

COUNTING UP

THE ASSESSMENT OF MATHEMATICAL
REASONING WITHIN GENERAL EDUCATION AT
CHICO STATE, 2014-2015

This document summarizes the activities and findings of the Mathematical Reasoning Assessment Committee of the Curriculum Advisory Board for the General Education Program at California State University, Chico, during the fall of 2014 through the fall of 2015.

COUNTING UP

THE ASSESSMENT OF MATHEMATICAL REASONING WITHIN GENERAL EDUCATION AT CHICO STATE, 2014-2015

TABLE OF CONTENTS

BACKGROUND	2
DATA COLLECTION	4
2014-15 Academic Year.....	4
Spring 2015.....	4
Fall 2015	4
PARTICIPATION IN QLRA.....	5
RESULTS	6
LOWER DIVISION ASSESSMENT – SPRING 2015.....	6
UPPER DIVISION ASSESSMENT – FALL 2015.....	8
DISCUSSION.....	12
What do the data tell us?	12
How should we measure of quantitative reasoning?.....	13

Appendix A – Lower division assessment instrument

Appendix B – QLRA question conceptual coverage

Appendix C – QLRA results by item

MATHEMATICAL REASONING ASSESSMENT COMMITTEE MEMBERS

AY 2014-15

- Don Miller
- Julie Holland
- Maria Giovanni
- Chris Nichols
- Kent Sandoe
- Mark Wilpolt

Fall 2015

- Chris Nichols
- Don Miller
- Jason Clower
- Jacque Chase
- Kent Sandoe





BACKGROUND

At California State University, Chico, mathematical reasoning is taught in both the lower and upper division through a combination of the required foundation course and through pathway courses that teach mathematical reasoning as a student learning outcome. In the lower division, students are required to complete a one semester foundation course to satisfy their A4 requirement. This requirement is a CSU system wide requirement specified in Executive Order 1100 (March 16, 2016). Currently, a prerequisite to this course is completion of an entry-level math (i.e., intermediate algebra) requirement.

Courses in subarea B4 shall have an explicit intermediate algebra prerequisite, and students shall develop skills and understanding beyond the level of intermediate algebra. Students will not just practice computational skills, but will be able to explain and apply basic mathematical concepts and will be able to solve problems through quantitative reasoning.

EO-1100, Article 4, Area B4, Pg 8.

Among general education pathway courses at Chico State, 14 lower division and 13 upper division courses list mathematical reasoning as one of their student learning outcomes (see Table 1).

DEMONSTRATES KNOWLEDGE OF AND APPLIES MATHEMATICAL OR STATISTICAL METHODS TO DESCRIBE, ANALYZE AND SOLVE PROBLEMS IN CONTEXT.

***Student Learning Outcome for Mathematical Reasoning
Executive Memorandum 10-001, May 15, 2013***

DEPT	Pathway(s)	GE_Area
ABUS 101	Food	D--1 IS
ANTH 111	Div, STV	B--2
APCG 110	STV	C--1 Arts
BIOL 104	H&W	B--2
BIOL/SCED 102	STV, SUST	B--2
CHEM 100	STV, SUST	B--1
CIVL 175	SUST	B--2
ECON 101	Food	D--2 SI
ECON 102	Gdev, IS	D--2 SI
ECON 103	E,J&P	D--1 IS
GEOS 101	Food, GBI, IS	B--1
GEOS 105	GBI, STV	B--1
GEOS 130	Gdev, H&W, STV	B--1
PHYS 100	GBI	B--1
BIOL 303	E,J&P, STV	UD--B
BIOL 323	Div, G&S	UD--B
BIOL 334	SUST	UD--B
GEOS 330	SUST	UD--B
GEOS 355	Div, EJ&P, IS	UD--B
HCSV 323	IS	UD--B
NFSC 303	H&W	UD--B
NURS/MCGS 326	G&S	UD--B
PHYS 376	IS	UD--B
NSCI 300	GBI	UD--B
ECON 340	Div, EJ&P	UD--D
ECON 352	H&W	UD--D
ECON 365	SUST	UD--D

Table 1—General Education pathway courses (by pathway and area) that have **mathematical reasoning** as an SLO

DATA COLLECTION

Data on mathematical reasoning were collected by CAB's Mathematical Reasoning Assessment Committee (MRAC) in three distinct phases occurring (1) during the 2014-15 academic year, (2) the spring semester of 2015, and (3) the fall semester of 2015.

2014-15 Academic Year

During the 2014-15 academic year, the MRAC attempted to gather qualitative data consisting of student work samples that demonstrated quantitative reasoning abilities from upper division GE pathway courses.

