Volume 8 | Issue 1   Article 6

2015

Assessment for Improvement: Two Models for Assessing a Large Quantitative Reasoning Requirement

Mary C. Wright
University of Michigan, mcwright@umich.edu
Joseph E. Howard
University of Michigan, josephehoward@gmail.com

Assessment for Improvement: Two Models for Assessing a Large Quantitative Reasoning Requirement

Abstract

We present two models for assessment of a large and diverse quantitative reasoning (QR) requirement at the University of Michigan. These approaches address two key challenges in assessment: (1) dissemination of findings for curricular improvement and (2) resource constraints associated with measurement of large programs. Approaches we present for data collection include convergent validation of self-report surveys, as well as use of mixed methods and learning analytics. Strategies we present for dissemination of findings include meetings with instructors to share data and best practices, sharing of results through social media, and use of easily accessible dashboards. These assessment approaches may be of particular interest to universities with large numbers of students engaging in a QR experience, projects that involve multiple courses with diverse instructional goals, or those who wish to promote evidence-based curricular improvement.

Keywords

assessment, quantitative reasoning requirement, learning analytics

Creative Commons License

Cover Page Footnote

Mary C. Wright is Director of Assessment and an Associate Research Scientist at the University of Michigan's Center for Research on Learning and Teaching (CRLT). Her research interests include teaching cultures, evaluation of engaged and experiential learning initiatives, and educational development to promote scholarly approaches to teaching, such as learning analytics and the Scholarship of Teaching and Learning (SoTL).

Joseph E. Howard is a doctoral candidate with the University of Michigan's Center for the Study of Postsecondary and Higher Education. His research interests focus on issues of organizational assessment and improvement in college and university settings, particularly in the areas of admissions, financial aid, and student affairs.

Introduction

Advanced math skills are required by more college majors than ever, prompting universities to explore ways to use general education requirements to build competencies needed for success in the major, in college, and on the job market. A survey of American Association of Colleges and Universities (AAC&U) member institutions indicates that most (71%) colleges and universities include quantitative reasoning as one of their central learning outcomes (Hart Research Associates 2009). But can these classes effectively build student skills? Two evaluations of the quantitative reasoning requirement at the University of Michigan's College of Literature, Science and the Arts (LSA) are presented here. These evaluation models address two key challenges in quantitative reasoning (QR) assessment: (1) encouragement of use of the study results for curricular improvement, and (2) resource constraints associated with measurement of large programs.

One common challenge of assessment is that results typically are not disseminated or utilized for curricular change (Banta 2007; Wehlberg 2008). Therefore, a critical step of assessment includes "closing the loop," or implementing changes suggested by data collection (Welsh and Metcalf 2003a, b). The assessment models we present here were undertaken by staff in the university's teaching and learning center; collaboration with such offices can help encourage use of data for institutional change (Cook et al. 2011). We present three approaches we used to share data to promote curricular improvement: meetings with instructors to share findings and disseminate best practices, use of social media, and development of a learning analytics-based dashboard tool.

A second key challenge – noted frequently in QR assessment – involves resource constraints, balanced with the need to collect valid and reliable data. For those conducting program assessment, there is a clear tension between the desire to have the validity of direct measures of learning and the very real time and financial constraints that suggest indirect measures, or student self-reports. Some QR assessment models include direct measures, or demonstrations of student learning, which can be very beneficial given possible shortcomings in students' estimates of their own learning (Chun 2002; Porter 2013; Ehrlinger and Shain 2014). Examples of these include students' analysis of a quantitatively oriented newspaper article in a single course, as well as James Madison University's Quantitative Reasoning Test, administered on specially designated assessment dates (Grawe 2012; Boersma and Klyve 2013; see Hathcoat et al. this issue). However, others have noted that most examples of assessments using direct

1

measures are time-consuming and resource-intensive, perhaps explaining why key studies are located at small colleges or universities. Some have argued, "Possibly in contrast to smaller institutions, midsize or larger universities need an assessment of QL [Quantitative Literacy] that is not time- or resource-intensive" (Ward et al. 2011).

Given current resource constraints facing all institutions of higher education, however, there is a clear need for cost-effective assessment processes, such as those using indirect measures, albeit with prior testing to validate instruments and give greater confidence in the data. Indeed, at many large research universities, university-wide surveys are one of the most frequently used sources of assessment data (Kuh and Ikenberry 2009). Clearly, use of direct measures for QR assessment is preferable. However, what options may be available to those who must utilize an indirect approach, for reasons that may range from time and resource constraints to lack of knowledge about other measures?

Here, we present two approaches for assessment of a QR requirement that primarily use indirect measures, but not blindly. Instead, they first validate the indirect assessment by examining its convergent validity with direct measures of learning. Convergent validity examines the alignment between two measures, or, as Pike (2011: 43-44) explains,

> Self-report measures [should] be related to other measures that theory posits are related to the construct. For example, if theory states that certain programs of study (for example, English) should be related to acquisition of certain skills (writing), then English majors' reports of their writing abilities should be greater than the reported writing abilities of students in other majors.

There are certainly legitimate criticisms of the use of student self-reports for assessment, as well as criticisms of convergent validation (see for example, debates between Porter 2013 and Pike 2013). However, in general, prior validation of any survey can greatly increase confidence in resulting data.

The first study described below evaluates a University of Michigan QR requirement using a tailored version of a survey that analyzed skills related to quantitative reasoning. The second study describes a more efficient approach to evaluate the same requirement, making use of QR-related items on an existing, annually administered, university-wide survey. Despite differing methodologies, both studies achieved relatively consistent results. We present both options to offer a greater range of approaches for campuses to consider in assessing QR and disseminating the findings for course and curricular improvement. (In both cases, human subjects approval was granted, permitting us to publish findings.) These assessment approaches may be of particular interest to universities with large numbers of students engaging in a QR experience, projects that involve multiple courses with diverse instructional goals, or those seeking resource-efficient ways to support evidence-based curriculum design.

Quantitative Reasoning Requirement

In Fall 2009, the University of Michigan's teaching center, the Center for Research on Learning and Teaching, collaborated with LSA to evaluate its QR requirement. LSA is the University of Michigan's largest academic unit, with nearly 20,000 active undergraduates. According to *The LSA Bulletin*, "The goal of the Quantitative Reasoning requirement is to ensure that every graduate of the College achieves a certain level of proficiency in using and analyzing quantitative information" (College of LSA 2009).

Unlike quantitative literacy requirements at some other universities, all LSA students must complete a QR course (i.e., students cannot "test out" or use Advanced Placement credit); in very limited cases, they may use transfer credit. Currently, students have the option of fulfilling the distribution requirement by taking one of many different possible QR classes situated in 13 disciplines. Courses fulfilling the requirement include mathematics courses such as *Calculus*, as well as courses from a wide range of other disciplines, such as *Social Systems, Energy and Public Policy* (Program in the Environment); *From Quarks to Cosmos* (Physics), *Introduction to Symbolic Logic* (Philosophy); and *Quantitative Analysis and Writing in the Disciplines* (English). With many course options, students are enabled to engage in "multiple courses of study," which ideally will "develop quantitative strengths keyed to their actual interests," consistent with the recommendations made by AAC&U president Carol Geary Schneider (2004).

