Assessing Quantitative Reasoning in GE (Owens, Ladwig, and Mills)

Introduction

Many students at CSU, Chico, receive much of their college-level mathematics education from the one MATH course they complete to satisfy the GE Core A4 requirement. Students who receive additional, significant exposure to college-level mathematical principles and applications are typically enrolled in academic programs having some additional quantitative requirements.

EM 99-05 states that "[i]n every course, relevant skills of the Core must be applied as essential to the process of mastering content and making applications." This objective is further reinforced in the capstone requirement, which dictates that "[t]hemes will incorporate, build upon, and nurture skills from Area A…" Since mathematics is part of the Core (i.e., Area A), this requirement strongly implies that mathematical content should be distributed throughout the GE program, not just in MATH courses and a few science courses. However, this implication has not been systematically evaluated during any previous assessment of the GE program nor has it been encouraged or enforced other than within Sub-Area A4.

Other academic areas housed in the GE Core – written communication, oral communication, and critical thinking – have to varying degrees received greater attention to interdisciplinary application than has mathematics. This can most readily be explained by the relative newness of mathematics to the Core. Previously, the mathematics component of GE was grouped with the science requirement, as is the case with the statewide description of general education. With the adoption of EM 99-05, mathematics at CSU, Chico has now been aligned as a core area in order to recognize the central nature of mathematics in general education. That is, mathematics is not a culminating experience. Rather, it is a basic yet essential component of knowledge having potential to facilitate understanding and ability in a breadth of academic subjects. Indeed, mathematics forms the foundation for quantitative reasoning, which can be defined as the application of mathematics to describe, analyze, and solve authentic problems in context.

Mathematics in general education at CSU, Chico has previously been assessed in two different ways. The most significant and continuous assessment means has consisted of regularly scheduled GE course reviews. Area A, including Sub-Area A4 Mathematics, was last reviewed during academic year 1999-2000. No significant deficiencies in Sub-Area A4 Mathematics were discovered during the last review. Additionally, during AY 2001-02, a pilot assessment project was conducted, focusing on one course in the GE program, MATH 005, Statistics. The Learning Environment Assessment Project – Mathematics (LEAP-M) involved both direct and indirect measures of student learning. Lessons learned in the LEAP-M project proved valuable in designing the probability component of the current project.

Process

1. **Planning the Process**

    a. **Develop a list of Area A4 GE student learning outcomes.**

In the late 1990s, a CSU-sponsored project solicited input from three campuses, CSU, Chico, CSU, Sacramento, and San Francisco State University, in developing a comprehensive set of student learning outcomes in mathematical reasoning (Mills, et al., *Learning Outcomes for Mathematical Reasoning for the Baccalaureate Degree*, Learning Outcomes Project Final Report, California State University, January 15, 1999). That study identified nineteen different student learning outcomes deemed desirable for graduates of the CSU. (See appendix.) These nineteen learning outcomes provided a valuable starting point for this study. Recognizing that nineteen outcomes were too many to target initially, input was solicited from current GE mathematics faculty at CSU, Chico to help rank these outcomes by relevance and suitability for assessment.

Ultimately, a total of six student learning outcomes were identified as most significant and pertinent to this assessment project (Table 1, with highlighted goals corresponding to the indicated outcome). The first three outcomes are suitable for assessment by indirect means, i.e., by way of a student response survey. The last three are suitable for direct and embedded assessment. (See the appendix of *Learning Outcomes for Mathematical Reasoning for the Baccalaureate Degree* for a full description of the significance of each learning outcome.)

These student learning outcomes were particularly appropriate to this study because they are strongly correlated with goals of the GE program, as expressed in EM 99-05. These goals fall into three different categories: GE program goals, GE Core skills, and GE Mathematics (Sub-Area A4) goals. Goals relating specifically to quantitative reasoning, either directly or indirectly, are summarized in Table 2.

