EDITOR'S NOTES

MAJ Joseph Shehan, USNA

Greetings from West Point!

The purpose of *Mathematica Militaris* has been to provide a forum for the faculties of the Mathematics Departments of the several service academies. Over the last several years, there has been an exciting exchange of interesting methods, developments, and philosophies. We have learned much from each other.

This issue continues this trend of learning from each other. Our theme is the general concept of benchmarking. By benchmarking, we mean the process of comparing our performance against the performance of others who we believe are doing an excellent job. This process requires some hard thinking and answers to tough questions. First, what should we measure about our performance and how should we measure it? Do we focus on student learning or faculty teaching? Do we consider the average cadet/midshipman or the success of our outstanding students? What other performance measures (such as publications, intramural and extramural consulting, et cetera) are appropriate? How do we measure these things? What do they tell us?

Then there is the difficult issue of selecting other institutions for comparison. Even among the service academies, there are distinct differences in emphasis and approach. The service academies are markedly different from the majority of other civilian colleges. Still, we can learn from others and among ourselves. This issue represents an attempt to do so.

These articles provide varied and illuminating perspectives on the issue of benchmarking at and among the service academies' mathematics departments. The editorial staff thanks the many contributors for their work.

We encourage others to submit articles on this theme in the future. We especially encourage authors to submit data which is useful for comparisons. As an example, the editorial staff includes a trend chart (see Figure 1, page 2) of the number of mathematics majors and operations research majors at USMA since our curriculum reform. We draw no conclusions from the data, but hope it informs your discussions about your own programs.The theme for the spring issue will be "***How are Multimedia and the World Wide Web changing the teaching and learning of mathematics at the service academies***?" Does this expensive new emperor have any clothes? We strongly encourage your submission of articles and URLs for this issue, which promises to be extremely exciting.

Finally, I'd like to thank the associate editors, Professors **Judy Holdenor** and **Joseph Wolcin** and Major **Joe Shehan**, and the managing editor, Major **Garry Lambert**, for their efforts to keep the quality of submissions so high and so timely. This bulletin could not be published without their hard work.

We also thank the Association of Graduates, USMA, for continuing to provide the funds for our publication.

Dave Olwell

CONTENTS

EDITORIAL STAFF

EDITOR IN CHIEF:
MAJ (P) David Olwell, USMA

MANAGING EDITOR:
MAJ Garrett R. Lambert, USMA

ASSOCIATE EDITORS:
Dr. Judy A. Holdener, USAFA
Dr. Joseph Wolcin, USCGA

Overview

In this issue, Captain **Len Kelley** outlines the Coast Guard Academy's academic outcomes. Those ten goals "represent the targets" for

SUBSCRIPTIONS TO *MATHEMATICA MILITARIS*:

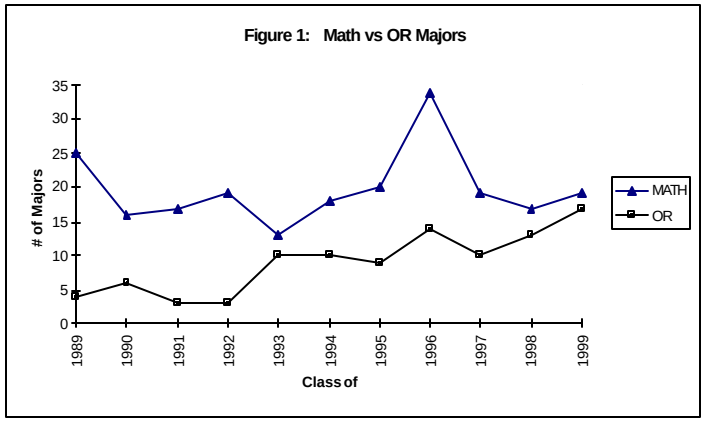
If you would like to be on our mailing list, please send your name, address, and affiliation to:

MAJ Garrett R. Lambert

Department of Mathematical Sciences
United States Military Academy
ATTN: MADN-A
West Point, New York 10996

academic efforts at USCGA. Captain Kelly politely suggests that a more appropriate focus for (at least) his school is contemplating what constitutes the best mathematics programs for each service, and not what the world class programs are. He outlines the methods his department uses to assess its performance against the USCGA outcomes. His department is working towards setting internal benchmarks for self-assessment.

Majors **Russ Paulsen** and **Joe Shehan**, USMC, provide the first contribution from the Naval Academy. They surveyed their fellow faculty members, and have captured several interesting insights. They discuss some possible benchmarks, including number of students selecting mathematics majors, what graduate schools are attended by the students, and the number of students who go on to graduate study in mathematics. They believe that "Nobel" quality faculty distinguish world-class departments. They also discuss their department's three fold mission as providing a focus for theirassessment activities.



