Teaching and Learning Project Assessment Report

Program or Unit: Transfer Math Program
Submitted by: Kwadwo Poku, Nick Knueppel, and Shaw Matthews
Date: January 25, 2009

What we wanted to learn about our students:

Background: New accreditation standards for community colleges require the assessment of learning at three levels: course, program, and degree. As part of program review, the Math Department was required to define Student Learning Outcomes for the Transfer Math Program and design an assessment plan. We choose to begin our assessment plan with a focus on Statistics since so many of our transfer-bound students take this course.

This project began in spring 2007 with the goal of developing course-level SLOs for Math 34 and updating the course outline. The Math Department offered a series of eleven 3-hour retreats over the next year that were attended by 11-14 statistics instructors. Activities included:

- Reading and discussing the implications of statistical education literature, specifically "How Students Learn Statistics" by Joan Garfield from the *International Statistical Review* and "Mathematics, Statistics, and Teaching" by Cobb and Moore from the *American Mathematical Monthly.*
- Collaboratively defining course SLOs aligned with Transfer Math Program SLOs
- Designing, sharing, and critiquing problems that elicit student work relevant to the SLOs
- Analyzing a national exam for introductory statistics courses
- Rewriting the course outline to reflect the agreements reached by the group and revising the course outline based on feedback from math faculty at the main campus and Brentwood
- Collaboratively writing a rubric that captured agreed-upon standards for assessing student work

This project assessed student attainment of the following Transfer Math Program-level Student Learning Outcomes for students completing a transfer-level introductory Statistics course (Math 34):

1. **Mathematical Literacy:**
   Communicate using mathematics:
   - Clearly articulate mathematical information accurately and effectively, using a form, structure and style that suit the purpose (including written and face-to-face presentation).

2. **Problem-solving ability:**
   - Reason with and apply mathematical concepts, principles and methods to solve problems or analyze scenarios in real-world contexts relevant to their major;
   - Use technology effectively to analyze situations and solve problems;

3. **Modeling ability:**
   - Construct and interpret mathematical models using numerical, graphical, symbolic and verbal representations with the help of technology where appropriate in order to draw conclusions or make predictions;
   - Recognize and describe the limits of mathematical and statistical methods.

To assess these program-level learning outcomes, we focused on analyzing student work on two parts of the final exam that addressed the following course-level learning outcomes:

Statistical Literacy (PSLOs: literacy and problem-solving)

CSLO 1: Based on statistical reasoning and supported by critical thinking, students should be able to read and critique simple statistics-based studies in order to make an informed judgment on the reliability of the statistical presentation or argument.

Data Production

CSLO 2: Students should be able to apply the basic principles of study design to develop and analyze the validity of simple experiments and sampling plans related to a given situation and goal.

Data Exploration and Representation (PSLOs: modeling and communication)

CSLO 3: Students will be able to examine raw data using graphical, tabular, and analytical exploratory tools in order to investigate and describe patterns in data with the goal of describing shape, center, and spread within a quantitative data set, making comparisons among data sets, and looking for relationships between data sets.

Modeling and Inference (PSLOs: modeling and problem-solving)

CSLO 4: Students will analyze data to identify an appropriate statistical model, use technology to perform statistical tests or find confidence intervals, explain the concepts underlying inference, and interpret results in a context. Students will also use correlation coefficients and scatterplots to determine if a linear regression model is appropriate, then find, use, and interpret linear regression models when appropriate.

The Role of Probability in Inference (PSLOs: modeling and problem-solving)

CSLO 5: Students will be able to explain in layman's terms how variability and probability are connected to statistical inference, as well as be able to interpret and apply basic laws and concepts of probability to sampling distributions.

What we did to assess student learning:

In Fall 2007 five instructors submitted seven class sets of student work on a common final exam problem. This included 4 of the 5 instructors teaching Math 34 on the main campus (comprising 5 of 6 sections), 2 of 4 instructors teaching at Brentwood (comprising 2 of 5 sections; note that one instructor was teaching at both sites.). We selected a random sample of 50 papers from the 139 papers submitted. The sample contained approximately 35% of the students from each section.

