DEPARTMENT OF MATHEMATICS

TECHNICAL REPORT

DO CALCULUS STUDENTS EVENTUALLY LEARN TO SOLVE NON-ROUTINE PROBLEMS

Annie Selden, John Selden, Shandy Hauk, Alice Mason

May 1999

No. 1999-5

TENNESSEE TECHNOLOGICAL UNIVERSITY
Cookeville, TN 38505

Do Calculus Students Eventually Learn to Solve Non-routine Problems?

ANNIE SELDEN, JOHN SELDEN, SHANDY HAUK, ALICE MASON

ABSTRACT. In two previous studies we investigated the non-routine problem-solving abilities of students just finishing their first year of a traditionally taught calculus sequence. This paper[1] reports on a similar study, using the same non-routine first-year differential calculus problems, with students who had completed one and one-half years of traditional calculus and were in the midst of an ordinary differential equations course. More than half of these students were unable to solve even one problem and more than a third made no substantial progress toward any solution. A routine test of associated algebra and calculus skills indicated that many of the students were familiar with the key calculus concepts for solving these non-routine problems; nonetheless, students often used sophisticated algebraic methods rather than calculus in approaching the non-routine problems. We suggest a possible explanation for this phenomenon and discuss its importance for teaching.

1. Introduction

Two previous studies demonstrated that C and A/B students from a traditional first calculus course had very limited success in solving non-routine problems [12, 13]. Further, the second study showed that many of these students were unable to solve non-routine problems for which they appeared to have an adequate knowledge base. This raised the question of whether more experienced students, those towards the end of a traditional calculus/differential equations sequence, would have more success; in particular, would they be better able to use their knowledge in solving non-routine problems? Folklore has it that one only really learns material from a mathematics course in subsequent courses. The results reported here in part support and in part controvert this notion. As will be discussed, the differential equations students in this study often used algebraic methods (ideas first introduced to them several years before their participation in the study) in preference to those of calculus courses taken more recently. These students, who had more experience with calculus than those in the first two studies, appealed to sophisticated arithmetic and algebraic arguments more frequently than students in the earlier studies. While somewhat more accomplished in their problem-solving ability, slightly more than half of them still failed to solve a single non-routine problem, despite many having an apparently adequate knowledge base.

As in the previous two studies, what we are calling a non-routine or novel problem is simply called a problem, as opposed to an exercise, in problem-solving studies [10]. A *problem* can be seen as comprised of two parts: a task and a solver. The solver comes

[1]We would like to acknowledge partial support from the Exxon Education Foundation and Chapman and Tennessee Technological Universities.

equipped with information and skills and is confronted with a cognitively non-trivial task; that is, the solver does not already know a method of solution. Seen from this perspective, a problem cannot be solved twice by the same person, nor is a problem independent of the solver's background. In traditional calculus courses most tasks fall more readily into the category of exercise than problem. However, experienced teachers can often predict that particular tasks will be problems for most students in a particular course, and tasks that appear to differ only slightly from traditional textbook exercises can become problems in this sense.

In this study we used the same two tests as before: a five-problem non-routine test and a ten-question routine test (see Sections 2.3.1 and 2.3.2). The second, routine, test was intended to assess basic skills sufficient to solve the corresponding non-routine problems on the first test. This allowed some distinction to be made between the lack of routine skills and the inability to access such skills in order to develop a solution method for the associated non-routine problem.

2. The Course, the Students, and the Tests

2.1 The Calculus / Differential Equations Sequence

The setting is a southeastern comprehensive state university having an engineering emphasis and enrolling about 7500 students -- the same university of the earlier studies of C and A / B first-term calculus students [12, 13]. The annual average ACT composite score of entering freshman is slightly above the national average for high school graduates, e. g., in the year the data was collected the university average was 21.1, compared to the national average of 20.6.

A large majority of students who take the calculus / differential equations sequence at this university are engineering majors. The rest are usually science or mathematics majors. A separate, less rigorous, three-semester-hour calculus course is offered for students majoring in other disciplines.

Until Fall Semester 1989, the calculus / differential equations sequence was offered as a five-quarter sequence of five-hour courses. Since then, it has been offered as a four-semester sequence. Under both the quarter and the semester system it has been taught, with very few exceptions, by traditional methods with limited, if any, use of technology and with standard texts (Swokowski, Berkey, or Stewart, for calculus and Zill for differential equations [2, 14, 15, 18]). Class size was usually limited to 35-40 students, but some sections were considerably smaller and a few considerably larger. All but three sections were taught by regular, full-time faculty of all ranks; the three exceptions were taught by a part-time associate professor who held a Ph.D. in mathematics. All instructors taught according to their normal methods and handled their own examinations and grades.

2.2 The Students

The pool of 128 differential equations students considered in this study came from all five sections of differential equations taught in Spring 1991, omitting only a few students who had taken an experimental calculator-enhanced calculus course or who were participants in the previous two studies. All students in this pool had passed their first term of calculus with a grade of at least C.

In the middle of the Spring semester, all of the 128 beginning differential equations students were contacted by mail and invited to participate in the study. As with the previous study of A / B calculus students [13], each student was offered $15 for taking the two tests and told he or she need not, in fact should not, study for them. The students were told that three groups of ten students would be randomly selected according to their first calculus grades (A, B, C), and in each group there would be four prizes of $20, $15, $10, and $5. The latter was an incentive to ensure that all students would be motivated to do their best. Altogether 11 A, 14 B, and 12 C students volunteered and ten were randomly selected from each group. Of those, 28 students (10 A, 9 B, and 9 C) actually took the tests: three mathematics majors and twenty-five engineering majors (nine mechanical, five chemical, four civil, four electrical, two industrial, and one undeclared engineering concentration). These majors reflected the usual clientele for the calculus/differential equations sequence.

At the time of the study, all but one of the 28 students tested had taken the third semester of calculus at this university; the one exception was enrolled in Calculus III and Differential Equations simultaneously. Their grades in Calculus III were 5 A, 8 B, 8 C, 4 D, and 2 F. Of these, one D student and one F student were repeating Calculus III while taking Differential Equations. Twenty-three of the students took Calculus III in the immediately preceding semester (Fall 1990). Their grades and the grades of all students who took Calculus III that semester are given in Table 1, which indicates that the better mathematics students are over-represented in this study.

