2006-1902: USING RAPID FEEDBACK TO ENHANCE STUDENT LEARNING

John Chen, Rowan University

John Chen is an Associate Professor of Mechanical Engineering. He has been a faculty member since 1994, when he began his career as an Assistant Professor in the Department of Mechanical Engineering at North Carolina A&T State University. He joined Rowan University in his current position in 1998. He is an active member of ASEE and is currently the Chair of the Mechanical Engineering Division.

Jennifer Kadlowec, Rowan University

Jennifer Kadlowec is an Associate Professor in Mechanical Engineering at Rowan University. She began as an Assistant Professor in 1999 after she received her M.S. and Ph.D. in Mechanical Engineering from the University of Michigan and a B.S. in Physics from Baldwin-Wallace College. She has been a member of ASEE since 1998 and regularly presents at the ASEE annual conference.

Dexter Whittinghill, Rowan University

Dexter Whittinghill is an Associate Professor in the Mathematics Department at Rowan University. He is in his ninth year at Rowan, and has been a professor since 1984 when he received a Ph.D. in Statistics from Purdue University. His research interests have migrated from the design of experiments to statistical education, and for many years he has enjoyed consulting with fellow faculty. He has held office in statistical education groups within the statistics and mathematics community.

© American Society for Engineering Education, 2006

Using Rapid Feedback to Enhance Student Learning

Abstract

In this project our goal is to improve student learning in foundation engineering courses. Our hypothesis is that learning is improved by providing rapid feedback to students of their understanding of key concepts and skills being taught. This hypothesis was tested through experiments in which student performance on quizzes was measured after classes in which they were provided rapid feedback or not. The feedback system acts as a catalyst to encourage students, working in pairs, to assist each other in correcting misconceptions or deepening each other's understanding of the concept or skill at hand. The feedback is enabled through wireless-networked handheld computers (PDAs).

In each of the past two years, the feedback system was implemented in two sections of Statics. One author taught both sections of the course in order to minimize any differences in teaching style and in the content or pace of coverage. A crossover design of experiment was used. In such experiments, each student acted as his or her own control to eliminate the non-correctible confounders. Student performance on a quiz at the end of each treatment period provided the data for comparison between the two groups. A general linear statistical model was used to analyze the treatment factor while controlling for the other 'nuisance' or confounding factors.

Our findings from Fall 2004, when both sections were provided rapid feedback and the PDA-enabled feedback system was compared with using flashcards for feedback, showed that there was no significant difference between student performance. In Fall 2005, we compared the PDA-enabled feedback system with having no feedback. We found a significant and positive effect when students received feedback. This is a noteworthy finding in that it confirms the value of providing rapid feedback and the usefulness of the currently popular 'clickers' that many professors are employing to promote classroom interaction.

Introduction

Core engineering courses, such as Statics, are comprised of key concepts and skills that students need to master in order to succeed in follow-on courses. Students must comprehend these concepts at sufficient depth (as opposed to rote memorization of procedure) and transfer this understanding to other courses and contexts. In this multiyear project, our hypothesis is that such learning is facilitated in an active, peer-assisted environment in which the students are provided frequent and rapid feedback of their state of learning.

Background and Motivation

Providing feedback to students of their current level of understanding of concepts is critical for effective learning. It is also important for the professor. This feedback is typically realized through homework sets, quizzes and tests. All of these techniques, however, suffer the faults of being too slow, too late, and too tedious to apply frequently. Freeman and McKenzie[1] discuss several issues that inhibit better student learning in higher education. For students, there is a lack of individual feedback on learning; few opportunities for dialogue to improve learning; and a feeling that the subject is impersonal. From the faculty members' perspective, the difficulties lie

in knowing what students are really learning, providing individualized feedback, addressing students' specific misconceptions, attending to diverse learning styles, and engaging students in learning.