Professor **W. D. Withers**, also from USNA, argues that a simple objective assessment tool is needed. He says this will take time, but it will be time well spent: "We currently lose much more time by adjusting and re-adjusting our program without any objective data on where our students stand and where they are headed."

Major **David Cribb**, USAFA, provides an example of how a change in a program can be assessed longitudinally. He presents the effects of changing the pre-calculus placement strategy for entering cadets. He also provides data which is useful for comparisons. For example, at USMA, our percentage of students enrolling in precalculus is about 5%, compared to USAFA's 28% for the class of 1999. What can USMA learn from this comparison? Major Cribb also provides data for student success as measured by mathematics core courses averages. This, too, is very useful for comparison. These benchmarks can be used by other schools to derive insights about their own approaches, and illustrate the power of the idea of benchmarking.

Major **Steve Hadfield** discusses the internal assessment procedures used by the engineering mathematics division at USAFA. He discusses the formal instructor observation process, and the student course critique. He also makes the excellent point that attendance at national conferences and meetings provides opportunities to learn about the best practices of other schools.

Our third USAFA contribution comes from Major **Marie Revak**. Her review of benchmarking highlights the difficulties of deciding what to measure. She provides a good summary of the features of assessment instruments.

Lieutenant Colonel **Rich West** discusses the general framework for assessing change at the USMA Department of Mathematical Sciences. His doctoral dissertation discussed the assessment of the sweeping changes in the core mathematics curriculum at USMA in 1991. In particular, he discusses soft criteria such as student attitudes and beliefs as important components of an assessment plan. He talks about the key decisions about content and goals which must precede selection of specific benchmarks. The internal department self-assessment done by the senior faculty each year can be brutally frank, and LTC West talks about how those self-assessment results are incorporated into long range planning.

Assessing the Impact of The Department of Mathematics at The United States Coast Guard Academy

CAPT Len Kelley, Department Head
US Coast Guard Academy

Within the past few years the Academic Division at the Coast Guard Academy has established ten

Desired Academic Division Outcomes which cut across the educational spectrum and seek to focus our efforts within the military college function of the Academy. These outcomes are:

Outcome 1: A cadet shall be able to read and understand a variety of written materials, listen critically to oral arguments, and formulate penetrating questions.

Outcome 2: A cadet shall be able to write clear, concise, persuasive and grammatically correct passages on general or professional topics, from a paragraph to several pages in length.

Outcome 3: Apply the basic skills of critical analysis, quantitative reasoning and problem solving to complex tasks in a broad range of contexts.

Outcome 4: Prepare and deliver well organized and polished oral presentations to a variety of audiences on topics within their fields of competence.

Outcome 5: Gain access to a broad range of information systems and locate desired data reliably.

Outcome 6: Integrate knowledge and information efficiently into a working conceptual framework that lends itself to continued expansion and refinement.

Outcome 7: Show evidence that they (graduates) are capable of honest, realistic, and constructive self-evaluation, that they can devise successful and creative strategies to develop their strengths and correct their weaknesses, and that they demonstrate the intellectual, moral, and physical stamina to follow through.

Outcome 8: Function effectively as a member of a team or working group that is charged with studying a complex problem or a significant policy issue and arriving at a solution or recommendation.

Outcome 9: Comprehend the interrelationship of the diverse social, economic, political, cultural, technological, and environmental forces that shape the world in which the Coast Guard operates.

Outcome 10: Articulate their personal values and those of the Coast Guard and public service in general, recognize conflicts in value systems when they exist, and formulate reasoned arguments to support their resolutions of the conflicts.

Obviously, each course offered by the Department of Mathematics may not contribute to all ten of these outcomes, and, in fact, the collective mathematics curriculum may not address the entire group of ten desired outcomes. However, these outcomes provide a structure within which individual courses and departmental curricula can be developed--a focus for assessing the manner in which individual components of the academic experience contribute to these division-wide outcomes. These outcomes represent the targets upon which we must focus our energies.

We in the Department of Mathematics at the Coast Guard Academy have not contemplated what constitutes the best mathematics program in the world or a world class undergraduate engineering mathematics program. Frankly, our focus is more narrow. We are interested in providing the best mathematics program for the United States Coast Guard. The ten Desired Academic Division Outcomes, which were endorsed by the Academy's Board of Trustees as needful products for the good of the Coast Guard, represent the critical ends which our academic program ought to address. The extent to which activities in the Department of Mathematics contribute to these ends provides a tangible measure of our ability to provide the best mathematics program for the Coast Guard.

