Mathematics Education Research Journal 2001, Vol. 13, No. 3, 165-184

Retention of Concepts and Skills in Traditional and Reformed Applied Calculus

Bradley E. Garner
Boise State University

Lynn E. Garner
Brigham Young University

Striving to make calculus more meaningful for students, recent calculus reform initiatives have shifted emphasis from rote memorisation and symbol manipulation to conceptual understanding and practical application. But is reform calculus making a difference? This study compares outcomes of a traditional and a reform calculus course in terms of students' retention of basic concepts and skills after the passage of time. Although traditional and reform students did not differ statistically in overall performance, reform students retained better conceptual knowledge and traditional students retained better procedural knowledge. Reform calculus students also demonstrated that concepts can be understood before computational competence is achieved.

Over the past decade there has been much change in secondary school and collegiate calculus teaching. The traditional teaching of calculus has been seriously challenged because of technological advances, the rote way it has been taught, and ill-prepared students (White, 1995). It has also been questioned because it has left gaps in students' conceptual understanding (Thomas & Hong, 1996). Persistent high dropout and failure rates in traditional courses and mounting evidence that even apparently successful students acquire limited understanding of calculus concepts have convinced many instructors that there must be better curriculum materials and teaching strategies (Selden, Selden & Mason, 1994, p. 19). Recent calculus reform initiatives have sought ways to make the subject more meaningful for students and to lead them to better understanding of key ideas and techniques. These initiatives have tried to present key calculus concepts in more depth with less stress on memorisation of algorithms and symbol manipulation (Schoenfeld, 1997, p. 2).

In this study, we will refer to calculus for mathematics, physical science, and engineering majors as "theoretical calculus." We refer to calculus for business, social science, and life science majors as "applied calculus." Most of the calculus reform efforts have focused on theoretical calculus. Little reform work has been done in applied calculus (Narasimhan, 1993, p. 254).

There are several themes underlying much of the calculus reform movement. One of these is the shift in emphasis from rote memorisation and symbol manipulation to conceptual understanding and practical application (Cavanagh, 1996). In other words, reformers are less interested in whether a student knows how to compute a given definite integral and more interested in whether the student understands the meaning of that definite integral and can use it appropriately in a real-life situation.

A second major theme of the calculus reform movement is the use of technology (computers and calculators). Technology is having an impact on mathematics teaching and learning for many reasons, not the least of which is its universal proliferation. As one instructor put it, "Technology is changing the way we teach. Not because *it's here*, but because *it's everywhere*" (Kenelly, 1996, p. 24). A second reason for the prominence of technology in mathematics education is that instructors have recognised the potential power of computers and calculators (e.g., Moursund, 1985).

A third theme of calculus reform is the changing landscape of classroom instruction, where traditional focus on the individual student is giving way to increasing emphasis on group work and social interaction. Anderson, Greeno, Reder and Simon (2000) declare that for educational research, both the cognitive and the situative perspectives are valuable.

We shall call textbooks that have developed in consideration of the foregoing themes reform textbooks. For purposes of this paper, a course using a reform textbook shall be considered a reform course (provided the instructor's implementation of the course adheres to the goals envisioned in the textbook). In other words, a reform course is one which, through use of a reform textbook, emphasises concepts at least as much as procedures, requires frequent and appropriate use of technology, and employs a style of instruction where students are required to interact with each other in group work.

As new developments in the teaching of mathematics compete for attention, it is important for researchers to evaluate their effectiveness. We must evaluate curricula that exemplify the competing theories. Thus simple curriculum evaluation can be interpreted as a comparison of the effectiveness of differing theories of how calculus should be taught. The goal of this study was to provide such a comparison in the applied calculus arena, where very little evaluation has been done. Specifically, our guiding research questions are: How well do students of reform and traditional applied calculus courses understand key concepts and perform key procedures of applied calculus seven months after taking their courses, and how do these groups of students compare to each other?

An additional highlight of this study is its retention aspect. Some studies have investigated the effectiveness of reform courses, but few have sought to establish whether the reform courses have any lasting effect on their students. The "residual" knowledge of calculus after the passage of time is the key question.

This study focuses on commonalities in content between reform and traditional applied calculus courses. It compares the outcomes of the courses in terms of student understanding and the retention of basic calculus concepts and skills.

Relevant Literature

This study is connected to two main bodies of research literature: conceptual versus procedural knowledge and calculus reform.

Conceptual versus procedural knowledge. Mathematical knowledge can be thought of as being composed of two types: procedural and conceptual. Procedural knowledge is an understanding of the rules for completing mathematical tasks, including the use of symbolic language (Hiebert & Lefevre, 1986, p. 6). Conceptual knowledge is an understanding of mathematical ideas, including reasons why algorithms work and when they should be applied. According to Hiebert and Lefevre (1986), conceptual knowledge is "rich in relationships.... in fact, a unit of conceptual knowledge cannot be an isolated piece of information; by definition it is a part of conceptual knowledge only if the holder recognizes its relationship to other pieces of information" (pp. 3-4).

Both types of learning are important to competency in mathematics, and linking the two helps to create the kind of networked information that is better retained (Hiebert & Carpenter, 1992; Hiebert & Lefevre, 1986). On choosing which of the two types should come first, theory generally recommends concepts before procedures (Hiebert & Carpenter, 1992, p. 78). Byrnes and Wasik (1991) found support for the "dynamic interaction view" of the relationship between procedural and conceptual knowledge, which claims that, although conceptual knowledge must precede procedural knowledge at first, each type of knowledge subsequently can be enhanced by the other. Yet, as Hiebert and Lefevre (1986) note, "Formal mathematics instruction seems to do a better job of teaching procedures than concepts or relationships between them" (p. 22). Tempering this emphasis on procedures with deeper conceptual understanding is one of the central goals of curricular reforms in calculus.