Bransford et al.[2] state: "Learners are most successful if they are mindful of themselves as learners and thinkers. In order for learners to gain insight into their learning and their understanding, frequent feedback is critical: Students need to monitor their learning and actively evaluate their strategies and their current levels of understanding." Freeman and McKenzie support this idea, noting that "Feedback is fundamental to learning... Students may receive grades on tests and essays, but these are summative assessments… What are needed are formative assessments, which provide students with opportunities to revise and improve the quality of their thinking and understanding. If the goal is to enhance understanding and applicability of knowledge, it is not sufficient to provide assessments that focus primarily on memory for facts and formulas."

Our project addresses these issues by providing students with timely feedback and opportunities to improve learning. Our goal is to combine rapid feedback with conceptual learning and skills development and to evaluate our methods through rigorous experimental design and data analysis.

Project Design and Implementation

Course Description

At Rowan University, Statics is a required course for sophomores in three of the four engineering disciplines (Civil & Environmental, Electrical & Computer, and Mechanical Engineering). The course content is similar to that of most engineering programs in the US, although the pace and length of the course is unusual. Rowan students take Statics in a compressed, half-semester (7.5 weeks) format, with classes meeting for three 75-minute periods each week. Students receive two semester-hour credits upon passing the course. The format dictates a faster-than-usual pace of coverage of the material with little time spent in reviewing course material from previous lectures. Statics is delivered in the first half of the Fall semester, followed in the second half-semester by Dynamics. In the first half of the Spring semester, Civil & Environmental and Mechanical Engineering students continue in the engineering mechanics sequence by taking Solid Mechanics (also known as Mechanics of Materials).

In Fall 2003, we began this study with one of the authors teaching two sections of this course. We collected some data to practice for what we might expect in the following years and focused on the details of implementing this project. Essentially, we treated the year as a 'trial run.' For example, we acquired all the personal digital assistants (PDAs) that were to be used for this study, set up, tested and practiced with the software used to collect data, and developed most of the quizzes for which rapid feedback would be provided to students. In Fall 2004 and Fall 2005, we repeated what was implemented in 2003 except that data was taken for subsequent analysis. What differed in the two latter years was the control group used to compare with the treatment group (the group that received rapid feedback with the PDAs), as will be explained later. One of the authors always taught two sections of Statics as part of this study. This was done in order to minimize any differences in teaching style or content between the two sections. Having a single

professor also ensured that the two sections maintained the same pace through the course from day to day.

The in-class portion of this study is conducted in a similar manner to that described by Mazur.[3] The professor presents a new topic or concept for no more than 10-15 minutes, using traditional lecture, demonstration, or sample problem solution. Thereafter, he poses a 'concept question' or a 'skill quiz' to gauge the students' understanding. If the student responses from the feedback method (if feedback was used) show that a high percentage of students do not understand the concept or have not mastered the skill, the professor elaborates on or further explains the topic. If the responses show that a reasonable fraction of students understands (a distribution of answers, but a plurality with the correct answer), the professor directs the students to take time and explain the concept or skill to each other. Thereafter, the students are asked to either respond again to the same question, or a different question on the same topic may be posed. The final scenario occurs when the student response shows a high percentage of correct answer, indicating that students understand the topic. In this case, the professor simply continues to the next topic.

Traditional assessment methods were used to determine a student's course grade. In addition to assigned homework sets, which were completed by students in two-person teams, quizzes and tests were given. In the 7.5-week period of the course, nine homework sets were assigned, and eight quizzes and two non-cumulative examinations were given. Identical homework sets were assigned to the two sections. Whenever a homework set was submitted by the students, a brief quiz based on a concept covered in the homework was given. Quizzes were designed to be similar, but not identical, between the two sections. The scores on the quizzes were analyzed, as described later, to assess for any treatment effect due to the feedback provided.

A crossover design of experiment is used in this study.[4] The method is intended to eliminate potential confounding factors that cannot be controlled for using a standard analysis of variance model. For example, students may not be randomly assigned to each of the two Statics sections (for example, one section may have mostly electrical engineering students, who have a different motivation level than the other section, which might be populated mainly with mechanical engineering students), or the time at which each section is held may affect student performance. Without the crossover a potential treatment effect would have been indistinguishable from a section effect.

