

**AC 2010-770: ASSESSING ENGINEERING STUDENTS' ABILITY TO USE THE
MATHEMATICS THEY HAVE LEARNED**

Maria Terrell, Cornell University Math Dept.

Robert Terrell, Cornell University

Lisa Schneider, Cornell University

Assessing Engineering Students' Ability to Use the Mathematics They Have Learned

Abstract

A Mathematics Applications Inventory (MAI) is being developed by engineering and mathematics faculty at Cornell University to assess students' ability to apply the mathematics they learn in freshman calculus to engineering-related problems. This paper reports on three aspects of this work: (1) Development of the first draft of the MAI, (2) Pilot testing the MAI, and (3) Preliminary analysis of the pilot test data.

To develop the MAI, faculty of second- and third-year engineering courses were surveyed about how key concepts and techniques from single variable differential and integral calculus are used in intermediate-level engineering courses. Based on their feedback, as well as feedback from advanced undergraduate engineering students, an initial set of test items was developed. The resulting MAI consists of five open-ended questions with eleven sub-questions. The test is designed to be administered during one hour in paper-and-pencil format.

The MAI was administered during the first week of the Fall 2009 semester as a pre-test to a sample of first-year engineering students (N=51). Based on a review of their performance, a subset of 14 participants was selected to participate in follow-up cognitive interviews to probe their thought processes relative to the MAI items.

The pilot test data are being analyzed to determine patterns in students' abilities to apply the mathematics they have learned, to gain insight into common gaps in understanding related to the application of specific mathematical topic areas, and to understand patterns of responses and performance by other background and status variables such as gender, race, SAT scores, and level of mathematics preparation. These data are being used to refine the MAI items and to develop a multiple-choice version of the MAI. Ultimately, the MAI is intended to be used as a pretest/posttest instrument to determine whether curricular changes affect students' ability to apply mathematics to problems set in engineering contexts.

Background

Entering engineering students at Cornell University have median SAT Math scores of approximately 750 and Verbal scores of approximately 685. Credit for the equivalent of first semester calculus is expected at entrance (i.e., the equivalent of a 4 or 5 on the AB Advanced Placement examination). The first mathematics course for half of the entering class, Math 1910, is the equivalent of second semester calculus, for the other half it is multivariable/vector calculus, Math 1920, or higher (Linear Algebra for Engineers, or Differential Equations for Engineers). By all the usual measures we have a very able and motivated group of students. Yet engineering faculty at our institution have reported consistently that students in introductory engineering courses have difficulty using even elementary mathematics to represent quantities and relationships between them. This inability to use the mathematics that they have apparently

learned has been all the more perplexing since the core engineering mathematics courses are taught jointly by mathematicians and faculty in engineering.

The mismatch between students' apparent mathematical background knowledge and their ability to apply concepts to new problem situations is by no means unique to our institution. Many schools face it. The Mathematics Applications Inventory (MAI) is being developed as an instrument to assess students' ability to apply their knowledge of mathematics up to and including first semester calculus, that is material covered in an Advanced Placement AB examination. The goal is to develop a test that is valid, reliable, and sensitive to measuring gains from interventions designed to enhance students' ability to use mathematics to represent mathematical relationships in applied settings. This paper is a preliminary report on our efforts to develop the alpha version of the instrument, to pilot test it, to study its validity and reliability, and to modify and improve it .

Development of the MAI

In the Spring of 2008 a preliminary survey was done to determine the mathematical concepts most widely used in sophomore level Engineering courses required for affiliation with a major. The open ended survey questions yielded responses and final exams from engineering faculty in five key engineering distribution courses: ENGRD 2020 Mechanics of Solids, ENGRD 2030 Dynamics, ENGRD 2610 Mechanical Properties of Materials, ENGRD 2620 Electronic Materials for the Information Age, and ENGRD 2700 Engineering Statistics and Probability. Based on a study of the exams and information provided by the faculty of these courses, a questionnaire was developed to gather information about how key concepts and techniques from single variable calculus are applied in 2000 and 3000 level engineering courses.

