

2019

Quantitative Literacy in the Core Curriculum of Hood College: Chapter II, Outcomes and Assessment

Betty Mayfield

Hood College, mayfield@hood.edu

Ann Stewart

Hood College, stewart@hood.edu

Follow this and additional works at: <https://scholarcommons.usf.edu/numeracy>



Part of the [Educational Assessment, Evaluation, and Research Commons](#)

Recommended Citation

Mayfield, Betty, and Ann Stewart. "Quantitative Literacy in the Core Curriculum of Hood College: Chapter II, Outcomes and Assessment." *Numeracy* 12, Iss. 1 (2019): Article 9. DOI: <https://doi.org/10.5038/1936-4660.12.1.9>

Authors retain copyright of their material under a [Creative Commons Non-Commercial Attribution 4.0 License](#).

Quantitative Literacy in the Core Curriculum of Hood College: Chapter II, Outcomes and Assessment

Abstract

In a previous article, we described our college's new core curriculum, which included a Quantitative Literacy (QL) component for the first time. We explained how we defined QL in the college catalog, and how we used that definition to choose courses to satisfy the new requirement. We then discussed our early efforts at assessing the effectiveness of the QL program and described our plans for the future. Here we report on our progress towards those goals, including working with faculty from other departments and with our institutional research office to develop a more sophisticated assessment plan, as well as creating and implementing easier-to-use surveys and assessment instruments.

Keywords

QL, assessment, core curriculum, learning outcomes

Creative Commons License



This work is licensed under a [Creative Commons Attribution-Noncommercial 4.0 License](https://creativecommons.org/licenses/by-nc/4.0/)

Cover Page Footnote

Betty Mayfield has recently retired after teaching mathematics for thirty-eight years at Hood College in Frederick, Maryland. She caught the QL bug at an MAA PREP workshop at the Sleeping Lady Mountain Resort in 2004, and she learned about assessment in a series of MAA SAUM workshops at around the same time. She has enjoyed weaving those two strands together in her work at Hood.

Ann Stewart is an Associate Professor of Mathematics at Hood College in Frederick, Maryland. She learned about both QL and assessment on the job at Hood, and oversees both in her current role as department chair. She is also a co-Principal Investigator on Hood's Noyce STEM Teacher Education Partnership (NSTEP), a project funded by the NSF.

Erratum

In the originally published version, a cited author's name was mis-spelled. That has been corrected here.

Introduction

Three years ago, we reported in this journal on Hood College's new core curriculum and its inclusion of a Quantitative Literacy (QL) component (Mayfield and Dunham 2015). We described our first efforts at assessment of the program by asking three types of questions:

- What are the characteristics of a QL course? How will we recognize one? How can we make sure all instructors are on the same page?
- What are the overall student learning objectives for a QL course? How can we tell if students are meeting them?
- Did students' attitudes towards mathematics and their confidence in doing mathematics change after taking one of those courses?

We ended by describing our preliminary results and plans for the future in those three areas.

In this paper, we report on our progress towards those plans by addressing each of the three questions above. Our paper will follow this general outline:

- QL courses: What are they?
- Setting goals for QL courses and assessing them
- A word about attitudes and confidence
- Current discussions and future plans.

We will focus primarily on learning objectives and assessment. In our earlier paper, the term "assessment" referred to program assessment. Here we will also broaden our focus to include the evaluation of individual student performances. Our objective is to help and encourage other institutions as they develop and implement their own QL assessment plans.

Assessing Quantitative Literacy: Background

As the editors of this journal have pointed out, educational assessment of quantitative literacy is of great interest to academics involved in numeracy efforts (e.g., Vacher 2015). How do we define quantitative literacy? How do we teach it? What do we want students to be able to do once they have completed a QL course? How can we tell what they know, and what they can do? Many of the articles we cited in our first paper (Gold 2006; Grawe 2011; Sikorski et al. 2011; Ward et al. 2011; Boersma and Klyve 2013; Wright and Howard 2015) were devoted to this topic, and that is just the tip of the iceberg. We note especially the multi-institution assessment instrument described in Gaze et al. (2014) and all of the articles in the Assessment Theme Collection in Volume 8, Issue 1 of this journal. Since the publication of our first paper, we have attempted to better

articulate our QL goals and objectives and to create a more robust assessment plan.

QL Courses: What Are They?

In our first approach to the “What is a QL course?” question, the mathematics faculty brainstormed over several department meetings to create a list of desirable characteristics of any course that belonged in the QL section of the core curriculum – not just *what* we taught, but *how* we taught it. Those characteristics included the use of problem-solving and working with data, as well as collaboration, active learning, and multiple forms of assessment. Many of those characteristics made their way into the official college catalog description (Hood College 2017, 36-37) of Quantitative Literacy:

Quantitative Literacy (QL) is a habit of mind. It involves using elementary mathematical tools to interpret and manipulate quantitative data arising in a variety of contexts. It is marked by computational fluency, and by competence and comfort in working with numerical data. Those who are quantitatively literate can create arguments supported by data and can communicate those arguments in many ways – using tables, graphs, mathematical expressions, and words.

A course that satisfies the QL section of the Core Curriculum should have as its main focus the use of mathematics to solve real-world problems. In those courses, using data and appropriate technology, students will collaborate to solve multi-step problems and effectively communicate their reasoning to others.

During the 2015-2016 academic year, we surveyed faculty members who were teaching mathematics courses that had been proposed for the QL section of the core curriculum, and we asked if their courses in fact included many of those characteristics. Although we were satisfied with the results, we realized that our instructions were too vague. Operationally, the exercise proved to be much too time-intensive for the faculty involved.