Students have a further option for fulfilling the requirement by passing one course designated for QR credit (QR1) or two courses designed for half-credit (QR/2). The LSA Curriculum Committee makes decisions about which courses may fulfill QR1 or QR/2. In the 2014 academic year, 52 courses were available to fulfill QR1 and 37 to fulfill QR/2. However, the vast majority (95%) of students fulfill the requirement through the QR1 option. Calculus I (28%) and Introduction to Statistics (23%) are the most popular options for doing so.

Because students cannot test out of the QR requirement, this assessment avoids most selection issues that might confound findings. For instance, other explorations may involve study settings where "students are usually assigned to participate on the basis of attributes (like baseline test scores) correlated with lower performance…[but] such comparisons are almost preordained to yield disappointing and biased assessments of program effects" (Butcher et al. 2010: 188). In this manner, we are able to evaluate the assessment on students with a diverse range of quantitative backgrounds. However, a possible limitation of the methodology is that it does not involve random assignment, because students were able to choose their approach for fulfilling the requirement, as well as the timing at which they did so. It should be noted that most general education requirements offer considerable course selection choice.

3

2009 Assessment: Tailored Survey

In Fall 2009, the first author conducted an assessment at the request of LSA's Associate Dean for Undergraduate Education, as part of an initiative to gather more systematic data on all of the College's key general education requirements. The assessment focuses primarily on QR1 courses because the vast majority of students fulfill the requirement in this manner.

Study 1 Research Questions

In collaboration with the Associate Dean (Robert Megginson), some of the questions[1] that this evaluation of the QR requirement sought to answer were:

1. Do LSA students perceive that QR1 courses contribute to gains in their quantitative reasoning skills?

    a. Are there statistically significant differences between students who take QR1 courses and those who do not (overall and by specific skill)?

    b. Do these reported learning outcomes differ by subgroups of students, e.g., by gender, race/ethnicity, incoming math proficiency, class rank, and discipline of course taken?

2. Do students who took a QR1 course report that their course met the goals of the LSA Quantitative Reasoning requirement?

3. Are QR1 takers able to offer an example of how they are able to transfer analytical skills learned in their QR course to another context?

4. What instructional activities do students find useful for the development of QR skills?

Study 1 Methods

Pilot instrument. A custom survey was developed to evaluate the requirement, which –in alignment with recommendations about QR assessment instruments – included mixed multiple-choice items with open-ended prompts (Grawe 2011). LSA first- and second-year students were asked to reflect on gains in their quantitative reasoning ability in either their QR1 (sample of interest) or Fall Term (control sample) coursework. Most of the survey was derived from a University of Wisconsin assessment study of one of its QR requirements (Halaby 2005). The

---

[1] A full list of questions and a summary of findings can be found here: http://www.crlt.umich.edu/assessment/lsaqrassessment. (This and all other links in this paper were accessed November 26, 2014.)

Wisconsin study used two methods: (a) pre- and post-tests of intermediate and college algebra problems and (b) a post-only survey of student self-reports about learning gains on 14 different QR skills. These skills are ones promoted by multiple scholars of Quantitative Reasoning and Literacy (e.g., Dwyer et al. 2003; Wolfe 1993), including using quantitative information to evaluate an argument and expressing ideas using quantitative information.[2] In addition to assessment of student learning, Halaby (p. 24) notes that a key objective of his study is to understand if the results can be corroborated, noting "self-report data collected through surveys, … are less expensive to mount than laboratory tests and allow for the measurement of a broader range of assessment-relevant outcomes." Indeed, the study does find an alignment between the two measures, concluding that "the findings of study 2, though based exclusively on student self-reports rather than on laboratory testing, nicely corroborate the conclusions of study 1 even as they identify the limits of the [QR course] effect" (p. 38).

Given the large number of students taking U-M's QR requirement and the diversity of course objectives, it was decided relatively early that a survey-based measurement of student learning was the most feasible option for conducting the study in a resource-efficient manner. The Wisconsin survey was attractive because it was validated through direct measures of student learning (i.e., convergent validation), could be applied to a variety of course contexts, and was tested on a population relatively similar to Michigan's.

In addition to the 14 gains items from the Wisconsin assessment, we added several open-ended prompts to mix multiple-choice and open-ended survey formats, as recommended in Grawe (2012). An integral piece of QR learning involves transfer of knowledge and analytical skills to real-world contexts outside of the classroom (Rhodes 2010). Therefore, the U-M survey also asked students, "Are you able to describe an example of the way in which you leave [course] being BETTER able to use and analyze quantitative information?" and if so, to provide an example.

A second added question pertained to instructional methods. Challenges in teaching QR are many and include addressing the diverse range of mathematical preparation of individual students, assuaging student anxieties, teaching technical language, and incorporating QR processes in humanities and other traditionally non-quantitative disciplines (Bahls 2012; Baloglu 2003; Lutsky 2008; Latiolais and Laurence 2009; Moore et al. 2010). Therefore, students were invited to note the instructional activities they found beneficial for their learning, choosing from a number of options that were derived from the literature on student learning of QR skills. Finally, the U-M survey asked students if they felt that the course met LSA's goals for the QR requirement and why they took the course.

---

[2] See Appendix 1 (additional files) for the full list of 14 items.

The survey was pilot tested on U-M students in six QR1 courses, as well as a non-QR course (first-year writing), in Spring/Summer 2009. The full instrument used for the pilot is presented in Appendix 1 in the additional files material for this article. The seven courses had a total enrollment of 384 students, and the response rate for the pilot was 54%.

Pilot results. The 14 QR skills items had a Cronbach's alpha of 0.95, indicating the high consistency of the items in measuring the same underlying concept.

When asked to give their feedback about the survey, students reacted very favorably. To gather feedback about the survey, we asked four Likert-scale questions at the end of the instrument:

- I was able to understand the survey directions.

- I was able to understand the survey questions.

- The survey's length was appropriate.

- I understood how this survey explained LSA's goals for the Quantitative Reasoning requirement.

Each of these items received a mean response over four (range of 4.1 to 4.5), with 1=strongly disagree and 5=strongly agree. Additionally to gain additional feedback, we asked respondents the following open-ended question: *"When I think about my experience with the LSA Quantitative Reasoning requirement, what I really wish this survey had asked me about is..."* The most common suggestions for questions were: (a) How well was the QR course taught? (noted by 28 students) and (b) Is there a need for a QR requirement and if so, what should its goals be? (14 students).