Research in calculus reform. Much has happened in calculus reform since 1986, yet the number of published research studies investigating the effectiveness of reform curricula is relatively small in comparison to the number of testimonials recommending use of reform calculus curricula. In the United States, the research that has been done has been mainly in the form of doctoral dissertations and master's theses (Ganter, 1997, p. 11). Some of these have sought to evaluate the theoretical calculus curriculum developed by the Calculus Consortium based at Harvard (CCH); others have dealt mostly with the use of technology in calculus reform. In Australia, a reform calculus curriculum based on the work of Mary Barnes has been deemed fairly successful. As White and Pegg (1996) report,

> Cavanagh (1996)....found that Barnes' approach not only made calculus more relevant and interesting for his students, but that also their general understanding of calculus was enhanced. Further, the students' ability to perform standard differential calculus techniques was not adversely affected (p. 10)

Do other reform calculus curricula have these positive effects? More evaluation work is needed.

Studies conducted by Heid (1988) and Judson (1988) showed that students do not need to learn calculus procedures before they can understand calculus concepts. The main component in both studies was a re-sequencing of skills and concepts (concepts and applications were taught first, before focusing on algorithms for computation). Both studies also showed no substantial differences in computational abilities between experimental and control groups of students. Heid found better conceptual understanding in the experimental group of students; Judson did not (although the re-sequencing did increase motivation and interest). Park and Travers (1996) found similar results in their study of the Calculus and Mathematica curriculum; students experiencing the reform curriculum exhibited deeper conceptual understanding and more positive attitudes towards mathematics.

Several studies on the CCH reform of theoretical calculus (Allen, 1995; Brunett, 1995; Hadfield, 1996; Ratay, 1993) report mixed results. Ganter (1997), in her analysis of all 127 NSF-funded calculus reform projects, concluded that reforms have had an overall positive effect on student achievement, with some negative results. In general, those projects involving technology led students to increased conceptual understanding in comparison with traditional curricula; levels of procedural understanding achieved in reform courses were usually the same as or slightly lower than those achieved in traditional courses (*ibid*, pp. 12-13). Ganter reported further that the students most likely to do well in reform calculus courses using long-term projects or group work were above average mathematics students, those who do poorly on traditional tests, and engineering majors (*ibid*, p. 13). In contrast, Ratay (1993) found that weaker students benefited most from reform calculus. In summarising student achievement results, Ganter says,

> Student achievement has been affected by the calculus reform efforts. What is perhaps less clear is the degree to which achievement has been affected and the appropriate 'mixture' of reform ideas that should be implemented at various institution types to achieve the greatest positive effect. (1997, p. 13)

In light of this, it is reasonable to expect that some positive effect will be observed in studies such as the present one.

Because student and faculty attitudes both improve when modifications are made to reform courses on the basis of student input, Ganter concludes that "the instructor and the instructional methods—not reform textbooks—are the real crux of reform" (Ganter, 1997, p. 14). Yet textbooks can greatly influence instructors and their methods, thus having a significant impact on the outcome of reform courses:

> Commercially published curriculum materials dominate teaching practice in the United States (Goodlad, 1984). Unlike frameworks, objectives, assessments, and other mechanisms that seek to guide curriculum, instructional materials are concrete and daily. They are the stuff of lessons and units, of what teachers and students do. That centrality affords curricular materials a uniquely intimate connection to teaching. (Ball & Cohen, 1996, p. 6)

While the influence of textbooks on teaching at the college level may not be exactly the same as it is at the primary or secondary levels, certainly the connection exists—and thus it is worthwhile to evaluate courses using reform textbooks.

Methodology

For purposes of this study, we chose to separate all knowledge and understanding of applied calculus into two main (sometimes overlapping) categories: conceptual knowledge and procedural knowledge. While this view may be limited, it forms the basis of a very common approach to the teaching and learning of applied calculus. It also provides a way of analysing the content of most calculus textbooks, be they reform or traditional. To test the ideas of the reform movement in the applied calculus arena, we used a written test and interviews to compare the conceptual and procedural understandings of two groups of students who had completed either a reform or a traditional one-semester applied calculus course seven months after they had completed it.

The students and the course. Subjects for this study were students who were enrolled in the applied calculus course at a large private university in the western United States. Most were middle to upper class American citizens, with a few minority Americans and a few foreign students. Admission at their private university was competitive, so these students were above average in academic preparation.

At this particular university, the applied calculus course is a terminal course in calculus for business, biological science (including pre-medical and pre-dental) and computer science majors, having college algebra (a third course in algebra) as a prerequisite. The various departments that require this course of their majors deem the course important for their students, either directly for their careers or as necessary prerequisite material for topics they will encounter later in their major courses.

During the winter semester (January to April) of 1997, a total of 782 students were enrolled in applied calculus. Four hundred and eight of the students were enrolled in a total of 12 reform sections of 25 to 40 students each. These students are referred to as "reform students" in this paper. These sections met with their instructor (the second author) four times per week (six sections at a time in a large lecture hall) for 50 minutes. Each section also met once per week separately in a small classroom where a teaching assistant was available to answer questions and review course material. These students were tested periodically and had a comprehensive final examination at the end of the course. Each of the tests had two components: group problems and individual problems (students took the tests home to complete); most of the problems were presented in some application context (story problems).

The remaining 374 students (referred to as "traditional students") were enrolled in one of two large traditional sections (one of 165 students and one of 209), each taught by a different instructor. The smaller section met four times per week for one hour; the larger met twice per week for one hour and 50 minutes. Both sections met in the same large lecture hall as the reform sections. A tutoring centre was also available to all maths students, where they could receive individual help. The students in the traditional course were tested periodically and had a comprehensive final examination at the end of the semester. The tests these students received were much more traditional tests—they consisted mostly of drill exercises (problems devoid of any application context).