The questionnaire was sent the following spring to Engineering faculty who regularly teach 2000 and 3000 level courses, asking them to complete a survey of the concepts and techniques from single variable calculus that are used in their courses. We also asked for copies of the most recent final examination from the course, with annotations indicating where the mathematics was used. We received feedback from nineteen courses/instructors. There were some unexpected patterns in the responses and the examination materials, that informed the development of the MAI pilot test items. In particular, slightly more faculty than we expected responded that they used the concepts of the derivative, definite integral, and the fundamental theorem of calculus conceptually rather than computationally. In addition we found a greater than expected use of derivatives and definite integrals of functions for which the independent variable was not time. A review of the examination questions showed that there were numerous demands on students to use pre-calculus mathematics –ratios, differences, reasoning from graphs, reasoning about parameters from implicit relations.

An initial set of test items was written, modifying contexts and problems from the engineering exams so that the questions would be accessible to students who had completed the equivalent of first semester single variable differential and integral calculus, equivalent to an AB Advanced Placement course in high school. The writing team was careful to avoid contexts and problems similar to those that are heavily represented on the AB or BC level Advanced Placement examinations. Some items were included that are similar to AP questions to provide a link

between performance on the MAI and data that we have on students' AP scores. The initial set of questions was reviewed by five faculty from engineering and mathematics who regularly teach engineering mathematics, and who were not members of the writing team. The questions were also reviewed by five advanced engineering undergraduates in their junior or senior year. Based on the feedback from faculty and students the questions were revised and pared down to five open ended questions with a total of 11 sub-questions. Student responses to the open ended questions are being used to refine the questions and to develop a set of distracters for the multiple choice version of the instrument.

Characterizing MAI Items

The mathematical concepts represented on the alpha version of the MAI were taken from pre-calculus and introductory calculus. We have found it helpful to develop a short taxonomy of numerical content topic codes ¹ and letter cognitive process codes ². Each of the eleven items on the MAI has at least one and no more than 2 topic codes, each paired with a cognitive process code.

Content Areas Topic Codes

100	Algebra and Arithmetic
101	Sums or Differences
102	Quotients, difference quotients, ratios
200	Function and Pre-calculus (includes limits, continuity)
201	Defined/determined contextually, or implicitly
202	Given through visual display (graph, table, charts)
203	Defined though an explicit formula
300	Derivative
301	Definition of derivative--limit of difference quotients, instantaneous rate of change
302	Graphical representation / interpretation
400	Definite Integral
401	Definition if definite integral, limit of Riemann sum, accumulation of a quantity over an interval
402	Graphical representation/interpretation

Cognitive Processes Codes

K	Knowing
	Recalls definitions, mathematical conventions , properties, recognizes key words
	Recalls and carries out algorithms for routine computations
	Retrieves information form tables, graphs, other sources
**	Recalls solution strategies/methods for solving routine problems
A	Applying
	Selects the appropriate method for solving a problem
	Generates an appropriate model
	Selects the mathematical facts necessary to solve a problem
	Determines/describes relationships between objects/variables
R	Reasoning

	Generalizes/extends the domain of mathematical thinking and problem solving/ makes conjectures
	Makes connections between different elements of knowledge
	Solves problems in contexts where the student is unlikely to have encountered similar items
	Apply mathematical procedures in unfamiliar or complex contexts.

**Note: “solve routine problems” is in the TIMMS applying domain. We have included solving familiar/routine problems in the knowing domain for the MAI. Solving problems that are straightforward, but not routine, are included in the applying domain. Solving very complex problems or problems in a very unfamiliar setting are in the reasoning domain.