Since there were a relatively small number of these courses, the committee divided up the list among its members and, following a personalized email sent to these instructors, one-on-one meetings were arranged. At these meetings, instructors were asked to select an assignment or exam question that would demonstrate learning in this area. A relatively small number of example assignments were collected and reviewed by the committee. The small number of examples and the diversity of disciplines and approaches to quantitative reasoning that these represented gave rise to much discussion about the selection and adaptation of a rubric that could be used to assess student work. Nonetheless, the committee decided to move forward with data collection during the spring of 2015.

Despite ample notification and offers of assistance with tasks related to the collection and reproduction of student work, very few instructors cooperated with the data collection efforts, and committee abandoned this line of inquiry at the end of the spring semester.

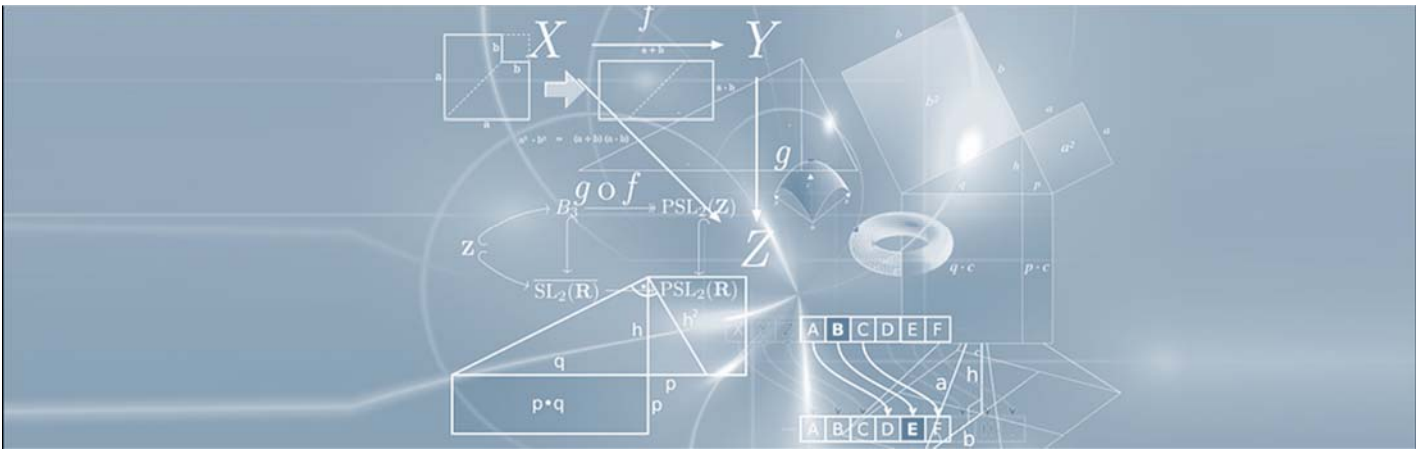
Spring 2015

Overlapping with the AY 2014-15 data collection effort, lower division students were assessed in a very large section of a GE foundation course, MATH 105 (Statistics) during the spring term. At the beginning of the semester, students were given a 6-question multiple-choice pre-test (see Appendix A). The tests were administered anonymously, so individual students' improvements cannot be analyzed. As a post-test, these same questions were embedded in either one of the students' midterm exams or their final exam.

A total of 892 students took the pre-test which was administered in a paper-and-pencil format. Between 734 and 860 students attempted each of the questions on their midterm/final exams. The post-test numbers are smaller because some instructors neglected to include the questions on one or more of their exams, and some students had dropped the course before taking one or more of the exams.

Fall 2015

Over the summer of 2015, CAB chair Chris Nichols reflected on the challenges that the committee faced during the prior year as it attempted to collect evidence of student learning of a qualitative nature. After researching the area, he located a nationally normed quantitative reasoning assessment developed by a team led by Professor Eric Gaze of Bowdoin College (see Appendix B). Chair Nichols proposed shifting to this quantitative assessment early during the fall semester to the reconstituted mathematical reasoning assessment committee. His plan was approved.



The Quantitative Literacy and Reasoning Assessment (QLRA) is a 20-question multiple choice quiz (with 5 choices available for each question) that has been administered to over 3800 students at a variety of institutions in 2012 and 2013. The questions in the QLRA examine students' mathematical understanding and reasoning skills in a variety of areas. The researchers clustered these into four categories including Number Sense (NS), Visual Representation (VR), Probability and Statistics (PS), and Reasoning (R). Professor Gaze kindly allowed CSU, Chico to use this assessment instrument and has been very helpful in providing details of his data for comparison with ours.

