An Overview of the Mathematics Departmental Final and Data Analysis

1. Why does the Mathematics Department administer common final exams for developmental mathematics courses?

The Mathematics Department established formal departmental goals and objectives in 2000-01, one of which was to increase the cohesiveness of the developmental mathematics program in order to ensure that students who successfully complete developmental mathematics courses (Math 20, Math 35, Math 45, Math 60 and Math 70) exit with a more consistent skill level. (attachments 1 and 2) The Department asserted that achievement of this goal would translate into increased student success in subsequent mathematics courses. The School of Mathematics, Science and Engineering (MSE) and the Mathematics Department identified objectives that would contribute to achieving this goal, including establishment of departmental final exams for all developmental mathematics classes.

Prior to implementation of the departmental final and with upwards of 100 full-time and adjunct faculty teaching mathematics, the mathematics faculty observed that the differences between one instructor's class and another instructor's class were often quite significant in terms of content emphasis, degree of rigor, and grading standards. Over 3400 students enroll in developmental mathematics classes each semester, so potentially many students could be affected by this variation. A large percentage of these students continue on to the next sequential mathematics class, where they would be expected to have the entering skills necessary for success.

For the benefit of the students, the Mathematics Department agreed to create a common standard, the departmental final, so that all instructors would have a clear guideline about departmental standards and expectations, what content to emphasize, and the level of difficulty of problems for each topic. Every step in the implementation of this enormous project was the result of research, deliberation, consensus, implementation, analysis, evaluation, and revision. The William and Flora Hewlett Foundation recently recognized SWC as Leaders in Student Success, specifically citing the mathematics departmental final as an example of faculty-driven strategies that have strengthened the mathematics curricula. (attachment 3)

Management of the mathematics department finals is a faculty-run operation. The process is year-round and impacts about 12,000 students per year. The photo-copying, packaging, instructions, and solution keys are entirely overseen by a committee of mathematics faculty. In addition, this team re-writes the Math 20, Math 35, Math 45, and Math 60 and Math 70 finals each year. The number of hours involved in this operation is large - and the work is done without reassigned time or stipends.

The departmental final has gone through several modifications since it was first implemented in fall 2000. It began as a free response exam, followed by a free response exam given to all students in a course at the same time with grading done on a departmental level, followed by the multiple choice format currently in place. Instructors are not required to use the departmental final exam grade in their calculation of course grades and may choose to administer a different final exam of their own making for that purpose. However they are required to administer the departmental

final exam to their students and to return all materials for data collection. Instructors are permitted to weight the value of the departmental exam, if they use it, in any way that they please, although the department recommends a range of 10-25% of course grade.

2. How is student content knowledge, skill, and critical thinking measured?

The departmental final measures student content knowledge, skill, and critical thinking via 40 to 50 multiple choice questions with five possible responses from which students choose. The questions correspond to specific learning objectives contained in the course outline. Students choose the best answer and fill in the appropriate oval on the NCS form provided. NCS was chosen because it provides the complexity and capacity to meet our data-processing needs. It's not Scantron, but it's based on the same optical scan technology as Scantron. The forms are run through a machine that compares the student responses to a standard exam key.

Some of the benefits of the multiple choice assessment of student content knowledge, skill, and critical thinking include:

- It ensures that all instructors use the same criteria and method to evaluate students. The resulting scores can be legitimately compared student to student, section to section, instructor to instructor, term to term, short-course to full-term course, etc.
- Scores can be determined quickly, so that students' final grades can be calculated quickly and reported to the office of admissions quickly. This permits student prerequisites to be confirmed in a timely way before the start of the next term.
- Many professions (as well as transfer institutions) require students to demonstrate mastery of a subject matter by taking multiple-choice exams: law school, medical school, the foreign (diplomatic) service of the US government, even the military. The SWC mathematics department provides workshops to assist students in taking multiple-choice exams so that they will be successful not only on the departmental final exams, but also in these other exams they may take.
- Data collection is automated, so that we can provide the necessary documentation to WASC for accreditation of SWC.
- Students can be taught to prepare effectively for a multiple choice exam by learning to be careful about all the details and to identify the distractors (wrong answers) that have been designed into the test.