Table 1
Surveyed Courses in the Core Curriculum (AY 2015-2016)

MATH 111A	Mathematics of Daily Life (two different instructors)
MATH 111B	Mathematics of Democracy
MATH 111G	Mathematics of Games and Sports
MATH 112	Applied Statistics (for non-majors)
MATH 201	Calculus I (two different instructors)
MATH 213	Statistical Concepts and Methods (for math and science majors)
ECON/MGMT 212	Statistics for Economics & Management
ENSP 103	Intro to Geographic Information Systems
SOC 261	Quantitative Methods for the Social Sciences

After that initial experience, we developed a brief survey (Text Appendix A¹) and deployed it electronically, via Survey Monkey, to faculty members in all the

¹ Text Appendices A-D are in Supplemental File 1.

departments involved during the 2015-2016 academic year. We received responses from eleven sections of nine courses taught by ten different instructors (Table 1). We will repeat this exercise every few years, to make sure that we all agree on – and perhaps need to fine-tune – what we consider to be important in the content and pedagogy of a QL course.

Results from the brief survey are in Tables 2-4, representing the responses to multipart questions 3-5, respectively, of the survey (questions 1 and 2 were to identify the responder and course section; see Text Appendix A). Questions 3 and 4 sought information about characteristics of the courses, and question 5 asked about pedagogy.

Table 2
Instructor Survey Results, AY 2015-16. Question 3, Course Characteristics

For each of the following characteristics, please indicate the extent to which it is incorporated into your course.				
Data in table: raw number of responses, percent of all responses.				
<i>n</i> = 11	Not at all or not much (1)	A moderate amount (2)	Often or a lot (3)	Weighted average
Problem solving: applying mathematics to real-world problems	0 0%	3 27%	8 73%	2.73
Working with data	0 0%	1 9%	10 91%	2.91
Using (and knowing when to use) appropriate technology	0 0%	1 9%	10 91%	2.91
Examining quantitative arguments in the media, or in professional journal articles	4 36%	5 45%	2 18%	1.82

The only responses to four-part question 3 that surprised us were those of its fourth part (last row of Table 2). The instructors of those courses explain:

Although most of the homework problems that I gave the students used data drawn from reports in the media or from professional articles, I did not have the students look directly at those sources.

I would like to incorporate more focus on reading quantitative studies than is in the course currently.

Additionally, some instructors pointed out that, since textbooks have already done much of this work for us, especially in freshman-level courses, they did not look for outside articles themselves.

Similarly, the responses to the first part of three-part question 4 stood out (first row of Table 3). Most of these QL classes apparently did not involve defending one's opinion about an issue as much as we expected – something we thought would be important at a liberal arts college dedicated to encouraging critical thinking. Drilling down (middle row of Table 3), it was also of interest that the two classes whose instructors used a modified Rule of Four (graphs,

charts, tables, equations) only “a moderate amount” were mathematics classes – both sections of MATH 111. But we realize we gave the instructors no guidance on what constituted “moderate” use and what “often” meant; these responses may just represent individual interpretations of those terms. The one class that did not use a long class assignment or project that semester (last row of Table 3) was The Mathematics of Games and Sports. After reviewing these results, the current instructor decided to include a final project.

Table 3
Instructor Survey Results, AY 2015-16. Question 4, More Course Characteristics.

How often does your course incorporate...				
<i>n = 11</i>	Not at all or not much (1)	A moderate amount (2)	Often or a lot (3)	Weighted average
Using quantitative skills to defend one's opinion	0 0%	9 82%	2 18%	2.18
Presenting data in useful ways: graphs, charts, table, equations	0 0%	2 18%	9 82%	2.82
Solving multi-step problems, as in a long assignment or class project	1 9%	3 27%	7 64%	2.45

Table 4
Instructor Survey Results, AY 2015-16. Question 5, Pedagogical Strategies

How often does your course involve...				
<i>n = 11</i>	Not at all or not much (1)	A moderate amount (2)	Often or a lot (3)	Weighted average
Active or discovery learning	1 9%	4 36%	6 55%	2.45
Collaborative learning	0 0%	3 27%	8 73%	2.73
Students' writing about quantitative issues in everyday life	2 18%	5 45%	4 36%	2.18

Finally, three-part question 5 (Table 4) followed up on the description of quantitative literacy in the Hood College Catalog, where it makes clear that we expect these courses to include active learning, collaboration, writing, and the appropriate use of technology. We agree with Larry Cuban (2001) that quantitative literacy and progressive pedagogy are inextricably linked. We found that instructors tend to use these “reform” teaching strategies at least a moderate amount in these classes, especially collaboration. We realized from instructor comments on the survey that we should look carefully at how these questions are written and express them more clearly in the future. For example, one instructor

worried whether drawing conclusions from real data supplied in a textbook and gathered with careful experimental design counted as “everyday life.” Another commented “I wasn’t sure if the emphasis was on *problem-solving* or on *real-world*.”

Setting Goals and Assessing Them

Our previous foray into the assessment of the QL requirement consisted of the mathematics faculty’s developing student learning outcomes and then looking at examples of student work which seemed to represent those students’ mastery of the outcomes. The outcomes we developed then were that *students who successfully complete a QL course should be able to*:

1. Demonstrate computational fluency.
2. Understand and interpret data presented in a variety of formats, and convert from one format to another.
3. Draw conclusions based on numerical data and assess the limitations of those conclusions.
4. Evaluate quantitative arguments in a variety of settings.
5. Communicate their understanding of the usefulness of mathematics.

We did not attempt to measure results in the aggregate with any sort of data collection. We also later wondered if we should have included an explicit outcome related to the use of technology, and if we should mention that we hoped students’ attitudes towards and confidence in using mathematics should improve. We planned to address those questions in the next assessment plan.