Instructors of sections of upper-division GE courses, which had identified mathematical reasoning as a student learning outcome, were contacted by email early in the fall semester and asked to participate in the assessment. Their participation consisted of setting aside 30 minutes of one class period near the end of the semester during which time MRAC members would administer the assessment. No extra preparation was required, but instructors were asked to offer differential extra credit (credit in proportion to quiz scores) as a motivator for the students.

PARTICIPATION IN QLRA

The GE Program at CSU, Chico includes 80 upper-division courses, 13 of which have self-identified Mathematical Reasoning as one of the course SLO's. Those courses are listed in Table 2 below.

DEPT	Pathway(s)	GE_Area	Participation
BIOL 303	E,J&P, STV	UD--B	NONE
BIOL 323	Div, G&S	UD--B	1 (of 3 sections)
BIOL 334	SUST	UD--B	1 section
ECON 340	Div, EJ&P	UD--D	NONE
ECON 352	H&W	UD--D	CANCELLED
ECON 365	SUST	UD--D	1 section
GEOS 330	SUST	UD--B	1 (of 2 sections)
GEOS 355	Div, EJ&P, IS	UD--B	4 (of 5, one online)
HCSV 323	IS	UD--B	NONE
NFSC 303	H&W	UD--B	6 sections
NSCI 300	GBI	UD--B	1 section
NURS/MCGS 326	G&S	UD--B	Online only
PHYS 376	IS	UD--B	1 (of 2 sections)

Table 2—Participation among upper-Division GE pathway courses

Of the 13 upper-division GE courses with a mathematical reasoning SLO, one (ECON 352) was cancelled for the Spring 2016 semester, and another (NURS/MCGS 326) was excluded since the assessment tool was a pencil-and-paper test and all of its sections were online. Of the remaining 11 courses, 8 of them had one or more sections participate in the QLRA assessment. There were a total of 17 participating sections (out of 27 eligible sections) with 11 different instructors.

Those 8 courses include at least one course from 7 of the 10 GE Pathways (all except Food Studies, Global Development, and Science/Technology/Values). 7 of the 8 courses were UD-B (Science) courses; the other, ECON 365, is a UD-D (Social Science) course. Enrollment and participation are summarized in Table 3, below.

Course	Sections	Total Enrollment	Total Participation	Instructor(s)
NFSC 303	6	373	314	A. Alger, S. Boek-Dominguez, T. Berman
GEOS 355	4	178	115	S. Riggins, J. Nuester
BIOL 323	1	62	59	R. Brunelli
GEOS 330Z	2	48	39	E. Willard, D. Brown
BIOL 334	1	36	31	S. Kirn
ECON 365	1	37	30	J. Stone
PHYS 376Z	1	28	24	S. Omarzai
NSCI 300Z	1	23	19	S. Omarzai
TOTAL	17	785	631	

Table 3—Enrollments and participation among upper-Division GE pathway courses in QLRA

RESULTS

LOWER DIVISION ASSESSMENT – SPRING 2015

The results of the lower division assessment are shown in Tables 4 and 5, below, and represented graphically in Chart 1 (A and B).

Question	1	2	3	4	5	6
Pre %	30.3% (n=892)	78.1% (n=892)	74.7% (n=892)	41.3% (n=892)	44.1% (n=892)	36.9% (n=892)
Post %	49.9% (n=794)	94.5% (n=794)	98.2% (n=734)	88.7% (n=860)	89.2% (n=859)	55.9% (n=860)

Table 4—Pretest and posttest results from the math quiz administered to students of MATH 105, spring 2015.

Statistic	Mean Score (out of 6)	Median (out of 6)	Std. Dev.
Pre	3.05	3	1.42
Post	4.82	5	1.01

Table 5—Summarized pretest and posttest results from the math quiz administered to students of MATH 105, spring 2015.