Study 1: Instrument. In response to the changes suggested by the pilot, small changes were made to the QR assessment instrument. Because other data sources (e.g., student ratings) provide the College with student feedback about teaching, the evaluators felt it unnecessary to add another question about how well students thought the course was taught. However, to address the need for a QR requirement, another question was added to the fall instrument that asked students if they would take a QR course even without a requirement in place.

A second instrument, for distribution to a comparison group of non-QR takers, was created for the main study. This instrument consisted of only the 14 QR skills items, with a slightly modified prompt. (The wording was modified to read, "Compared to when I started the Fall Term, I am now BETTER able to…").[3]

---

[3] Both full instruments can be found in Appendix 1 (additional files).

Study 1: Methods. In Fall 2009, all first- and second-year LSA students who were taking a QR1 course for the first time (N=2,624) were surveyed at the end of the term. Additionally, a parallel survey was distributed to a control group of 900 first- and second-year students who had never taken a QR1 course. To boost response rates, we mailed an advance letter from the LSA Associate Dean to the entire sample, noting the importance of the study. We also offered significant incentives to respondents (a chance at one of six $100 gift cards). The size of the incentive was chosen based on findings presented in Porter and Whitcomb (2004) that suggested that this was the most effective prize amount for undergraduate surveys.

Both groups of students were asked, "Compared to when I started ['QR course' or 'the Fall Term' for control group], I am now BETTER able to: …" in regard to 14 different skills, which were aggregated into an overall mean QR score (1=Strongly Disagree to 5=Strongly Agree, N/A). Higher values of the QR score suggest greater self-reported acquisition of QR skills.

Study 1 Results

The Fall Term assessment was slightly revised from the pilot survey, and therefore, levels of internal consistency were again measured. The Cronbach's alpha was computed for the 14 gain items, and the value was again found to be 0.95.

Response rates. The final survey was distributed to 2,624 QR takers and 1,419 responded, a 54% return rate, which has generally been found to be adequate for reporting. For the control group, 384 students replied, a 43% response rate. For both the intervention and control group, the survey was first distributed on December 23, 2009. (This date was chosen because it was the first date after exams were finished and online student ratings forms were closed, to avoid any potential conflicts between E&E surveys and student perception that study participation was linked to grades.) Two follow-up attempts were made to non-respondents, with a close date of January 5, 2010.

In both the QR and control groups, women composed the majority of respondents (59% and 68%, respectively), and non-underrepresented minorities also made up most respondents (83% and 86%). Unfortunately, it was not possible to uniformly compare students by SAT or ACT math, because applicants have the option of taking one or the other.[4] However, most respondents (84%) had an ACT math score. As noted above, a majority of students fulfill their QR

---

[4] Concordances exist to convert ACT to SAT scores (and *vice versa*) (e.g., Dorans, 1999). However, personal correspondence with ACT at the time indicated that no official concordances were published to reflect newer updated standardized test formats (email correspondence with ACT, January 10, 2010).

7

requirement in their first year. Therefore, it is appropriate that in both the QR and control groups, the vast majority of the survey respondents were first-year students (75% QR, 72% control).

Comparing respondents and non-respondents in both the QR and control samples, there were no statistically significant differences by gender nor underrepresented minority status. However, a potential source of bias is that survey takers had slightly higher ACT math scores than non-respondents, by a margin of 0.59 ($t=3.94$, $p<.001$). Additionally, among respondents, QR takers' mean scores (29.11) were nearly a point higher on their ACT Math than the control (28.43) at a statistically significant level ($t=2.77$, $p<.01$). Although we did not have significant selection effect issues due to enforcement of the requirement (i.e., no test-out opportunities), differences on ACT math scores suggest that a limitation of this study is that there may be residual selection issues. Students with lower ACT math scores may choose to delay fulfillment of their requirement. Implications of this potential bias are discussed further below.

Additionally, there was a slightly higher proportion of sophomores among respondents (25%) than non-respondents (21%) ($\chi^2(1, N=3527)=9.6$, $p<.01$). Because later analyses indicate that sophomore status is a statistically significant predictor of students' reported quantitative reasoning gains, it may be that the estimates of QR gains derived by the survey are upwardly biased. However, given the small differences in proportions (25% sophomores in the survey and 21% among non-respondents), it is likely that this effect is minimal. More importantly, among survey respondents, there is no significant difference in class rank between QR1 takers and the control group. Therefore, findings are less likely to be influenced by class rank in a manner that could completely account for any observed effects of QR enrollment.

Research Question #1a: Do LSA students perceive that QR1 courses contribute to gains in their quantitative reasoning skills? Are there statistically significant differences between students who take QR1 courses and those who do not (overall and by specific skill)?

For QR1 takers, the overall mean QR score was 3.52 (SD=0.82), indicating that these students tend to agree that they gained key quantitative reasoning skills in their Fall Term QR courses (scale is 1=strongly disagree to 5=strongly agree). QR takers reported slightly higher gains (+.10) than the control group, a small yet statistically significant difference ($t=2.38$, $p=.018$, Cohen's d=0.13). Although a small effect, we anticipated that there would be a relatively modest impact from just one required course taken by students with heterogeneous backgrounds, and so this result actually was confirmatory of small, but positive benefits.

On individual items, QR1 students scored significantly higher than the control group on seven of the 14 measures of QR skills (Table 1). Non-QR students rated

their gains significantly higher on three items (use of evidence, logically sound arguments, inference of causation) and there was no statistically significant difference between the two groups on four items.

Table 1. Comparison of QR and Control Group on 14 QR Items

	QR Mean	(SD)	Control Mean	(SD)	Mean Dif.	t	(S.E)	p-value	Cohen's d
Q1. Recognize logically sound arguments	3.41	(1.01)	3.84	(0.73)	-.43	-9.26	(.047)	***	.49
Q2. Use quantitative information to evaluate an argument	3.57	(1.00)	3.57	(0.83)	.00	0.01	(.051)		.00
Q3. Understand the difference between correlation and causation	3.16	(1.14)	3.27	(1.03)	-.11	-1.67	(.067)		.10
Q4. Solve problems using arithmetic, algebra or statistics	3.83	(1.00)	2.89	(1.01)	.95	15.24	(.062)	***	.94
Q5. Understand randomness, uncertainty and risk	3.18	(1.13)	3.27	(0.98)	-.09	-1.51	(.062)		.09
Q6. Use statistics to evaluate factual claims	3.22	(1.12)	3.08	(1.02)	.14	2.14	(.064)	*	.13
Q7. Understand charts and graphs showing quantitative information	3.70	(1.04)	3.25	(1.01)	.45	7.11	(.063)	***	.44
Q8. Express ideas using quantitative information	3.69	(0.99)	3.35	(1.00)	.34	5.75	(.059)	***	.34
Q9. Recognize when arguments use evidence well	3.49	(1.06)	4.03	(0.78)	-.54	10.84	(.050)	***	.58
Q10. Know when it is valid to infer that one thing causes another	3.41	(1.09)	3.72	(0.91)	-.31	-5.39	(.057)	***	.30
Q11. Understand rates and percentages	3.54	(1.05)	3.09	(1.02)	.45	7.13	(.063)	***	.43
Q12. Understand how data can be used to test a hypothesis	3.43	(1.12)	3.35	(1.02)	.08	1.32	(.063)		.08
Q13. Use quantitative information to solve problems	3.79	(0.99)	3.24	(1.01)	.55	9.14	(.061)	***	.55
Q14. Solve problems using formal logic	3.62	(1.07)	3.48	(1.02)	.14	2.18	(.063)	*	.13

Note: All items use a 5-point scale (1=Strongly Disagree…5=Strongly Agree).
*$p<0.05$; **$p<0.01$; ***$p<0.001$*

Research Question #1b: Do these reported learning outcomes differ by subgroups of students, e.g., by gender, race/ethnicity, incoming math proficiency, class rank, and discipline of course taken?