As we moved into the next phase of this work, we knew we would need to include faculty members from other departments who were teaching QL courses and to develop a more comprehensive assessment plan. Our new efforts took place in the context of a much larger campus-wide assessment of the College’s core curriculum.

The C⁴ Plan

In the Fall 2014 semester, the Hood College Office of Institutional Research and Assessment (OIRA) began work on a comprehensive program to evaluate learning outcomes for the Core Curriculum. Known thereafter as the Core Curriculum and College Competencies (C⁴) Assessment Plan (OIRA 2014), the program defined a process for determining student learning outcomes and identifying “key assignments and assessment tools for capturing and evaluating college-wide student performance.” In particular, OIRA determined an assess-

ment cycle for the QL component of the Core Curriculum, to be repeated every four years (Table 5).

As a first step, the chairs of the departments of mathematics and sociology (and, later, psychology and economics/ business) met in January 2015 with the College's assessment coordinator to define common student learning outcomes for QL classes. The initial four outcomes agreed upon by that group were that *students would*:

Table 5
OIRA Assessment Cycle

Academic Year	Activity
2014-2015	Plan
2015-2016	Pilot
2016-2017	Gather data
2017-2018	Analyze and report

1. Demonstrate computational fluency using numerical data and appropriate technology.
2. Interpret quantitative data presented in a variety of formats.
3. Create arguments supported by data.
4. Communicate arguments using tables, graphs, mathematical expressions, and/or words.

Note that technology now appears in the outcomes, and we have combined some of the earlier statements, but we lost the battle on the usefulness of mathematics. Also, there is no mention of attitudes or confidence.

Those faculty members also discussed performance criteria: How will we know if students are achieving those outcomes? A preliminary criterion for demonstrating competence was agreed upon: 70% of students should achieve a performance level of at least 70% on a specific assignment or activity.

The next step was for the chairs of the four departments to take the preliminary student learning outcomes back to their departments for discussion and approval, and to make sure that the learning outcomes were appropriate for all courses in the QL section of the core curriculum. After much discussion and negotiation, the departments agreed on a final list of student learning outcomes – specifically, *students would*:

1. Interpret quantitative data arising in a variety of contexts.
2. Demonstrate computational fluency, including the use of technology as appropriate.
3. Create arguments supported by data.
4. Communicate arguments using quantitative tools such as tables, graphs, and mathematical expressions.
5. Communicate arguments through the narrative analysis.

We retained the focus on the appropriate use of technology, and we separated out the *verbal* description of arguments from the use of quantitative tools – something the assessment coordinator suggested. *Post-hoc*, we recognized that these learning outcomes are aligned much more closely with the description of QL in the College Catalog and are expressed in a way that is easier to assess.

In fact, these learning goals are used as criteria to evaluate whether proposed courses would satisfy the QL section of the Core Curriculum. The Biology Department, for example, recently proposed that their new course Environmental Science and Policy (ENSP) 103, Introduction to Geographic Information Systems, be approved by the College's Curriculum Committee as a course satisfying the QL component of the Core Curriculum. The Curriculum Committee consulted with the Department of Mathematics, who worked with the ENSP faculty to ensure that their course would cover the five student learning outcomes approved by the interdisciplinary group as well as the characteristics of a QL course measured by our faculty survey. After some adjustments, the course was approved as a QL course:

An introduction to Geographic Information Systems for students of all disciplines. This course will provide a suite of tools for creating, manipulating, analyzing, visualizing, and illustrating spatial data. Concepts presented in lecture will be put into practice through hands-on laboratory exercises utilizing appropriate GIS software. The culmination of the course is the presentation of discipline-specific original research projects employing the methods learned.

As for the focus on appropriate use of technology, we tend to use the computer in all of our QL courses. The specific technology depends on the subject matter and is determined by the instructor. We use Excel extensively in all of our MATH 111 courses; we know that use of spreadsheets will be helpful to students in subsequent courses across the curriculum. Statistics courses may use Excel, Minitab, R, or SPSS, depending on the discipline and the level of the course. Students in calculus learn to use Maple since it is used throughout the major.

The World of Assessment: A Rubric and an Assessment Map

We were then led by OIRA's assessment coordinator to develop a rubric (Text Appendix B) associated with this outcome set and an assessment map (Text Appendix C), which indicates key assignments in each QL course that would be used to exhibit mastery of each outcome. The rubric, similar in structure to that of other sections of the Core Curriculum, defines novice, emergent, proficient, and advanced levels of student achievement for each student learning outcome, where "proficient" is the minimum desired performance level in each case, corresponding to the 70% level agreed on by the faculty committee.²

² For both the statements of the student learning outcomes and the associated rubric, we relied heavily on the AACU QL VALUE Rubric: <https://www.aacu.org/value/rubrics/quantitative-literacy> (accessed Nov. 24, 2018). See also Boersma et al. 2011.

For instance, for the first outcome, “Interpret quantitative data,” the levels are:

- 1. Novice:** Attempts to explain information presented in mathematical formats, but draws incorrect conclusions about what the information means or uses incorrect terminology. For example, attempts to explain the trend of data shown in a graph, but misinterprets the nature of that trend, perhaps by confusing positive and negative trends or misinterpreting the scales used on the axes.
- 2. Emergent:** Provides somewhat accurate explanations of information presented in mathematical formats, but occasionally makes minor errors related to mathematical computations or units. Use of appropriate terminology is inconsistent. For example, accurately explains trend of data shown in a graph, but may miscalculate the slope of the trend line.
- 3. Proficient (our target):** Provides accurate explanations of information presented in mathematical formats. For example, accurately explains the trend of data shown in a graph.
- 4. Advanced (our dream):** Provides accurate explanations of information presented in mathematical formats. Makes appropriate inferences based on that information. For example, accurately explains the trend of data shown in a graph and makes reasonable predictions regarding what the data suggest about future events.