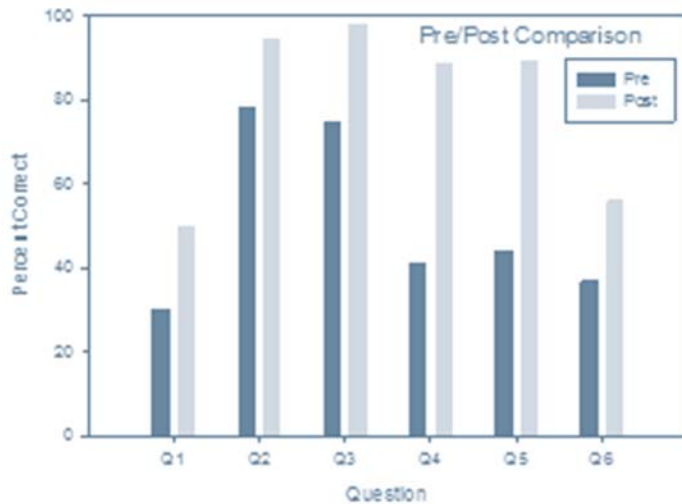


Chart 1-A: Scores on Individual Questions

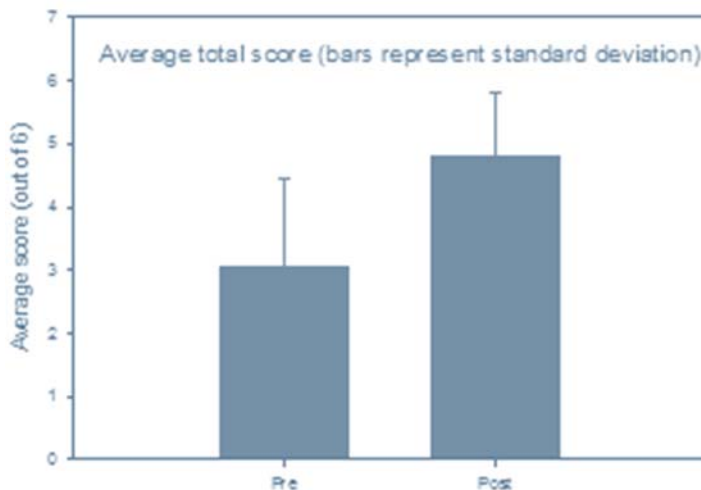


Chart 1-B: Average total score (out of 6)

Students' scores improved on each of the 6 questions they were given (Table 4). The greatest improvements were for questions 4 and 5, which went from below 50% correct (on a question with only two choices) to over 88% correct on each question.

The average total score went from 3.05 out of 6 (52%) to 4.82 out of 6 (81%) (see Table 5). Based on the standard deviations, and by a student's t-test, this is considered a "highly statistically significant" improvement. ✓



UPPER DIVISION ASSESSMENT – FALL 2015

The results of the upper division assessment are summarized in Table 6, below.

Class Enrollment	785
Number of surveys	631
Average Correct	6.51
Average incorrect	11.71
Average not answered	1.78
Std. Dev. (Correct)	3.18

Table 6—QLRA results

Data were collected and tabulated at the end of the fall semester of 2015. As results were tabulated, scores were reported immediately to instructors so that they could allocate appropriate credit. QRLA scores are reported by item in Appendix C.

Because the QLRA instrument captured student identifiers, it is possible to summarize the data according to various student demographics stored in the University's enrollment database. The following tables summarize the results by college enrollment (Table 7), class standing (Table 8), remediation requirement (Table 9), and native versus transfer status (Table 10).

COLLEGE	Average Correct	<i>n</i>
Agriculture	5.0	9
Behavioral and Social Sciences	5.9	191
Business	7.5	99
Communication and Education	6.1	169
Engin, Comp Sci & Const Mgmt	10.4	14
Humanities and Fine Arts	6.7	35
Natural Sciences	7.0	105
Undergraduate Education	6.0	8
Grand Total	6.5	630

Table 7—QLRA results by college

CLASS LEVEL	Average Correct	<i>n</i>
Freshman	5.7	15
Sophomore	5.9	92
Junior	6.5	336
Senior	7.0	187
Grand Total	6.5	630

Table 8—QLRA results by class standing

REMEDICATION	Average Correct	<i>n</i>
English	6.0	29
English and Math	4.2	29
Math	5.1	35
None	6.8	537
Grand Total	6.5	630

Table 9—QLRA results by remediation requirement

STUDENT TYPE	Average Correct	n
01 - Continuing UG Students	6.7	427
03 - Returning Students	5.8	6
04 - UG Transfer Students	6.1	193
05 - First-Time Freshmen	5.0	1
09 - Transitory Students	4.7	3
Grand Total	6.5	630

Table 10—QLRA results by native vs transfer

In addition, results are cross tabulated and graphed for first generation status by college (Table 11, Chart 2), and Pell eligibility by college (Table 12, Chart 3).

COLLEGE	NO	YES
Agriculture	4.5	5.4
Behavioral and Social Sciences	6.5	5.5
Business	7.5	7.6
Communication and Education	6.5	6.0
Engin, Comp Sci & Const Mgmt	12.2	9.7
Humanities and Fine Arts	7.3	5.8
Natural Sciences	8.2	6.3
Undergraduate Education	6.7	4.0
TOTAL	7.1	6.1