The sampling procedures developed for this study involved the selection of subjects by one author, concealing important characteristics of the subjects from the other, who analysed the data, thus ensuring a high degree of objectivity.

During the fall semester (September to December) of 1997, all students who had taken applied calculus the previous winter were sent a message that briefly introduced them to the study and invited them to participate. Those selected to take the written test were offered $10, and those selected to participate in the follow-up interviews would receive an additional $20. Students chosen for interviews (half reform, half traditional) were randomly selected from among study participants who received an A or B grade in their applied calculus course.

Hypothesising that the effect size of eventual differences found between the two groups would be medium (0.25), we originally sought 64 students from each group to participate in the study. This is the number per group required to perform a one-way analysis of variance (ANOVA) with power = 0.80 and with $\alpha = 0.05$ (Cohen, 1992, p. 158). A total of 93 students from the reform course and 58 from the traditional course responded to the invitation. We invited all responding traditional students to take the written test and trimmed the sample of reform students randomly to 58 also (with a grade distribution matching that of the traditional group of students). Having 50 students per group would still allow for power = 0.80 with $\alpha = 0.10$ (Cohen, 1992, p. 158). For this exploratory study, we decided that a 0.10 probability of Type I error was acceptable. Only those students who had received a grade of A, B, or C in their applied calculus course were invited to participate so that we would be dealing with those who were actively involved in their course and put forth a somewhat substantial effort. Of those selected, one reform student and seven traditional students failed to participate in the study. Since the grade distribution of the 108 students who did participate in the study was weighted more in favor of higher grades than the grade distribution of all applied calculus students during the winter semester (see Table 1), one would expect the calculus knowledge of those in the study to be slightly higher than the target population in general.

Table 1
Grade Distribution of Applied Calculus Students, Winter Semester 1997

	A	B	C	Total
Reform Sample	27	17	13	57
Reform Population	111	202	74	408
Traditional Sample	29	11	11	51
Traditional Population	136	87	85	374
Total Sample	56	28	24	108
Total Population	247	289	159	782

The principal author conducted interviews with 16 students (8 from the reform group and 8 from the traditional group) after they had taken the written test. They were chosen randomly from among those taking the written test who had received an A or B grade in their applied calculus course--4 A's and 4 B's in each group of 8 students.

The textbooks. The Calculus Consortium based at Harvard (CCH) published in 1996 a new text for applied calculus entitled *Applied Calculus for Business,* *Social Sciences, and Life Sciences* (Hughes-Hallett, et al., 1996). Although fairly new, this textbook was in use at 100 different colleges and universities at the time this study was conducted. Courses using this book as its authors intended fit this paper's definition of a reform applied calculus course (see preface of Garner & Hughes-Hallett, 1996). The CCH book was the one used by the instructor teaching the reform applied calculus at the study site.

A "traditional" textbook by Goldstein, Lay and Schneider (GLS) was also in use at the study site: *Calculus and its Applications* (Goldstein, Lay & Schneider, 1996). The GLS text continues to be widely used and was in its seventh edition at the time of this study (it is now in its ninth edition).

Analysis of the two textbooks reveals that they are indeed good representatives of their respective types. Although they contain essentially the same material, they differ in philosophy. The CCH book has a long preface in which fairly specific goals for student learning are listed; the GLS preface is a more general discussion of topics. The CCH authors tend to use applications to introduce and motivate a concept or skill, treating the application as the reason for calculus as well as the end result of using calculus (Hughes-Hallett, et al. 1996, p. vii). The GLS text often presents facts or theoretical ideas first and embodies them in applications thereafter. The CCH text leads students along gradually, spending more time on the development of the reasoning behind calculus; the GLS text develops the calculus concepts and techniques more quickly, allowing less time for student exploration. The CCH text is designed to cover fewer topics than the GLS text, but it explores them in greater depth and uses a greater number of approaches to each topic. The CCH book expressly states its goal of treating topics in four different ways ("the rule of four"): geometrically, numerically, algebraically, and verbally (Hughes-Hallett et al., 1996). The GLS text also uses all four of these ways at various points, but in a much less focused way. Finally, the CCH text is aimed more at student activity than the GLS text (Garner & Hughes-Hallett, 1996, pp. 5-6).

The instructors. Three instructors were involved in this study. The instructor who taught the reform course (the second author) has taught calculus for 29 years, teaching applied calculus occasionally during that time. In 1991, he began teaching reform theoretical calculus, and from fall 1994 to winter 1997 he taught reform applied calculus exclusively. The instructor who taught the smaller traditional section has taught calculus for about 30 years, teaching applied calculus most of that time. He used the traditional approach only. The instructor who taught the larger traditional section has taught applied calculus for about ten years, using only the traditional approach. He has also taught theoretical calculus occasionally. All three instructors had much experience in both the course and the teaching approach. Each was committed to the approach he used, and each seemed equally preferred by the students. All three instructors were interviewed.

The second author assisted in making arrangements for the study to take place (including sampling, keeping important characteristics of the subjects concealed from the principal author, etc.) and assisted in the analysis, but he had little or no part in the design of the study. The study was designed after the courses had terminated.

Interviews. The student interviews had two parts: the first part afforded insights into the way students selected their classes and the way the classes were taught; the second part clarified students' understandings about calculus ideas and the written test they had taken. The interviews lasted approximately 30 minutes and were audio recorded and later transcribed. In the analysis, each time a new idea was mentioned in a transcript, it was summarized and listed as the heading for a spreadsheet column. No ideas mentioned by students were omitted. There were 16 rows in the spreadsheet, one for each student interviewed. When students' transcripts contained one of the ideas, the appropriate row and column intersection was marked. The resulting counts allowed for summary of the opinions expressed by students in terms of both variety and strength.