For each course offered in the Department of Mathematics, activities which contribute to the various Desired Academic Division Outcomes have been identified, enumerated, and linked to specific outcomes. Collectively, these documents provide an overview of the extent to which our courses contribute to the division-wide outcomes. Just as importantly, these documents enable individual faculty members to know how their particular course supports the various outcomes and allows them to make adjustments in their course activities to target a particular outcome either more or less intensely. Taken together these documents permit an instructor, or any interested party, to know how the various courses and activities within the Department of Mathematics support the Academic Division's

aim to produce the kinds of graduates our Board of Trustees have indicated the Coast Guard needs.

We have only recently completed listing activities in the Department of Mathematics, by course, which support the Desired Division Outcomes. Now that we have identified activities which we believe to be related to the Desired Division Outcomes, we are attempting to determine the degree to which each of these activities supports the outcome to which they have been linked. In essence, we are attempting to develop methods to quantify or qualify the extent to which an activity contributes to a particular outcome. For example, nearly every course offered by the Department of Mathematics requires a daily reading assignment, and it seems logical that these reading assignments ought to contribute to Outcome 1: A cadet shall be able to read and understand a variety of written materials, listen critically to oral arguments, and formulate penetrating questions. The current challenge is to determine whether or not reading assignments in courses offered by the Department of Mathematics really contribute to this outcome, and, if they do, to what extent. Similar investigations need to be carried out for every activity-outcome link.

These kinds of evaluations seem to be essential if we are to assess effectively the manner in which the Department of Mathematics supports the Desired Division Outcomes and utlimately the needs of the Coast Guard. Likewise, the ability to evaluate the manner in which particular course activites contribute to the development of individual cadets relative to the Desired Division Outcomes will enable us to assess the impact such activities have on the growth of all of cadets in dimensions of considerable importance to the service. A critical first step in this endeavor is to establish benchmarks with which comparisons can be made over time. Unfortunately, we lack such benchmarks for many of the Desired Division Outcomes and must address this issue before any meaningful evaluation of the impact of our activities on such outcomes can proceed.

Presently, we rely on a combination of specific criterion referenced standards and the professional judgment of our faculty to determine the level at which our cadets are meeting the objectives of particular courses and, by extrapolation, the Desired Division Outcomes. While we are responding to the issue of assessment in a proactive manner and certainly wish to improve our ability to evaluate the impact of our activities as mentioned above, it is probably fair to say that no formal evaluative mechanism ought to replace the judgment of competent faculty--the kind of faculty we have in the Department of Mathematics at the U.S. Coast Guard Academy.

Benchmarking for Better Mathematics at a Service Academy

MAJ Russ A. Paulsen, Instructor
US Naval Academy
MAJ Joe M. Shehan, Instructor
US Naval Academy

Organizations of any form that remain competitive within their particular specialty, are usually engaged in several common activities that ensure their viability. One shared activity is the process of continued organizational improvement. To this end, service academies have remained competitive and viable in our system of higher education, as well as in quality officer accessions.

Successful businesses in the private sector measure their products, services and practices against top competitors with an eye towards identifying and adopting practices which enhance future success. Although this practice of "benchmarking" has its origins deep within the business community, there is also tremendous utility for its application in the field of education. In particular, Math Department faculty at the Naval Academy generally agree that this approach to educational improvement has merit. Still, any attempt at improvement through benchmarking must be consistent with what seems to be the tri-fold mission of the mathematics department. That is, (1) to support the Academy's primary mission of producing young officers possessing critical thinking skills, (2) to offer a sequence of course-work which not only leads to a baccalaureate in mathematics but also to success in graduate school, and (3) to equip non-math majors with the necessary mathematical skills required for their chosen majors.

Faculty responses to a questionnaire regarding this topic illustrate the challenge of implementing ideas borrowed from other non-academy world class departments.