Grade	Participants	All Students
A	4 (17%)	8 (5%)
B	6 (27%)	24 (15%)
C	7 (30%)	49 (30%)
D	4 (17%)	41 (25%)
F	2 (9%)	33 (20%)
W	0 (0%)	10 (6%)
Total	23 (100%)	165 (100%)

TABLE 1. Calculus III grades for the participants compared with those of all students taking the course the previous semester

In Table 2 we give the mean ACT scores and the mean cumulative grade point averages (GPA) at the time they took the test for the 28 students in this study. The same information is given for the students in our two earlier studies [12, 13]. The numbers for the Differential Equations (DE) students are quite close to those of the A/B calculus students [13] but considerably above those of the C calculus students [12].

Study	Mean ACT	Mean Math ACT	Cumulative GPA
DE	26.26	27.74	3.145
(A/B) Calculus	27.12	28.00	3.264
(C) Calculus	24.18	25.65	2.539

TABLE 2. Mean ACT and GPA of students in all three studies

Eleven of the 28 differential equations students in this study had already taken additional mathematics courses. Of the three math majors, two had completed, and the third was currently taking, a "bridge to proof" course, and the third had also taken discrete structures.

In addition, five students had taken complex variables, and another was enrolled in that course at the time of the study. One of these five had also taken an introductory matrix algebra course, as had two other students. Except for one C, all grades for these students in these additional mathematics courses were at least B. In our analysis of the results we will compare the students who had studied mathematics beyond the calculus/differential equations sequence with those who had not.

Of the 28 students in this study, 22 (79%) graduated within six years of their admission to the university as freshmen. In comparison, for the university as a whole over the same time period, the average graduation rate within six years of admission was 41%.

As of May 1999, it is known that all but two of the 28 students tested had earned bachelor's degrees at this university, three in math and the others in engineering. In addition, it is known that five students had earned master's degrees in engineering and one had earned an MBA, all at this university. One student had earned a master's degree in mathematics at this university and a Ph.D. in mathematics at another university. There may be additional accomplishments of these kinds among the 28 students, but they could not all be traced.[2]

The students in this study represented 33% of the A's, 26% of the B's, and 6 % of the C's awarded in Differential Equations that semester, and none of the 30 D's, F's and W's. In addition, after a minimum of three semesters at this university, these 28 students had a mean GPA of 3.145 for all courses taken. Their graduation rate was almost double that of the university as a whole, and at least 25% of them went on to complete a graduate degree. By all these indicators, at the time of the study and subsequently, these students were among the most successful at the university.

2.3 The Tests

Students were allowed one hour to take the five-problem non-routine test, followed by half an hour for the ten-part routine test comprised of associated algebra and calculus exercises. Prior to the non-routine test, the students were told they might find some of the problems a bit unusual. No calculators were allowed. They were asked to write down as many of their ideas as possible because this would be helpful to us and to their advantage. They were told A students (in first calculus) would only be competing against other A students for prizes, and similarly, for B and C students. They were assured all prizes would be awarded and partial credit would be given.

Each non-routine problem was printed on its own page, on which student work was to be done. All students appeared to be working diligently for the entire hour. As in the previous studies [12, 13], each problem was assigned 20 points and graded by one of the authors and checked by the others. The mean score on the non-routine test was 21.3, as compared with 20.4 for the A/B and 10.2 for the C calculus students [12, 13]. The lowest non-routine score in all three studies was 0.

[2] Because the data were laid aside for some time, we are able to include this long-range information.

2.3.1 The Non-routine Test

1. Find values of a and b so that the line $2x+3y=a$ is tangent to the graph of $f(x)=bx^2$ at the point where x=3.
2. Does $x^{21} + x^{19} - x^{-1} + 2 = 0$ have any roots between –1 and 0? Why or why not?
3. Let $f(x)=\begin{cases} ax, & x \le 1 \\ bx^2+x+1, & x>1 \end{cases}$. Find a and b so that f is differentiable at 1.
4. Find at least one solution to the equation $4x^3 - x^4 = 30$ or explain why no such solution exists.
5. Is there an a such that $\lim\limits_{x \to 3} \dfrac{2x^2 - 3ax + x - a - 1}{x^2 - 2x - 3}$ exists? Explain your answer.

2.3.2 The Routine Test

1. (a) What is the slope of the line tangent to $y=x^2$ at x=1?
   (b) At what point does the tangent line touch the graph of $y=x^2$?

2. Find the slope of the line $x+3y=5$.
3. If $f(x)=x^5+x$, where is f increasing?
4. If $f(x)=x^{-1}$, find $f'(x)$.

5. (a) Suppose f is a differentiable function. Does f have to be continuous?
   (b) Is $f(x)=\begin{cases} x, & x>0 \\ 2, & x \le 0 \end{cases}$ continuous?
6. Find the maximum value of $f(x)=-2+2x-x^2$.
7. Find $\lim\limits_{x \to 1} \dfrac{x^2-1}{x-1}$.
8. Do the indicated division: $x-1\overline{)x^3-x^2+x-1}$.
9. If 5 is a root of $f(x)=0$, at what point (if any) does the graph of $y=f(x)$ cross the x-axis?
10. Consider $f(x)=\begin{cases} x^2, & x \le 1 \\ x+3, & x>1 \end{cases}$.
    (a) Find $\lim\limits_{x \to 1^+} f(x)$.
    (b) What is the derivative of $f(x)$ from the left at x=1 (sometimes called the left-hand derivative)?

As soon as the non-routine tests were collected, the students were given the two-page routine test. They were told answers without explanations would be acceptable, but they could show their work if they wished. Most students worked quickly, taking from 12 to 17.5 minutes to complete their routine exercises. None stayed the allotted half hour. Each question was assigned 10 points. The highest score was 100 (out of 100) and the lowest score was 50, as compared to a high score of 90 and a low score of 53 for the A/B calculus students [13]. The mean score on the routine test was 75.3.

Each problem on the non-routine test could be solved using a combination of basic calculus and algebra skills. The correspondence between routine questions and non-routine problems is given in Table 3.

Non-routine Problem:	1	2	3	4	5
Corresponding Routine Questions:	1, 2	3, 4, 9	5, 10	6,9	7, 8

TABLE 3. Correspondence between routine questions and non-routine problems

3. The Results

We present the test results from several different perspectives. We examine the students' ability to solve non-routine problems, their knowledge base (of associated basic calculus and algebra skills) and whether they used their resources effectively.

3.1 Non-Routine Problem-Solving Ability

Slightly more than half (57%) of the differential equations students (16 of 28) failed to solve a single non-routine problem. This is somewhat better than the two-thirds of A/B, and all of C, first-year calculus students who failed to solve a single non-routine problem in the previous studies [12, 13].