Then the instructor of each course identified at least one key assignment that gave students the opportunity to demonstrate achievement of multiple learning outcomes. In fact, most instructors chose assignments that would address all five QL outcomes. Here are some examples:

- **MATH 111A Mathematics of Everyday Life:** An in-class Excel lab on one-variable statistics, as described in our earlier paper (Mayfield and Dunham 2015). Typical topics include buying a home, looking at the NFL passer rating, and cracking the scratch lottery code.
- **MATH 112 Applied Statistics:** The final project. As we described in our first paper, students choose a topic, form a hypothesis, gather and analyze data (with Minitab), write a convincing report, and present their results to the class.
- **MATH 201 Calculus I:** A question on the final exam. Students use historical data to create, evaluate, and describe a model. In a supplemental file,³ we show how this assignment has in fact evolved from an exam question to an in-depth homework assignment to better reflect our learning outcomes.
- **SOC 261 Quantitative Methods for the Social Sciences:** A take-home final exam in which students select a research project provided in their textbook, run appropriate statistical analyses (with SPSS) with national representative data, and summarize their process in an abbreviated research paper format.

³ Text Appendix E in Supplemental File 2. The problem explores the fit of some basic modeling functions to a data set from D’Arcy Wentworth Thompson’s classic *On Growth and Form* (rev. 1942, now available from Dover Press; Thompson 1992), which pioneered the use of mathematics in biology (first edition, 1917).

Implementation: Chalk & Wire

For some years, students in the Education Department at our college have used the Chalk & Wire online portfolio software⁴ to submit assignments and create electronic portfolios of their work. The OIRA staff determined that this software could also be used effectively to collect and assess student work across the curriculum. In fact, the company now advertises on their website that they provide “learning assessment and credentialing solutions for the forward-thinking institution.” The College purchased an institutional license and provides students and faculty access to the software free of charge.

In its most basic form, the submission and assessment process works as follows:

- A faculty member creates the key assignment for a course and enters it into Chalk and Wire via the course management software Blackboard.⁵
- Students submit their work electronically, using Blackboard on a desktop computer or Microsoft OneDrive⁶ on an iPad, for instance.
- The faculty member is notified when work has been submitted and is available to be assessed.
- Using the QL rubric, the faculty member accesses the student work on Chalk and Wire and assigns a score of 1, 2, 3, or 4 for each rubric item. He or she may also leave comments for the student and may release the results to the student.
- The software collects all of the data and makes it available to both the faculty member and OIRA.
- Ideally, OIRA performs an analysis and provides it to appropriate faculty and staff.

Other options are available for faculty members who prefer not to use electronic submission of student work – when the key assignment is a question on a final exam, for instance. Faculty members can also collect assessment scores in an Excel template, and submit the results to OIRA for inclusion in Chalk & Wire.

Murphy’s Law: What Can Possibly Go Wrong? Especially in these early years when students and faculty are getting accustomed to this software, there are many, many things that can – and do – go wrong at each step. Faculty, students, Information Technology staff, and OIRA have faced many challenges in implementing this process. For example, Chalk & Wire requires students to essentially hit a submit button twice, but many students only complete the first step. They believe they have submitted the required assignment, only to be contacted later by their instructor informing them they have not. Instructors have

⁴ <http://www.chalkandwire.com/> (accessed Nov. 24, 2018)

⁵ <http://www.blackboard.com/index.html> (accessed Nov. 24, 2018)

⁶ <https://onedrive.live.com/about/en-us/> (accessed Nov. 24, 2018)

reported needing to use valuable class time to verify that student submissions were successful. Likewise, the multistep process of linking the posted assignment on Blackboard to the correct assessment group on Chalk & Wire proves challenging for faculty. Once things are set up and assignments submitted, faculty members find the Chalk & Wire interface quite confusing to navigate as they attempt to find the assignments waiting to be assessed. And, finally, both the Director of the College Office of Instructional Research and Assessment and its Director of Assessment left the institution during the period described in this article. Those losses obviously affected our ability to collect and interpret assessment data.

Results: What Are Students Learning?

In the pilot phase (Table 5) of our program, we collected data in Fall 2015 for:

- Mathematics of Daily Life (2 sections)
- Mathematics of Games and Sports
- Applied Statistics (3 sections)
- Calculus I (3 sections)
- Quantitative Methods for the Social Sciences.

The aggregate data for all students in all courses in Fall 2015 yielded “success” percentages shown in Table 6 for each outcome, where “success” on the outcome was meeting the learning outcome goal by achieving a rating of proficient or advanced on the outcome using the common rubric.

We noted that students were weakest in creating and communicating their arguments through the narrative analysis, which gave us something to work on. Instructors worked to address this deficit with more intentional instruction and additional assignments to develop their skills in creating arguments and communicating them in writing.

Some instructors also made changes to their assessment instruments. In Math 201, as noted above, students shifted from writing a short summary of a math problem to writing a more comprehensive “report” on their process. The report/assessment tool was also deployed earlier in the semester to avoid assessing students when they are exhausted and overwhelmed at the end of term.

Table 6
Aggregate Data for All QL Courses,
Fall 2015

QL Learning Outcome	n	Percentage of students who met goal*
Interpret Quantitative Data	143	81%
Demonstrate Computational Fluency	143	90
Create Arguments	143	76
Communicate Arguments: Tools	143	82
Communicate Arguments: Narrative	143	74

* Proficient or Advanced

Additionally, in MATH 111A, the instructor during AY 2017-18 moved toward using multiple assignments for assessment.

On the surface, our job seemed to be done! We met our goal in the very first semester (Table 6) – each learning goal was met by 70% or more of the students. If we look at data for all classes for the first few semesters of assessment (Table 7), things look even better.