Six instructors in Mathematics and Engineering (four at Cornell including one of the project PIs, one at New York University, and one high school teacher with experience teaching AP calculus) were asked to characterize the items and to apply the codes. The project PI’s description and codes, which were the *intended* content and cognitive dimension, was used as the “key”. Each of the coder’s responses was compared to the key to check for agreement on the content, and cognitive dimension of the items. There was substantial (90-100%) initial agreement between the five coders and the key on the content codes. The coders’ cognitive codes agreed with the key on more than 80% of the items. Each item had agreement among at least 4 of the coders, as to whether the question was a K, A, or R. Differences were reconciled through email discussions and consensus was reached on both content and cognitive dimension codes. The composite characterization and consensus coding of each item are as follows:

MAI 1—Standard application of the definite integral to find an accumulated amount given the rate of change. Coded 401 K

MAI 2a—Function, average value. Standard application of the definite integral. Coded 401 K

MAI 2b—Average rate of change problem-- pre-calculus, difference quotient. Coded 102 and 201 K

MAI 2c—Average of instantaneous rates of change. Very non-standard calculus question that requires careful interpretation of relationships. Coded 101 and 301 R

MAI 3a—Algebra/use of variables to represent a change in length. Coded 101 and 201 A.

MAI 3b—Algebra/use of variables to represent a fractional change. Highly unfamiliar use of difference quotient. Coded 102 and 201 R

MAI 3c—Derivative as limit of a difference quotient, i.e., fractional change at a point. Highly unfamiliar use of the derivative. Coded 301 R

MAI 4a—Functions-reasoning from graphical information, asymptotic behavior, arithmetic of positive/negative numbers. Coded 202 R

MAI 4b—Derivative interpreted graphically. Standard calculus problem. Coded 302 K

MAI 5a—Function given graphically in a non standard way; independent variable is on the vertical axis, and dependent variable is horizontal. Definite integral as accumulation. Coded 202 A and 401 A.

MAI 5b—Computation of the definite integral-graphically (using areas). Coded 402 A

Fall 2009 Pilot Test

The MAI was administered as a pre-test to 51 Cornell Engineering freshmen in the first week of class, 27 enrolled in Math 1910 (second semester calculus assumes a score of 4 or 5 on the calculus advanced placement AB test, or the equivalent), and 24 enrolled in Math 1920 (multivariable vector calculus assumes a score of 4 or 5 on the Advanced Placement BC test, or the equivalent). One hundred and forty-two students (71 from each course) were initially selected for inclusion in the sample and invited to participate. Fifty from each course were randomly selected to participate, while 21 others from each course were selected to assure a demographically diverse sample with respect to gender, race, ethnicity, and score on the quantitative part of the Scholastic Aptitude Test (SAT M). Students were informed this was a research project, they were compensated \$20 to take the test, and asked to do their best and show their work. Students' papers were read and scored within 24 hours of the test. Fourteen test takers were interviewed about their responses within 48-72 hours of the completing the test. Their responses are being used to refine the test questions and to obtain insight into the cognitive processes they used in answering the questions.

The same group of 51 students was invited to take the identical exam as a posttest at the end of the term. Students were paid \$30 to take the post-test. 35 of the original 51 took the posttest - 18 students from the Math 1910 group and 17 from the 1920 group. One student changed his enrollment from 1920 to 1910 during the term. His scores are included in the 1910 pre and posttest data. The chart below shows pre and post mean scores on each item for the students who took both exams. The MAI has item scores ranging from 0 (incorrect) to .5 (partially correct) to 1.0 (fully correct).

Table 1. Pretest and Posttest Means

Item	Codes	1910 Pretest	1910 Posttest	1920 Pretest	1920 Posttest
MAI 1	401 K	.81	.89	.94	.94
MAI 2a	401 K	.50	.44	.71	.82
MAI 2b	102 201 K	.56	.72	.59	.76
MAI 2c	101 301 R	.31	.61	.65	.71
MAI 3a	101 201 A	.47	.69	.76	.82
MAI 3b	101 201 R	.36	.31	.41	.56
MAI 3c	301 R	.06	.14	.29	.35
MAI 4a	202 R	.39	.72	.71	.82
MAI 4b	302 K	.69	.83	.94	.88
MAI 5a	401 202 A	.28	.31	.56	.59
MAI 5b	402 A	.28	.39	.44	.62
Valid N		18	18	17	17

Preliminary analysis of the pilot pretest data

The pretest data that we have is from fifty-one students. The number is far too small to assure the stability of the sample and the test item statistics. The posttest data is even smaller. Although it is tempting to speculate about the pre and posttest gains, we have focused our preliminary analysis of the pilot data on the pretest data. The multiple choice version of the MAI will be given as a pretest and posttest to 400 students in Spring 2010 . At that time we anticipate analyzing the differences in gain on particular items.