Question: *What are the best or what distinguishes the best undergraduate mathematics programs in the world and why are they the best?*
Response: Several faculty categorize specific departments as world class based on student performance on standardized tests (e.g., GRE). Others felt that Nobel quality faculty distinguish world class departments. Still others suggested that more abstract factors provide a good measure, such as those institutions which furnish positive classroom experiences possibly due to smaller class sizes. Also included might be those programs that successfully impart a degree of mastery of the material such that situational application of mathematics is a reality for the student.
Question: *What features do and do not apply to the service academies and why?*
Response: Several faculty members believe that standardized test scores may not be an effective assessment tool at the academies given the many diverse and unique demands placed upon the academy student. Extensive professional development training requirements, which civilian students do not encounter, necessarily limit the time available for in-depth study and reflection for the academy math major. While it is true that Nobel quality mathematicians are not typically attracted by the pay and grant opportunities offered by the academies, there are still dedicated and talented mathematicians who oversee small classes which make very personalized instruction possible. Finally, due to the realities of academy life, the study of fundamental, less esoteric mathematical concepts overshadows a more theoretical treatment and aims at creating "functional" mathematicians.
Question:*: How do world class mathematics departments measure the success of their average students?*
Response: Faculty responses unanimously agreed that departmental success is measured by which graduate schools students attend as well as the number of students who attend.
Question: *How do service academy mathematics departments measure the success of their average students?*
Response: Although faculty opinions varied, some suggested that the grades earned in courses is the best yardstick. Others felt that the number of students who become math majors and successfully complete degree requirements, as well as the number of these students who are accepted into graduate programs provide some measure of success. Additionally, the success of the non-math majors who are pursuing highly technical degrees is a reflection, at least in part, on the math department.
Clearly the direct benchmarking of some ideas from world class institutions may not be feasible--or even desirable--at a service academy. Arguably, the implementation of programs designed solely for the purpose of preparing students to achieve top scores on skill assessment tests or for accessions into top graduate schools would involve tradeoffs at the expense of the academies' primary mission of producing skilled junior officers. Still, the success of academy students on skills assessment tests and in graduate programs as compared with students from universities and colleges with renowned math departments would provide valuable insight into the viability of academy programs. The Math Department at the Naval Academy addresses these challenges by implementing a tailored academic program which strives to maximize student test performance and graduate school competitiveness while satisfying all mission requirements. For instance, a text is used for beginning course work that boasts of developing critical thinking vice rote mathematical abilities; departmental course offerings are designed to cover topics supportive of both technical and non-technical majors alike; and finally, a strong core of required courses within the math major provides an excellent base for skills assessment testing and graduate school admissions.

Constrained implementation of benchmarks, as the Naval Academy's program illustrates, is a reality of the academy environment. Still, that which works well for world class organizations will probably apply to other similar organizations so it would seem benchmarking does have its place at a service academy math department. For the Naval Academy Math Department, it is suggested that borrowed ideas are best applied singularly to one of its three mission facets. In the case of departments at other academies or institutions, borrowed ideas should suit their existing paradigm.

Perfect Assessment Test Not Required

W. D. Withers, Associate Professor
US Naval Academy

The Mathematics Department at USNA has a dual mission. We teach service courses to

provide every midshipman with a basic knowledge of calculus, probability, and/or differential equations; and engineering and science majors with more advanced mathematical tools needed for their own disciplines. We also support our own math majors with courses in even more advanced and abstract mathematics and operations analysis.

There are ongoing discussions and proposals within the department on changing our approaches to each side of our mission. These are probably motivated largely by a pervasive sentiment within the department that neither side of the mathematics program is yielding all the results we would hope for.

Hitherto these discussions have been grounded in our various direct experiences with teaching math courses. The service courses account for a large portion of almost every faculty member's teaching load each year, while the various math-major courses circulate from one member to another with more or less rapidity. The result is that each of us sees the mathematics program through the prism of his or her own limited experience.

I believe it is all too easy to overreach for conclusions based on what we see in our own classes. For example, four years ago I taught a second-semester calculus course in the fall; my students had validated the first semester by taking a test during plebe summer. Two years ago I taught the same course under the same circumstances. It seemed to me that the students performed distinctly worse then their predecessors two years earlier.

Was this a sign that entering plebes were declining in quality? Or---different people were now handling validation---that validation criteria had changed? Or---we had also changed calculus texts in the interim---that the new text was inferior to the old? Or perhaps that the new text was superior to the old? Or was the teacher getting worse? Or was it just the luck of the draw? I believe each faculty member faced with this situation, based on his or her own philosophical predilections, might have reached a different conclusion and proposed some different means of redress.

Discussions and planning for improving all facets of our math program can only be more effective if underpinned by some sort of uniform assessment of where our students stand, year after year. I belong to a small subcommittee charged with seeking ideas for such. The focus currently is on the math-majors program, but such an assessment might be equally useful for our service courses. In this role I have collected reports from a dizaine of schools assessing their own math majors by various means. Originally I was hoping to find a quick and dirty solution, but everyone seems to have elected thoughtful, nonquantitive approaches, using tools such as essays and portfolios.

Nonetheless, I believe the greater danger for us on this issue is to focus too much attention on quality. We shall be better served by quickly implementing a few simple, imperfect tools (such as tests) and then using the results with an awareness of their limitations than by spending years perfecting the world's best assessment tool. Developing even a simple assessment tool will take time and effort; administering it will take time both on our part and the mids'. But we currently lose much more time by adjusting and re-adjusting our program w thout any objective data on where our students stand and where they are headed.