Table 1: Selected Student Learning Outcomes

Student Learning Outcomes in Quantitative Reasoning	GE Program Goals			GE Core Skills		GE Math Goals	
Baccalaureate graduates of CSU, Chico will be able to:	1	2	3	1	2	1	2
Outcomes 1 - 3 are suitable for indirect assessment, e.g., a student response survey.							
1. view mathematics with heightened interest, increased confidence, and less anxiety as a result of their educational experiences.		■					
2. regard mathematics as a way to think, reason and conceptualize, not simply as a set of techniques.		■			■		
3. understand and appreciate the connections between mathematics and a variety of quantitative and non-quantitative disciplines.		■	■		■		
Outcomes 4 - 6 are suitable for direct assessment, e.g., student performance on a problem.							
4. develop and apply measurement techniques to data collection, and evaluate potential sources of error, including variability and bias.	■			■	■	■	■
5. interpret, make appropriate judgments, and draw reasonable conclusions based on numerical, graphical and symbolic information.	■			■	■	■	■
6. critically evaluate quantitative information, and identify deceptive or erroneous reasoning.	■			■	■		■

Each of the selected student learning outcomes is associated with one or more of the GE goals (Table 1). For example, Outcome #2, *Regard mathematics as a way to think, etc.*, is associated both with GE Program Goal #2 and with GE Core Skill #2. It is noteworthy that all six learning outcomes together encompass all of the GE goals, i.e., assessment of all six student learning outcomes would concurrently evaluate all mathematics-related aspects of the GE program.

Much of the prior assessment of mathematics in general education has focused on indirect measures applied to individual MATH courses. In the current study, the focus was to be on direct and embedded means of assessing student performance. That is, students in a course would be evaluated by their ability to perform a task relevant to that course, where the task would also be associated with one or more overarching student learning outcomes.

Table 2: GE Goals Relevant to Quantitative Reasoning (EM 99-05)

Relevant GE Program Goals are: 1. to improve … mathematical reasoning, analysis and problem solving, and the ability to access, evaluate, and apply information. 2. to enhance general knowledge and attitudes … 3. to provide … coherence, connectedness, and commonality within broad areas of undergraduate education. *Relevant GE Core (Area A) Skills* The principal charge to this area is to provide students opportunities to learn and demonstrate: 1. effective mathematical reasoning. 2. fundamental links between thinking … and mathematical reasoning. *GE Mathematics (Sub-Area A4) Goals* Students must demonstrate: 1. understanding of one or more mathematical fields … 2. understanding of more than computational skills; they must also demonstrate understanding of basic mathematical concepts and apply these concepts to complex real world activities.

b. Determine which SLOs will be assessed, and identify course locations, tasks, and rubrics for assessment of the SLOs.

With the list of SLOs reduced to six items, the Task Force planned to seek broad input from faculty across campus to determine campus SLO priorities and possible locations for measuring the SLOs. With the goal of identifying locations for embedded assessment of the SLOs, the Task Force initially targeted non-mathematics courses likely to involve quantitative reasoning, focusing on upper-division general education courses, but also including service courses for technical majors. However, several events caused the Task Force to modify this plan.

- Informal communication with upper-division GE instructors indicated that, with just a few exceptions, there is very little quantitative reasoning in these courses. In most of the exceptions, the instructor-identified tasks were primarily low-level arithmetic, not quantitative reasoning.
- In a few cases, upper-division GE instructor-identified tasks involving the use of formulas, but even here, the tasks seldom involved even algebraic reasoning. Students simply substituted values into a formula and evaluated the result.
- It was much easier to identify higher-level quantitative reasoning tasks in service courses for technical majors (including some GE courses) and in major courses in technical disciplines. However, the Task Force and others had serious concerns about what would be measured in such courses. Would we be measuring the impact of the GE program or understanding developed in major coursework or through some other mechanism?

Consequently, by the end of fall 2005, it was determined that the assessment would be performed primarily in lower-division MATH courses with instructors willing to embed a common task. This effort was undertaken with the assumption that this would be a first step in a multi-year project. Consequently, it would be better to start with a fairly uncomplicated approach with the expectation that subsequent studies would be more sophisticated. Early in spring 2006, the Task Force crafted a probability task and a calculus task and sought faculty volunteers to imbed these tasks in their courses. This two-pronged approach allowed the Task Force to collect data from a broad student population. With very few exceptions, most students spend a part of their mathematics coursework studying either probability or calculus; in fact, Liberal Studies is the only major with students who would not typically be enrolled in at least one course in the study. To assist with future efforts, embedded assessment was also piloted in three upper-division courses, two in GE and one outside of GE. While data from the two upper-division GE courses could not be included in the study (for reasons described below), data from the upper-division non-GE course, a course taken by all engineering majors, is included. Table 3 provides details about the participating courses.

