The Pyramid Exam

David Cohen and James Henle

Calculus reform has changed both the content of the calculus course and the methods we use to teach that content. We now can use computers and graphing calculators to allow students to work on problems that are more meaningful and complicated than the time-worn max-min problems we once assigned as the prime examples of the utility of calculus. We often put students to work in groups so that they can experience the challenge and rewards of collaborating on hard problems to arrive at a deeper understanding than might be achieved in isolation.

When it's time to test how much the students have gained from our teaching different material differently, however, we fall back on the same old tool of the in-class or take-home examination. And since we must assess individual achievement, we don't let students collaborate on exams. So we reduce our expectations and cook up problems that can be solved in a class period, or perhaps in a few days in the case of a take-home. After emphasizing the value of collaboration, we send students the contradicting message that what counts when push comes to shove is what they can do in isolation. And after helping them develop the ability to organize a complicated situation and make a calculus problem out of it, we avoid testing them for that ability because we're afraid they'll be stymied at the outset and get nowhere. So not only do our tests send confusing messages to students, they fail to give us feedback on how well we've achieved some of the most important objectives of calculus reform.

The pyramid test is an assessment tool well suited to the content and pedagogy of reformed calculus courses. It combines individual work and group work, and also tests for students' abilities to assemble and probe complicated sets of ideas. It works best after student groups have worked together long enough to form productive working relationships. While we describe here a few different formats for pyramid tests, it should not be difficult to see how the basic formula can be altered to suit variations in class schedules and sizes.

Variation 1	Variation 2	Variation 3
1 hr alone	1 hr alone	1/2 hr alone
1 hr with group	2 days with group	1/2 hr with group
2 days alone	1 class session	2 days alone
2 days with group	2 days with group	
1 class session	1 class session	

The test itself consists of some routine exercises as well as some sophisticated problems requiring time to understand the question, and more time to investigate the answer. Computers or graphing calculators are usually required for the more challenging problems. The test is taken in stages over the course of two or three days, or up to a week.

If possible, the first stage is scheduled for the first hour in a two-hour time block. (As we discuss later, we've also had success with an alternate schedule when a two-hour block is not feasible.) During the first hour the students work on the test individually without books, notes or computers. Their task is to go through the entire test, answering whatever questions they can, and writing suggestions for tackling those questions they cannot answer but have ideas about. At the end of the hour they pass in a paper which is their first effort. This part counts as a portion of their final grade—anywhere from 20 to 50 percent. It is made clear to them that they are not expected to complete the test at this stage. Rough estimates and partial proofs are welcome.

If a second hour is available, the students work on the test again, this time working in their groups, still without books, notes or computers. The task during this second stage is to share the ideas they've come up with during the first stage, share answers on the easier problems, and help each other develop methods of attack on the more difficult problems. At the end of this stage each group passes in a second effort. If a second hour is not available, this stage and the next are omitted, and the fourth stage becomes the second.

The third stage takes place between this class period and the next. Students work individually again, this time using books, notes, computers and calculators to come up with answers as complete as they can make them. At the beginning of the next class period the students pass in a paper on their individual work. It counts 20 to 40 percent.

The fourth stage then begins: for the next two days they work on the test in groups to come up with a

See Pyramid Exam on page 15

UME TRENDS

Published by the Joint Policy Board for Mathematics, on behalf of the American Mathematical Society, the Mathematical Association of America, and the Society for Industrial and Applied Mathematics. UME Trends appears 6 times per year: January, March, May, July, September, and November. The annual subscription rate is $16 in the United States and $24 in other countries.

Editor
Ed Dubinsky, Purdue University

Department Editors
Martin E. Flashman, Humboldt State Univ.
Rogers J. Newman, Southern Univ.
David A. Smith, Duke Univ.
Annie A. Selden, Tennessee Tech.
Bonnie Gold, Wabash College
Sharon Cutler Ross, Dekalb College
Alan Tucker, SUNY at Stony Brook
John Selden,Jr., Mathematics Education Resources Co.