With the small sample size in mind a variety of group statistics were computed by Dr. Howard Everson, an independent statistical consultant for our project.³ The table and analysis that follow are taken from Dr. Everson's preliminary report that presents the results of the descriptive analyses of the students' performance on each of the eleven items. Item difficulty is most often measured by calculating the percentage of test takers, p , who answer the item correctly. The MAI has item scores that range from 0 (incorrect) to 0.5 (partially correct) to 1.0 (fully correct). As a result $p = (\text{sum of students' scores on the item}) / (\text{number of students taking the test})$. In this sample item MAI 1 ($p = .85$) and item MAI 4b ($p = .81$) appear to be the easiest items. This is consistent with the design of the items, that they be familiar applications concerning the definite integral and the derivative. The most difficult item appears to be MAI 3c ($p = .18$). Item MAI 3 was designed to test students ability to apply their understanding of algebra and the derivative to the problem of stretching a rubber band, in particular the notion of fractional change and fractional change at a point (the concept of "strain"). Most test developers seek to develop tests where the average item difficulty score is between 0.2 and 0.8. We are interested in capturing information about students ability to apply mathematics in straight forward and highly challenging and unfamiliar contexts, so very high and very low level scores are desirable. The remaining items exhibit item difficulty ranging from .36 to .69

Table 2. Item Level Statistics

Item	N	Minimum	Maximum	p values	Std. Deviation
MAI 1	51	.00	1.00	.8529	.33607
MAI 2a	51	.00	1.00	.6863	.46862
MAI 2b	51	.00	1.00	.6275	.48829
MAI 2c	51	.00	1.00	.4412	.49646
MAI 3a	51	.00	1.00	.6275	.46737
MAI 3b	51	.00	1.00	.4216	.47300
MAI 3c	51	.00	1.00	.1765	.34385
MAI 4a	51	.00	1.00	.6471	.48264
MAI 4b	51	.00	1.00	.8137	.38679
MAI 5a	51	.00	1.00	.4314	.45847
MAI 5b	51	.00	1.00	.3627	.45911

Tables 3a and 3b present data on the item difficulty as indicated on the pretest for students enrolled in Math 1910 and 1920. As expected the students enrolled in Math 1920 (who have scores of 4 or 5 on the Advanced Placement BC level exam) performed better than the students enrolled in Math 1910 (who have scores of 4 or 5 on the Advanced Placement AB level exam).

Table 3a. Item-Level Statistics: Course 1910

Item	N	Minimum	Maximum	p Values	Std. Dev.
MAI 1	27	.00	1.00	.8148	.37076
MAI 2a	27	.00	1.00	.6667	.48038
MAI 2b	27	.00	1.00	.6667	.48038
MAI 2c	27	.00	1.00	.3519	.47666
MAI 3a	27	.00	1.00	.5926	.48113
MAI 3b	27	.00	1.00	.4630	.45838
MAI 3c	27	.00	1.00	.0926	.24167
MAI 4a	27	.00	1.00	.5556	.50637
MAI 4b	27	.00	1.00	.7222	.44578
MAI 5a	27	.00	1.00	.3148	.41944
MAI 5b	27	.00	1.00	.2593	.40121