When we controlled for QR course discipline, we saw relatively consistent gains across departments, suggesting that the menu of choices offered to students

9

presented them with essentially similar outcomes, albeit in different disciplinary contexts.

To control for multiple background factors, a multivariate regression also was performed, with ACT math, gender, race/ethnicity, class rank, QR1 enrollment (first-time enrollment in at least one QR1 course) and QR/2 enrollment (enrollment in at least one QR/2 class) regressed on the overall mean QR score. For the regression, we used block entry, with simultaneous forced entry of all variables. Taking the independent variables separately, enrollment in a QR1 course is a significant predictor of the overall mean QR score, even when controlling for gender, race/ethnicity, class rank, and math proficiency (Table 2). Taking at least one QR1 course associates with a small (0.14 or 4%) increase on the five-point mean overall QR score, and it is the strongest predictor of students' reported quantitative reasoning gains.

Table 2.

Prediction of Overall Mean QR Scores: Descriptives and Multiple Regression Analysis Results (N=1412)

	Descriptives		B	S.E. of B	Beta (Standardized B)	t-statistic	Sig.
Enrolled in QR1?	Not enrolled (coded 0) = 22.4%	Enrolled (coded 0) = 77.6%	0.14	0.050	0.076	2.82	**
Enrolled in QR/2?	Not enrolled (coded 0) = 11.0%	Enrolled (coded 0) = 89.0%	0.10	0.049	0.055	2.06	*
Female Status?	Male (coded 0) = 38.9%	Female (coded 1) = 61.1%	-0.06	0.044	-0.039	-1.434	
Underrepresented Minority Status?	Non-URM (coded 0) = 9.1%	URM (coded 1) = 10.9%	0.11	0.072	0.043	1.508	
Sophomore Status?	1st-Year (coded 0) = 74.2%	Soph. (coded 1) = 5.8%	0.13	0.049	0.070	2.594	*
ACT Math Score	Mean = 28.96 Std Dev = 4.0		-0.01	0.006	-0.028	-0.961	
(Constant)			3.50	0.178		19.63	***

* $p<0.05$ ** $p<0.01$ *** $p<0.001$

ACT math was not statistically significant, suggesting that after controlling for gender, URM status, and sophomore status, prior math proficiency was not

significant associated with students' reported learning gains. Because survey participation was positively linked to ACT math, the lack of a statistically significant finding in this analysis is important, suggesting that any potential bias from differences in math proficiency test scoring would have a limited effect on self-reported QR gains. Sophomore status was significant, and among QR1 takers, sophomores did have a significantly higher QR score (mean=3.66), compared to first-year students (mean=3.47) (t = 3.72, p<.001). Most strikingly, gender and underrepresented minority status were not statistically significant, controlling for other covariates.

Research Question #2: Do students who took a QR1 course report that their course met the goals of the LSA Quantitative Reasoning requirement?

When presented with the goals of the requirement, the vast majority (76%) of students agreed or strongly agreed that their course fulfilled these aims, with very little difference between the departments offering courses.

Research Question #3: Are QR1 takers able to offer an example of how they are able to transfer analytical skills learned in their QR course to another context?

The assessment suggested two areas of concern, first with application. Students were asked, "Are you able to describe an example of the way in which you leave [course] being BETTER able to use and analyze quantitative information?" and if so, to provide an example. A third of respondents (33%) responded, "no," or that they were not able to offer such an application. Less than half (45%) indicated they could provide an example, and only 42% of all respondents actually gave one. (22% of students skipped the question.)

In examples given, there was some variation by discipline of QR course, but common themes included evaluation of research claims, development of problem-solving skills, and transfer of learning to other courses. For example, one calculus student noted,

> One aspect of the course that I enjoyed was that there was an emphasis on word problems that used real life examples that students can relate to real life instances where using calculus could be used. For instance, a couple of friends and I used optimization to find out the maximum number of rooms that could fit into a dorm if you needed to have both single and double dorm rooms.

Likewise, a student in physics indicated, "[The class was applicable to] anything from hanging a picture frame to hitting a billiards ball. After this course, I am able to be more successful and efficient in many everyday tasks." Although many students gave very clever applications like these, the relative lack of response to this prompt – and the fact that a third of respondents indicated that they could not provide an application – was concerning.

11

Research Question #4: What instructional activities do faculty and students find useful for the development of QR skills?

A second concern identified by the study was possible gaps in instructional approaches. Students who had just taken a QR course were asked to identify activities that they perceived contributed to their being able to use and analyze quantitative information. Learning activities that students found particularly helpful were preparing for exams (selected by 77%), completing regular problem sets (64%), reviewing notes (63%), listening to lectures (62%), and listening to Graduate Student Instructor (TA) explanations (54%) (Table 3).

Table 3.
Activities Contributing to Use of Quantitative Information

Instructional Activity	% of Respondents
Studying Course Material	
Preparing for exams	77%
Reviewing your notes	63%
Listening to lectures	62%
Interaction with Instructors	
Listening to explanations from your GSI in section	54%
Receiving feedback on your exams/quizzes	49%
Asking questions in office hours	35%
Receiving feedback on your homework	34%
Homework and Activities	
Completing regular problem sets	64%
Participating in group work in class	41%
Being part of a homework group outside of class	38%
Solving real-world problems	35%

Note: Students could check all that apply.

Given the importance of applied problem-solving in the literature for teaching quantitative reasoning (Atkinson et al. 2006; Bok 2006), it is notable that only about a third (35%) of respondents pointed to "solving real-world problems" as an instructional technique that helped them. It is unclear whether these responses stem from this approach being infrequently used in U-M classrooms, or from students feeling that the use of the strategy is unhelpful to their learning. Either interpretation raises concerns.