But in fact we quickly realized that these aggregate data are pretty useless. If we look at a couple of individual courses, we see a very different picture.

In Table 8, for instance, are the scores for one section of MATH 111A Mathematics of Daily Life. Compare those data with the results from one section of MATH 112 Applied Statistics (Table 9), and it quickly becomes apparent that there is something else going on here.

We may in fact have a real difference in the way in which this assessment rubric is being used by different instructors. For an even more startling example, see Table 10, which displays scores from three different sections of MATH 112 from Spring 2018, taught by two different instructors, but using similar assignments.

Table 7
Aggregate Data for All QL Courses,
Fall 2015 – Fall 2017

QL Learning Outcome	n	Percentage of students who met goal*
Interpret Quantitative Data	513	82%
Demonstrate Computational Fluency	509	89
Create Arguments	509	79
Communicate Arguments: Tools	522	85
Communicate Arguments: Narrative	522	80

* Proficient or Advanced

Table 8
MATH 111A Mathematics of Daily Life,
Fall 2015

QL Learning Outcome	n	Percentage of students who met goal*
Interpret Quantitative Data	11	55%
Demonstrate Computational Fluency	11	82
Create Arguments	11	36
Communicate Arguments: Tools	11	36
Communicate Arguments: Narrative	11	55

* Proficient or Advanced

Table 9
MATH 112 Applied Statistics, Fall 2015

QL Learning Outcome	n	Percentage of students who met goal*
Interpret Quantitative Data	21	100%
Demonstrate Computational Fluency	21	100
Create Arguments	21	100
Communicate Arguments: Tools	21	100
Communicate Arguments: Narrative	21	95

* Proficient or Advanced

Table 10
MATH 112 Applied Statistics, Spring 2018

QL Learning Outcome	Percentage of students who met goal*	
	Instructor I (two sections) n = 32	Instructor 2 (one section) n = 22
Interpret Quantitative Data	91%	82%
Demonstrate Computational Fluency	94	50
Create Arguments	97	27
Communicate Arguments: Tools	100	82
Communicate Arguments: Narrative	94	32

* Proficient or Advanced

We have learned that assessing student work, even in areas involving computation, can be wildly subjective. Our next steps include training instructors to use the rubric and addressing issues with inter-rater reliability (Hallgren 2012, Saxton et al. 2012).

A Word about Attitudes and Confidence

In addition to OIRA's assessment of our five learning outcomes, we in the mathematics department were still curious as to whether the attitudes of possibly math-averse students could be improved by taking a QL course. In our earlier paper, we reported the (mostly inconclusive) results for our classes as a whole. Our reviewers suggested we might get more interesting results if we were able to track changes in individual students' responses over the course of the semester. And so in the Spring 2017 semester we administered the same attitude survey (Text Appendix D), this time electronically via Google Forms – and this time making an attempt to save identifying information with each response, while preserving the anonymity of the respondent. (We asked each student to type in her mother's middle name and the numerical day of her birthday, as in Jones17.)

Students apparently had a difficult time following these instructions: It appears that some students spelled a name one way at the beginning of the semester and another way at the end of the semester; a student might also interpret the phrase “middle name” in different ways on different days – middle name at birth? Current middle name as in “maiden name”? Something else? Some students filled out the survey only at the beginning of the semester; others filled it

out only at the *end* of the semester. We administered the survey to students enrolled in The Mathematics of Democracy, Calculus I, Statistical Concepts and Methods, and in four different sections of Applied Statistics. Of a total enrollment of 126 students in those seven classes, 89 students completed a pre-class survey – and in fact we suspect that only 85 of those responses represent distinct students. A mere 39 of those students also completed a post-class survey – but we somehow got 31 additional post-class surveys. Despite these glitches, and instead of attempting to apply a formal statistical analysis of the scant data, we present a snapshot for those 39 students.

If we assign a value of 1 to Strongly Disagree, 2 to Disagree, 3 to Agree, and 4 to Strongly Agree, and make the necessary adjustments for negatively-phrased statements, we can compute an average score for each student for both the pre-class and post-class surveys. The difference between those two averages is an indication of the student's change in confidence in and attitude towards mathematics: a positive difference corresponds to a positive change in attitude.

If we graph those values (Fig. 1), we see that most values (29 of 39) are within half a point of zero. There are slightly more positive changes (in score, and in confidence and attitude) than negative ones, but in general we really do not see much change.

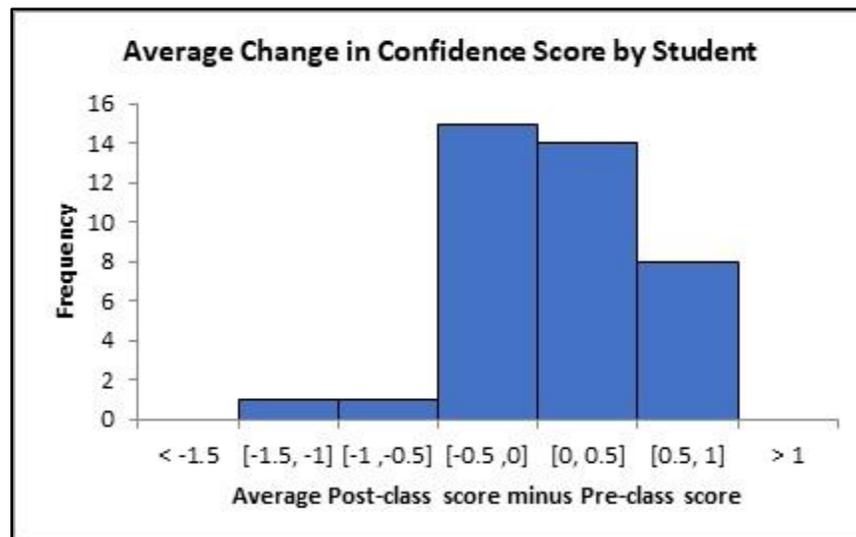


Figure 1. Differences in confidence scores over the course of a semester, Spring 2017, by student.