Math Placement Procedures at the United States Air Force Academy

Major David W. Cribb, Instructor
US Air Force Academy

Abstract:
This article discusses the recent modifications and some results of the Air Force Academy's new mathematics placement procedures that have been implemented for the past two incoming freshmen classes. These new placement procedures have tried to deal with the noted common trend of significant weaknesses in algebra and/or trigonometry skills in many of our new students. The increased enrollments in our precalculus course have produced encouraging results in the average GPAs in math courses for the class of 1998 (the first class placed under these procedures).

The Air Force Academy's Mathematical Sciences Department altered its math placement procedures for the selection of the initial math course for incoming freshmen during the summer of 1994 after an extensive review of the process.

This review was initiated because of the perception of declining algebra and/or trigonometry skills noted by many of our instructors as well as feedback received from other science and engineering departments.

We administer four placement examinations during the first week of cadet basic training - usually around the Fourth of July weekend. These four tests cover:
(1) Algebra/Trig, (2) Calculus I, (3) Calculus II, and (4) Calculus III. The timing and stressful environment at the time of testing is obviously not ideal, but consideration is given in a somewhat lower 'acceptable' passing range since no thought of mathematics has probably taken place in the 6-12 weeks since their high school graduation. We have found, in many instances, accelerated high school students may not have had any math during their entire senior year of high school! This was almost always the case if a student took calculus during his/her junior year of high school.

The main concern of our instructors is that many of the incoming cadets needed extensive review/strengthening of their basic skills. Many of these students possess quite adequate differentiation and integration skills. It was noted that many of our students did quite well on the Calculus I exam, but did not obtain a passing score on the Algebra/Trig exam. In several cases, a student may have validated both the Calculus I and II exams but did not pass the Algebra/Trig exam.

We consider the decision of where to place an incoming cadet as one of our most important. It can be a critical decision if the cadet has aspirations of becoming an engineer or pursuing some similar major. If a cadet is placed in the remedial precalculus course with no other validations, it becomes almost impossible to major in Astronautical or Aeronautical Engineering. All other majors should be possible strictly dependent on the finishing of the calculus sequence during the next summer school period.

The major result of the new placement procedures was an increased enrollment in our remedial precalculus course. This change is demonstrated in Table 1 which shows the increased percentage in the precalculus course for the classes of 1998 and 1999.

Table 1. Math Placement for the Classes of 1997, 1998, and 1999

	CLASS of 97		CLASS of 98		CLASS of 99	
	# cadets	% of class	# cadets	% of class	# cadets	% of class
Precalc	138	11.9	365	27.3	373	27.8
Calc I	533	45.8	528	39.5	516	38.5
Calc II	399	34.3	320	23.9	345	25.7
Calc III	86	7.4	113	8.5	99	7.4
> CalcIII	2	0.2	7	0.5	4	0.3
No Math	5	0.4	4	0.3	4	0.3
Total	1163		1337		1341	

The increased percentage resulted when we raised the acceptable passing standard for the algebra/trig test, which was something that was strongly encouraged by the other science and engineering departments. 'No math' indicates those cadets that validated Calculus I and II and had no desire for a technical major and were thus finished with the calculus sequence.

The major change in procedure was the initial placement of cadets strictly by their performance on our placement exams. Cadets had several opportunities to disagree with their initial placement and request a different course. A request to move down to a lower level math course was almost always granted after a lengthy discussion with the cadet about his/her background and desires. When a request to move up was received, many factors were considered. These factors included SAT/ACT scores, previous collegiate experience, two different composite scores: the Prior Academic Record (PAR) and Academic Composite (ACOMP), and their desire for a technical major. The PAR combines a student's GPA, high school class standing, class rank, and the strength of the high school (measured by the success of a high school's past graduates in college). The ACOMP considers a student's PAR score with his/her SAT/ACT scores.

We have found that many times there is a lack of correlation between SAT/ACT scores and a student's actual performance in class. It was interesting to examine the bottom nine students in the remedial precalculus course for this semester (Fall, 95). Two had SAT scores of 690 and 680, while four of the nine had SAT scores over 600 with the mean of all nine at 599. Their percentages in the remedial class ranged from 47% to 14% after the first graded review which

had an overall mean of 80.2% with a standard deviation of 13%. Their performance was remarkably predicted from their algebra/trig placement scores of 27.5% to 10%. Even though the SAT may be a good measure for potential, it certainly isn't a good predictor for achievement.