In order to analyze the non-routine test results, we make a distinction between a solution *attempt*, a page containing written work submitted in response to a non-routine problem, and a solution *try*, any one of several distinct approaches to solving the problem contained within a single solution attempt. In only four instances did a student not attempt a non-routine problem; thus there were 136 attempts by the 28 students on the five non-routine problems. On these attempts there were a total of 243 solution tries. Of the 136 attempts, 20 were judged *completely correct*, that is, correct except possibly for a minor computational error. Twelve other solution attempts were found to be *substantially correct* because they exhibited substantial progress toward a solution, that is, the proposed solution could have been altered or completed to arrive at a correct solution.

Of the 20 completely correct solutions, five were for Problem 1, three for Problem 2, five for Problem 3, none for Problem 4, and seven for Problem 5. These completely correct attempts were from 12 of the 28 students; the 12 substantially correct attempts came from 11 students. Altogether, 18 of the 28 students provided at least one substantially or completely correct solution attempt. That is, 36% of the differential equations students (10 of 28) were unable to make substantial progress on any non-routine problem; this is lower than the 71% and 42% reported previously for C and A/B first-calculus students [12, 13].

Thus, the differential equations students did perform somewhat better than the first-year calculus students.

3.2 Comparison of Non-Routine and Routine Test Results: Did Students Have Adequate Resources and Use Them?

The routine test was designed to determine whether the students' inability to do the non-routine problems was related to an inadequate knowledge base of calculus and algebra skills (i.e., "resources" [10, 11]). Did the students lack the necessary factual knowledge or did they have it without being able to access it effectively? Scores on the corresponding routine questions (Table 3) were taken as indicating the extent of a student's factual knowledge regarding a particular nonroutine problem. As in [13], a student was considered to have *substantial factual knowledge* for solving a nonroutine problem if that student scored at least 66% on the corresponding routine questions. A student was considered to have *full factual knowledge* for solving a nonroutine problem if that student's answers to the corresponding routine questions were correct, except possibly for notation, for example, answering (1, -1) instead of -1 to Question 6. All others were considered as having less than substantial factual knowledge.

In Table 4 we give the number of students who exhibited substantial or full factual knowledge on the corresponding routine questions with the number of students whose solution to a particular non-routine problem was completely correct or substantially correct. For example, on Problem 1, 15 (of 28) students had full factual knowledge; of these 5 gave completely correct and 2 gave substantially correct solutions. An additional ten (of 28) students had substantial factual knowledge for Problem 1, but none of them gave completely or substantially correct solutions. That is, the performance of these ten students on the routine questions seemed to indicate they had sufficient factual knowledge to solve, or at least make substantial progress on, Problem 1; yet they either did not access it or were unable to use their knowledge effectively to make progress. The remaining three students demonstrated less than substantial factual knowledge, making a total of 7 students giving completely or substantially correct solutions on Problem 1.

Taking another perspective, in the 59 routine question solution attempts in which students demonstrated full factual knowledge, they were able to solve the corresponding non-routine problem 14 times (24%) or make substantial progress towards its solution 6 times (10%). Thus, on slightly more than a third of their attempts (34%), these students accessed their knowledge effectively. Students with substantial, but not full factual knowledge, did so on less than a quarter of their attempts. These results are summarized in Table 5. Furthermore, six students showed *no* factual knowledge of the components necessary for a non-routine problem and they each had a score of zero on the corresponding non-routine problem.

	Non-Routine Problem 1	2	3	4	5
Full Factual Knowledge	15	3	5	14	22
Non-Routine Problem					
Completely Correct	5	1	1	0	7
Substantially Correct	2	0	1	0	3
Substantial Factual Knowledge	10	19	12	1	2
Non-Routine Problem					
Completely Correct	0	2	4	0	0
Substantially Correct	0	3	1	0	0
Less than Substantial Factual Knowledge	3	6	11	13	4
Non-Routine Problem					
Completely Correct	0	0	0	0	0
Substantially Correct	0	0	1	0	1
Total Completely or Substantially Correct	7	6	8	0	11

TABLE 4. A problem-by-problem view of the number of students having full, substantial, or less than substantial factual knowledge along with the number who either completely or substantially correctly solved each non-routine problem.

	Full Factual Knowledge (59)	Substantial, but not full Factual Knowledge (44)
Completely correct non-routine solution	24% (14/59)	14% (6/44)
Substantially correct non-routine solution	10% (6/59)	9% (4/44)

TABLE 5. Overview, giving the percentage of correct solutions by those having the requisite factual knowledge

In order to compare overall student performances on the routine and non-routine tests, we introduce the notion of a *score pair*, denoted $\{a, b\}$, where a is the student's score on the routine test and b is the student's score on the non-routine test. In every case, $a > b$. Figure 1 shows students' routine test scores in descending order (from left to right); superimposed below each student's routine test score is her/his non-routine test score.

The three students with the highest non-routine scores had score pairs of {90, 69}, {85, 59}, and {83, 59}; the first of these subsequently obtained a B.S. in civil engineering, summa cum laude with a cumulative grade point average of 3.948. The three mathematics majors in the study, all of whom had completed at least one additional mathematics course at the time of the study, had score pairs {95, 34}, {89, 18} and {87, 21}. That is, they scored in the top quarter on the routine test, but the latter two scored slightly below than the mean non-routine score of 21.3. The last of these three subsequently obtained a Ph.D. in mathematics from a major state university.

Student performance on the routine and non-routine tests was not improved by having studied additional mathematics. The respective mean scores for the eleven students who had done so were 73.1 (vs. 75.3 for all of the students) and 15.3 (vs. 21.3 for all of the students).

Figure 1 shows a generally positive relationship between factual knowledge (resources) and the ability to solve novel problems. Nonetheless, having the resources for a particular problem is not enough to assure that one will be able to solve it. Two students had score pairs of {86, 4} and {80, 3}suggesting that having a reasonably good knowledge base of calculus and algebra skills (resources) is not sufficient to make substantial progress on novel calculus problems. One student, score pair {83, 59}, lacked substantial factual knowledge on only those routine questions associated with Problems 2 and 4 (on which he scored zero) and solved the three remaining non-routine problems (1, 3 and 5) completely correctly. This student and one other, the {90, 69}score pair, were the only students whose performance on the routine tests could have predicted precisely which non-routine problems they would be able to solve. In addition, the literature [1, 7] suggests that misconceptions are more likely to surface during non-routine problem solving, a phenomenon observed in this study and discussed below.