Editorial Committee
Lida K. Barrett
James A. Donaldson
Sandra Keith
James W. Lea
David A. Smith
Karl J. Smith
David O. Lomen
Glenda Lappan

Director of Publications
Donald J. Albers

Managing Editor
Harry Waldman

Production Editor
Amy S. Fabbri

Pyramid Exam from page 2

complete answer for the examination. This includes carefully written explanations for required essays, complete descriptions of the computer programs or packages they've used, and discussions of accuracy for estimated answers which can't be found exactly. They work together to write one paper, which is submitted as the group's effort, each member of the group receiving the same grade for that effort. That paper is due at the beginning of the next class period. (They are asked to make copies of it before they pass it in to use during the final stage.) It counts as another portion of their overall grade for the test—anywhere from 40 to 70 percent.

The final stage is a class effort. Again, it is helpful, but not essential, to schedule this during an extended class period. With the instructor out of the room, the entire test is put on the blackboard by the class and discussed until everyone agrees that it is all correct, or until half the class period is over, whichever comes first. The instructor is then invited back into the room to read the test, comment on it and correct it in front of the entire class. This also counts as part of the final grade, perhaps 10 percent. It is usually entirely correct by this time.

In the four years we've been using it at Smith College, students have found the pyramid test more compatible with the stated objectives of our new calculus curriculum than standard tests, and instructors have found it more valuable for determining which ideas and skills have been mastered by students and which have not. By the final stage of this challenging exercise it becomes clear to both students and instructors precisely which ideas have caused students the most difficulty. Students recognize the contrast between this form of test and the experience of struggling in isolation with a problem they cannot solve until the instructor reveals the answer, too late for them to receive credit for the learning. The pyramid test is truly a learning experience, providing students with a spirit of shared adventure and accomplishment, the sense of joy a mathematician feels after mastering challenging mathematics.

Administration and evaluation of a pyramid test requires more effort than a standard test, but not as much as it might seem at first. Depending on the variation used, the instructor might have to read one extra first-try individual effort. The other extra papers are group papers, and there are only as many of those as there are groups. These papers are usually so much better than individually written exams that it is often easy to read through them. Since the pyramid test is so comprehensive, and provides a format for even weaker students to learn and perform satisfactorily, fewer tests need be given during the semester.

The instructor can score each stage separately, and compute the final grade by taking a weighted average for the stages. It's possible to imagine a type of pyramid test devised for very large classes which now use mass short-answer exams, but we haven't had experience with that.

In summary, we believe that examination and assessment reform is an essential ingredient of calculus reform, and we believe that the pyramid exam works well with the goals now inherent in many of the variations of calculus reform.

David Cohen and James Henle are at Smith College.

Alternative from page 13

(in descending order) Macroeconomics, Fundamentals of Mathematics (for elementary and middle school teachers), Microeconomics, Social Statistics, Social Research, Music Theory II, Introduction to Spaceflight, Statistical Analysis I.

Grade Comparisons in Subsequent Courses

Six quantitative courses drew enough students who had passed either the traditional course or the alternative course to allow for statistical comparisons. These courses were Macroeconomics, Microeconomics, Principles of Chemistry I, Topics in Finite Mathematics (for business students), Social Statistics and Principles of Accounting I. In Microeconomics, traditional students earned significantly higher grades (0.001 level) than did the alternative students even though the two groups were not significantly different in all other characteristics examined. In Topics of Finite Math, traditional students earned marginally higher grades (0.07 level) than did the alternative students. Again the two groups taking this course were not significantly different in any other characteristic. In the other four courses, no significant differences were found in the grades earned.

Conclusions

Grades earned in subsequent quantitative courses by traditional and alternative students differ only slightly with the advantage going to the traditional course. This may be due to the additional credit hour of the traditional course, to the content of the course or to the better mathematical background of students entering the course. Thus this study does not definitively show which is the better preparation for a student intent on taking additional quantitative courses.

However, the student profiles show that students with better ACT and SAT math scores select the traditional course whereas students with better ACT English and SAT verbal scores select the alternative course. Thus we intend to continue offering both courses to our student population.

William Bosch, Robert Heiny and Tabitha Mingus are at the University of Northern Colorado.

Letter from page 14

This may not be the proper question. Computers would certainly change the course of progress. They would lead to new areas such as computational mathematics and computer graphics that are important in their own right. However, the use of computers would not help to produce mathematicians like Newton or Euler. These mathematicians could develop extremely complicated theories in their heads.

Some ways of using a computer in teaching mathematics are proper, some are not. Using a computer as a tool to simulate the real world or, at least, its quantifiable part is an appropriate application of technology. Using it merely as a supercalculator is improper. The first application would help in learning mathematics. The second would only make life easier.

Reza Noubary

Bloomsburg University