Table 11—QLRA results for first generation status by college

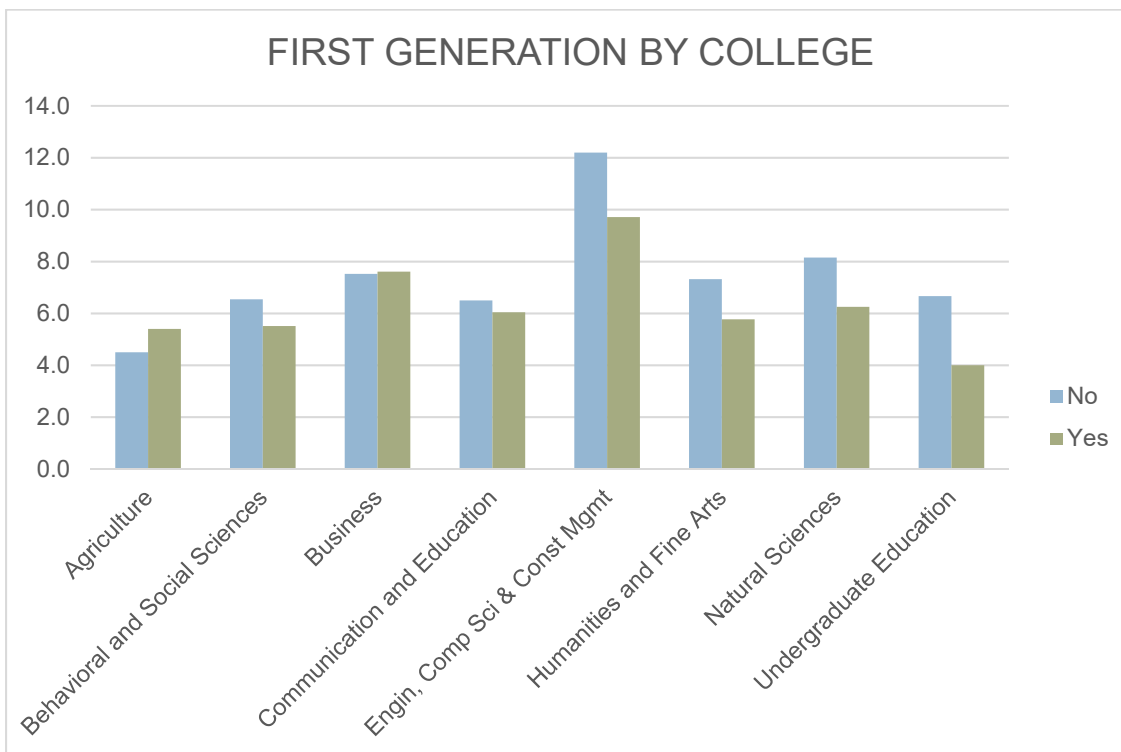


Chart 2—QLRA results for first generation status by college

COLLEGE	NO	YES
Agriculture	4.5	6.0
Behavioral and Social Sciences	6.6	5.1
Business	7.4	7.9
Communication and Education	6.3	5.9
Engin, Comp Sci & Const Mgmt	9.1	11.7
Humanities and Fine Arts	7.5	6.0
Natural Sciences	7.4	6.4
Undergraduate Education	6.7	4.0
TOTAL	6.8	6.1