We have had some very encouraging results in the average GPAs from the class of 1998 in our core courses. Table 2 shows the general trend of increases in the GPAs throughout the courses in comparison with the class of 1997. These indicators alone are not a validation of the new placement procedures. Since the raising of standards to validate a math course would result in the bottom tier of students 'washing back' to the next lower course (supposedly a course that would be easier for them in some regards), one would expect the grades to be somewhat higher.

Table 2. Math Sciences Core Courses Average GPA

		Precalc	CalcI	CalcII	CalcIII
Remedial	CL97 CL98	2.63 2.7	2.31 2.44		
Regular CalcI	CL97 CL98		2.69 2.73	2.24 2.66	
Advanced CalcII	CL97 CL98			2.98 2.97	2.79 2.95
Dbl Adv CalcIII	CL97				3.1

But since our exams are criterion referenced, using test bank questions and anchored finals, this should give some indication that the improvement is real. One interesting test will be in the comparison of grades in Calculus III after students have completed the entire calculus sequence. Hopefully, a more solid foundation in the basic skills (calculus as well as algebra/trig) will produce the more solid foundation for the upper level courses that we desire.

Guidance and Assessment within Engineering Mathematics at USAFA

Major Steve Hadfield, United States Air Force Academy

Within the Engineering Mathematics Division at the U.S. Air Force Academy's Department of Mathematical Sciences, the focus of direction and assessment are more introspective than that espoused by the concept of "Benchmarking". While benchmarking certainly has significant merits and we do not ignore its use, we feel that a "service to our clients" perspective works more effectively for our situation. In this article, I will briefly outline our guidance and assessment activities within the Engineering Mathematics Division.

GUIDANCE:

Guidance is provided by both the department chain of command and from our client departments, composed of the other science and engineering departments. Within the Engineering Mathematics division, course directors and instructors are provided a description of course content, course objectives, and an overview of critical concepts and skills (organized by topic and subtopic). These provide the general guidance for the course and are driven by client needs and our seven Dean's educational outcomes. Course content is defined by the curriculum handbook while course objectives provide more general direction on how to focus class activities to include such aspects as student-oriented learning activities, use of computers in the classroom, and an emphasis on relevance by incorporating realistic applications. The overviews of concepts and skills are organized as matrices. Rows correspond to course subtopics which are clustered under three to five main topics for each course. Within each subtopic are identified zero to five critical concepts and a like range of critical skills. This overview guides and stabilizes course content and provides a key outline for both student and teacher activities.

These guidance materials also provide an effective mechanism for communications with our client departments. Once every year this data is provided formally to our client departments (also more frequently on an informal basis). From this data our clients are provided a good understanding of our courses' focus and what to expect of the students when they reach their higher level majors courses. The clients' review of this data stimulates informal discussions of course composition and concludes with the formal discussions of the yearly Math

Curriculum Review occurring in the spring semester.

ASSESSMENT:

The formal course documents also provide a basis for assessment of our courses' effectiveness. Besides the inter-department assessment provided by our clients' review of these documents, we use three key mechanisms for our own internal assessments. First, baselined final exam questions are tied directly to critical skills and concepts. Use of these test items across semesters allows for trend analysis of student performance. While these baselined final exam questions provide excellent objective measures, some of the key course initiatives, such as student-oriented classroom activities, incorporation of technology, emphasis on relevance, and critical thinking, are difficult to measure with such tools. For measuring these and other objectives, we use formal classroom observations and student end of course critiques. Our formal classroom observation form uses quantifiable scales for 14 measures of the classroom environment. Ten of these are oriented towards the instructor and four towards the students. The specific areas measured are administration, organization, role model, knowledge of material, presentation, promoting critical thinking, use of student-oriented techniques, rapport/interaction, establishes relevance, use of technology, cadet preparation for class, cadet interaction, cadet questions & answers, and cadet problem solving. A cadre of six to eight senior department members conduct two formal observations of each instructor each semester. The data is then sanitized of personal information and accumulated by course for use in trend analysis of how well certain initiatives are being met. Similar analysis is done using the end of course student critique results accumulated by course.

So, within the Engineering Mathematics Division at USAFA, we use a set of formal course guidance documents to direct our courses and to serve as a vehicle for both internal assessment and communications with our client departments for their assessment of how well we meet their needs. This does not mean, however, that we completely disregard the concepts behind benchmarking. Several department members each year attend conferences and workshops across the nation both presenting details of our programs and more importantly obtaining information on other programs. Information from these sources is used to gauge and refine our programs.

For more information and details, please contact the author at (719) 333-3725 (FTS), 333-3725 (DSN), or hadfieldsm.dfms@usafa.af.mil (email).