In a crossover design, one of two study groups (course sections in this case) will be randomly chosen to receive instruction with the PDA-enabled feedback system (the 'treatment' group) while the other group will act as the 'control' for a fixed period of time (or 'treatment period'). For the next treatment period, the two sections simply swap the roles of treatment and control, and this continues for the duration of the course. In this manner, each student acts as his or her own control to eliminate the non-correctible confounders. This design has the additional advantages of eliminating any bias that may be introduced by the professor in course delivery in the two sections, and minimizing any attitude bias that may be displayed by students of either section due to receiving a single method of feedback for the entire course if swapping did not occur. The treatment periods generally lasted from two to five class meetings, as was determined to be logical based on the skills or concepts being covered during the period.

In Fall 2004 the control group used a flashcard system, similar to that described by Mehta,[5] to provide rapid feedback. In Fall 2005 the control group used no feedback as a comparison with the treatment group. Although the students could not respond to the concept question or skill quiz using a rapid feedback method, the problem was still presented to the control group and the instructor used active-learning methods in these sessions. The students were instructed to work cooperatively on each problem and were encouraged to provide answers, which were recorded on the whiteboard for the class. We emphasize that regardless of the feedback method or its absence, the instructor used identical teaching methods in both sections of the course, which included various active learning techniques.

Rapid Feedback Methods

The flashcard method for providing feedback to students was developed by Mehta.[5] In short, double-sided and color-coded cards are used by students to display their answer to a question posed by the professor. Each card can display one of six possible responses. The cards provide a quick means for the professor to scan the class's response and qualitatively determine the distribution of answers. The students are also able to see the class's response by a quick visual scan. Because of this, the professor asks the students to respond simultaneously to the posed question so that students who are uncertain of their answer cannot vote with the majority response.

A fleet of PDAs is used for the PDA-enabled feedback method. In 2004, half of the PDAs were Palm OS-based and half were Pocket PC-based. In 2005, all of the PDAs were upgraded to Pocket PC-based. All of the PDAs have wireless networking capabilities (802.11b or WiFi) and communicate with the professor's Windows XP Tablet PC using a peer-to-peer networking mode. The software that is used to manage the intercomputer communications and to record and display student responses from the PDAs is a pre-beta version of OptionFinder VP, which is being developed by Option Technologies Interactive.

Regardless of the feedback method used each time, the concept question or skill quiz is posed by the professor through his Tablet PC and is projected to the front of the class along with the possible solutions. The correct solution is embedded with incorrect answers, also known as 'distractors,' which are derived from common student mistakes or misconception. Students are given time to reflect on the question posed, discuss it with their peers, and then must select from the possible solutions. The major differences between the two feedback methods are that the PDA/software-based method allows for 1) quantitative and permanent recording of the student responses for future review and 2) a display of the tallied student responses, which is projected up on the screen nearly instantaneously after the students respond. As mentioned previously, in Fall 2005, when the control group received no rapid feedback through PDAs or flashcards, the concept question or skill quiz was still presented to the students, and the professor used active learning techniques to encourage students to solve the problem and to share the solution with the class.

Data Analysis

This project is comprised of three major components: the development of a suite of concept questions and skills quizzes for the course, the use of rapid feedback and peer-assisted learning

in the classroom, and a comparison between the two methods of providing rapid feedback (or not) to students. The third component required the bulk of the statistical analysis. The goal of this analysis was to see if the method of implementing the rapid-feedback – using PDAs or flashcards or nothing – had an effect on the students' learning. The response variable tested is the score on a quiz for the corresponding period of instruction where one section had the treatment and the other the control. This would be done while controlling for factors (or variables) other than the treatment factor which might affect the scores.