Table 3b. Item-Level Statistics: Course 1920

Item	N	Minimum	Maximum	p Values	Std. Dev.
MAI 1	24	.00	1.00	.8958	.29411
MAI 2a	24	.00	1.00	.7083	.46431
MAI 2b	24	.00	1.00	.5833	.50361
MAI 2c	24	.00	1.00	.5417	.50898
MAI 3a	24	.00	1.00	.6667	.45842
MAI 3b	24	.00	1.00	.3750	.49454
MAI 3c	24	.00	1.00	.2708	.41649
MAI 4a	24	.00	1.00	.7500	.44233
MAI 4b	24	.00	1.00	.9167	.28233
MAI 5a	24	.00	1.00	.5625	.47348
MAI 5b	24	.00	1.00	.4792	.49955

A principal concern when developing an instrument like the MAI is whether or not the measurement method is reliable. Reliability and replicability are central to measuring change in a pretest/ posttest instructional framework. Dr. Everson's preliminary analysis has used

Cronbach's alpha (a measure of internal consistency of measurement based on the inter-correlations of the eleven items on the MAI). His estimate of reliability based on this method yielded $\alpha = .77$ which is lower than the range of .85 but is not alarmingly low given the brevity of an 11 item scale like the MAI. We plan to modify the test by adding items and explore the relationships between the items, to improve the reliability of the test.

The next phase of test development is to write multiple choice versions of the questions and alternate forms of the exam. Distracters have been written by analyzing students' responses to the open ended questions on the pretest and posttest. We have also gathered information about how students responded to the questions through in depth interviews with 14 of the test takers.

Preliminary results of the interviews.

45 minute interviews were conducted with 14 test takers to obtain more detailed information about students responses to the test items. Students were shown their original test paper (no marks were made by graders on the papers) and asked four questions about their response to each sub-questions:

- a. How confident were you in your response to this question?
- b. Is this question similar to problems you have solved in some other setting? If yes, please describe the setting.
- c. Talk me through your answer to this question.
- d. Did you have other ideas about how to solve the problem that you did not write down?

Our review of the interviews reveal some interesting patterns. We were surprised to find that students who did well on the test were often less certain about their responses than students who did not do well, but who were nonetheless quite confident in their responses. Interviews revealed that more students in the 1920 class had taken Advanced Placement Physics, and that some of the problems or methods they used to solve the problems were similar to those they used in physics. In particular students mentioned using an analysis of the units to determine what the relationships might be. Another unexpected pattern emerged in students responses to question b, "Is this question similar to problems you have solved in some other setting?" Students who did not do well on many problems were much more likely to say that they had seen problems like these before in their AP calculus class, but could not recall how to solve them. Whereas, students who did well on the most difficult problems were aware that these problems were unfamiliar, that they were unlike any problems they had seen before, but that they did the best they could to figure them out.

Concluding remarks

The pilot test data are being analyzed to determine patterns in students' abilities to apply the mathematics they have learned. We are using student interviews to gain insight into students' application of specific mathematical topic areas. We are using demographic data to understand patterns of responses and performance on the test by other background and status variables such as gender, race, SAT scores, and level of mathematics preparation. These data are being used to refine the MAI items and to develop a multiple-choice version of the MAI. Ultimately, the multiple choice version of the MAI is intended to be used as a pretest/posttest instrument to

determine whether curricular changes affect students' ability to apply mathematics to problems set in engineering contexts.

Bibliography

1. *Survey of Enacted Curriculum (SEC) K-12 Mathematics Taxonomy*
<http://seconline.wceruw.org/Reference/K12MathTaxonomy2008.pdf>.
2. *TIMSS Advanced 2008 Assessment Framework* http://timss.bc.edu/PDF/TIMSS_Advanced_AF.pdf pg. 17-22.
3. Everson, Howard *Mathematics Applications Inventory Results of preliminary Pilot Study*, technical report November 2009. . <http://www.math.cornell.edu/~maria/>
4. Ganter, Susan and William Barker editors, *Curriculum Renewal Across the First Two Years: Curriculum Foundations Project: Voices of the Partner Disciplines* http://www.maa.org/cupm/crafty/cf_project.html 2004,