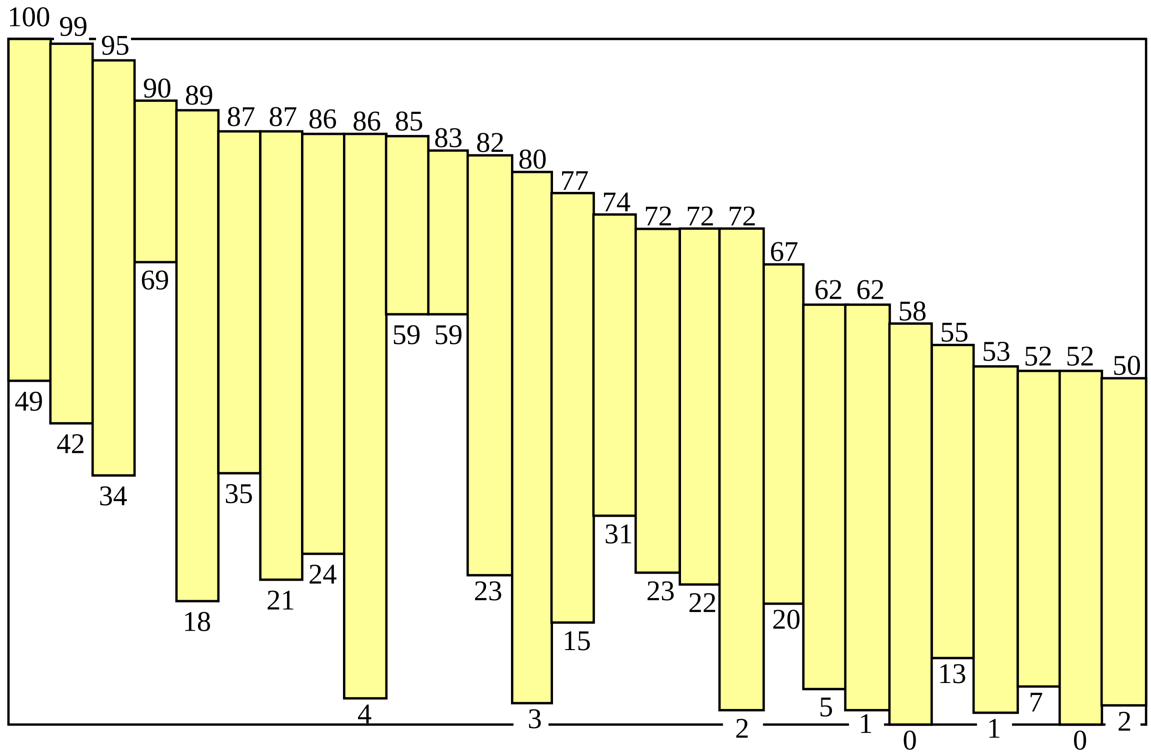

FIGURE 1. Upper score is for routine/factual knowledge test.
Lower score is for non-routine test.

4. Favored Solution Methods

4.1 Non-routine Problem 1.

Find values of a and b so that the line $2x+3y=a$ is tangent to the graph of $f(x)=bx^2$ at the point where $x=3$.

Fifteen students set the equation of the line equal to that of the parabola and solved for either *a* or *b*. Eight of these ignored the tangency of the line to the curve. The remaining seven also set the derivative of *f* equal to the slope of the line to obtain a second equation so *a* and *b* could be fully determined. These were the seven with completely or substantially correct solutions. An additional seven students took a derivative, but abandoned it at some point.

The two most frequently occurring misconceptions on Problem 1 involved the meaning of tangent line. Six students conflated the ideas of the equation of the tangent line, the slope of the tangent line, and the derivative of the function. Two other students incorrectly claimed that the tangent line was perpendicular to the curve at the point of tangency and used the negative reciprocal of the derivative for the slope of the tangent line. Missing among these students was the error found among A/B calculus students of confusing a secant line, calculated using two points on the parabola, with a tangent line [13].

4.2 Non-routine Problem 2.

Does $x^{21} + x^{19} - x^{-1} + 2 = 0$ *have any roots between –1 and 0? Why or why not?*

On Problem 2, the three correct, and three substantially correct, solutions made no use of calculus. Instead, they all used sophisticated arithmetic and algebraic reasoning to compare the relative sizes of the component terms in the given polynomial. The correct solutions made use of the observation that, for values of x in (-1, 0), both $-x^{-1}$ and $x^{21} + x^{19} + 2$ would be positive and hence their sum must be positive. This type of first-principles argument was also used, although less successfully, by about the same proportion of the A/B first-calculus students [13].

Other solution attempts on Problem 2 suggest that these differential equations students were relatively comfortable with, and knowledgeable about, algebraic techniques. Yet their knowledge included some common misconceptions. Two students inappropriately applied Descartes' Rule of Signs, and eight erroneously concluded that no roots could exist when the Rational Root Test produced no solutions. Even this flawed use of algebra is an improvement upon the favored solution method of the C calculus students [12]: substitute values for x until becoming convinced that no guess is ever going to work and hence no root can exist (a method also used by four of the differential equations students). Only five of the differential equations students used any calculus on Problem 2, four taking derivatives but not using them. The fifth student did make effective use of the derivative but made an unrelated error.

4.3 Non-routine Problem 3.

Let $f(x) = \begin{cases} ax, & x \le 1 \\ bx^2 + x + 1, & x > 1 \end{cases}$. *Find a and b so that f is differentiable at 1.*

Ten students set $ax = bx^2 + x + 1$, eight of them substituting $x = 1$ to get the relationship $a = b+2$, and then stopped. Two of these ten students expressed doubts about the completeness of their solutions. One student had seven solution tries, all variations on the theme of matching for continuity. Eleven of the 26 who attempted a solution to Problem 3 took a derivative. Several differentiated ax and got x. Of the eleven who used calculus, eight made at least substantial progress towards a solution; five used the derivative to arrive at a completely correct solution and three used it to obtain a substantially correct solution.

4.4 Non-routine Problem 4.

Find at least one solution to the equation $4x^3 - x^4 = 30$ *or explain why no such solution exists.*

There were no completely correct or substantially correct solutions. (Since these traditionally-taught calculus students were not allowed to use graphing, or other, calculators in this study, no easy graphical methods were available to them.) Twelve of the 27 students who attempted this problem used the same method: narrow the domain of possible x values by eliminating those for which $4x^3 - x^4$ cannot be close to 30. Most of these students determined that no solutions could exist outside of the interval (1, 4) and some also eliminated $x = 1, 2, 3$. Thus 44% of the students in this study used this method, as compared to only 26% (5 of 19) of the A/B calculus students [13], and their algebraic and arithmetic reasoning was quite sophisticated.