Differences between the reform and traditional courses gleaned from the interviews are outlined below.

In the reform course, students worked in groups or did homework in class, spending a lot of time talking with each other while the instructor roamed the classroom and visited each group. This did not occur in the traditional classes.

Students from the reform course had more homework than the other students in the sense that they were assigned more difficult problems (rather than a greater number of problems). Two students said the homework problems were all "word problems", which probably meant they were descriptions of real-life applications such as were emphasised in the CCH textbook. Traditional course students said things such as "the homework wasn't too strenuous" and it took "maybe an hour a night", whereas one or two of the reform course students remarked that they spent 2 or 3 hours each night. One student described the homework this way: "It wasn't doing the problems over and over again, it was each problem I had to think harder about."

When asked whether the way the class was taught agreed with their own learning styles, all eight of the traditional course students interviewed responded in the affirmative. In contrast, only four of eight reform course students replied in like manner—and two of them specifically said that the teaching disagreed with their preferred learning style. Similar findings were reported by Bookman and Friedman (1998), who said that a reform course can "violate students' deeply held beliefs about the nature of mathematics and how it should be learned" (p. 117). Students who normally have been taught in a traditional lecturing style may naturally be uncomfortable with new emphasis on working in groups or discovering more of the concepts on their own. As one student put it, "I'm more of a numbers person as opposed to, like, a story type person." Another student said it this way: "...he always said this isn't a plug and chug class, and I'm like, 'but that's what I'm used to, that's what I'm good at!' " Students generally agreed that the teaching styles they encountered in class matched their respective textbooks.

The instructors' comments about their own teaching help illuminate what happened during class time. The reform instructor talked about trying "to get the students engaged in the process" and having the students work together in small groups, whereas the traditional course instructors encouraged group work outside of class. The reform course instructor said he "...had the classroom noisy a good part of the time" and tried to wander around and interact with the groups. The reform course instructor mentioned that because the class was so large, he had to let the students work more in groups than usual. In contrast, the traditional course instructors said that, although they encouraged questions from students, a lot of interaction was difficult and the class had to be run more on a lecture format because of class size. Both traditional course instructors said they preferred traditional teaching, and the reform course instructor said he preferred a reform style. Reform ideas were probably new to most students, so there is a possibility that there was a teacher effect for the reform group, despite the fact that the two groups of students liked their instructors equally well.

Calculators were allowed in the traditional courses, but graphing capabilities were not utilised and not much time was deliberately spent on using them. Graphing calculators were not allowed on exams. In the reform course, graphing calculators were required, were frequently used, and were allowed on exams. The instructor gave students supplementary software for various calculus topics.

All of the foregoing confirms that the traditional course was indeed traditional, and the reform course was indeed a reform course. The interviews also showed that students did not register for a particular section of applied calculus based on whether it was traditional or reform (14 of the 16 students interviewed said that scheduling was the reason for their choice of sections). Thus we can reasonably attribute differences between the two groups of students to the nature of the courses rather than to differences between the students themselves.

The test. We created a written test containing 10 items (a total of 18 parts worth 67 points; see the Appendix) that included items in three categories: primarily conceptual, primarily procedural, and both conceptual and procedural. Items were structured in this fashion because the interplay of conceptual and procedural knowledge was of primary interest in the study; more and deeper conceptual understanding is the main thrust of the CCH reform. The test items were also designed to represent the main topics of instruction addressed by both textbooks. The key concepts and procedures distilled from a content analysis of the CCH and GLS texts were: description of graphs of functions (both with and without using derivatives); differentiation of functions (both graphically and analytically); interpretation of the derivative (mostly in economic situations); integration of functions (both graphically and analytically); interpretation of the definite integral; and the relationship between the derivative and the definite integral (the fundamental theorem of calculus). We believe that someone receiving a perfect score on the written test would have a good basic understanding of the main concepts and procedures of applied calculus.

In regard to the validity of the written test, we note that it was piloted with maths students (students who had completed an applied calculus course) in an effort to refine clarity and to determine whether it assessed the desired concepts and procedures, after which slight modifications were made. The test was administered in practically identical conditions to all students who took it through the use of the testing centre on campus (calculators were allowed).

Written test items. We provide here a brief rationale for the selection of each test item (the full test can be found in the Appendix). Item 1 was designed to assess the graph-analysing capabilities of the students without using derivatives (describing monotonicity and extrema by looking at graphs). We consider item 1 to be primarily conceptual, given that it deals with how various parts of a graph relate to each other rather than with symbols and manipulation.

Item 2 asked students to use calculus to answer the same questions asked in item 1. This involves symbol manipulation and recall of the facts that a positive derivative means increasing, a negative derivative means decreasing, and a zero derivative means a possible local extreme value. Since no explanation is required, one could score well on this item with little conceptual understanding; thus we label it a procedural item. The function given was the same function used to create the graph in Item 1, which, if recognised, gave students a way to check their work against the graph.

Item 3 was a procedural question, asking students to perform some simple differentiation. The first function given required use of the chain rule; the second required the product rule as well. Fairly simple functions were chosen for this item in the hope that many students would remember how to differentiate them.

Item 4 was a word problem about the maximisation of a function. It gave a typical parabolic function for revenue in terms of price. The goal of this question was to require students to compute a derivative to find a critical value, and then to understand how to use that value to answer another question. Since this item involves both manipulating symbols and interpreting what the results mean, we consider it to be both conceptual and procedural.

The goal of item 5 was to get students to explain the derivative as marginal cost. Marginal cost was one of the main applications found in both the GLS and CCH texts. This item is very open and conceptual.

Item 6 is conceptual in that it requires students to think about the sign and magnitude of the derivative of a function from every-day life. Students needed to realise that the slope of $f(x)$ was positive for all positive x, and to make the inference that the motor home used more gas, and thus its graph would be above the other graph.