3. What key issues needed to be addressed prior to implementing the departmental finals?

Faculty expressed reservations about how the data from the departmental final would be used, thus prior to implementation it was agreed:

- only aggregate data from the departmental finals would be accessible to the Dean, Chair, and peer discipline faculty
- data from the departmental finals relating to individual faculty members would be made available only to that individual
- data from the departmental final would not be used to evaluate faculty members

Assurance of confidentiality and trust was a non-negotiable element for the department. Thus it was agreed that the data-analysis should be conducted by an outside entity rather than by an in-house faculty member or administrative unit.

4. How is student performance on the Mathematics Department Finals analyzed?

The Mathematics Department was allocated Basic Skills Initiative (BSI) funding to outsource the analysis of student performance on the departmental finals. R & K Enterprises (R & K) was chosen because of their long-standing track record analyzing student performance for the Sweetwater Union High School District. At the completion of the project, the analysis of student performance on the mathematics departmental finals will be completed for two academic years: 2007-08 and 2008-09.

The data analysis for the mathematics departmental finals, which involves thousands of students, was out-sourced to R & K for several important reasons:

- Capacity to handle the magnitude of the workload
- Quick turn-around time
- Experience with similar projects
- Responsiveness to requests
- Quality of deliverables
- Reliability of service
- Cost-effectiveness
- Security and confidentiality

The purpose of the data analysis is:

- To provide faculty with a means to self-assess, reflect upon, and improve their teaching effectiveness using student data specific to the topic taught in conjunction with aggregate course data
- To incentivize improvements in teaching
- To evaluate student achievement of mathematics skills, specifically those skills listed in course objectives
- To discover areas that should be targeted in a more systematic (e.g. course-wide) way to improve student achievement of mathematics skills
- To improve the test and test results by aligning exam questions with course objectives and by detecting and correcting poorly written questions, miscoded answer keys, cultural bias, etc.
- To meet accreditation standards requiring that the department final be validated

Confidential individual instructor data reports that include average student performance scores specific to the section taught in conjunction with aggregate course data and item analysis have been provided to all instructors for every term (including summers) since R &K began processing our data in Spring 2006. These data reports were expanded in Spring 2008 to include item analysis with course objectives noted, as well as percentage success on each course objective and a histogram of section averages with instructor's average noted.

The following data analysis is provided to each faculty member teaching a specific mathematics course in a sealed envelope (attachment 4):

- Histogram Analysis:
  - Histogram of individual student scores in each section taught
  - Aggregate histogram of student scores in all sections of course
  - Aggregate histogram of section averages of all sections of course.

  Section averages of each instructor are noted with grey bars and the section average for a specific instructor is noted with a black bar
- Item Analysis:
  - Item analysis of student responses to each question in each section taught with corresponding objective from course outline noted
  - Aggregate item analysis of student responses to each question in all sections of course with corresponding objective from course outline noted
- Objective Analysis
  - Analysis of course objectives in each section taught with percent correct student responses to aggregate of test questions that correspond to each objective
  - Aggregate analysis of course objectives of all sections of course with percent correct student responses to aggregate of test questions that correspond to each objective
- Instructor Summary Report
  - Average score of section, course average, etc.
- Answer Sheet Tracking Report
  - The number of exams with pre-printed names and number of exams scored for each section

In addition, the department chair and dean are provided with the following data (attachment 5):

- Faculty Return Data for Test Materials
- Point Biserial Correlation

Point biserial correlation data has been provided for two administrations of the test – Summer 2006 and Spring 2008, and is expected for every future administration. Point biserial correlations measure the effectiveness of each test question compared to the overall test by indicating whether a question is more likely to be answered correctly by students who perform well on the test (implying that it's a good question) or more likely to be answered correctly by students who perform poorly on the test (implying that it's not a good question). The biserial correlation for a question is positive when students who did poorly on the whole test did poorly on the question and students who did well on the whole test did well on the question. Similarly, a negative result suggests that the question has some bias which is "tricking" poor students into correct answers, while trapping good students into giving wrong answers.