The next most popular solution method, used by six students (22%), was the Rational Roots Test. However, all six students incorrectly concluded that if a rational root could not be found from the factors of 30 then no roots existed. In our earlier study of A/B calculus students, this approach was also used by 20% of the students. Only two students used the method favored by C calculus students in the earliest study [12]: factor $4x^3 - x^4$ and set each factor equal to 30. Four students took the first derivative, set it equal to zero and stopped. Several students used synthetic division to check whether particular values were roots.

4.5 Non-routine Problem 5.

Is there an a such that $\lim_{x\to 3} \frac{2x^2 - 3ax + x - a - 1}{x^2 - 2x - 3}$ *exists? Explain your answer.*

On Problem 5, there were 39% (11 of 28) substantially or completely correct solutions, whereas for the A/B calculus students, 37% (7 of 19) made substantial progress towards a solution. Of the 28 attempts, 15 involved the use of L'Hôpital's Rule. Nine of these attempts were at least partially successful. In the other six instances students failed to note that the numerator as well as the denominator must have limit zero before applying L'Hôpital's Rule. Five students substituted 3 for *x*, found the denominator of the expression to be zero and asserted that no limit could exist since the denominator was zero at the limiting value of the variable (two of these were math majors). This was the favored method of the C calculus students (47% of them used it) and was used by 26% of the A/B calculus students [12, 13]. Here it was found in just 18% (5 of 28) of the solution attempts. Six students struggled, algebraically, with finding a way to factor the numerator so that the (x-3) in the factored denominator could be cancelled and the limit taken; two of these resulted in completely correct solutions.

4.6 Summary

4.6.1 Use of calculus

Since Non-routine Problem 5 involved evaluating a limit, so that any solution attempt could be considered as involving calculus, we omit it from the following analysis. On the first four non-routine problems, the differential equations students used calculus on fewer than half (39%) of all solution attempts (42 of 108), for example, taking a derivative. Furthermore in fewer than half of these attempts (16 of the 42), they used calculus to produce useful information which enabled them to make progress, but these led to 15 substantially correct or completely correct solutions. That is, about three-fourths of the 21 completely or substantially correct solutions to Problem 1 through Problem 4 made effective use of calculus.

We observe that the differential equations students were no more inclined to use calculus than were the A/B and C calculus students of the previous two studies [12, 13]. They failed to use calculus on 61% of these attempts (66 of 108); this is essentially the same percentage as in the previous studies (61% for the A/B and 59% for the C calculus students).

However, it appears that as students proceed through calculus to differential equations, the number of calculus misconceptions decreases. Slowly they become more proficient (or perhaps the less competent students drop out). For example, incorrectly asserting on Problem 5 that the limit of a quotient cannot exist when the denominator is 0 went from 47% for C calculus students, to 26% for the A/B calculus students, to just 18% for the differential equations students.

4.6.2 More sophisticated, but still flawed, use of algebra

On 56% of the solution attempts calculus was not used; rather a combination of guessing, trial-and-error, arithmetic techniques, and algebra was used. Eleven of the 32 substantially or completely correct solutions were almost purely algebraic, seven on Problem 2 and three (once the student observed the need to reduce the fraction in order to take the limit) on Problem 5. While these solutions demonstrated a level of algebraic competence not found in the first-year calculus students of our earlier studies [12, 13], quite a few other attempts to use algebra were flawed. For example, fourteen attempts made improper use of the Rational Root Test and two attempts used Descartes' Rule of Signs in an inappropriate setting.

4.6.3 Use of graphs

Most graphing was done by students on Problem 1 -- only four other graphs, three incorrect, appeared in all of the solution attempts for the other problems. Of the 27 students who attempted to solve Problem 1, 22 (81%) sketched at least one graph and six (22%) sketched three or more; this is substantially higher than the 12 of 19 (63%) who used graphs in the A/B study, [13], where only 1 (5%) sketched more than two graphs. Fourteen of the 22 who used graphs in the present study first drew some version of the graph pictured in Figure 2. Of those 14 students, six also produced a correct graph, Figure 3. Three of these six rejected the incorrect graph (by striking through it) and gave one substantially correct

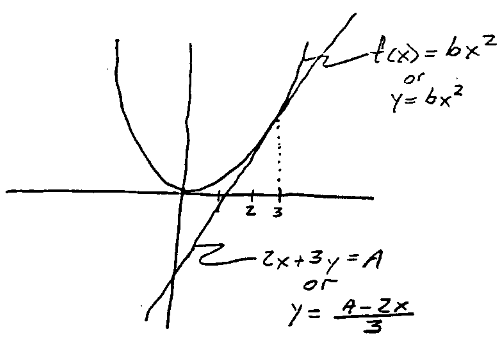

FIGURE 2. Example of a frequently sketched incorrect graph for Problem 1.

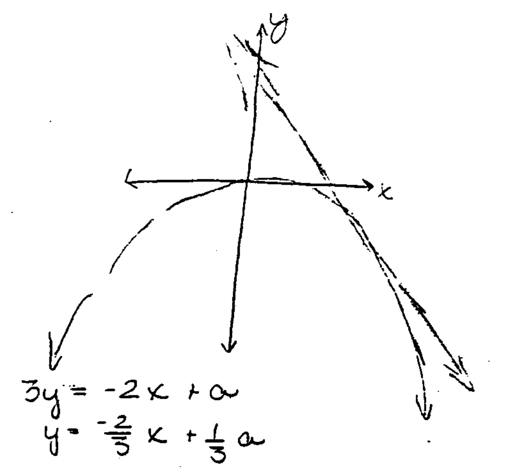

FIGURE 3. Example of a correctly sketched graph for Problem 1.

and two completely correct solutions to Problem 1, perhaps an example of monitoring their work [10, 11]. The fourth completely correct solution for Problem 1 came from a student who drew no graphs at all. Hence, three of the four completely correct solutions showed a graph in this study while only one of the three completely correct solutions in the previous study of A/B students was accompanied by a graph. The other three who drew both graphs used little calculus and did not indicate any rejection of incorrect graphs; in fact, all three seemed to be considering various cases for values of *a* and *b* by drawing a variety of graphs (between them, these three students produced ten graphs). The others who drew graphs were either misled by the assumption that the parabola was concave up (Figure 2) or gained no useful information from their graphs.

Where students who successfully solved Problem 1 in the present study were willing to draw and reject graphs, those in the A/B study who chose to use graphs generally had only one -- either it was the right one and the solution was substantially or completely correct or it was an incorrect graph and the solution was incorrect. In the A/B study, 16 graphs were produced by 12 students on Problem 1 whereas in the present study, 37 graphs were drawn by 22 students. Three of the 12 A/B students who used graphs (25%) rejected (crossed out) one of their graphs (including one who rejected a graph which was correct) while six of the 22 (27%) who used graphing on Problem 1 in this study rejected a graph. It would appear, then, that though the more experienced students were more willing to posit graphical ideas (81% versus 63%) than the less experienced students they were about equally likely to reject the graphs they produced.