Benchmarking: Let's Look Before We Leap

Major Marie Revak, United States Air Force Academy

Benchmarks can help an institution set and measure progress toward its goals, including standards of academic achievement. Because the Air Force Academy has a high rate of teacher turnover and a high percentage of inexperienced teachers, benchmarks can be essential.

Before a benchmarking program can be established, several key decisions must be made. How will we choose the benchmarks? Will our benchmarks move beyond the cognitive domain and assess students' beliefs and attitudes? How will we measure attainment of the benchmarks? Will norm-referenced or criterion-referenced measures (or both) be used? How can we ensure that our assessments are valid and reliable?

There is great dissension among calculus educators as to what topics are essential to calculus. There is further disagreement regarding the role of technology and the importance of routine drill. Unless two institutions agree on common norms and criteria, common benchmarks should probably not be used.

Normative measures compare the achievement of an individual examinee with that of a relevant group of examinees. The choice of a norm group is an important decision. National norm groups have the advantage of being simple and definite. But, for a unique group of students like military academy cadets, national groups may be too general. We may want to consider collecting our own normative data for comparisons within our own institution. Or, we could employ a user-selected norm group. For example, we at the Air Force Academy may want to compare our students with West Point and Annapolis cadets. We could make arrangements to share this data

directly. Another possibility would be to use a norm group provided by a standardized test publisher. A key disadvantage of norm-referenced measures is that the norm group is likely to become obsolete with time. A prime example is the "recentering" of the Scholastic Achievement Test (SAT) by the College Board. The SAT had to be re-normed because the norm group used by the college board no longer represented the population of students taking the test.

Criterion-referenced measures are divorced from any normative meaning. Instead of comparing students to a norm group, students are tested against a standard set of criteria. Although there is no need to rescale criterion-based tests, criteria may change over time. For example, as college mathematics courses move from traditional to reform curricula, we must begin to test for complex conceptual understanding in addition to testing the ability to perform mathematical manipulations. Criteria may also change as technology becomes more pervasive. It may be impossible to compare two groups of students if one group uses graphing calculators and the previous group did not.

Since we will be using our benchmarks to measure progress toward our goals, it is imperative that our tests be valid and reliable. A valid test is one that measures what we intend to measure. For a test to be valid, it must be appropriate, meaningful, and useful. The content validity of a test is usually confirmed by comparing the test items to the course objectives. Those teaching the course are usually the best judges of a test's content validity. A reliable test is one that produces consistent scores. Reliability comes into play when open-ended items are scored by more than one individual. If open-ended items are used as benchmarks, we must ensure that all tests are scored using the same scoring rubric. In addition, difficulty and discrimination indices can be used to check the quality of multiple choice test items. The difficulty index for a test item is the percentage of students that answered an item correctly. The discrimination index of a test item indicates the power of a test item to differentiate between persons possessing or not possessing the ability being measured. Validity and reliability coefficients and difficulty and discrimination indices should be used to design and improve our benchmarks.

Benchmarks could play a big part in an institutional assessment program, but only after we decide what we want to measure and design valid and reliable instruments to do the measuring.

Assessing Change at the US Military Academy

LTC Richard West, Associate Professor
United States Military Academy

Background. In this forum, you have heard much about the changes to the West Point core math curriculum, four sequential courses covering 16 credit hours and required of all students. Before heading off on this endeavor, the Department wrote down in very soft terms its expectations (goals) of all these students after the 16-hour experience. After dealing with this curriculum for a year, we found two things: nobody was evaluating the change and there were some disconnects between courses that were furthered by semester breaks. I got the task of evaluating the curriculum change, or at least setting up the plan of how to do it. We also felt like we needed something to connect the courses and enable us to establish intermediate (at least by course) objectives. We got together during the spring and summer of 1991 and came up with five educational threads, which allowed us to articulate measurable learning objectives for each course (growing toward our established goals). The details of these five threads have been addressed in many forms since. They are math reasoning, communication, scientific computing, math modeling, and history of math. We have found these threads very valuable in articulating measurable objectives to accomplish our stated goals. And, I have found these very handy to evaluate the curriculum change as it takes place.