12

Study 1 Dissemination of Results

After data analysis, findings were distributed to each department in person, which involved a 45-60-minute meeting with QR course instructors and department administrators. During the resulting follow-up conversations, the first author described the strengths and challenges observed in the department's data, providing a tailored report for each meeting. She asked instructors for their own interpretations of reasons for findings and took notes on the instructional strategies used by instructors in the department. This was particularly helpful for learning about best practices, especially for identifying ways that some instructors were effectively teaching real-world problem solving. For example, a statistics instructor described how she uses highly structured note-taking tools – a course pack with "Try It!" problems embedded throughout lecture notes – to promote problem-solving skills in a 1700-person course. Follow-up interviews with other U-M instructors teaching QR courses supplemented these initial conversations, and these ideas were disseminated in a blogpost,[5] which has received over 500 hits to date (51 local hits). It was also distributed more broadly through a national education listserv (*Tomorrow's Professor).*

Study 1 Discussion

Compared to non-QR takers, students taking a QR course reported small statistically significant gains in their estimated quantitative reasoning skills. Generally, the results of the QR assessment were confirmatory of the relatively modest associations we expected one required course to have.

One of the most striking findings was a lack of statistical significance for gender, controlling for other covariates. This finding is in contrast to other assessments that have found differences by gender on quantitative literacy for entering students (Sikorski et al. 2011) or lower exposure to QR activities among female students (Rocconi et al. 2013). One possible reason for this difference is that math students composed a large proportion (73%) of the QR1 sample, and a large proportion of math courses at the University use Inquiry Based Learning (IBL) methods. Other research has found that men and women in IBL math classrooms report similar learning gains (Laursen et al. 2011).

Another interesting finding was that in analysis of the 14 QR skills making up the composite measurement, non-QR takers rated their gains higher on three items. These three skills – use of evidence, logically sound arguments, inference of causation – appear to be most applicable to general analytical reasoning capacities, rather than quantitative reasoning. Interestingly, these findings are consistent with the Wisconsin study, as the non-QR group of that study also scored higher than the QR sample on the same three items (Halaby 2005).

---

[5] Available: http://www.crlt.umich.edu/node/1010

13

Additionally, as noted by Grawe (2012: 48), it seems likely that professors outside of QR-identified courses would also be teaching these skills, given the improbability that only faculty in a small subset of courses would represent the "full range of context-rich practice such as QR."

Although we did not have significant selection-effect issues due to enforcement of the requirement (i.e., no test-out opportunities), one possible limitation of this study is that there may be residual selection issues. Indeed, although ACT math did not have a statistically significant impact on the composite QR score, students with lower ACT math scores were more heavily represented in the control group. This pattern suggests that some students – especially those with more anxiety or less math preparation or proficiency – would choose to delay fulfillment of their requirement. Indeed, 2009 LSA statistics indicate that 17% of students complete their QR requirement as juniors or seniors. Because the focus of this study was on first- and second-year students, the assessment does not capture the outcomes nor experiences of all students meeting the requirement. Therefore, we implemented a second, focused study in 2013 that focused on the experience of all class years, albeit with a more limited measurement of QR.

2013 Assessment: Existing Student Survey

The 2009 assessment provided useful information to LSA, generally supporting the discipline-rooted, "multiple courses of study" (Schneider 2004) approach to a QR requirement. Because of the affirmation of this approach, since that time, there has been an increase in the number of courses fulfilling the QR1 requirement (from 40 in 2009 to 52 today), as well the range (from 8 to 13 departments offering courses). While the assessment was generally confirmatory of the approach to teaching QR, it was admittedly resource-intensive, involving survey incentives and considerable staff time to achieve and analyze a large number of student responses. Additionally, we also hoped to address a key limitation of the 2009 study, which was its focus on students in their two first years at U-M.

Consequently, we sought to see if we might generally replicate the 2009 findings in a more efficient – albeit broader – research study and disseminate findings in a way that would have even greater reach. To do this, we employed a learning analytics approach whereby we leveraged existing data sources: university-wide student surveys and Registrar data. Learning analytics is based on the idea that datasets already generated through normal administrative, teaching or learning activities – such as Registrar data or interactions with learning management systems – can be analyzed to enhance student learning and teaching practice. The Society for Learning Analytics Research (SOLAR) describes it as

the "measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs" (Siemens et al. 2011).

It is important to note that the two studies are not directly comparable, given the differing methodologies and, therefore, they cannot be used to document impact of any intervening interventions. Instead, we offer the second approach as a relatively resource-efficient way to achieve assessment feedback about a QR requirement.

Study 2 Research Questions

Because of the more-streamlined nature of this second study, fewer research questions were utilized:

1. As a validation of the items, is there a statistically significant positive correlation between ACT math score and starting estimate of quantitative skills?

2. Are there associations between students' reported changes in mathematical and statistical skills and QR completion?

    a. Are there statistically significant differences in reported mathematical and statistical gains between students who take QR1 courses and those who have not yet completed the requirement?

    b. Compared to "partial completers" (i.e., students who have taken only one of the required two QR/2 classes), do students who have taken their QR1 course report higher gains in mathematical and statistical skills?

Study 2 Methods

Study 2: Instrument. The SERU (Student Experience in the Research University) survey[6] was developed by an AAU Consortium, and it is utilized by more than 25 research universities across the world (Douglas et al. 2012). SERU is a retrospective pre-test survey. In other words, it asks students to recollect their starting levels of skills/knowledge/attitudes when they enrolled at U-M, as well as estimate their current proficiency. Other research suggests that retrospective self-report survey data correlate better with objective measures of learning, compared to traditional pre-/post-survey designs (Coulter 2013), and the SERU instrument has been validated against traditional markers of educational performance (Douglas et al. 2012). Another advantage of such a format is that it avoids one potential problem of traditional pre/post designs, namely, response shift, where

---

[6] http://www.cshe.berkeley.edu/SERU

the intervention (e.g., a course) changes the respondents' evaluation standard by which they rate themselves (Howard 1980; Pratt, McGuigan and Katzev 2000). For example, Pohl (1982) describes an example of a seminar designed to raise students' awareness of leadership behaviors. A student may rate herself average at the pre-test, but then at the end of the course, realize that with her newfound understanding of leadership competencies and challenges, her initial rating was inflated. Armed with this greater understanding of the concept of leadership behaviors, the student may again rate herself as average on the post-test. Because there would be no difference in pre- and post-test scores, Pohl notes that we would mistakenly conclude that the workshop had no impact, when, in fact, the student's evaluation standard was changed by it.

At the University of Michigan, SERU is called "UMAY" (University of Michigan Asks You), and it is distributed to all undergraduates by the Office of Budget and Planning.[7] The UMAY survey data used here stems from survey distribution in Winter-Summer Terms 2013.