Basic statistics have been calculated for each version of each of the exams by faculty members Alex Juden and Melanie Branca. (attachment 6) Initially, these statistics exposed a procedural flaw in the processing of the keys provided to each instructor, and that flaw has since been repaired. The corrected data were suggestive of normal distributions with mean scores and standard deviations as expected. In addition, the maximum and minimum scores were as expected by the department. This statistical analysis is ongoing, each term.

More intensive statistical analysis was proposed by Branca and Juden, but has not yet been finalized and approved by the department. (attachment 7) For example, the Mathematics Department is considering two additional analyses in an effort to satisfy the Accreditation requirement to validate the exam:

- Cronbach Alpha Coefficient - to determine reliability or the extent to which an instrument consistently measures whatever it purports to measure.
- Factor Analysis – to analyze the correlations among the items and group together those that seem to be performing similarly.

5. What are the specific accreditation requirements for departmental finals and how have they been addressed?

The Western Association of Schools and Colleges (WASC) Accreditation Standard II. A. 2. g. states, "If an institution uses departmental course and/or program examinations, it validates their effectiveness in measuring student learning and minimizes test biases.

The effectiveness of the mathematics departmental final exam in measuring student learning will be validated in Spring 09. Mathematics knowledge will be measured by conducting pre- and post-tests using a state-approved assessment instrument in selected developmental math classes (five sections per course). An appropriate control group of students who took the pre-and post tests but did not enroll in a math class during the Spring 09 semester will be identified. Learning will be measured by determining whether there is a statistically significant difference between the pre- and post- test scores after adjusting for the control group. Thereafter we will measure the correlation between the post test results and the departmental final exam results. If the pre- and post-test results demonstrate learning and the post-test is reasonably correlated with the math departmental final, then we will conclude the math departmental final also measures learning. (attachments 8 and 9)

The mathematics department established content validity of their departmental final by assembling a committee of two to three experts (i.e. full time SWC mathematics faculty members) for each course for which a departmental exam has been written. One member had recently taught the course and one member did not usually teach the course. The committee members independently matched the items to the course outline, paying attention to both the content and the relative weight assigned to the content. Committee members made judgments as to which items should be added/deleted, which topics require additional items, and which topics are covered too heavily. Thereafter Committee members met to exchange ideas and come to consensus regarding the adequacy and appropriateness of the items relative to the course outline.

The Mathematics Department is still deliberating on whether and how to establish predictive validity. Establishment of predictive validity requires that test results are highly correlated with an independent measure purporting to measure the same construct.

In effort to minimize test biases of a single instructor's style or cultural background, all the departmental final exams are written by a team of two or more fulltime faculty members. In addition, all exams are checked by a logistics team of three additional

faculty members, who work all the problems on all of the exams, check the multiple-choice answers for clarify, accuracy, and lack of duplication, proofread the instructions for each question, and provide suggestions and feedback to exam writers where appropriate. After exam administration, feedback from all mathematics faculty, including adjunct faculty, is conveyed to the exam writers to further reduce bias toward a particular instructional or cultural approach and to ensure that the exam adheres strictly to the material described in the course outline.