In Spring 2008 our instructors submitted six class sets of student work on a common final exam problem. We selected a random sample of 30 papers from the 117 papers submitted.

Each final exam was assessed holistically relative to each outcome using a rubric written collaboratively by 11 faculty members. For each outcome we conducted a benchmarking exercise in which each instructor graded the same paper. We then discussed the scores and reached consensus. Next, for each outcome each final was assessed independently by two instructors and the scores were averaged. If the facilitating instructor judged that two scores differed by more than was useable, that student's work was assessed by a third instructor. The closest two scores were then averaged. 8 instructors participated in the grading and one facilitated.

In addition to the holistic scoring of written work on the final, we also analyzed the results of our students' performance on a 40-question multiple choice test, called the Comprehensive Assessment of Outcomes in a first Statistics Course test (CAOS). This test was written by a group of statistical educators and statistical education researchers collaborating on an NSF project. The group's goal is to "help teachers assess statistical literacy, statistical reasoning, and statistical thinking in first courses in statistics. Their website is https://app.gen.umn.edu/artist/index.html. The acronym ARTIST stands for Assessment Resource Tools for Improving Statistical Thinking. A report of normative statistics for the CAOS test based on a sample of 1470 undergraduate students enrolled at 33 United States institutions who took the CAOS test in Fall 2005 or Spring 2006 is available on their website. We used this report to analyze our students' performance.

In Fall 2007, five instructors submitted class sets of student work on the CAOS exam. The sample size for this analysis was 100 students (all of the work submitted.) In Spring 2008, six instructors submitted class sets of student work on the CAOS exam. The sample size for this analysis was 117 students (all of the work submitted.)

What we learned about our students:

CSLO 1: Statistical Literacy
Relevant CAOS questions: 1, 11-13, 19, 23-31, 33

	Percent correct						
	Min	Q1	Med	Q3	Max	Range	Mean
LMC sample Fall 2007, n=100	37.0%	45.5%	54.0%	74.5%	85.0%	48.0%	58.0%
LMC sample Spring 2008, n=117	38.5%	54.7%	61.5%	76.9%	88.9%	50.4%	64.4%
National sample n=1470	41.2%	55.8%	65.4%	74.1%	89.0%	47.8%	64.7%

Term	Number in LMC sample	Correct responses	Total possible
Fall 2007	100	870	1500
Spring 2008	117	1130	1755

A two-proportion z-test of the data sets showed an improvement between the two terms that was statistically significant at the 0.05 level. There is statistically significant evidence that the Spring 2008 proportion of correct responses was greater than the Fall 2007 proportion of correct responses.

Analysis of the problem (part c) on the final exam:
Fall 2007: 10/52 = 19% proficient (score of 3 or greater on the rubric)
Spring 2008: 7/30 = 23% proficient (score of 3 or greater on the rubric)
This improvement was not significant at the 0.05 level. There was not significant evidence that the proportion of proficient student performance on part c of the final problem in Spring 2008 was greater than the Fall 2007 proportion of proficient student performance.

CSLO 2: Data Production
Relevant CAOS questions: 7, 34-35, 38

	Percent correct						
	Min	Q1	Med	Q3	Max	Range	Mean
LMC sample Fall 2007, n=100	7.0%	22.0%	31.5%	43.3%	65.0%	58.0%	33.8%
LMC sample Spring 2008, n = 117	7.7%	29.5%	41.5%	50.4%	63.2%	55.6%	38.5%
National sample n=1470	14.7%	31.7%	42.2%	52.5%	69.2%	54.5%	42.1%

Term	Number in LMC sample	Correct responses	Total possible
Fall 2007	100	135	400
Spring 2008	117	180	468

A two-proportion z-test of the data sets showed that the improvement between the two terms that was not significant at the 0.05 level. There was not significant evidence that the Spring 2008 proportion of correct responses was greater than the Fall 2007 proportion of correct responses.