One section of UMAY, a retrospective pre-test, focuses specifically on students' reported learning gains, ranging from analytical/critical thinking to quantitative reasoning and appreciation of cultural and global diversity. For each learning outcome, students were asked to apply a Likert-like scale to evaluate their abilities at these two time points, "when you started" and "current." The items we used for quantitative reasoning appear like this:

> Please rate your level of proficiency in the following areas when you started at this institution and now: <u>Quantitative (mathematical and statistical skills)</u>
>
> • When you started here [Scale: Very poor to Excellent]
>
> • Current ability level [Scale: Very poor to Excellent]

Study 2 Results

The UMAY survey is confidentially administered; therefore, by using U-M Registrar records, the U-M Office of Budget and Planning was able to align survey responses and students' demographics and QR completion records. Definitions used below are:

• "Completers" fulfilled the full requirement with one QR1 class or two QR/2 courses.

• "Partial completers" have taken only one of the required two QR/2 classes.

• "Non-completers" have not yet taken any QR courses.

Response rates.   LSA's 2013 UMAY response rate was 32%, with 3,590 respondents. Demographics of respondents were relatively similar to the 2009

---

[7] Survey items are available here: http://obp.umich.edu/root/surveys-research/undergraduate-surveys/umay/survey-instrument

assessment, with most being women (68%) and non-underrepresented minorities (89%). A key difference was that there was much more representation from upper-level students. Seniors (36.2%) composed the largest group of respondents, followed by sophomores (25.4%), juniors (24.8%), and first-year students (13.7%).

Research Question #1: Is there a statistically significant positive correlation between ACT math score and starting estimate of quantitative skills?

As an initial validation of the UMAY, we analyzed "when I started" scores for students by entering ACT math score. Our expectation was that higher ACT math scores should be associated with higher starting estimates of quantitative skills. We did find that there is a moderate correlation ($r = 0.41$) with the score and students' self-assessed beginning quantitative skills. There are statistically significant differences in starting estimate by quartile, and students' self-ratings are linearly related to ACT math quartile. Students in the lowest quartile, with ACT scores of 14-26, had the lowest starting mean, and students in the top quartile (ACT = 33+) had the highest average (Fig. 1). These differences suggest that students approximate their starting quantitative skills in a manner that aligns with real-world performance, giving us greater confidence in the students' retrospective evaluations.
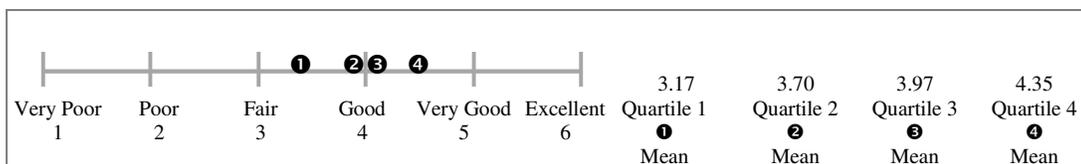

Figure 1. : Means of answers to the "When I started" question, by ACT math quartile.

Research Questions #2: (a) Are there statistically significant differences in reported gains between students who take QR1 courses and those who have not yet completed the requirement? (b) Compared to "partial completers," do students who have taken their QR1 course report higher gains in mathematical and statistical skills?

Here, the definition of a *gain* is the difference between a student's: (a) current self-assessed proficiency and (b) recollection of proficiency when entering the university ("when you started here"). We focus only on students' estimates of their beginning and current proficiency for "quantitative (mathematical and statistical) skills."

Consistent with results from the 2009 assessment, completers reported a small estimated gain (0.24) in quantitative skills, rating their mean current proficiency as fair to good (Figure 2; Table 4). However, even though the reported growth is small, any improvement is important considering that partial and non-completers

17

reported *losses* in their quantitative skills. When comparing QR completers and non-completers, there were statistically significant differences in current levels ($t=7.06$, $p<0.001$) and gains ($t=5.11$, $p<0.001$).
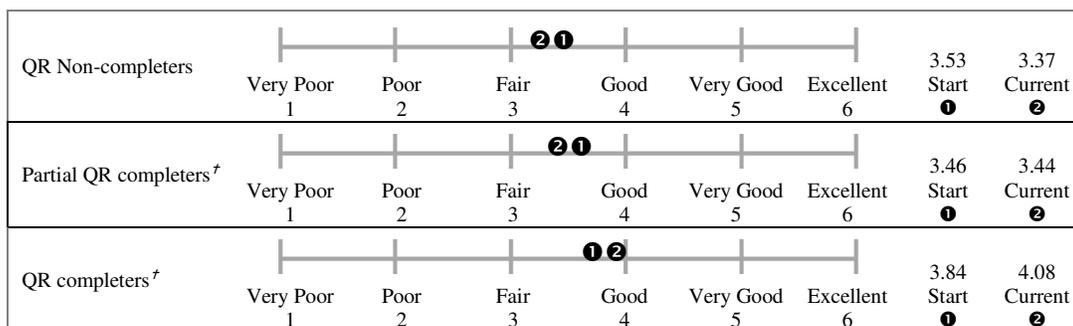

Figure 2: Change in self-reported quantitative (mathematical and statistical) proficiency between "when I started" and "current ability" for…. Students used a Likert-like scale to evaluate their abilities: 6=Excellent, 5=Very good, 4=Good, 3=Fair, 2=Poor, and 1=Very poor. Gains are measured individuals by subtracting a student's starting evaluation from the student's current evaluation. †"Completed" means the student took one QR1 or two QR2 courses; "Partially completed" means the student has taken one QR2 course but has not yet finished the second QR2 course.

Table 4.

"Current" and "Gain" Measures for Quantitative (Mathematical and Statistical) Skills, by QR Completion

	Current mean (s.d.)	Gain (s.d.)	n (current/gain)
QR Non-completers	3.37 (1.04)	-0.16 (0.71)	188/187
Partially completed QR†	3.44 (1.04)	-0.02 (0.77)	169/167
Completed QR†	4.08 (1.12)	0.24 (1.04)	2759/2754

Note: Students used a Likert-like scale to evaluate their abilities: 6=Excellent, 5=Very good, 4=Good, 3=Fair, 2=Poor, and 1=Very poor. Gains are measured individual by subtracting a student's starting evaluation from their current evaluation. †Completed means the student took one QR1 or two QR2 courses; Partially completed means the student has taken one QR2 course, but has not completed the second QR2 course.

Study 2 Dissemination of Results

In the 2009 study, we sought to disseminate assessment data through a series of in-person meetings, followed up by wider dissemination through Web and blogposts. In short, several in-person meetings served to bring results back to the departments and instructors who were responsible for the implementation of the QR requirement. While this was valuable for gathering a good deal of information about promising instructional strategies that were associated with student outcomes, it was resource-intensive (e.g., setting up and holding multiple

meetings) and the in-person data-sharing component had more limited reach.

In keeping with a learning analytics approach, we are in the process of developing a simple online toolkit for administrators and faculty to access UMAY data directly, with a display similar to Figure 2. This tool was only very recently made available to department chairs and key administrators, so it is too soon to measure its use and usefulness. However, our hope is that such a tool will provide greater accessibility to information that could be used for course and curricular change. Because the results of both the 2009 and 2013 studies suggest that the QR requirement was generally working well, the reaction to the results was to encourage small revisions to QR courses and assignments, rather than wholesale change of the requirement.