We can also look at pre- and post-class responses by question (Fig. 2). We see that most of the changes are small but positive, indicating an increase in confidence or attitude over the semester. The biggest increase was for Questions 1 (“Mathematics is very interesting to me, and I enjoy math courses.”) and 2 (“I feel

confident in my ability to complete math problems.”), which naturally encourages us.

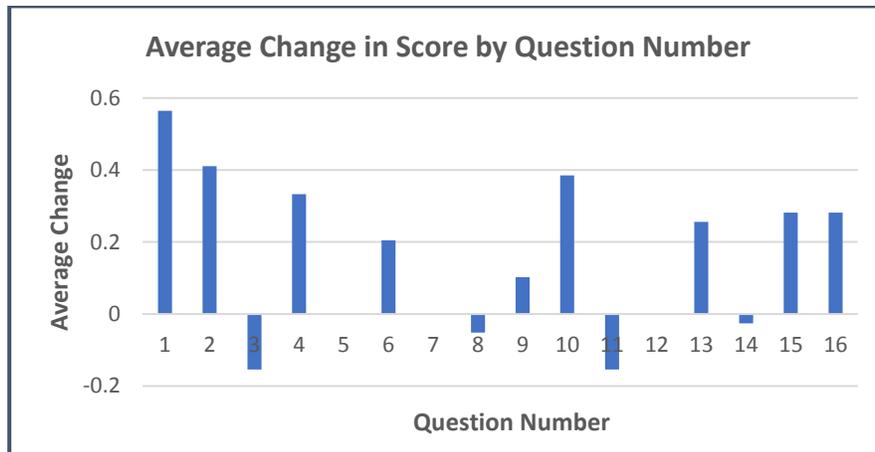


Figure 2. Differences in confidence scores over the course of a semester, Spring 2017, by question.

The four statements that suggested a slight *decrease* in confidence in or attitude towards mathematics were:

- I do not feel that I have a good understanding of the mathematics courses I have taken so far.
- I enjoy working in groups in class.
- If I work hard, I can succeed in math.
- In mathematics you can be creative and discover things for yourself.

The first of those statements depends on things over which we do not have much control; the second one reflects what we know about students’ resistance to group work. We would particularly like to see an increase in the score for the third and fourth statements, which gives us something to work on in the future.

In fact, we cannot draw many conclusions from the two times we have used this survey. As we mentioned in our first paper, it is difficult to draw meaningful conclusions about possible changes in student attitudes, especially after just one class. If we decide that we want to try to measure changes in student attitudes after taking one QL course, we will undoubtedly use a different instrument. A recent discussion on the National Numeracy Network Listserv has given us some good ideas for resources and rubrics regarding the affective dimensions of QR, including the Dartmouth College Mathematics Across the Curriculum Survey (similar to our survey but more comprehensive), the Select Numeracy Scale

(Fagerlin et al. 2007), and a student writing assignment described in a note in this journal (Ricchezza and Vacher 2017).

Current Discussions and Future Plans

Data Collection

In the future, we hope to address obstacles we have encountered with data collection. At the most basic level, it can be difficult to motivate some instructors to participate in the assessment process at all, for example, if the instructor has already decided not to return to the institution or if he/she is the lone instructor of a QL course in a department other than mathematics. But sometimes even well-intentioned efforts to collect data are not successful. The Fall 2017 data set (Table 7) does not include any data at all from MATH 112 Applied Statistics because of scheduling issues that affected the type of assignment being assessed. When the instructors ran out of time to assign the original assessment instrument as homework, they decided to include a version of the same material as a multi-part question on the in-class final exam. While this idea was appealing, it did not work in practice. Many students, including some who had otherwise been quite successful in the course, submitted either incomplete or nonsensical responses. In fact, we believe there are more reliable assessment instruments than an in-class final exam, such as papers and take-home assignments, with which to measure student learning (Plakans and Gebril 2015; Berkeley 2018).

We have also learned how important it is to save copies of the student artifacts, either electronically or on paper. In our previously mentioned scenario with Spring 2018 data from MATH 112 Applied Statistics, we had two instructors using similar instruments but with drastically different assessment results (Table 10). Because the work was submitted by the students on paper and returned to them without making copies, it is impossible to evaluate the artifacts now to determine if the issue is inter-rater reliability or something else.

The Rubric Itself and Choice of Key Assignment(s)

We have agreed that it is time to revisit the design of our rubric. Experienced instructors in several departments have reported challenges with differentiating among the three learning outcomes related to the creation and communication of arguments using quantitative tools and the narrative analysis. Users felt that it was difficult to separate their evaluation of student work into these distinct outcomes, because the processes were naturally overlapping.

Some Hood instructors have also expressed discomfort with using only one key assignment for assessment. One instructor decided to use multiple assignments for assessment, aiming for more of a portfolio design. This same

instructor has had some success experimenting with preliminary peer assessment of student work: after a student submits an assignment, two peers read it and offer advice. The student may then choose to re-submit the assignment (and about half of them do so).

We acknowledge that the selected key assignments in different departments and courses vary widely in terms of type, length and difficulty. More discussion is required to determine if this is in fact problematic.