Analysis of the problem (part c) on the final exam:
Fall 2007: 14/52 = 27% proficient (score of 3 or greater on the rubric)
Spring 2008: 14/30 = 47% proficient (score of 3 or greater on the rubric)
This improvement was statistically significant at the 0.05 level. There was statistically significant evidence that the proportion of proficient student performance on part c of the final problem in Spring 2008 was greater than the Fall 2007 proportion of proficient student performance.

CSLO 3: Data Exploration and Representation
Relevant CAOS questions: 1-15, 18, 20-21, 33, 36

	Percent correct						
	Min	Q1	Med	Q3	Max	Range	Mean
LMC sample Fall 2007, n=100	6.0%	30.3%	48.0%	74.3%	95.0%	89.0%	49.9%
LMC sample Spring 2008, n = 117	7.7%	35.3%	44.9%	74.4%	94.0%	86.3%	51.4%
National sample n=1470	28.9%	51.8%	63.8%	77.1%	93.5%	64.6%	62.8%

Term	Number in LMC sample	Correct responses	Total possible
Fall 2007	100	997	2000
Spring 2008	117	1202	2340

A two-proportion z-test of the data sets showed an improvement between the two terms that was not significant at the 0.05 level. There was not significant evidence that the Spring 2008 proportion of correct responses was greater than the Fall 2007 proportion of correct responses.

Analysis of the problem (part a) on the final exam:
Fall 2007: 16/52 = 31% proficient (score of 3 or greater on the rubric)
Spring 2008: 16/30 = 53% proficient (score of 3 or greater on the rubric)
This improvement was statistically significant at the 0.05 level. There was statistically significant evidence that the proportion of proficient student performance on part c of the final problem in Spring 2008 was greater than the Fall 2007 proportion of proficient student performance.

CSLO 4: Modeling and Inference
Relevant CASO questions: 21-32, 39, 40

	Percent correct						
	Min	Q1	Med	Q3	Max	Range	Mean
LMC sample Fall 2007, n=100	18.0%	37.0%	45.0%	60.0%	80.0%	62.0%	48.0%
LMC sample Spring 2008, n = 117	18.8%	46.2%	54.7%	67.7%	76.9%	58.1%	53.5%
National sample n=1470	18.6%	50.5%	55.9%	64.5%	83.3%	64.7%	55.5%

Term	Number in LMC sample	Correct responses	Total possible
Fall 2007	100	672	1400
Spring 2008	117	877	1638

A two-proportion z-test of the data sets showed an improvement between the two terms that was statistically significant at the 0.05 level. There was significant evidence that the Spring 2008 proportion of correct responses was greater than the Fall 2007 proportion of correct responses.

Analysis of the problem (part b) on the final exam:
Fall 2007: 13/52 = 25% proficient (score of 3 or greater on the rubric)
Spring 2008: 16/30 = 25% proficient (score of 3 or greater on the rubric)
This improvement was statistically significant at the 0.05 level. There was statistically significant evidence that the proportion of proficient student performance on part c of the final problem in Spring 2008 was greater than the Fall 2007 proportion of proficient student performance.

CSLO 5: Probability in Inference
Relevant CASO questions: 16-19, 25-31, 34, 35, 37

	Percent correct						
	Min	Q1	Med	Q3	Max	Range	Mean
LMC sample Fall 2007, n=100	9.0%	37.0%	44.5%	61.3%	76.0%	67.0%	47.4%
LMC sample Spring 2008, n = 117	12.8%	45.5%	57.3%	62.8%	76.9%	64.1%	52.6%
National sample n=1470	22.4%	48.0%	55.8%	67.3%	80.0%	57.6%	55.8%

Term	Number in LMC sample	Correct responses	Total possible
Fall 2007	100	664	1400
Spring 2008	117	861	1638

A two-proportion z-test of the data sets showed an improvement between the two terms that was statistically significant at the 0.05 level. There was significant evidence that the Spring 2008 proportion of correct responses was greater than the Fall 2007 proportion of correct responses.