The goal of Item 7 was three-fold: to require students to explain how antidifferentiation can transform one function into another; to find out whether they remembered how definite integrals can be used to compute area; and to require them to give the meaning of the area of a shaded portion of a graph in economical terms. All these are conceptual ideas important in applied calculus.

Item 8 tests the students' ability to estimate derivatives and integrals. For part (a) they needed to use the formula for the slope of the line joining two points to get an estimate for the derivative. For part (b) they needed to add up areas approximating the area under the function curve. Since this item involved both symbol manipulation and interpretation of results, we classified it as both conceptual and procedural.

Item 9, involving computation of definite integrals, was considered to be procedural. The functions chosen called upon knowledge of integrating with the power rule and exponentials.

The connection between the derivative and the integral was the subject of test item 10. Students needed to remember the fundamental theorem of calculus. It was postulated that not many students from either course would be able to state this theorem entirely correctly, since most applied calculus classes focus on it only briefly. However the question was designed to see what students remembered about the relationship of the derivative and the definite integral, the two central ideas of calculus. While one might argue that this item is conceptual in that it prompted students to think about relationships between calculus concepts, the way the question was stated was primarily procedural in that it could be answered through pure recall of the statement of the theorem.

The test questions were grouped into three categories: primarily conceptual (items 1, 5, 6, and 7, which require explanation and understanding of principles underlying computational algorithms); primarily procedural (items 2, 3, 9 and 10, which require simple recall of facts or knowledge of how to manipulate symbols when computing a derivative or integral); and both conceptual and procedural (items 4 and 8).

Data analysis. A grading key was prepared which gave in detail the answers expected and the number of points allotted for various responses. The principal author graded each problem in turn and recorded test scores. Next, the grading key was refined and clarified. A subset of 10 tests was chosen (5 reform, 5 traditional), the scores were concealed, and copies were distributed (along with the original key) to three of the principal author's colleagues. Once their grading was complete, the author met separately with each one and compared scores, discussed any differences, and decided upon changes to be made in the key. All 108 tests were searched for places where scores needed to be changed and final scores for each test were recorded.

Statistical comparisons were performed on the scores from the written tests in three ways. First, mean total test scores from the traditional and reform groups were compared using standard one-way analysis of variance (ANOVA) procedures on the total scores of all students. Three new scores of interest were tallied for each student: first, a total of all points earned on conceptual problems; second, a total of all points earned on the procedural problems; third, the total of all points earned on problems classified as both conceptual and procedural. The ANOVA procedure was performed on the averages (reform and traditional) of each new score.

Strictly speaking, the ANOVA procedure is only to be applied to completely random samples. Thus any significant results obtained through ANOVA testing in our study can only be interpreted as indicators of where possible differences may lie; we cannot make firm conclusions or generalisations to the larger population.

Results

Although the group of students from the reform class had a higher overall mean score (25.9) on the written test than their counterparts from the traditional course (25.1), the difference was not statistically significant ($F = 0.103, p = 0.749$). It was disappointing that both groups scored so low (73% of all students received a score at or below 50%).

The most interesting results came from comparing the groups of students by their mean procedural and conceptual subtotals. The difference in mean conceptual scores (see Table 2), favoring the reform course students, was statistically significant ($F = 6.296$, $p = 0.014$), as was the difference in mean procedural scores (see Table 3), which favored the traditional course students ($F = 7.2, p = 0.008$). Note that for all statistical tests reported, the two sections of the traditional course are considered as one group; this is justified because the two sections were not significantly different from each other statistically (that is, the instructor factor was negligible when considering total scores, conceptual scores, and procedural scores). Note also that students' total scores on the written test were consistent with grades earned in the courses, as summarised in Table 4.

Table 2
Conceptual Score Distributions Summary (Max. 32)

	Mean	SD	Maximum	Minimum
Reform	14.9	5.6	28.0	2.0
Traditional	12.4	4.7	26.5	4.5

Table 3
Procedural Score Distributions Summary (Max. 21)

	Mean	SD	Maximum	Minimum
Reform	5.3	4.8	16	0
Traditional	7.9	5.5	17	0

Table 4
Test Average by Course Grade (Max. 67)

Course Grade	Reform Group Average	Traditional Group Average
A	34.1	29.6
B	20.5	24.6
C	15.9	13.8

Results of the calculus portion of the student interviews added some insights into the test results. While the results of the written test seem to show the two groups of students very close together, particularly in terms of overall performance, the interviews revealed that there were differences between them.

Students from the reform course seemed more confident in their ability to explain derivatives. There were six students who said they could not explain what a derivative was; five of these were from the traditional course. The notion that one can understand and use derivatives without being able to explain them was clearly brought out in the interviews. As one student from the traditional course put it, "...you focus so much on just doing the problems that I didn't really think about what it meant."

The reform course students also mentioned graphs more in reference to derivatives than did their counterparts from the traditional course. When asked how they would explain the derivative to someone, only students from the reform course (four of eight) said they would use a graph. More reform students than traditional also mentioned the change in a curve or function, the slope of a line or graph, or the slope of a line tangent to a graph. Students from both groups actually used graphs in their explanations, but the reform students seemed more comfortable with graphs as their primary vehicle for explanation.

In general, integration meant anti-differentiation to the students from the traditional course. They seemed to have a clearer idea of the connection between the derivative and the integral; they spoke of the derivative and the integral being linked as opposites (all eight traditional students mentioned this idea at one point or another in their interviews; only four of the reform students did). For the students from the reform course, integration had more to do with finding area in various applications. Most of them said that the integral had to do with the area under a curve. We note that the reform students studied Riemann sums (and the traditional students did not). Why didn't the reform students understand the connection between derivatives and integrals better? It could be that the reform course stressed applications so much that students lost sight of the connections.