Table 12—QLRA results for Pell eligibility by college

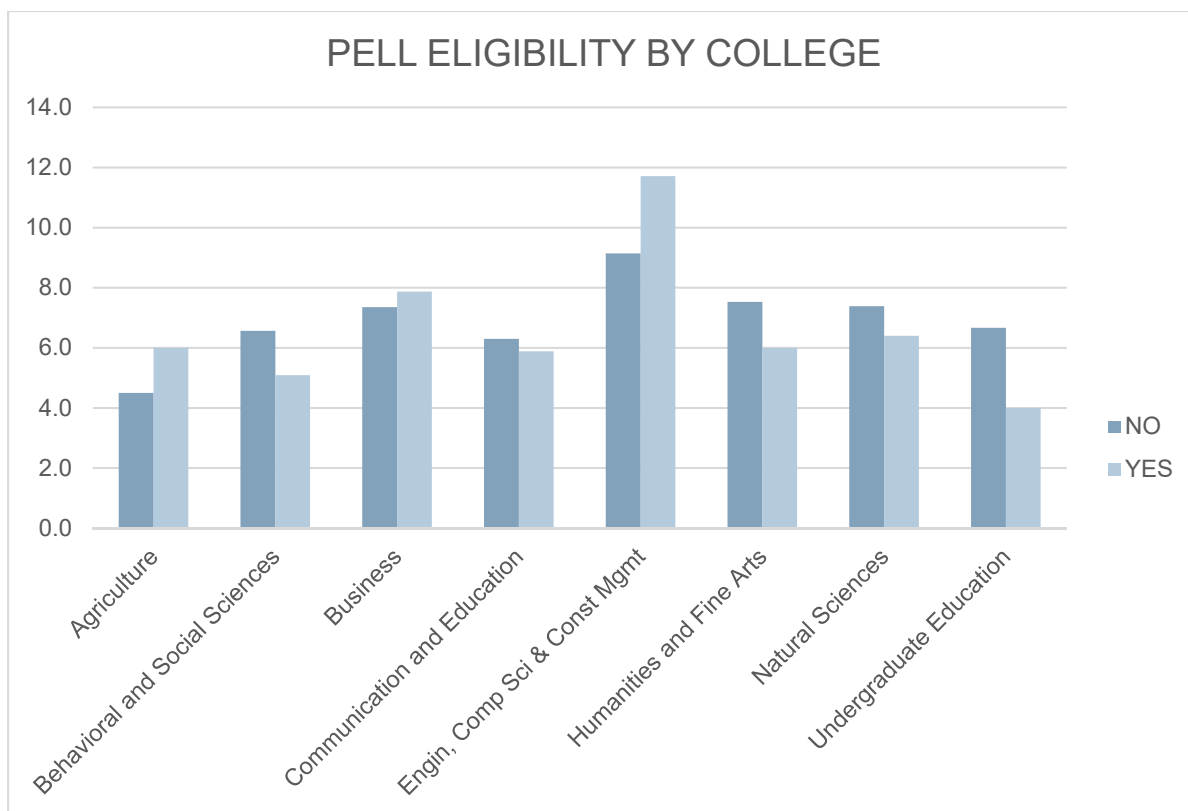


Chart 3—QLRA results for Pell eligibility by college



DISCUSSION

The purpose of this report is to summarize the findings that the Mathematical Reasoning Assessment Committee (in its two incarnations) achieved over nearly two years of activity. While at times frustrating, the work undertaken on behalf of CAB was by no means fruitless. The committee gained a better understanding of the quantitative reasoning skills of Chico State students at the same time as it developed a certain competence in the process of assessment of these skills. In this section, these findings will be discussed along with directions for further exploration.

What do the data tell us?

It does appear that when students are taking a math course, they are developing some quantitative reasoning skills. Based upon the assessment of learning in a foundation course over one semester, it is safe to conclude that students' understanding of statistical definitions and calculations did improve by a significant margin.

What we don't know is whether this learning lasts. Certainly, we have some evidence that this learning decays quite significantly based on the relatively poor performance on the QLRA by upper division students. However, until we can match student data samples by ID, we cannot say definitively that these skills are or are not lost.

When we compare the QLRA results to those of other institutions, the data are mixed. Overall, Chico State students performed worse than students at other universities who took the assessment in 2012 and 2013. However, when compared against nonselective institutions (arguably a more comparable group), Chico State students underperformed those from 2012 while they outperformed those from 2013. It is possible that this shift has less to do with Chico State students than with the composition of the overall pool of institutions that were drawn on in the 2013 data (see Table 13).

Samples	Average Correct	<i>n</i>
ALL (2012)	11.1	1659
Nonselective (2012)	8.7	334
ALL (2013)	9.3	2172
Nonselective (2013)	6.0	811
Chico (2015)	6.5	631
TOTAL	8.3	5607

Table 13—QLRA results for all institutions nationwide (2012, 2013), nonselective institutions (2012, 2013), and Chico State 2015)

How should we measure quantitative reasoning?

In addition to what we learned about the quantitative reasoning capabilities of our students, we also learned quite a bit about the process of assessment within general education. We began our assessment efforts with the intention of using qualitative measures. The plan was to collect a broad cross-section of student work samples that would be assessed using the AAC&U VALUE rubric for quantitative reasoning. We considered the possibility of adapting this rubric to work with evidence coming from a wide range of disciplines within upper division GE at Chico State.

Unfortunately, we didn't get very far before encountering significant faculty resistance. It was not possible to adapt the VALUE rubric to local evidence when so little evidence was available. In hindsight, we probably should have abandoned this approach early on; collecting actual student work was unlikely to become any easier in the spring than collecting assignments and exemplars was in the fall.

In terms of faculty acceptance, asking them to give up 30 minutes or one class period is clearly superior to asking them for student work samples. It is somewhat counterintuitive, but our committee learned that faculty teaching mathematical or quantitative reasoning ask for digital submissions of student work much more rarely than do faculty who are teaching written communication or critical thinking. So, one barrier to the collection of work samples relates to the burden of collecting and reproducing this work for assessment. Moving to the quiz was well-received by faculty. Participation increased from less than 10% to over 60% of faculty teaching mathematical reasoning in upper division GE.

Apart from faculty acceptance, two other issues with our assessment process merit discussion: (1) ensuring consistent motivation for students and (2) overall instrument quality.

