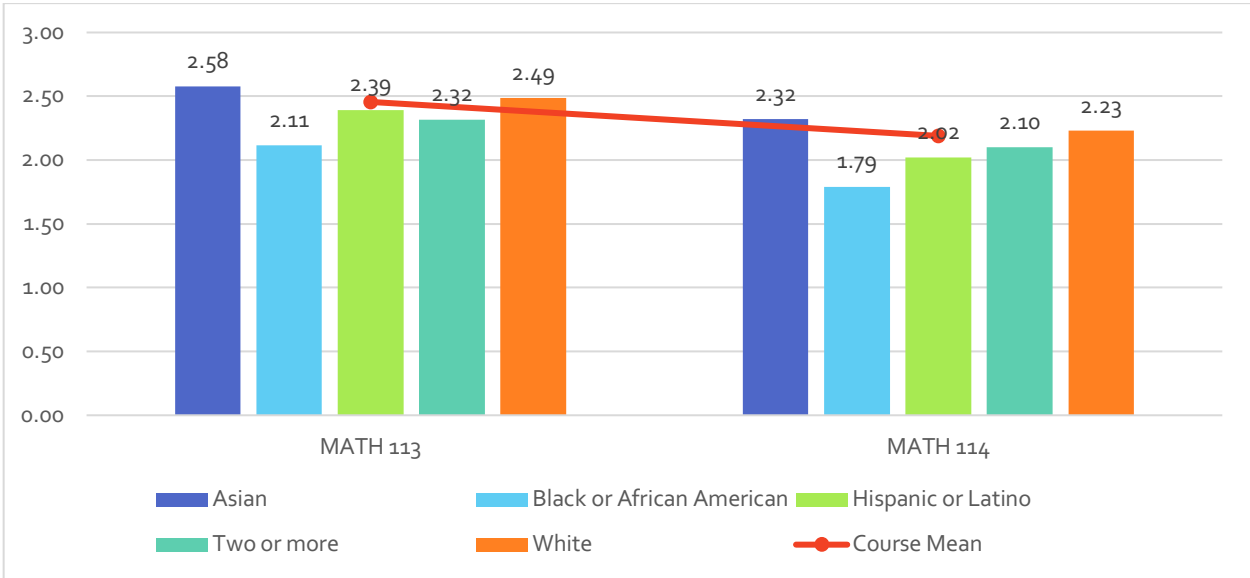


Figure 56. Mean Grades in MATH 113/114 by Course and Race, AY17-19



QUANTITATIVE LITERACY VALUE RUBRIC

for more information, please contact value@aacu.org



The VALUE rubrics were developed by teams of faculty experts representing colleges and universities across the United States through a process that examined many existing campus rubrics and related documents for each learning outcome and incorporated additional feedback from faculty. The rubrics articulate fundamental criteria for each learning outcome, with performance descriptors demonstrating progressively more sophisticated levels of attainment. The rubrics are intended for institutional-level use in evaluating and discussing student learning, not for grading. The core expectations articulated in all 15 of the VALUE rubrics can and should be translated into the language of individual campuses, disciplines, and even courses. The utility of the VALUE rubrics is to position learning at all undergraduate levels within a basic framework of expectations such that evidence of learning can be shared nationally through a common dialog and understanding of student success.

Definition

Quantitative Literacy (QL) – also known as Numeracy or Quantitative Reasoning (QR) – is a "habit of mind," competency, and comfort in working with numerical data. Individuals with strong QL skills possess the ability to reason and solve quantitative problems from a wide array of authentic contexts and everyday life situations. They understand and can create sophisticated arguments supported by quantitative evidence and they can clearly communicate those arguments in a variety of formats (using words, tables, graphs, mathematical equations, etc., as appropriate).

Quantitative Literacy Across the Disciplines

Current trends in general education reform demonstrate that faculty are recognizing the steadily growing importance of Quantitative Literacy (QL) in an increasingly quantitative and data-dense world. AAC&U's recent survey showed that concerns about QL skills are shared by employers, who recognize that many of today's students will need a wide range of high level quantitative skills to complete their work responsibilities. Virtually all of today's students, regardless of career choice, will need basic QL skills such as the ability to draw information from charts, graphs, and geometric figures, and the ability to accurately complete straightforward estimations and calculations.

Preliminary efforts to find student work products which demonstrate QL skills proved a challenge in this rubric creation process. It's possible to find pages of mathematical problems, but what those problem sets don't demonstrate is whether the student was able to think about and understand the meaning of her work. It's possible to find research papers that include quantitative information, but those papers often don't provide evidence that allows the evaluator to see how much of the thinking was done by the original source (often carefully cited in the paper) and how much was done by the student herself, or whether conclusions drawn from analysis of the source material are even accurate.

Given widespread agreement about the importance of QL, it becomes incumbent on faculty to develop new kinds of assignments which give students substantive, contextualized experience in using such skills as analyzing quantitative information, representing quantitative information in appropriate forms, completing calculations to answer meaningful questions, making judgments based on quantitative data and communicating the results of that work for various purposes and audiences. As students gain experience with those skills, faculty must develop assignments that require students to create work products which reveal their thought processes and demonstrate the range of their QL skills.

This rubric provides for faculty a definition for QL and a rubric describing four levels of QL achievement which might be observed in work products within work samples or collections of work. Members of AAC&U's rubric development team for QL hope that these materials will aid in the assessment of QL – but, equally important, we hope that they will help institutions and individuals in the effort to more thoroughly embed QL across the curriculum of colleges and universities.

Framing Language

This rubric has been designed for the evaluation of work that addresses quantitative literacy (QL) in a substantive way. QL is not just computation, not just the citing of someone else's data. QL is a habit of mind, a way of thinking about the world that relies on data and on the mathematical analysis of data to make connections and draw conclusions. Teaching QL requires us to design assignments that address authentic, data-based problems. Such assignments may call for the traditional written paper, but we can imagine other alternatives: a video of a PowerPoint presentation, perhaps, or a well designed series of web pages. In any case, a successful demonstration of QL will place the mathematical work in the context of a full and robust discussion of the underlying issues addressed by the assignment.

Finally, QL skills can be applied to a wide array of problems of varying difficulty, confounding the use of this rubric. For example, the same student might demonstrate high levels of QL achievement when working on a simplistic problem and low levels of QL achievement when working on a very complex problem. Thus, to accurately assess a student's QL achievement it may be necessary to measure QL achievement within the context of problem complexity, much as is done in diving competitions where two scores are given, one for the difficulty of the dive, and the other for the skill in accomplishing the dive. In this context, that would mean giving one score for the complexity of the problem and another score for the QL achievement in solving the problem.