The traditional students also seemed to know the quick way of finding the slope of a graph at a particular point (differentiate and plug in the x-value), whereas the reform course students seemed more inclined to use estimation techniques (usually finding the slope between two points on the graph). This difference relates to the traditional course's emphasis on calculating derivatives of given functions and the reform course's emphasis on applications, where the functions aren't always given (or are given in tables, which traditional students never saw) and estimation is required.

The reform students also recognised the limit definition formula for the derivative more readily than the traditional students. The reason for this may be that there is a greater emphasis placed on the definition in the reform course. One reform student mentioned that the formula was "on every quiz," whereas one traditional student said that the class had used that formula early in the semester before finding a shortcut method.

Finally, many of the students from both groups gave imprecise descriptions of mathematical ideas. Sometimes this seemed to be due to confusion on their part about calculus concepts; other times they seemed to lack the vocabulary they needed to describe adequately what they were thinking. In terms of educational significance, neither course was successful in helping students achieve a lasting facility with the mathematical vocabulary.

Discussion

In overall comparison, there was no statistically significant difference between the reform course students and the traditional course students in terms of mean scores on the written test. This suggests that if an instructor or an institution is interested only in the total overall learning and retention attained by its applied calculus students, it might not pay to invest in the change to a reform curriculum. In other words, the ideas mentioned earlier for teaching reform applied calculus, at least in this case, were no more effective in general than traditional teaching ideas as measured by this test.

Yet the knowledge of calculus retained by students in the reform course was clearly different from that of students in the traditional course. The statistically significant difference between groups in mean conceptual and procedural scores is evidence of this fact. Students in the reform course gained in conceptual understanding of calculus as compared to students in the traditional course, but they did not acquire or retain the same procedural ability. This trade-off between concepts and procedures means that the choice of curricula made by instructors should depend upon the goals they envision for students. If the desire is for students to gain increased conceptual understanding, it might well be advisable to choose some type of reform curriculum. If, on the other hand, facility with calculus computations is desired, then there is reason to choose a more traditional curriculum. In other words, the question "What curriculum should I use?" is best answered by "It depends on what you want students to learn." This brings us to the ongoing debate about the relative merits of procedural and conceptual knowledge in mathematics.

Hiebert and Lefevre (1986, p. 8) point out that procedural knowledge can be learned by rote (i.e., without meaning). A student from the traditional group interviewed in this study said something that exemplified this idea of rote learning. When asked for any final comments, he remarked:

> Well, it occurred to me that I can still remember how to do some things...but I don't know what exactly they are.... I know how to get the derivative of a function, I know how to get the integral and find areas between two curves at certain points...but I just can't remember what it means—how to use it....

Proponents of calculus reform argue that being able to get the right answers to procedural questions without understanding how or why the algorithms work, or without being able to correctly interpret methods or answers in authentic problem contexts, is of little value. They also point out that people will increasingly use calculators or computers to perform mathematical calculations. Therefore we should focus more on conceptual teaching than we have in the past, so that people will be able to use technology appropriately and effectively.

Opponents of calculus reform argue that one cannot fully understand or appreciate the concepts until they gain proficiency in skills. They believe that knowing the algorithms and being able to use them is the most important part of mathematics, and that reformers are trying to take a short cut that leads nowhere. As George Andrews (one critic of reform) put it, "The reformers believe that they will get around the roadblocks of basic arithmetic so students can get to higher-order skills. But to learn piano, you must learn scales and chords before you move to the 'Moonlight Sonata' " (Wilson, 1997, p. A13). This study has shown that the "scales before Beethoven" analogy is not necessarily true. In fact, it has shown that it is possible to understand concepts better than procedures or vice versa.

It is also not true that everyone who learns mathematics will become a mathematician, just as not all musicians will become concert pianists. As noted earlier, the things we require students of mathematics to learn depends on what use they will have for mathematics.

This study also suggests that the differential effects of both the CCH and GLS curricula are not transient, although a lot of forgetting occurred. This forgetting may be disappointing, but is not surprising, given that forgetting occurs very quickly at first. Both curricula were strong enough that their students remembered measurable amounts of material seven months later. It would be informative to investigate the difference between the groups after additional passage of time.

Interviews with students in this study generally supported the results of the written test in that the reform students seemed to be more conceptually oriented than the traditional students. The reform group seemed more confident in trying to explain things, talked more about applications of calculus, and used graphical explanations more than the traditional group. This agrees with the research of Selden, Selden, and Mason (1994), in which they found that students of traditional calculus have weaker graphical knowledge than analytical knowledge. They also found that these students knew the skills but could not apply them to solve problems.

Are the tenets of reform applied calculus better than those of traditional applied calculus? We do not know yet. It appears that the CCH applied calculus, at least, is achieving its goal of increasing the conceptual understanding of its students. Although losses in computational facility may not have been desired by the CCH authors, they may have decided that some loss was inevitable, and were willing to accept that in exchange for deeper conceptual understanding. The question about the apparent trade-off between concepts and procedures still remains: must one form of knowledge always be diminished in order to make room for growth of the other, can they both be enhanced at the same time, or can they mutually enhance each other?

There is a tendency on the part of some educators to discard current reform efforts in calculus. Cursory reading of studies such as this show no big gains, and it may be decided that it is not worth the effort to make big changes in teaching. As one researcher put it:

> The lack of studies to indicate that [reform] efforts are having a positive impact on students, together with the increase in workload brought on by reform, is creating an environment of uncertainty that could result in the withdrawal of support for such activities by funding agencies, institutions, faculty and students. (Ganter, 1997, p. 3)

We must remember that current reform efforts are still relatively young, and there are still too many unanswered questions. It is premature to judge the calculus reform movement. Those supporting it need to keep developing and evaluating it. Not only does more research in this area need to be done; it also needs to be brought to the attention of those making decisions about teaching and curricula.