A challenge for non-embedded assessment, such as a math quiz, is assuring that students are motivated to perform at their best. It is important therefore to offer incentives that are consistent across courses and tied to the level of performance. Data from our assessment may indicate some inconsistencies in the levels of incentives provided by faculty. Specifically, there was a marked decline in completion of questions in the later portion of the quiz (see Chart 4).

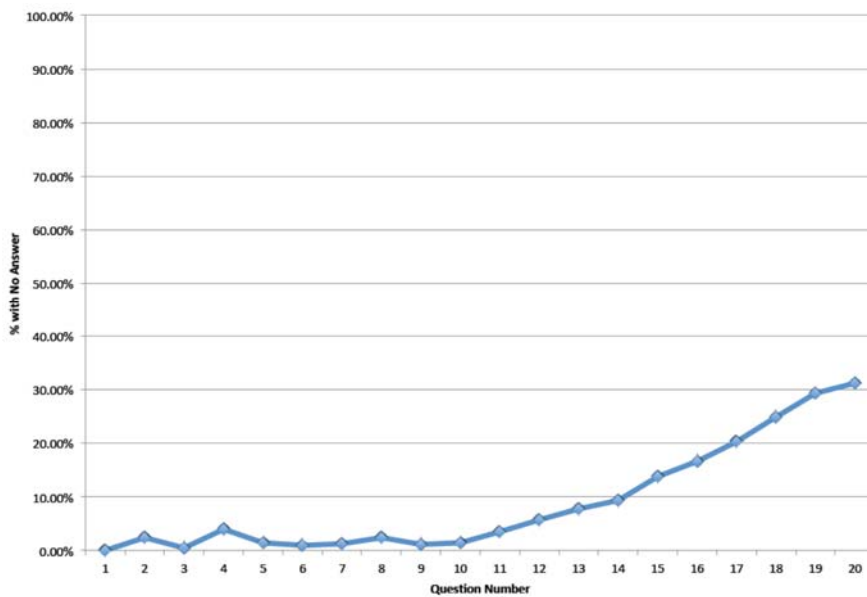


Chart 4—Rise in percentage of unanswered questions in QLRA results

Selecting a high-quality instrument for the assessment of quantitative reasoning is critical to developing meaningful and actionable results. While Eric Gaze and his colleagues have developed a reliable, internally consistent, and valid instrument for measuring quantitative reasoning in their QLRA, it is not without limitations. In particular, the length of the instrument, while allowing administration within a 30-minute window, does not provide enough data to perform discrimination amongst traits that comprise quantitative reasoning. In a recent paper, they acknowledge this limitation: “Prior to administering the test, the 23 items on the 2012 test were divided into four subscales: Number Sense (NS), Visual Representation (VR), Probability and Statistics (PS), and Reasoning (R). Analysis of these areas did not reveal anything worth pursuing in this project.”¹

At the same time, a longer instrument that might provide greater ability to discriminate amongst the traits of quantitative reasoning might discourage faculty participation. In addition, there is often a cost associated with high quality instruments, and the QLRA is currently available for free.

¹ Gaze, E. et al. “Towards Developing a Quantitative Literacy/Reasoning Assessment Instrument,” *Numeracy*, Vol. 7 [2014], Iss. 2, Art. 4, pp 6-7. <http://scholarcommons.usf.edu/numeracy/vol7/iss2/art4>

Appendix A – Lower division assessment instrument

COUNTING UP

THE ASSESSMENT OF MATHEMATICAL
REASONING WITHIN GENERAL EDUCATION AT
CHICO STATE, 2014-2015

APPENDIX A: Math 105 pre-test (correct answers circled)

1. The ages for 200 members of a high school class are given in the frequency table to the right. The **mean** is equal to (circle one):

- (a) 15 (b) 15.4 (c) 16 (d) 16.5 (e) 17

Age	Frequency
14	65
15	56
16	41
17	18
18	10
19	10

2. A study in Switzerland examined the number of caesarian sections (surgical deliveries of babies) performed in a year by doctors. The data are reported below.

20 25 25 27 28 31 32 33 34 36 37 39

The **median** is:

- a) 30.6 b) 31 (c) 31.5 d) 32 e) 33.5

3. In Denali National Park, Alaska, the wolf population is dependent on a large, strong caribou population. In this wild setting, caribou are found in very large herds. It is thought that wolves keep caribou herds strong by helping prevent over-population. Let x represent the number of fall caribou herds and y represent the late winter wolf population in the park. A random sample of recent years gives the following results:

x	31	34	27	25	17	23	20
y	75	85	75	60	48	60	60

The least-squares regression equation is $\hat{y} = 15.35 + 2.0x$

What is the predicted number of wolves in the park when there are 34 caribou herds in Denali National Park?