DeFranco's paper on expert problem solvers with Ph.D.s in mathematics suggests that the skills possessed by the experts which are often lacking in the non-experts might include a willingness to create *and abandon* (reject) *and revisit* many ideas in the solution process [5]. Thus it might be useful to know how students develop the willingness to risk commiting graphical and other ideas to paper and to reject such ideas once they have been given life on paper.

5. Analysis

Reflecting on the three studies, one wonders when, if ever, students learn to use calculus flexibly enough to solve more than a few nonroutine problems. Furthermore, how could it happen that students, who were more successful than average by a variety of traditional measures and who demonstrated full factual knowledge for a non-routine problem, failed to access and use their knowledge successfully on 76% of their attempts (Table 5)? Many of these students (the engineering majors) had almost completed their formal mathematical educations, except possibly for one or two upper-division mathematics courses, leaving them limited opportunity in future mathematics courses to improve their non-routine problem-solving abilities. Finally, does it matter whether students can solve such non-routine problems?

5.1. When do students finally learn to apply calculus flexibly?

It would appear from this sequence of three studies that, at least for traditionally-taught calculus students in classes of 35-40 students, the ability to solve non-routine calculus problems develops only slowly. Performance for the best students went from one third who could solve at least one non-routine beginning calculus problem toward the end of their first year of college calculus to slightly less than half who could do so toward the end of the two-year calculus/differential equations sequence. In addition, the percentage of correct solutions increased modestly over the three studies (Table 6).

On the other hand, by the time these students were coming to the end of their calculus/differential equations sequence their algebra skills seem to be relatively sophisticated and readily accessible, albeit somewhat flawed. Such slow, incremental

Study	Completely Correct Solutions	Substantially Correct Solutions
DE	14% (20/140)	9% (12/140)
(A/B) Calculus	9% (9/95)	9% (9/95)
(C) Calculus	0% (0/85)	7%(6/85)

TABLE 6. Percent of correct solution attempts in all three studies

growth in mathematical capabilities may not be unusual. In a cross-sectional study of students' development of the function concept, Carlson investigated students who had just received A's in college algebra, second-semester calculus, or first-year graduate mathematics courses. She found that "even our best students do not completely understand concepts taught in a course, and when confronted with an unfamiliar problem, have difficulty accessing recently taught information. . . . Second-semester calculus students had a much more general view of functions [than college algebra students, but] . . . they were unable to use information taught in early calculus . . . ". [3]

In addition, it is not unusual for students to fall back on earlier mathematical techniques with which they are perhaps more familiar and more comfortable. In a study of 900 Australian high school students in Years 9 and 10 ("in their third or fourth year of algebra learning"), Stacey and MacGregor found that when asked to solve three simple word problems, a large proportion wrote no equations, even though specifically asked to, and others tried to write equations but then switched to non-algebraic methods including trial-and-error and arithmetic reasoning to solve the problems [9].

In an analogous fashion, the differential equations students in this study relied more often (in 76 of 136 solution attempts) on a variety of arithmetic and algebraic techniques than on calculus. Taken together, our three studies suggest the folklore that one only really learns a course's material in the next course appears to be not quite accurate, rather several courses may be necessary. The differential equations students seemed most comfortable with algebraic methods -- ideas first introduced to them several years before. As in the previous two studies [12, 13] over half of them made no use whatever of calculus. This negatively answers the question posed in [13]: Would students at the end of a traditional calculus/differential equations sequence be more inclined to use calculus techniques in solving problems? Most of the students in this study had not learned to apply beginning calculus flexibly by the end of the calculus-differential equations sequence and many might well never do so.[3]

[3]A. Robert and colleagues have introduced three levels of applying mathematical knowledge to tasks: (1). A *technical level* in which students are asked to apply calculus skills and use definitions, properties, and theorems directly. (2). A *mobilizable level* in which students adapt their knowledge to tasks which are not direct applications, require several steps, require some transformation or recognition that a property or theorem is to be applied. (3). An *available level* in which students solve problems without being given an indication of methods, or must change representations.

They tested one hundred French university students who had graduated (Licence) and were preparing for the CAPES competitive examination for teaching at secondary level (which just 1 in 8 pass); the curriculum of the exam is the same as that of mathematics majors in the first two years of university. They found that whenever problems were at anything but the technical level, the success rate was under 10% [6]. These findings seem comparable to those in this study regarding students, who at the end of their calculus-differential equations sequence, produced just 14% completely correct solutions to our non-routine problems (Table 6).

5.2 How does it happen that students can have the knowledge, but not be able to effectively use it to solve non-routine problems?

Part of the rationale for having the students take the routine test after the non-routine test was to determine whether they had the requisite algebra and calculus skills to solve the non-routine problems, but were unable to think of or use them; this was a concern raised by the study done with C calculus students [13]. Adding together the numbers in Table 4, one sees that, for most non-routine problem attempts (103 out of a possible 140), the differential equations students had an adequate knowledge base (i.e., full or substantial factual knowledge), yet just 32 of these attempts were successful (i.e., produced completely or substantially correct solutions to non-routine problems). The inability of many otherwise successful students to access and effectively use their factual knowledge in non-routine problem solving is perhaps the most striking feature of our data.

Editorial comment on our second study [13] of A/B calculus students raised the question of whether the parameterizations in some of these non-routine problems might have caused them to be viewed as questions about families of functions, thereby rendering them inordinately difficult. Indeed, three of the five problems *could* be viewed in this way (Problems 1, 3, and 5). However, rather than being interpreted as indicating families of functions, the letters a and b appeared to have been interpreted by these students as fixed unknowns whose values they were expected to find. Much of their written work supports this idea. The a and b seemed to play much the same "arbitrary constant" role for these students as m and b do in the slope-intercept form of a line, $y = mx + b$. In fact, it was on these three problems (especially Problem 5) that the students did the best. Problems 1, 3, and 5 accounted for 85% (17 of 20) of the completely correct solutions.