It may be that "reform vs. traditional" is not the most important question. Looking again at how low the test scores were after seven months, we may draw the conclusion that it does not matter what kind of course we teach. Although some comparisons between the two groups showed *statistical* significance, there was no evidence of *educational* significance: both reform and traditional students forgot most of what they supposedly had learned.

Perhaps those who teach calculus should spend less time thinking about which type of knowledge to emphasize or which topics to cover and spend more time thinking about how to help students retain more of what they study for a longer period of time. These questions are not usually the ones at the forefront of instructors' planning and teaching. They want students to learn well, but their assessment of student learning stops at the end of their course—at which time, students leave without much thought being given to what concepts or procedures they will remember at a later date. What could teachers do differently in their classes if they were focused on long-term retention? This focus in research would be beneficial to all mathematics instructors.

In summary, we highlight the following conclusions:

- There was no overall statistically significant difference in the performance of the traditional and reform groups of students.
- There was a statistically significant difference in the type of knowledge retained by the two groups: reform students retained better conceptual knowledge and traditional students retained better procedural knowledge.
- Choice of applied calculus curriculum should depend on desired outcomes.
- Students can understand calculus concepts without being able to perform the relevant procedures—in other words, concepts can be taught before computational competence is achieved.
- Calculus students remembered a disappointingly small proportion of what they had supposedly learned.
- Future research would do well to focus on long-term retention of mathematical concepts and skills.

References

Allen, G. H. (1995). A comparison of the effectiveness of the harvard calculus series with the traditionally taught calculus series (Doctoral dissertation, San Jose State University). *Dissertation Abstracts International, 33*(06), 1646.

Anderson, J. R., Greeno, J. G., Reder, L. M., & Simon, H. A. (2000). Perspectives on learning, thinking, and activity. *Educational Researcher, 29*(4), 11-13.

Ball, D. L., & Cohen, D. K. (1996). Reform by the book: What is – or might be – the role of curriculum materials in teacher learning and instructional reform? *Educational Researcher, 25*(9), 6-8, 14.

Bookman, J., & Friedman, C. P. (1998). Student attitudes and calculus reform. *School Science and Mathematics, 98*(3), 117-122.

Brunett, M. R. (1995). A comparison of problem-solving abilities between reform calculus students and traditional calculus students (Doctoral dissertation, American University). *Dissertation Abstracts International, 57*(01A), 0140.

Byrnes, J. P., & Wasik, B. A. (1991). Role of conceptual knowledge in mathematical procedural learning. *Developmental Psychology, 27*(5), 777-786.

Cavanagh, M. (1996). Student understandings in differential calculus. In P. C. Clarkson (Ed.), *Technology in mathematics education* (Proceedings of the 19th annual conference of the Mathematics Education Research Group of Australasia, pp. 107-114). Melbourne: Mathematics Education Research Group of Australasia.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159.

Ganter, S. L. (1997). Impact of calculus reform on student learning and attitudes. *AWIS Magazine, 26*(6), 10-15.

Garner, L. E., & Hughes-Hallett, D. (1996). *Instructor's manual, applied calculus for business, social sciences, and life sciences*. New York: John Wiley & Sons.

Goldstein, L. J., Lay, D. C., & Schneider, D. I. (1996). *Calculus and its applications*. Upper Saddle River, NJ: Prentice Hall.

Hadfield, K. F. (1996). *Conceptual and procedural understanding in students of Harvard and traditional calculus*. Unpublished master's thesis, Brigham Young University, Provo, UT, USA.

Heid, M. K. (1988). Resequencing skills and concepts in applied calculus using the computer as a tool. *Journal for Research in Mathematics Education, 19*(1), 3-25.

Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 65-97). New York: Macmillan.

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1-27). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hughes-Hallett, D., Gleason, A. M., Lock, P. F., Flath, D., Gordon, S. P., Lomen, D. O., et al. (1996). *Applied calculus, preliminary edition*. New York: John Wiley & Sons.

Judson, P. T. (1988). Effects of modified sequencing of skills and applications in introductory calculus (Doctoral dissertation, University of Texas at Austin). *Dissertation Abstracts International, 49*(06A), 1397.

Kenelly, J. (1996). Technology in mathematics instruction. In A. W. Roberts (Ed.), *Calculus, the dynamics of change* (MAA Notes Number 39, pp. 24-27). Washington, DC: Mathematical Association of America.

Moursund, D. (1985). Will mathematics education rise to the challenge of computers? *Mathematics Teacher, 78*(9), 660-661.

Narasimhan, C. C. (1993). Calculus reform for the non-science client disciplines. *Primus, 3*(3), 254-262.

Park, K., & Travers, K. J. (1996). A comparative study of a computer-based and a standard college first-year calculus course. In J. Kaput, A. Schoenfeld & E. Dubinsky (Eds.), *Research in collegiate mathematics education, II* (pp. 155-176). Providence, RI: American Mathematical Society.

Ratay, G. M. (1993). Student performance with calculus reform at the united states merchant marine academy. *Primus, 3*(1), 107-111.

Schoenfeld, A. H. (Ed.). (1997). *Student assessment in calculus* (MAA Notes Number 43). Washington, DC: Mathematical Association of America.

Selden, J., Selden, A., & Mason, A. (1994). Even good calculus students can't solve nonroutine problems. In J. Kaput & E. Dubinsky (Eds.), *Research issues in undergraduate mathematics learning, preliminary analyses and results* (MAA Notes Number 33, pp. 19-26). Washington, DC: Mathematical Association of America.

Thomas, M. O. J., & Hong, Y. Y. (1996). The Riemann integral in calculus: Students' processes and concepts. In P. C. Clarkson (Ed.), *Technology in mathematics education* (Proceedings of the 19th annual conference of the Mathematics Education Research Group of Australasia, pp. 572-579). Melbourne: Mathematics Education Research Group of Australasia.