- a) 2.0 b) 15.35 c) 68 (d) 83.35 e) 92

4-5: The significance level and p-value of a hypothesis test are given. Decide whether the null hypothesis should be rejected.

4. (2) $\alpha=0.01$, p-value=0.03

- a) Reject the null hypothesis (b) Fail to reject the null hypothesis

5. (2) $\alpha=0.05$, p-value=0.025

- (a) Reject the null hypothesis b) Fail to reject the null hypothesis

6. (2) Suppose a researcher **rejects** the null hypothesis $H_0: \mu=104$ in favor of $H_1: \mu>104$ at $\alpha=.05$. Which of the following 95% confidence intervals is consistent with the results of the hypothesis test.

- a) (103,107) (b) (105,110) c) (99,103)

Appendix B – QLRA question conceptual coverage

The QLRA researchers organized the questions into four subscales: Number Sense (NS), Visual Representation (VR), Probability and Statistics (PS), and Reasoning (R). Analysis of these areas has not been fruitful thus far. The questions were further coded by researchers¹ into 24 quantitative literacy/reasoning concepts show in Table B-1, below.

Key		Que	Question Coding																							
			PR	F	%	C	UC	T	R	Rnd	Sc	G	VR	Geo	A	LG	P	D	PC	H	M	St	L	Md	BC	Rsn
F	Fractions	1	1	1																						
%	Percentages	2	1		1																					
C	Complement	3		1		1																				
UC	Unit/Conversion	4	1				1	1	1																	
T	Tables	5	1				1		1																	
R	Rates	6		1						1																
Rnd	Rounding	7			1		1				1															
Sc	Scale/Large #	8			1			1				1														
G	Graph/ic	9									1	1														
VR	Verbal Representati	10		1							1															
Geo	Geometry	11	1								1		1													
A	Algebra	12												1												
LG	Linear Growth	13					1		1		1	1			1											
P	Probability	14							1							1										
D	Decimal	15	1	1	1												1	1								
PC	Pie Chart	16	1														1									
H	Histogram	17			1							1							1							
M	Mean	18		1	1												1	1								
St	Statistics	19										1								1	1	1				
L	Logic	20			1														1						1	
Md	Median	21																					1			
BC	Bar Chart	22																				1		1		1
Rsn	Reasoning	23	1		1							1														1
			8	6	8	1	4	2	4	1	1	4	6	1	1	2	3	2	2	1	1	2	1	1	1	2

Table B-1—QLRA question coding

¹ From the QLRA online resource portal developed and maintained by the Science Education Resource Center at Carleton College - <http://serc.carleton.edu/qlra/coding.html>

Appendix C – QLRA results by item

COUNTING UP

THE ASSESSMENT OF MATHEMATICAL
REASONING WITHIN GENERAL EDUCATION AT
CHICO STATE, 2014-2015

GLOBAL										
Question	A	B	C	D	E	Omitted	TOTAL	% correct	% incorrect	% omitted
1	384	13	10	45	179	0	631	28.37%	71.63%	0.00%
2	16	113	61	48	378	15	631	17.91%	79.71%	2.38%
3	427	36	22	7	136	3	631	67.67%	31.85%	0.48%
4	76	71	95	140	224	25	631	35.50%	60.54%	3.96%
5	66	135	184	177	60	9	631	29.16%	69.41%	1.43%
6	29	55	392	129	20	6	631	62.12%	36.93%	0.95%
7	55	52	169	95	252	8	631	39.94%	58.80%	1.27%
8	122	144	169	67	114	15	631	22.82%	74.80%	2.38%
9	145	75	260	81	63	7	631	41.20%	57.69%	1.11%
10	26	39	346	210	1	9	631	33.28%	65.29%	1.43%
11	20	35	482	67	5	22	631	76.39%	20.13%	3.49%
12	59	119	262	84	71	36	631	11.25%	83.04%	5.71%
13	37	118	96	156	175	49	631	27.73%	64.50%	7.77%
14	61	251	133	74	53	59	631	39.78%	50.87%	9.35%
15	71	247	132	43	51	87	631	39.14%	47.07%	13.79%
16	343	49	56	32	46	105	631	7.29%	76.07%	16.64%
17	113	72	108	142	68	128	631	17.91%	61.81%	20.29%
18	46	168	103	139	18	157	631	22.03%	53.09%	24.88%
19	39	89	209	80	29	185	631	14.10%	56.58%	29.32%
20	66	116	94	109	49	197	631	17.27%	51.51%	31.22%
CORRECT	4107					6.51		6.51	11.71	1.78
INCORRECT	7391					11.71				
OMIT	1122					1.78				