We conclude that these problems were no more difficult for our students than Problems 2 and 4. If anything, Problems 2 and 4 proved to be the most difficult ones, and it was on these problems that students used algebraic and numerical trial-and-error methods. While it was possible to solve Problem 2 without calculus and three students did so completely correctly, Problem 4 was particularly intractable without calculus. Some students may not initially have accessed their calculus knowledge because this problem brought to mind algebraic techniques for solving an equation, namely $4x^3 - x^4 = 30$. Developing a calculus-based solution would have entailed considering $4x^3 - x^4$, alternatively $4x^3 - x^4 - 30$, as a function to be maximized, something they did not do. This might suggest the students were lacking some general metacognitive abilities of expert problem solvers, e.g., control [10, 11]. However solutions to our non-routine problems are relatively "straightforward." For example, there is no need to consider multiple sub-problems and the unfruitfulness of false starts, such as attempting to factor $4x^3 - x^4 - 30$ in Problem 4, are not especially hidden. Thus we do not see metacognitive abilities, or lack thereof, as accounting for all, or perhaps even most, of our data. We suspect that for some of these students even seeing that a new idea is needed and having full factual knowledge may not bring the requisite knowledge to mind.

The question of access appears not to be a simple one. One can have the knowledge, but not think of using it. Doing so may often be supported by an additional kind of knowledge beyond what we have called full factual knowledge, i.e., beyond mathematical knowledge that is logically adequate to solve a given problem.

The nature of this additional knowledge cannot be fully established from an analysis of data such as ours, which does not emphasize the process aspect of problem solving. However, we will frame a discussion of it in terms that might be useful in later, more process-oriented research. A person who has reflected on a number of problems is likely to have seen (perhaps tacitly) similarities between some of them. He or she might be regarded as recognizing (not necessarily explicitly or consciously) several overlapping problem situations, each arising from problems with similar features. For example, after much exposure many students will recognize a problem as one involving factoring, several linear equations, or integration by parts, etc.[4] Such problem situations act much like concepts. While these situations may lack names, for a given individual they are likely to be associated with images, i.e., strategies, examples, non-examples, theorems, judgments of difficulty, and the like. Following Tall and Vinner's idea of concept image [16], we will call this kind of mental structure a *problem situation image* and suggest that some such images may, and others may not, contain what we will call *tentative solution starts*: tentative general ideas for beginning the process of finding a solution. The linking of problem situations with one or more tentative solution starts is a kind of (perhaps tacit) knowledge. For instance, the image of a problem situation asking for the solution to an equation might include "try getting a zero on one side and then factoring the other." It might also include "try writing the equation as $f(x) = 0$ and looking for where the graph of $f(x)$ crosses the x-axis," or even "perhaps the max of f is negative so $f(x) = 0$ has no solution." We suggest that the problem-solving process is likely to include the recognition of a problem as belonging to one or more problem situations, and hence, to bring to mind a tentative solution start contained in one's image of one of those problem situations. This may, in turn, mentally prime the recall of resources from one's knowledge base. Thus a tentative solution start may link recognition of a problem situation with recall of appropriate resources, i.e., what we have called full factual knowledge. For example, in this study a number of students tried factoring on Problem 4. When this approach did not work, had those students' images of such problems included "try looking at whether the graph crosses the x-axis," they might have recalled their knowledge of graphs and calculus to discover that the maximum was too small for the equation to be solvable. In viewing our data from this perspective, we are suggesting that problem situations, their images, and the associated tentative solution starts all vary from student to student and that the process of mentally linking recognition (of a problem situation) to recall (of requisite resources) through a problem situation image might occur several times in solving a single problem. We are not suggesting this is the only way access might occur, just that it could play a role in solving the kind of moderately non-routine problems discussed in this paper.

Except while taking tests the students in this study would have had worked examples (from textbooks and lectures notes) available to them during most of their previous problem-solving attempts. Those who habitually consulted such worked examples before attempting their own solutions would have had little occasion to reflect on multiple tentative solution starts. Such students might well have problem situation images with few tentative solution starts, thereby reducing the usefulness of their situation images in priming the recall of their factual knowledge; this could happen even if they realized a new idea is needed, that is, even if they did not lack (metacognitive) control [10, 11]. In summary, we are suggesting that some of the students who were unable to solve non-routine problems while having full factual knowledge may have lacked a particular kind of knowledge, namely tentative solution

---

[4]Although the kinds of features noticed by students in mathematical problem situations do not seem to have been well studied, the features focused on in physics problem situations have been observed to correspond to degree of expertise. Novices tend to favor surface characteristics (e.g., pulleys), whereas experts tended to focus on underlying principles of physics (e.g., conservation of energy). [4]

starts, and this might be due to a combination of the way homework exercises had been presented and the students' study habits.

Much of our data concerning those students who had adequate factual knowledge, but did not access it to solve non-routine problems, is consistent with the above analysis. However, our data do not adequately encode the problem-solving process for this proposed explanation to be more than a conjecture which might be examined in a future study.

5.3 Does it matter whether students are able to solve non-routine problems?

Perhaps surprisingly, the answer seems to be both yes and no. No, because the students in this study were among the most successful at the university by a variety of traditional indicators, both at the time of the study and subsequently, yet half of them could not solve a single non-routine problem. They had overall GPAs of just above 3.0 at the time of the study and almost double the graduation rate of the university as a whole. At least seven of them subsequently earned a master's degree and one a Ph.D. in mathematics. Furthermore, the idea that traditional success may not require very much non-routine problem-solving ability, including metacognitive control, is supported by De Franco's problem-solving study comparing mathematicians of exceptional creativity (e.g., Fields medallists) with very successful published Ph.D. mathematicians. He found that while both were content experts, only the former were problem-solving experts [4]. Thus, it seems possible to be academically successful in mathematics and related subjects without being able to consistently solve non-routine problems, especially the more difficult ones in which Schoenfeld's problem-solving characteristics (heuristics, control, beliefs [10, 11]) play a large role.

On the other hand, yes, it does matter. Most mathematicians seem to regard this kind of problem solving as a test of deep understanding and the ability to use knowledge flexibly. In addition, most applied problems that students will encounter later will probably be at least somewhat different from the exercises found in calculus (and other mathematics) textbooks. It seems likely that much original or creative work in mathematics would require novel problem solving at least at the modest level of the problems in this study.

In addition mathematicians often appear to view students' mathematical ability through the lens of problem solving, meant broadly to include a wide range from simple exercises to non-routine problems and the construction of proofs. They typically design problems whose solution requires deeper understanding rather than gauging such understanding directly, for example, through evaluating an essay on the nature of continuity and its relationship to differentiability.