Table 3: Participation

Course	# Sections	# Students Enrolled	# Student Participants
Probability Task Results			
MATH 101: Patterns of Mathematical Thought	7	294	196
MATH 105: Statistics	8	273	157
MATH 107: Finite Mathematics for Business	7	260	221
CIVL 302: Engineering Economy and Statistics	1	66	33
Upper-Division GE	2	82	0
Probability Task Totals	**25**	**975**	**607**
Calculus Problem Results			
MATH 120: Analytic Geometry and Calculus	**5**	**153**	**100**

c. Develop instruments to be used to gather student performance data, student demographic information, and student attitudinal data.

The probability task selected is a minor modification of a task used in LEAP-M in 2002; similar problems occur naturally in a variety of Area A4 courses, increasing the pool of potential faculty participants. The calculus task was selected in consultation with MATH 120 faculty; a problem of this sort is common in all sections of first-semester calculus and such a problem frequently appears on common final exams in calculus. Each task relates to only two of the identified three student learning outcomes (5 and 6), but these two in combination with the student response survey described in the next paragraph are associated with all applicable GE goals. See Appendix C for the problems, the problem solutions and the calculus scoring rubric.

The conversation with GE mathematics faculty indicated the high priority they place on the attitudinal outcomes 1 through 3. Therefore, with the assistance of Institutional Research at CSU, Chico, the Task Force developed and administered a survey to the same students participating in the mathematics assessment tasks. (See Appendix C for the survey.) The objectives of the survey were to learn something about the math-related experiences of Chico students, to attempt to gauge their attitudes and perceptions regarding mathematics, and to collect student demographics. Moreover, the embedded task and survey were designed so that the survey findings could later be correlated, as desired, with individual student performance on the mathematics tasks.

2. Implementation of Assessment

a. The Probability Task and Survey

In a series of two multiple choice questions, students are first asked to determine the most likely outcome of five flips of a fair coin and to then explain why they chose their first answer. In early March, faculty in spring 2006 sections of MATH 101 (Patterns of mathematical Thought), MATH 105 (Statistics), and MATH 107 (Finite Mathematics for Business) were given the probability task, asked to include the probability task on an exam, and asked to administer the survey to their students. The collection of possible participants was expanded to include two upper-division GE classes and CIVL 302 (Engineering Economy and Statistics). Faculty who agreed to participate were also given a recording sheet so that they could record student responses as they graded their exams. By asking for student identification information on both the faculty recording sheets and the survey, we were able to link student performance on the task with demographic and attitudinal information about the student.

There were 975 students enrolled in the participating courses. A total of 607 students completed both the probability task and the survey; data analysis is based on this sample of 607 students. In a few cases, entire sections of student data were not used because results of either the task or the survey were not available. In one case, an instructor of an upper-division GE course modified the probability task, resulting in a task that was substantially different from that used in other courses; we were not able to include that instructor's data in our analysis. See Table 3.

b. The Calculus Task and Survey

The calculus problem is a standard related rates problem; similar problems can be found in any calculus textbook and frequently appear on final exams. Many related rates problems require some geometric reasoning before calculus can be applied. The problem selected requires only minimal geometry, an application of the Pythagorean Theorem. Since this is an open-response computational problem, the task Force worked with calculus instructors to develop a 4-point scoring rubric. The demographic/attitudinal survey was also administered to the calculus students.

All five spring 2006 sections if MATH 120 participated, resulting in three faculty participants and a potential for 153 student participants. Every class set of work was independently graded by the instructor and one of the Task Force members. After meeting to discuss two of the class sets of scores, the Task Force determined that the rubric was clear enough to assure consistent scores across readers.

We received both problem scores and survey results for 100 students. The problem chosen was not a good fit for the CIVL 302 course. That instructor collected student performance information for a different calculus problem; those results are not used in the data analysis. See Table 3.