The Mathematics Department identified six faculty from outside the Mathematics Department to evaluate the departmental final for cultural bias and report back on their findings in Spring 2009. (attachment 10) Each question on each departmental final exam was evaluated to assure its fairness and lack of offensiveness to people of different:

- Backgrounds
- Cultures
- Ethnicities
- Countries and languages of origin
- Genders
- Ages

6. What are the outcomes of the data analysis of departmental finals?

- Outcome #1: Improvement of teaching effectiveness
  Faculty surveys (~25% returned) using a Likert scale (1 – 5) indicated (attachment 11):
  - that faculty members would modify their classroom instruction as a result of reviewing the final exam statistical analysis. (ave 3.7, s.d 1.2)
  - the final exam statistical analysis clarified the concepts on which to focus classroom instruction. (ave 3.6, s.d 0.99)
  - the final exam statistical analysis was useful in assessing instructional performance. (ave 3.6, s.d 1.1)
- Outcome #2: Evaluation of student achievement of mathematics skills, specifically those skills noted in course objectives
  - The analysis of course objectives was useful in evaluating how well students did in answering questions associated with course objectives. Further study of these results is on-going.
- Outcome #3: Discovery of areas that should be targeted in a more systematic (e.g. course-wide) way to improve student achievement of mathematics skills
  - Further study of student tendency to select specific exam distractors is on-going – especially those cases in which a large percentage of students select the same incorrect distractor.
- Outcome #4: Improvement of the test and test results by aligning exam questions with course objectives and by detecting and correcting poorly written questions, miscoded answer keys, cultural bias, etc. (attachment 12)
  - The point biserial data has been used to detect and correct poorly written questions in Math 20, 35, and 45. The point biserial data from Summer 2006 were provided to the writers of each exam, and exam writers were asked to reconsider any questions with negative point biserial data. On the

eight exams (20A, 20B, 35A, 35B, 45A, 45B, 65A, and 65B), totaling 400 questions, there were only 13 questions with negative point biserial values, and all but one negative were removed by exam edits prior to the Spring 2008 calculations. The last remaining question that had a negative point biserial was edited in the Fall 2008 revision of the exams, and the effectiveness of this edit will be measured in the next set of point biserial calculations.

- The objective analysis has been used to align exam questions with course objectives. For example, the objective analysis of Math 45 showed that some content areas were either overemphasized or totally omitted in the exam. These corrections were then made in the new exam.

- Outcome #5: To meet accreditation standards requiring the department final be validated.
  - Discipline faculty are currently determining the components for validation, which will include, but not necessarily be limited to, content validation and an assessment of cultural bias.
  - The mathematics department is considering requesting that R & K include two additional analyses in an effort to satisfy the Accreditation requirement to validate the exam:
    - Cronbach Alpha Coefficient - to determine reliability or the extent to which an instrument consistently measures whatever it purports to measure.
    - Factor Analysis – to analyze the correlations among the items and group together those that seem to be performing similarly.

7. Have the departmental final exams improved course consistency and quality?

The department chair whose term included both the ante and post departmental final eras believes very strongly that both consistency and quality have been improved.

8. Since implementation of departmental finals, do students who successfully complete developmental mathematics courses (Math 20, Math 35, Math 45, Math 60 and Math 70) exit with a more consistent skill level?

Most of the mathematics faculty members who have been here long enough to judge believe that the consistency level has improved. However our only source of knowledge for the skill level ante is anecdotal.

9. Have the implementation of departmental finals and data analysis improved student learning in mathematics?

Since the mathematics departmental final is re-written each semester, comparing student learning from semester to semester is problematic. We are investigating techniques that would enable such a comparison. One possible solution is to imbed a set of constant questions into the exam for a specific course that could be compared over time, similar to the SLO assessment questions which SD Mesa College uses each term.

Attachments:

1 Developmental Math Program Goals, 3/7/01
2 Developmental Math Program Needs, 6/7/04
3 Hewlett Leaders in Student Success
4 Data Analysis Provided to Faculty
5 Data Analysis Provided to Chair and Dean
6 Basic Statistics Calculated by A. Juden and M. Branca
7 Proposed Statistics by M. Branca
8 Communiqué and Validation Design re Mathematics Department Final
9 Accreditation Plan for Validation of Mathematics Department Final
10 Cultural Bias Study
11 Faculty Survey Results
12 Impact of Data Analysis on Math Department Final