Data Analysis

Once we have collected data, the big question is, of course, what do we do with it? The mathematics faculty met with the College's Faculty Assessment Liaison (a faculty member in a new administrative role) in Spring 2018 to document work currently being done with QL assessment. Attendees discussed the many issues associated with trying to make sense of the aggregate data, as illustrated above. There are many potential confounding variables, particularly when considering the entire data set. How should we calibrate for different grading styles from different instructors or the use of different assessment instruments? Do we assess learning outcomes in a 200-level statistics course for math and science majors the same as in a 100-level course for non-majors? Instructors agreed that from now on it will be most helpful to examine the data at the course level.

But, even if we solve these problems, questions still remain; for example, when the data appear to reflect a deficit at the course level, how does one address this finding effectively without knowing the cause? Is there really something that can be changed about how material is delivered in this particular course, or are there larger causes at the root of the problem? As we encounter difficulty upon difficulty, we must admit that the faculty at our small college have been fascinated by the recent series of articles (Gilbert 2016; Eubanks 2017; Gilbert 2018) and op-ed pieces (e.g., Worthen 2018), questioning the entire assessment process. Is there any reliability or validity to be found in this process? We must learn to deal with assessment fatigue and sinking morale.

Changes to Assessment of the Core Curriculum

Beginning in 2018-2019, assessment of the Core Curriculum will be overseen by the newly formed Core Curriculum Assessment Board. The board will be composed of one faculty coordinator from each of the twelve Core areas, the faculty assessment liaison, the assistant director of assessment, and the provost. The charge to the board is to examine "how well the Core Curriculum is meeting its stated purpose (Hood College Catalog) 'to provide students with the basic skills needed to pursue a liberal arts education, to expose them to a variety of modes of inquiry to different disciplines, and to promote critical reflection about global perspectives.'" The current plan is that the review schedule will be

organized in a two-year rotation, with three to four areas under review each semester. This schedule may shift to a three- or four-year rotation in the future. We anticipate that this board will be the natural place to discuss many of the questions about the assessment process that we have outlined in this paper, since we are sure these are not issues that are unique to assessment of quantitative literacy courses.

Conclusion: Lessons Learned

In this iteration of our assessment process, we have focused on individual student learning outcomes. We feel that we have made substantial progress since the first paper and are on the right track. However, we must take a step back and look carefully at the QL objectives, the rubric, and inter-rater reliability to refine our process before we can really tell if or what course modifications would be appropriate.

Our advice to others who may be contemplating a campus-wide QL assessment program, based on our experiences so far, follows. We realize that some or even all items on this list may be obvious and old hat to veterans of assessment experiences, but for us they were lessons learned or affirmed, and we pass them on to others who may find them helpful when planning or beginning a project such as ours.

- Expect to put a lot of thought and effort into assessment. Obtaining meaningful results involves much more work than the everyday grading process.
- Establish a relationship with your campus office of assessment. You share the same goal of student success, but you may need to learn a new language to effectively communicate.
- Work with other academic departments to establish a list of measurable student learning outcomes and associated rubrics. The mere act of discussing and setting up student learning outcomes has a positive effect on course organization.
- Use available resources such as the AACU VALUE Rubric. There is no need to reinvent the wheel.
- Be careful when writing learning outcomes. Unintentional overlap can make assessment more difficult than necessary.
- Before implementing your assessment plan, work with faculty to use best-practices to establish inter-rater reliability among assessors.
- Collect and save student artifacts, either electronically or on paper, at least until you are satisfied your plan is working as designed.
- Focus on results at the course level. Don't expect to get meaningful results from aggregate data.

Finally, we have benefited from becoming part of a larger QL community. This experience continues to be an opportunity for us to learn more about teaching and learning and for sharing our results with others.

Acknowledgments

The authors would like to acknowledge the assistance of our former colleagues Penny L. Weidner, now at Harrisburg University, and Jill Bigley Dunham, now at Chapman University, and all of the Hood College faculty and staff who have contributed to our QL assessment efforts.

References

- Berkeley Center for Teaching and Learning. 2018. "Alternatives to Traditional Testing." Accessed June 12, 2018. <https://teaching.berkeley.edu/resources/improve/alternatives-traditional-testing>.
- Boersma, S., and D. Klyve. 2013. "Measuring Habits of Mind: Toward a Promptless Instrument for Assessing Quantitative Literacy." *Numeracy* 6 (1): Article 6. Accessed April 9, 2014. <https://doi.org/10.5038/1936-4660.6.1.6>.
- Boersma, S., C. Diefenderfer, S. W. Dingman, and B. L. Madison. 2011. "Quantitative Reasoning in the Contemporary World, 3: Assessing Student Learning." *Numeracy* 4 (2): Article 8. Accessed November 24, 2018. <https://doi.org/10.5038/1936-4660.4.2.8>.
- Cuban, L. 2011. "Encouraging Progressive Pedagogy." In Steen 2001, 87-91.
- Eubanks, D. 2017. "A Guide for the Perplexed." *Intersection* Fall 2017 (4). Accessed August 1, 2018. https://c.ymcdn.com/sites/www.aalhe.org/resource/resmgr/docs/Int/AAHLE_Fall_2017_Intersection.pdf.
- Fagerlin, A., B. J. Zikmund-Fisher, P. A. Ubel, A. Jankovic, H. A. Derry and D.M. Smith. 2007. "Measuring Numeracy without a Math Test: Development of a Subjective Numeracy Scale." *Medical Decision Making* 27(5): 672–80. Accessed March 20, 2018. <https://doi.org/10.1177/0272989X07304449>.
- Gaze, E. C., A. Montgomery, S. Kilic-Bahi, D. Leoni, L. Misener, and C. Taylor. 2014. "Towards Developing a Quantitative Literacy/Reasoning Assessment Instrument." *Numeracy* 7 (2): Article 4. Accessed February 7, 2018. <https://doi.org/10.5038/1936-4660.7.2.4>.
- Gilbert, E. 2016. "Why Assessment Is a Waste of Time." *Inside Higher Ed*, Nov. 21 issue. Accessed August 1, 2018. <https://www.insidehighered.com/views/2016/11/21/how-assessment-falls-significantly-short-valid-research-essay>