3. Analysis

a. The Probability Task

Table 4 displays data from the probability task and the student survey of participants in the probability portion of the study. Students in the "right" category gave both the correct answer and the correct reason for choosing that answer.

Table 4: Probability Task Data

	M101		M105		M107		CIVL 302	
right	124	63.27%	128	81.53%	154	69.68%	28	84.85%
wrong	72	36.73%	29	18.47%	67	30.32%	5	15.15%
totals	196	100.00%	157	100.00%	221	100.00%	33	100.00%
	Chi-square = 17.514, p-value = 0.0005539							

	male		female		dev math		no dev math	
right	188	67.14%	246	75.23%	164	67.21%	270	74.38%
wrong	92	32.86%	81	24.77%	80	32.79%	93	25.62%
totals	280	100.00%	327	100.00%	244	100.00%	363	100.00%
Chi-square = 4.84, p-value = 0.0278069				**Chi-square = 3.68, p-value = 0.0551349**				

	freshmen		not freshmen		math frequently use in major		math never used in major	
right	266	73.08%	168	69.14%	115	74.68%	90	67.16%
wrong	98	26.92%	75	30.86%	39	25.32%	44	32.84%
totals	364	100.00%	243	100.00%	154	100.00%	134	100.00%

Chi-square = 1.11, p-value = 0.2918646	Chi-square = 1.62 , p-value = 0.2031

	confident		not confident		attitude +		attitude -	
right	362	71.83%	18	72.00%	247	75.08%	101	64.74%
wrong	142	28.17%	7	28.00%	82	24.92%	55	35.26%
totals	504	100.00%	25	100.00%	329	100.00%	156	100.00%
Chi-square = 0, p-value = 1					**Chi-square = 5.574 , p-value = 0.018229**			

The data indicate some statistically significant differences in populations:

- Females were more likely than males to have the right answer for the right reason.
- Students who reported having taken a developmental mathematics course did not score as well as those who did not require developmental mathematics.
- Students who reported a positive attitude toward mathematics (determined by responses to the prompt "I have a good attitude about math.") scored better than those with a negative attitude.
- Students in MATH 105 scored better than either MATH 101 students (p = 0.0001613, chi-square test) or MATH 107 students (p = 0.009126, chi-square test).
- Students in CIVL 302 also scored better than students in MATH 101 or MATH 107, but the difference was somewhat less significant. (E.g., for CIVL 302 versus MATH 107, p = 0.0713799, chi-square test).

There were no significant differences in performance between freshmen and other students. Students' mathematical confidence (determined by responses to the prompt "I can use math correctly.") had no statistical relation to their performance on the task. The difference in performance between students in MATH 101 and MATH 107 was not statistically significant. While there were some differences in performance between students in different colleges, these differences were probably a reflection of the courses the students took. For example, 75.82% of the 91 students from the College of Behavioral and Social Sciences answered both questions correctly versus 64.14% of the 145 students from the College of Business; however, almost all students from the College of Business were enrolled in MATH 107, while students from the College of Behavioral and Social Sciences were more likely to be enrolled in MATH 105.

About the chi-square test: The chi-square test is used to measure the degree of confidence that the difference in characteristics of the samples can be generalized to differences in the population, and are not due to random error. The p-value is the probability if a chi-square value at least as large as the given value. The smaller the p-value, the greater is the statistical significance of the difference in characteristics of the samples.

b. The Calculus Task

Table 5 displays some of the data collected from the calculus problem and student survey. In the displayed data, there were no statistically different levels of performance in the groups examined. While males performed slightly better than females (mean 2.92 versus 2.52, respectively), the difference was not significant. Similarly, students who reported that they would be taking additional mathematics courses in the future did somewhat better than students in their terminal mathematics class (mean 2.91 versus 2.5, respectively), but again the difference was not significant. There was little difference between the freshman and non-freshmen groups. Attitude did not appear to be related to performance on the problem.