Analysis of the problem on the final exam:
Fall 2007: 21/52 = 40% correctly interpreted P-value
Spring 2008: 18/29 = 62% correctly interpreted P-value
This improvement was statistically significant at the 0.05 level. There was statistically significant evidence that the proportion of proficient student performance on part c of the final problem in Spring 2008 was greater than the Fall 2007 proportion of proficient student performance.

General observations:

Both of our indicators, the CAOS test and the final problem, showed that our students improved in all 5 areas. The two indicators did not always agree that the improvements were statistically significant. However, at least one of the indicators showed statistically significant improvement for each of the 5 course-level learning outcomes. The exceptions were that the final problem did not show significant improvement for CSLO 1 and the CAOS test did not show significant improvement for CSLO 2.

In an attempt to identify more precisely the troublesome areas for our students, we analyzed all of the questions on the CAOS test for which LMC student performance fell below the national performance by more than two standard errors, i.e. questions for which the LMC performance was outside of the 95% confidence interval based on national performance. The general trends we noticed in missed problems include the following:

CSLO 3: Data exploration and representation

(1) interpreting histograms and boxplots (e.g. CAOS problems 2-5, 8-10): connections between different graphical representations of data, connections between graphical representations and standard deviation
(2) understanding standard deviation as a measure of spread (e.g. CAOS problems 8, 14)
(3) this area remains a troublesome area this term

CSLO 2: Data production - improved

(4) data production: purpose of randomization and sampling design (e.g. CAOS problems 7, 38 were identified as problem areas last term)
(5) only #7 remained outside of the 95% confidence interval this term

CSLO 4 and 5: Inference and probability - improved

(6) interpreting the results of inference: meaning of statistical significance, P-value, and confidence level (e.g. CAOS problems 23, 24, 25, 27, 28, 30 were identified as problem areas last term)
(7) only CAOS problem 28 remained outside of the 95% confidence interval this term

Students' written work on the common problem of the final highlighted difficulties in the following areas:

CSLO 3 Data Exploration and Representation: this area improved somewhat, but the majority of students in the sample still either chose an inappropriate graph, drew a graph that did not illustrate the distribution of the variable, or had problems with accuracy when constructing their graphs. Many students did not address the tasks noted in part a. They did not justify their choice of numerical summary or did not include a description of patterns in the data.

CSLO 1 and 2 Data production and Statistical Literacy: this area improved, but the majority of students still did not give suggestions for improving the study that addressed fundamental issues of quality of data production, such as selecting a random sample or controlling for factors that may be confounding the study.

CSLO 4 Modeling and Inference: the majority of students omitted or inaccurately performed portions of the significance test, such as not accurately stating the hypotheses symbolically or in words, not verifying the conditions for the test, performing the wrong test, choosing the wrong conclusion based on their P-value.

General: There is some anecdotal evidence that students find the step-by-step format of the Final Question confusing, Some students were able to correctly perform a complete test of significance in a problem which allowed them to provide the format for the problem themselves, but performed poorly within the step-by-step format of the Final Problem.

What we plan to do next to improve student learning:

The group of 9 instructors participating in the assessment made the following recommendations

(1) Incorporate more opportunities within the course for students to practice exploratory data analysis, including more work on interpreting graphs, seeing connections between graphical representations, and being able to work with the standard deviation in different situations.

(2) Focus on statistical literacy. Provide more opportunities for students to analyze articles and other real world statistical artifacts in which questions about data production, interpreting P-values and statistical significance can be explored more frequently and more deeply.

The group also recommends that the department do the following to address the above recommendations:

(1) conduct an additional assessment using student work on the common final from SP 09
(2) continue to offer statistics retreats in 2008-2009 for faculty to focus on the teaching and learning of introductory statistics with a focus on exploratory data analysis and statistical literacy
(3) continue to use the Blackboard for faculty to share course materials with additional emphasis on in-classroom techniques aimed at data exploration and representation
(4) discuss and consider ways of determining if the *format* for the Final Problem is giving students trouble, perhaps by allowing the students a choice of formatted or unformatted versions of the problem on the final.