Study 2 Discussion

Admittedly, the second analytics-based approach offers a much less nuanced assessment of the requirement, because of its longer timeframe, more limited analysis of covariates, and reliance on only two survey items. For example, for students more advanced in their academic careers, we do not tease out the impact of the QR course versus other courses they may have taken that developed related skills. Additionally, we did not control for other demographic and educational factors, as we did in the 2009 study. However, by harnessing existing data, we were able to offer a quick and efficient assessment of the QR requirement that was relatively similar to the more in-depth and resource-intensive 2009 assessment. We were also able to capture the experience of the full spectrum of class years.

Universities not using the SERU instrument are likely conducting a similar survey of student learning and experiences (for instance, national commercial instruments or a homegrown instrument that explores similar skills), given that the vast majority of colleges and universities are using some national survey for assessment (Kuh and Ikenberry 2009). In addition to SERU, other common examples include the National Survey of Student Engagement,[8] Cooperative Institutional Research Program surveys,[9] and the Higher Education Research Institute instrument.[10]

Concluding Remarks

Given the multiple contexts in which college graduates must utilize QR skills – as well as an increased emphasis on analytics-based approaches in the workplace (Davenport and Harris 2007; Sweet and Strand 2006) -- the need to teach

---

[8] http://nsse.iub.edu; see also Rocconi et al. 2013, and Dumford and Rocconi, this issue.
[9] http://www.heri.ucla.edu/abtcirp.php
[10] http://www.heri.ucla.edu/

transferable quantitative reasoning skills is more critical than ever. However, teaching and assessing QR is challenging. Likewise, assessment of a QR requirement with very large numbers of students and across diverse disciplinary homes is complex, as is the crucial engagement of faculty and administrators with assessment data.

To address these challenges, we present two models for assessing a QR requirement that involved a large number of students taking a wide variety of classes across many academic departments. These two models were (1) a validated survey of course-takers and non-course-takers that looked in-depth at instructional approaches and variation in outcomes by discipline, and (2) a more-global, resource-efficient approach that drew on existing university surveys supplemented with learning analytics data. Both methods had their limitations, such as non-random assignment to treatments and reliance on indirect measures of student learning.

The two approaches presented here present options for utilizing indirect methods in QR assessment, while also acknowledging that complementary use of direct measures for QR assessment is preferable. However, these methods and dissemination strategies offer different possibilities depending on resources available on campus and aims of the assessment. Those seeking a helpful review of ways to directly measure students' QR skills will find Grawe's (2011) review of "context-rich" assessment options to offer a useful overview. We also recommend these approaches as complementary, with periodic in-depth investigation supplemented by more frequent broad-based checks.

A key stage in the assessment cycle is "closing the loop," or the use of "the resulting information to document, explain, and improve performance" (Angelo 1995). Therefore, it is important to note that these two models offer multiple ways to disseminate assessment findings, namely meetings with key stakeholders to share results and learn best practices, use of online media (e.g., blogs), and development of an online toolkit that enables academic departments to analyze data about their own students more easily. As each of these dissemination approaches has its strengths and drawbacks, we again suggest a multi-channel approach to working with administrators and faculty, offering both opportunities for face-to-face meetings to explain the data and development of new electronic tools that enable departments to explore their own data. Future research might more systematically analyze QR assessment dissemination strategies to make comparisons about the most-effective approaches.

Based on our experience, administrators and faculty tasked with reviewing QR requirements at their own institutions would be well advised to begin their work reviewing data already collected by their institution to begin their process, and then looking for tools developed and validated by other institutions that can be adapted. This approach allows the institution to invest resources in time

devoted to analysis of student-learning data, rather than in new data-collection processes, or the redevelopment of instruments that already exist. We also suggest that local dissemination approaches also be a critical piece of an effective QR assessment plan, which could potentially involve meetings, social media, and analytics tools.

Acknowledgments

References

Angelo, T. A. 1995. Reassessing and defining assessment. *AAHE Bulletin 48*(3): 149.

Atkinson, M. P., R. F. Czaja, and Z. W. Brewster. 2006. Integrating sociological research into large introductory courses: Learning content and increasing quantitative literacy. *Teaching Sociology* 34(1): 54–64. http://dx.doi.org/10.1177/0092055X0603400105. (accessed March 15, 2013).

Bahls, P. 2012. *Student writing in the quantitative disciplines: A guide for college faculty.* San Francisco, CA: Jossey-Bass.

Baloglu, M. 2003. Individual differences in statistics anxiety among college students. *Personality and Individual Differences* 34(5): 855-865. http://dx.doi.org/10.1016/S0191-8869(02)00076-4

Banta, T. W. 2007. Can assessment for accountability complement assessment for improvement? *Peer Review* 9(2): 9–12. https://www.aacu.org/publications-research/periodicals/can-assessment-accountability-complement-assessment-improvement

Boersma, S. and D. Klyve. 2013. Measuring habits of mind: Toward a prompt-less instrument for assessing quantitative literacy. *Numeracy* 6(1): Article 6. http://dx.doi.org/10.5038/1936-4660.6.1.6. (accessed March 23, 2014).

Bok, D. 2006. *Our underachieving colleges: A candid look at how much students learn and why they should be learning more.* Princeton, NJ: Princeton University Press.

Butcher, K.F., P. J. McEwan, and C. Taylor. 2010. The effects of quantitative skills training on college outcomes and peers. *Economics of Education Review* 29(2): 187–199. http://dx.doi.org/10.1016/j.econedurev.2009.06.001 (accessed March 23, 2014).

Chun, M. 2002. Looking where the light is better: A review of the literature on assessing higher education quality. *Peer Review* 4(2/3): 16–25.

21

College of Literature, Science, and the Arts. 2009. Quantitative Reasoning. *LSA Bulletin.* http://www.lsa.umich.edu/UMICH/lsa_bulletin/Home/Archive/Bulletin2009.pdf (retrieved December 11, 2014)

Cook, C.E., D. S. Meizlish, and M. C. Wright. 2011. The role of a teaching center in curricular reform and assessment. In *Advancing the culture of teaching on campus: How a teaching center can make a difference*, ed. C. Cook and M. Kaplan, 121–136. Sterling, VA: Stylus.

Coulter, S. E. 2013. Using the retrospective pretest to get usable, indirect evidence of student learning. *Assessment & Evaluation in Higher Education* 37(3):321–334. http://dx.doi.org/10.1080/02602938.2010.534761

Davenport, T. H., and J. G. Harris. 2007. *Competing on analytics: The new science of winning*. Boston, MA: Harvard Business School Press.

Dorans, N. J. 1999. *Correspondences between ACT and SAT I scores.* College Entrance Examination Board, NY: NY.