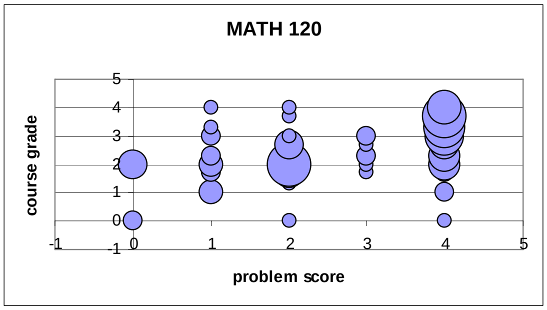
Table 5: Calculus Problem Data

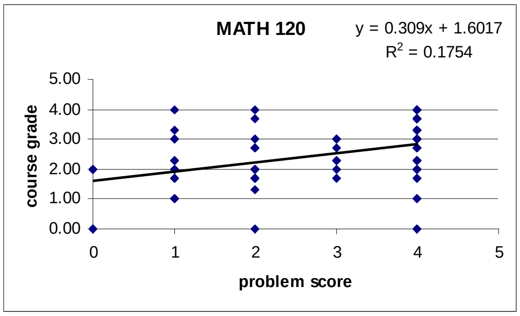
score	male		female		attitude +		attitude -	
4	41	56%	9	33%	39	51%	11	48%
3	3	4%	4	15%	7	9%	0	0%
2	16	22%	7	26%	15	19%	8	35%
1	8	11%	6	22%	13	17%	1	4%
0	5	7%	1	4%	3	4%	3	13%
totals	73	100%	27	100%	77	100%	23	100%
	mean = 2.92, σ = 1.36		mean = 2.52, σ = 1.28		mean = 2.86, σ = 1.32		mean =2.65, σ = 1.47	
	t-test, p = 0.189				t-test, p = 0.524			

score	no more math		yes more math		freshmen		not freshmen	
4	10	42%	40	53%	21	43%	29	57%
3	1	4%	6	8%	5	10%	2	4%
2	6	25%	17	22%	13	27%	10	20%
1	5	21%	9	12%	7	14%	7	14%
0	2	8%	4	5%	3	6%	3	6%
totals	24	100%	76	100%	49	100%	51	100%
	mean = 2.50, σ = 1.44		mean =2.91, σ = 1.31		mean = 2.69, σ = 1.33		mean = 2.92, σ = 1.37	
	t-test, p = 0.197				t-test, p = 0.401			

There was moderate correlation between a student's score on the calculus problem and the student's grade in the class; illustrating that assessment is different from grading. The bubble graph in Figure 1 illustrates frequencies of pairs of problem scores and course grades. The second graph in Figure 1 illustrates the pairs without frequencies, but includes the regression line for the entire data set. It is important to recall that only students who completed the survey *and* took the final exam were included in the sample, resulting in a sample containing few "D" or "F" course grades.

Figure 1: Calculus Problem Score versus Course Grade





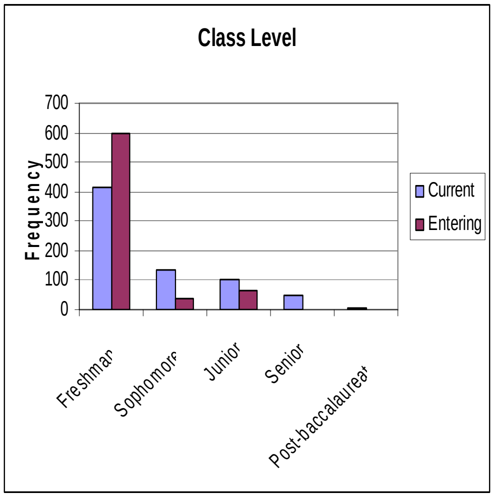
c. Survey Results

Figure 2 provides a summary of additional information collected in the student survey.

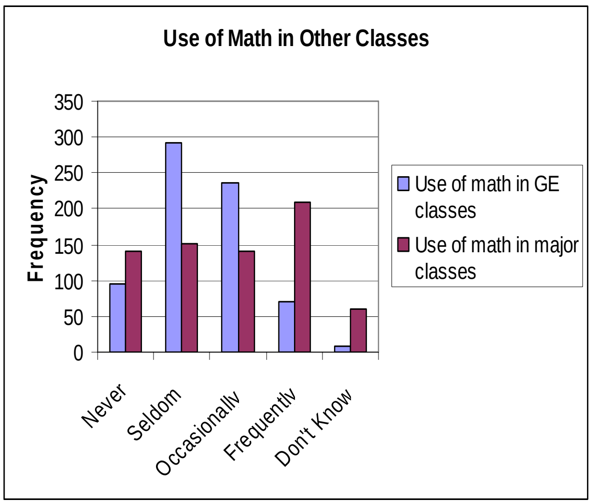
Observations:

- Since most of the students surveyed were freshmen, student perceptions about the use of mathematics in other classes are necessarily somewhat uninformed. In particular, students may think of low-level arithmetic as an example of "use of mathematics."
- Student perceptions confirm what faculty had suggested: mathematics plays at best a minor role in other general education courses.
- Excluding the MATH 120 students, most students in the sample experienced no previous college-level mathematics coursework.
- While many MATH 107 and MATH 120 students will take at least one more mathematics class, most other students are in their terminal mathematics class.

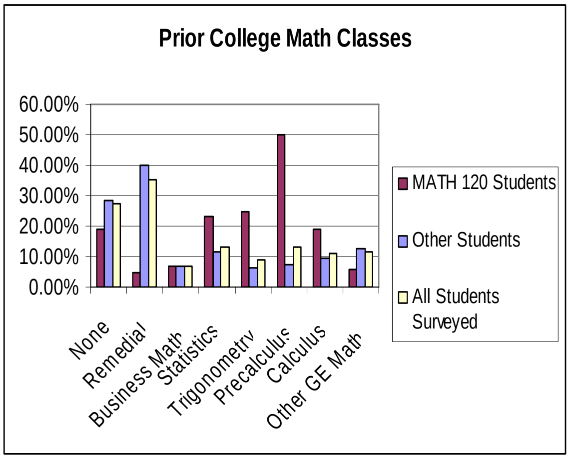
Figure 2: Survey Data

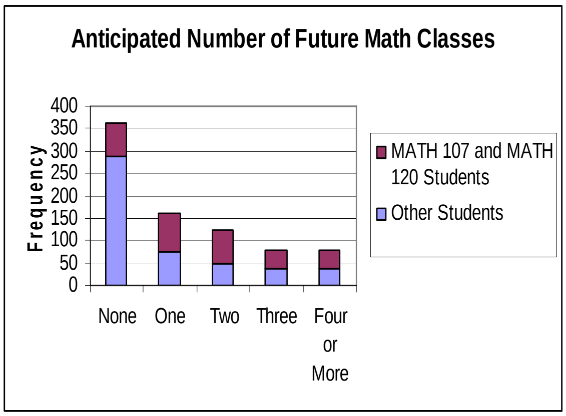


Class Level
Frequency
700
600
500
400
300
200
100
0
Current
Entering
Freshman
Sophomore
Junior
Senior
Post-baccalaureate



Use of Math in Other Classes
Frequency
350
300
250
200
150
100
50
0
Use of math in GE classes
Use of math in major classes
Never
Seldom
Occasionally
Frequently
Don't Know





A. Conclusions and Recommendations for Strengthening QR in GE

- **Establish benchmarks for quantitative reasoning expectations.**
  This is a campus-wide task. What are the campus expectations for quantitative reasoning and at which stages should we assess student progress?

- **Embed quantitative reasoning in courses across the curriculum.**
  Are we happy with the results? If not, faculty outside of math must take on their share of the responsibility. We do students a disservice when we avoid QR simply because students find it difficult. Like writing skills, quantitative reasoning skills need to be reinforced and developed over time and in a variety of settings, both in the majors and in GE.

- **Provide and support professional development opportunities in quantitative reasoning across the disciplines, including GE.**
  Develop a collection of resources for faculty wishing to incorporate significant QR into their courses. Provide workshops in which faculty can work together to develop such interdisciplinary resources, including appropriate assessment instruments.

B. Future Steps

- What is the mathematics enrollment history of a typical student? (Analyze a sample of student transcripts?)
- Begin a campus conversation about the results of this study.
- Help the campus understand that this study is only a first step is assessing quantitative reasoning and that quantitative reasoning cannot be assessed solely in mathematics courses.
- We don't know if our students leave college better able to reason quantitatively than when they entered. Do we wish to measure "value added?" How would we measure this?