Academy Assessment. In 1989, West Point had its Middle States Accreditation Review. As a result, the year prior was spent completing a self study. An outcome of this self study was the articulation (in very soft terms) of nine academic program goals. Further, one of the recommendations of Middle States was that we needed a system of outcomes assessment. For three years beginning in the spring of 1992, I served on a West Point committee to set up an academic assessment system. After a year of discussion of some very complex models, we settled on a very simple system of four phases. Taking one academic goal at a time, (1) establish

a learning model for accomplishing the goal, (2) assess the program designed (usually sequence of courses) to facilitate the learning model, (3) assess the implementation of this program as designed, and (4) assess whether the goal is being achieved. The purpose of this assessment system is to evaluate academic programs in terms of their stated goals towards primarily improving the programs and secondarily responding to external agencies. We conducted a pilot assessment of our academic program goal about the engineering thought process. The pilot was successful and the Academic Board adopted our proposed academic assessment system. For the last two years, the system has been in effect and operating.

National Assessment. From 1991 to 1996, I served on a Mathematical Association of America Subcommittee to the Committee on the Undergraduate Program in Mathematics (CUPM) which developed and published in FOCUS October 1995 Guidelines for the Assessment of the Mathematics Major. These guidelines reiterate a similar process to that taken at West Point to assess its core math program. Since all students take four courses in sequence over two years, we have a rare opportunity to look at student growth over time.

The hard work of this assessment process was done up front -- articulating program goals and learning models. Further, it is very difficult to develop goals which support them. In our particular pilot test, we picked a goal for which we already had an explicit learning model. Still, the phase one process of confirming this model took a committee of department heads six months to reach consensus. A question that comes to mind as I say this is: Do the departmental goals and objectives to be assessed only address student learning? What about student attitudes? Also, what about content, especially when programs such as ours provide a significant service for other disciplines?

Math Department Assessment. The current assessment situation is that one evaluation was completed and several department-level initiatives are on-going. I completed an evaluation of the initial cohort (Class of '94) to go through the new core curriculum in 1995. It was published as my dissertation: Evaluating the Effects of Changing an Undergraduate Mathematics Core Curriculum which Supports Mathematics-Based Programs, 1996, UMI. In this dissertation, I compared student achievement and attitudes of this cohort to those of the previous cohort (Class of '93). In the mean time I have collected similar data on all subsequent classes. With the graduation of the class of 1996, I now have one cohort that was observed with the same instruments throughout their four years at West Point. I plan to analyze the results for the January Joint meetings.

A department-level self-assessment has been ongoing for four years now. The purpose of the assessment is to provide a needs generating discussion among the senior faculty toward articulating department goals for a five-year strategic plan. It requires the assessment of the many facets of department life such as curriculum development, faculty development, faculty recruitment, advisement or student recruitment, facilities, resources, and image & reputation. After this self study, the departmental leadership comes to a consensus on goals for the next five years. Then projects and resources are identified and priorities are set for competing areas. All of this is written down and annually reviewed, thus requiring an annual reassessment. I think the process does several things for our program: provides a forum for evaluating our assessment products, develops an explicit plan for the future, builds consensus within the department, and facilitates growth.

One interesting result from this department level annual assessment was a study over the past two years to select a new calculus text and a new supporting computer software package. In the process of establishing criteria we identified nine content threads that we wanted to emphasize throughout our core math curriculum. These are vectors, limits, functions, rates of change, accumulation, models, approximation, visualization, and solution representations. In conjunction with the five educational threads stated earlier, we are trying to connect student growth in mathematical content terms. These threads enabled us to establish selection criteria based on the perceived needs of student learning over time. As a result, we selected a calculus text that will allow us to weave together three of the four courses and we selected a software package that students will use for all four years in their mathematics-based courses. Further, current efforts endeavor to link assessment instruments with content specific goals as part of the overall student growth model.

Conclusion. Thus, assessment is a vital tool to inform the senior faculty of West Point the direction for the future of the core math program and enhance student growth throughout their four-year stay.

Random Notes

6th Annual SASMC Hosted by USNA
USAFA Braces for the 7th!

The sixth annual Service Academy Student Mathematics Conference was hosted by the US Naval Academy in April, 1996. Over thirty cadets and midshipman participated, presenting the results of their senior research projects. The USNA Mathematics Department did an excellent job hosting and organizing the conference. The conference was so successful that parallel sessions were conducted so all presentations could be scheduled.

The seventh conference is tentatively scheduled for April, 1997, at the USAFA in Colorado Springs.

Problem of the Week Contest!

The West Point Chapter of Pi Mu Epsilon conducts a problem of the week contest. Cadet Chris Perry, '98, is the cadet in charge. These problems are accessible to students in our core mathematics curriculum. The problems are posted every Tuesday evening during the USMA academic semesters. A running score is kept of correct responses, and certificates are awarded each semester to winners in the USMA Cadet, USMA Faculty, and external to USMA categories. We invite you to participate. The contest is administered over the WWW, and the URL is:
http://www.dean.usma.edu/math/outreach/pme/pow.htm