White, P. (1995). An introduction to teaching calculus. In L. Grimison & J. Pegg (Eds.), *Teaching secondary school mathematics: Theory into practice* (pp. 165-185). Marrickville: Harcourt, Brace.

White, P. & Pegg, J. (1996). Advanced Mathematical Thinking. In B. Atweh, K. Owens & P. Sullivan (Eds.), *Research in Mathematics Education in Australasia 1992-1995* (pp. 151-168). Sydney: Mathematics Education Research Group of Australasia.

Wilson, R. (1997). A decade of teaching 'reform calculus' has been a disaster, critics charge. *The Chronicle of Higher Education, 43*(22), A12-A13

Authors

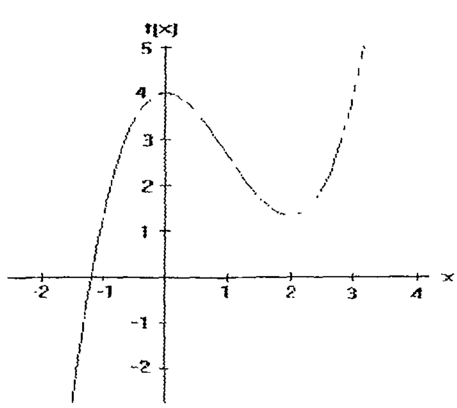
Bradley E. Garner, Boise State University, 1910 University Drive, Boise, ID 83725-1555, USA. Email: <garner@math.boisestate.edu>.

Lynn E. Garner, 292 TMCB, Brigham Young University, Provo, UT 84602, USA. Email: <lynng@math.byu.edu>.

Appendix

Below are the items included on the written test that was administered to all students in the study. The extra white space which was present on the actual test has been removed.

1. Given below is the graph of a function.
   a On what intervals is *f(x)* increasing? Decreasing?
   b. Estimate any points at which a local maximum or a local minimum occurs.



2. Consider the function:

$f(x) = \frac{2}{3}x^3 - 2x^2 + 4$

a. Using calculus, determine whether the above function is increasing or decreasing at $x = 1$.

b. Also using calculus, determine any values at which a local maximum or a local minimum occurs.

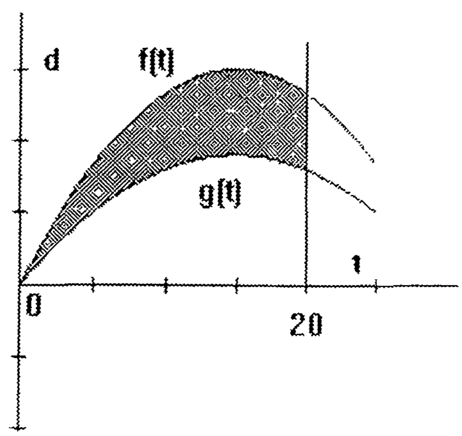
3. Find the first derivatives of the following functions:
   a. $y = (3x^2 + 1)^3$
   b. $y = xe^{2x}$

4. A candy company knows that the revenue, R, from sales of a certain product is a function of the selling price, p, and is given by $R = -500p^2 + 4500p$. What is the maximum revenue? (R and p are in dollars).

5. A publishing company knows that the cost (in dollars) of printing a book depends on the number of copies printed. The cost of printing N copies is $C = f(N)$. Suppose that when $N = 500$, the derivative of the function is 2 (that is, $f'(500) = 2$). What conclusions can you draw?

6. A compact car in motion on the freeway uses f(x) gallons of gas in driving x miles.
   a. For what values of x, if any, is $f'(x)$ positive? For what values of x, if any, is $f'(x)$ negative?
   b. A motorhome uses $g(x)$ gallons of gas in driving x miles. How would you expect the graphs of $f(x)$ and $g(x)$ to compare?

7. Suppose that on a typical day the rate at which the electric company accumulates revenue (in dollars per minute) is given for any time in the day by the function $f(t)$. The rate at which the company's operating costs accumulate (in dollars per minute) is given for any time in the day by the function $g(t)$. These functions are graphed below.

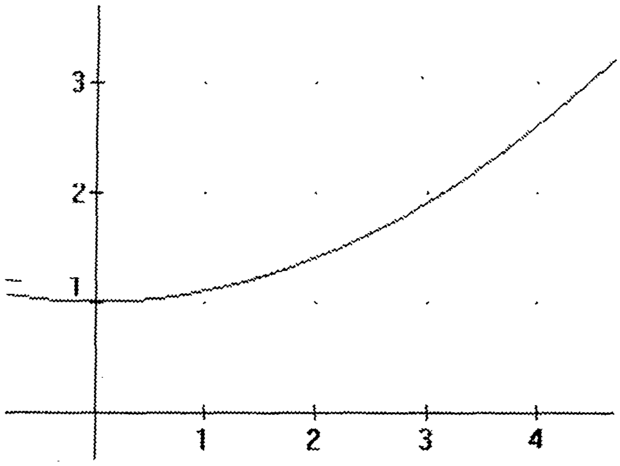


a. How could you use calculus to find the company's revenue and cost functions?

b How could you use calculus to compute the shaded area?

c. What does the shaded area represent economically?

8. Given the following graph and table of values for a function, $f(x)$,



x	0	1	2	3
y	1.0	1.1	1.4	1.9

a. approximate the derivative of the function at $x = 2$.
b. approximate the integral of the function from $x = 1$ to $x = 3$.

9. Compute the following definite integrals: a. $\int_1^2 4x^3\,dx$ b. $\int_0^{1/2} e^{2x}\,dx$

10. There is a theorem, called the Fundamental Theorem of Calculus, which is a statement about the relationship between the derivative and the definite integral. What is this theorem?