Douglas, J.A., G. Thompson, and C. Zhao. 2012. The learning outcomes race: The value of self-reported gains in large research universities. *Higher Education* 64(3):317–335. http://dx.doi.org/10.1007/s10734-011-9496-x

Dwyer, C. A., A. Gallagher, J. Levin, and M. E. Morley. 2003. *What is quantitative reasoning? Defining the construct for assessment purposes*. Princeton, NJ: Educational Testing Service.

Ehrlinger, J., and E. A. Shain. 2014. How accuracy in students' self-perceptions relates to success in learning. In *Applying science of learning in education: Infusing psychological science into the curriculum*, ed. V.A. Benassi, C.E. Overson, and C.M. Hakala, 142–151. American Psychological Association.

Grawe, N. D. 2011. Beyond math skills: Measuring quantitative reasoning in context. In *New Directions for Institutional Research, 149* (Spring),   ed. J. D. Penn, 41–52. Hoboken, NJ: Wiley.

———. 2012. Achieving a quantitative literate society: Resources and community to support national change. *Liberal Education* (Spring): 30–35.

Halaby, C. N. (2005, October). Two assessment studies of the general education quantitative reasoning "A" requirement at the University of Wisconsin-Madison. Report to the General Education Assessment Council Subcommittee on the Effectiveness of the Quantitative Reasoning "A" Requirement. Madison, Wisconsin. Available: http://www.ls.wisc.edu/gened/AssessmentReports.htm (accessed May 1, 2009).

Hart Research Associates. 2009. Learning and assessment: Trends in undergraduate education. A survey among members of the Association of American Colleges and Universities conducted by Hart Research Associates. Washington, DC: Author. Retrieved from http://www.aacu.org/membership/documents/2009MemberSurvey_Part1.pdf (accessed April 3, 2014).

Howard, G. S. 1980. Response-shift bias: A problem in evaluating interventions with pre/post self-reports. *Evaluation Review* 4:93–106. http://dx.doi.org/10.1177/0193841X8000400105

22

Kuh, G., and S. Ikenberry. 2009. More than you think, less than we need: Learning outcomes assessment in higher education. Report from the National Institute for Learning Outcomes Assessment. Available: http://www.learningoutcomesassessment.org/ (accessed April 7, 2014).

Laursen, S., M. L. Hassi, M. Kogan, A. B. Hunter, and T. Weston. 2011. *Evaluation of the IBL mathematics project: Student and instructor outcomes of inquiry-based learning in college mathematics*. Report to the Educational Advancement Foundation and the IBL Mathematics Centers. Boulder, CO: University of Colorado, Ethnography & Evaluation Research. Available at: http://www.colorado.edu/eer/research/steminquiry.html

Latiolais, M.P., and W. Laurence. 2009. Engaging math-avoidant college students. *Numeracy* 2(2): Article 5. http://dx.doi.org/10.5038/1936-4660.2.2.5

Lutsky, N. 2008. Arguing with numbers: A rationale and suggestions for teaching quantitative reasoning through argument and writing. In *Calculation vs. context: Quantitative literacy and its implications for teacher education*, ed. B.L. Madison and L.A. Steen, 59–74. Washington, DC: Mathematical Association of America.

Moore, G., J. R. Slate, S. Edmonson, J. P. Combs, R. Bustamante, and A. J. Onwuegbuzie. 2010. High school students and their lack of preparedness for college: A statewide study. *Education and Urban Society* 42: 817–838. http://dx.doi.org/10.1177/0013124510379619

Pike, G. R. 2011. Using students' self-reported outcomes in scholarly research. *New Directions for Institutional Research* 150: 41–58. http://dx.doi.org/10.1002/ir.388

———. 2013. NSSE benchmarks and institutional outcomes: A note on the importance of considering the intended uses of a measure in validity studies. *Research in Higher Education* 54: 149–170. http://dx.doi.org/10.1007/s11162-012-9279-y

Pohl, N. F. 1982. Using retrospective pre-ratings to counteract response shift confounding. *Journal of Experimental Education* 50(4): 211–214.

Porter, S. R. 2013. Self-reported learning gains: A theory and test of college student survey response. *Research in Higher Education* 54: 201–226. http://dx.doi.org/10.1007/s11162-012-9277-0

——— and M. W. Whitcomb. 2004. Understanding the effect of prizes on response rates. *New Directions for Institutional Research* 121: 51-–62. http://dx.doi.org/10.1002/ir.100

Pratt, C.C., W. M. McGuigan, and A. R. Katzev. 2000. Measuring program outcomes: Using retrospective pretest methodology. *American Journal of Evaluation* 21: 341–348. http://dx.doi.org/10.1177/109821400002100305

Rhodes, T., ed. 2010. *Assessing outcomes and improving achievement: Tips and tools for using rubrics*. Washington, DC: Association of American Colleges and Universities.

Rocconi, L. M., A. D. Lambert, A. C. McCormick, and S. A. Sarraf. 2013. Making college count: An examination of quantitative reasoning activities in higher education. *Numeracy* 6 (2): Article 10. http://dx.doi.org/10.5038/1936-4660.6.2.10 (accessed April 3, 2014).

23

Schneider, C.G. 2004. Setting greater expectations for quantitative learning. *Peer Review* 6(4): 26-–27.

Siemens, G., D. Gasevic, C. Haythornthwaite, S. Dawson, S. Buckingham Shum, R. Ferguson, E. Duval, K. Verbert, and R. S. J. D. Baker. 2011. *Open learning analytics: an integrated & modularized platform.* Society for Learning Analytics Research. Available: http://solaresearch.org/OpenLearningAnalytics.pdf (accessed April 13, 2014).

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. http://dx.doi.org/10.5038/1936-4660.4.2.5 (accessed April 3, 2014).

Sweet, S., and K. Strand. 2006. Cultivating quantitative literacy: The role of sociology, *Teaching Sociology* 34(1): 1-4. http://dx.doi.org/10.1177/0092055X0603400101

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. http://dx.doi.org/10.5038/1936-4660.4.2.4 (accessed March 31, 2014).

Wehlberg, C. 2008. *Promoting integrated and transformative assessment: A deeper focus on student learning.* San Francisco: Jossey-Bass.

Welsh, J.F., and J. Metcalf. 2003a. Faculty and administrative support for institutional effectiveness activities. A bridge across the chasm? *Journal of Higher Education* 74: 445–468. http://dx.doi.org/10.1353/jhe.2003.0032

———. 2003b. Cultivating faculty support for institutional effectiveness activities: Benchmarking best practices. *Assessment & Evaluation in Higher Education* 28(1): 33–45. http://dx.doi.org/10.1080/02602930301682

Wolfe, C. R. 1993. Quantitative reasoning across a college curriculum. *College Teaching* 41(1): 3–9. http://dx.doi.org/10.1080/87567555.1993.9926775