It may be helpful to consider problems arising in mathematics courses as arranged along a continuum according to their novelty relative to the course. At one end, we might have *very routine* problems which mimic sample problems found in the text or lectures, except for minor changes in wording, notation, coefficients, constants, or functions that students view as incidental to the way the problems are solved. Such problems are often referred to as exercises (and might not be counted as problems at all in the problem solving literature). Moving toward the middle of the continuum, we might place *moderately routine* problems which, although not viewed as exactly like sample problems, can be solved by well-practiced methods, e.g., ordinary related rates or change of variable integration problems in a calculus course. Sandra Marshall [8] has studied how schema can be developed to reliably guide the solution of such problems. Moving further along the continuum, one might have *moderately non-routine* problems, which are not very similar to problems that students have seen before and require known facts or skills to be combined in a novel way, but are "straightforward" in not requiring, for example, the consideration of multiple sub-problems. The non-routine test in this study consisted of problems of this kind. Finally, at the opposite end of the continuum from routine problems, one might place *very non-routine* problems which, while dependent on knowledge of the course, may involve more insight, the consideration of several sub-problems or constructions, and use of Schoenfeld's behavioral problem-solving characteristics (heuristics, control, beliefs [10, 11]). For such problems a large supply of tentative solution starts, built up from experience, might not be adequate to bring to mind the knowledge needed for a solution, while for moderately novel problems it probably would. Often the Putnam Examinations [17] include such very non-routine problems.[5]

Most university mathematics teachers would probably like students who pass their courses to be able to work a wide selection of routine, or even moderately routine, problems. In addition, we believe that many such teachers would expect their better students to be able to work moderately non-routine problems, and think of this as functionally equivalent to having a good conceptual grasp and understanding of the course. In other words, the ability to work moderately non-routine problems based on the material in a course is often part of the implicit curriculum. Thus, it matters if most students cannot work moderately non-routine problems because in a sense that means the "essence" of the material in the course is not being successfully mastered, even by good students.

The results reported here suggest that, at least in a traditional calculus/differential equations sequence, many good students may not reach the level of understanding and moderately non-routine problem-solving ability that their teachers expect. In order that good students reach this level, non-routine problem solving may need to become an explicit part of the curriculum, that is, to be in some way explicitly taught. Furthermore, our conjectured explanation of our data, that is, that students' problem situation images tend to lack a variety of tentative solution starts, suggests that the ability to solve moderately non-routine problems may depend partly on a richer knowledge of problem situations as well as on more general behavioral problem-solving characteristics (heuristics, control, beliefs [10, 11]). This, in turn, suggests that teaching aimed at improving non-routine problem solving ability should be integrated throughout a course, rather than saved for the end, and that encouraging students to consider and reflect on tentative solution starts might be useful.

---

[5]The following problem was on the 59th Annual William Lowell Putnam Mathematical Competition given on December 5, 1998: Given a point (a, b) with $0 < b < a$, determine the minimum perimeter of a triangle with one vertex at (a, b), one on the x-axis, and one on the line $y = x$. You may assume that a triangle of minimum perimeter exists.

This appears to be a calculus problem, but it only requires clever use of geometry. An elegant solution (posted by Iliya Bluskov to the sci.math newsgroup) involves extending the construction "outward" by reflecting across both the lines $y = x$ and the x- axis and noticing that the perimeter of the triangle equals the distance along the path from (b, a) to $(a, -b)$. Probably only very experienced geometry problem solvers could have previously constructed images of problem situations containing a tentative solution start that would easily bring this method to mind.

REFERENCES

1. Amit, M. and Vinner, S., *Some misconceptions in calculus: Anecdotes or the tip of an iceberg?* In G. Booker, P. Cobb, and T. N. de Mendicuti (Eds.), Proceedings of the 14th International Conference of the International Group for the Psychology of Mathematics Education, 1 (1990), CINVESTAV, Mexico, pp. 3-10.
2. Berkey, D. P., *Calculus*, 2nd ed., Saunders College Publishing, New York, 1988.
3. Carlson, M. P., *A cross-sectional investigation of the development of the function concept*, CBMS Issues in Mathematics Education: Research in Collegiate Mathematics Education III, 7 (1998), 114-162.
4. Chi, M. T. H., Feltovich, P., and Glaser, R., *Categorization and representation of physics problems by experts and novices*, Cognitive Science 5 (1981), 121-152.
5. DeFranco, T. C., *A perspective on mathematical problem-solving expertise based on the performance of male Ph.D. mathematicians*, CBMS Issues in Mathematics Education: Research in Collegiate Mathematics Education II, 6 (1996), 195-213.
6. Dorier, J. L., Pian, J., Robert, A., and Rogalski, M. *A qualitative study of the mathematical knowledge of French prospective maths teachers: Three levels of practice*, Pre-Proceedings of the ICMI Study Conference on the Teaching and Learning of Mathematics at University Level (1998), National Institute of Education, Singapore, 118-122.
7. Eylon, B. S. and Lynn, M. C., *Learning and instruction: An examination of four research perspectives in science education*, Review of Educational Research 58 (1988), 251-301.
8. Marshall, S. P., *Schemas in Problem Solving*, Cambridge University Press, New York, NY, 1995.
9. Stacey, K. and MacGregor, Mollie., *Multi-referents and shifting meanings of unknowns in students' use of algebra.* In E. Pehkonen (Ed.), Proceedings of the 21st International Conference of the International Group for the Psychology of Mathematics Education, 4 (1997), Gummerus, Finland, pp. 190-198.
10. Schoenfeld, Alan H., *Mathematical problem solving*, Academic Press, Orlando, FL, 1985.
11. Schoenfeld, A. H., *Learning to think mathematically: problem solving, metacognition, and sense making in mathematics.* In D. A. Grouws (Ed.), Handbook of Research on Mathematics Teaching and Learning (1992), Macmillan, New York, NY, pp. 334-370
12. Selden, J., Mason A., and Selden, A., *Can average calculus students solve nonroutine problems?*, Journal of Mathematical Behavior 8 (1989), 45-50.
13. Selden, J., Selden, A., and Mason A., *Even good calculus students can't solve nonroutine problems*, Research Issues in Undergraduate Mathematics Learning, MAA Notes 33 (1994), 19-26.
14. Stewart, J., *Calculus*, Brooks/Cole, Monterey, CA, 1987.
15. Swokowski, E. W., *Calculus with Analytic Geometry*, Alternate Edition, Prindle, Weber & Schmidt, Boston, MA, 1983.
16. Tall, D. and Vinner, S., *Concept image and concept definition with particular reference to limits and continuity*, Educational Studies in Mathematics, 12 (1981), 151-169.
17. The William O. Putnam Exam, *Mathematics Magazine*, 64 (1991), 143.
18. Zill, D. G., *Differential Equations with Boundary-Value Problems*, PWS-KENT, Boston, MA, 1986.

May 26, 1999