- Gilbert, E. 2018. "An Insider's Take on Assessment: It May Be Worse than You Thought." *The Chronicle of Higher Education* 64 (21). Accessed August 1, 2018. <https://www.chronicle.com/article/An-Insider-s-Take-on/242235>.
- Gold, B. 2006. "Assessment of Developmental, Quantitative Literacy, and Precalculus Programs." In Steen 2006, 29-35.
- Grawe, N. 2011. "Beyond Math Skills: Measuring Quantitative Reasoning in Context." *New Directions for Institutional Research* 149: 41–52. Accessed April 15, 2014. <https://doi.org/10.1002/ir.379>.
- Hallgren, K. 2012. "Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial." *Tutor Quant Methods Psychol.* 8(1): 23–34. Accessed June 13, 2018. <https://doi.org/10.20982/tqmp.08.1.p023>.
- Hood College. 2017. *Hood College 2017-2018 catalog*. Frederick, MD. Available in electronic form at <http://hood.smartcatalogiq.com/en/2017-2018/Catalog>.
- Howard, J. 2013. "Teaching Quantitative Reasoning: What's Working at U-M?" University of Michigan Center for Research on Learning and Teaching blog. Accessed May 26, 2015. <http://www.crlt.umich.edu/node/1010>.
- Mayfield, B., and J. Dunham. 2015. "Adapting to a New Core Curriculum at Hood College: From Computation to Quantitative Literacy." *Numeracy* 8 (2): Article 8. Accessed February 7, 2018. <http://dx.doi.org/10.5038/1936-4660.8.2.8>.
- Office of Institutional Research and Assessment, Hood College. 2014. Core curriculum and college competencies: A four-year assessment plan 2014-2018. Accessed March 8, 2018. Available in electronic format at http://www.hood.edu/uploadedFiles/Hood_College/Home/Academics/Office_of_Institutional_Research/Academic_Assessment/Hood%20College%20Core%20Curriculum%20and%20College%20Competencies%20Plan%20Oct%2016.pdf.
- OIRA. See Office of Institutional Research and Assessment. 2014.
- Plakans, L., and A. Gebriel. 2015. "Myth 2: A Comprehensive Final Exam Is the Best Way to Evaluate Students." *Assessment Myths: Applying Second Language Research to Classroom Teaching*. Ann Arbor: University of Michigan Press. Accessed June 13, 2018. <https://doi.org/10.3998/mpub.5056216>.
- Ricchezza, V. J., and H. L. Vacher. 2017. "Quantitative Literacy in the Affective Domain: Computational Geology Students' Reactions to Devlin's *The Math Instinct*." *Numeracy* 10 (2), Article 11. Accessed March 20, 2018. <http://doi.org/10.5038/1936-4660.10.2.11>.
- Saxton, E., S. Belanger, and W. Becker. 2012. "The Critical Thinking Analytic Rubric (CTAR): Investigating Intra-Rater and Inter-Rater Reliability of a Scoring Mechanism for Critical Thinking Performance Assessments."

- Assessing Writing* 17(4): 251-270. Accessed June 13, 2018.
<https://doi.org/10.1016/j.asw.2012.07.002>.
- Sikorskii, A., V. Melfi, D. Gilliland, J. Kaplan, and S. Ahn. 2011. "Quantitative Literacy at Michigan State University, 1: Development and Initial Evaluation of the Assessment." *Numeracy* 4 (2): Article 5. Accessed February 7, 2018.
<https://doi.org/10.5038/1936-4660.4.2.5>.
- Steen, L.A., ed. 2001. *Mathematics and Democracy: The Case for Quantitative Literacy*. Princeton, NJ: The National Council on Education and the Disciplines.
- Steen, L.A., ed. 2006. *Supporting Assessment in Undergraduate Mathematics*. Washington DC: The Mathematical Association of America. Accessed February 7, 2018. <http://www.maa.org/publications/ebooks/supporting-assessment-in-undergraduate-mathematics>.
- Thompson, D.W. 1992. *On Growth and Form: The Complete Revised Edition*. New York: Dover Publications. <https://doi.org/10.1017/CBO9781107325852>.
- Vacher, H. L. 2015. "Educational Assessment Is an Enduring Theme of *Numeracy*." *Numeracy* 8(1): Article 1. Accessed February 7, 2018.
<https://doi.org/10.5038/1936-4660.8.1.1>.
- Vacher, H. L., and D. Wallace. 2013. "The Scope of *Numeracy* after Five Years." *Numeracy* 6 (1): Article 1. Accessed February 7, 2018.
<https://doi.org/10.5038/1936-4660.6.1.1>.
- Ward, R. M., M. C. Schneider, and J. D. Kiper. 2011. "Development of an Assessment of Quantitative Literacy at Miami University." *Numeracy* 4 (2): Article 4. Accessed March 15, 2014. <https://doi.org/10.5038/1936-4660.4.2.4>.
- Worthen, M. 2018. "The Misguided Drive to Measure 'Learning Outcomes.'" *New York Times*, Feb. 23, 2018. Accessed August 1, 2018.
<https://www.nytimes.com/2018/02/23/opinion/sunday/colleges-measure-learning-outcomes.html>.
- Wright, M.C., and J.E. Howard. 2015. "Assessment for improvement: Two Models for Assessing a Large Quantitative Reasoning Requirement." *Numeracy* 8 (1): Article 6. Accessed February 7, 2018.
<https://doi.org/10.5038/1936-4660.8.1.6>.