To analyze the treatment factor while controlling for the other 'nuisance' factors that could affect scores but are not attributable to the treatment, we employed the following general linear model using the DataDesk statistical software package:

$$y_{ijklm} = \mu + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \alpha_i + \gamma(\alpha)_{j(i)} + \delta_k + \tau_l + \varepsilon_{ijklm} \quad (1)$$

where

y = the score on the quiz,

μ = the grand mean (average score with no factors taken in to account),

x_1 = the student's Calculus I grade,

x_2 = the student's Calculus II grade,

x_3 = the student's Physics I grade,

α = the Section in which the student is enrolled,

γ = the student nested in section, or Student-in-section,

δ = the Period (or quiz),

τ = the Treatment (PDA = 'treatment' and flashcard/no feedback = 'control'),

ε = random error.

Students' Calculus I, Calculus II and Physics I grades were treated as continuous covariates in the analysis. The Section factor was discrete, and the Student factor was discrete, and nested in section (student 1 in Section 01 is not the same as student 1 in Section 02). The Period (or quiz) factor was discrete and included because some quiz topics may be intrinsically more difficult than others. The Treatment factor was discrete as well. Although the quiz scores in both years were skewed towards zero (i.e., they were bunched toward the higher scores), the residuals were nearly normal, and no transformation of the data was needed.

Results and Discussion

We have previously described in detail the results from Fall 2004,[6] so only a summary will be provided here. During that semester, we conducted a crossover experiment in which two sections of students were provided rapid feedback and their performance on a series of quizzes was compared. The two rapid feedback methods were using the PDAs and using the flashcards, as described earlier. The most important finding was that there was no statistically significant difference between the student performance between these two groups. In other words, it does

not matter how one provides rapid feedback, so long as it is provided. Although we had thought that the 'coolness' of the PDA might affect a student's learning, it really would only affect their interest during the physical activity in class of reporting their answers. In the end their scores were not influenced by whichever of the two feedback methods used.

Student survey results from 2004 indicate that students overwhelmingly felt that having rapid feedback of their state of learning was helpful to them, regardless of the means of providing feedback. Specifically, a great majority of students felt that either method of feedback was at least 'somewhat helpful' to their learning, with a significant preference for the PDAs over the flashcards. Hence, although the use of PDAs versus flashcards did not affect the actual learning (as measured by the analyses of the quiz scores), the use of PDAs was *perceived* by students to be more helpful to their learning than the flashcards. Finally, 65% of the students believed that they would have performed worse in a course in which rapid feedback was not provided, while the remainder believed they would have performed at the same level.

The rapid feedback also had impacts on the authors as instructors. Regardless of the feedback method, we had to be more organized for each class and to plan ahead in preparing skill and concept questions and placing them appropriately in the lecture period. We also found that posing the feedback question was useful to get students to refocus or review, even if a question was created "on the spot" during class. We observed that the students took the feedback quizzes quite seriously and tried hard to answer them correctly even though no grade was involved. This was an additional benefit in that the students were forced to think about the concepts now rather than later (or perhaps much later) when they sat down to do homework. The results of the rapid feedback questions also allowed us to note what concepts were most difficult for students and thus improve future instruction. While technical difficulties and set-up time may be cumbersome for the PDA/software system, especially for the novice user, the authors believe that the benefits for the students and the faculty far outweigh these negatives.

Finally, based on the Fall 2004 implementation, we found evidence that rapid feedback use improves knowledge retention (durability) and knowledge application in a different environment (transferability) in subsequent courses of Dynamics and Solid Mechanics as shown by preliminary results in these courses.[6] We are currently collecting similar data for the Fall 2005 implementation and will report on this in the future.

Table 1 below presents results of our data analyses for the Fall 2005 cohort of Statics students. Recall that this cohort was subjected to a crossover comparison between having rapid feedback with the PDAs versus having no feedback. This comparison would allow us to determine the effect of having feedback or not on student learning, as a complement to the Fall 2004 comparative study. Each row within the table represents a different statistical model used to analyze the data (scores from eight quizzes administered at the end of each treatment period). The most noteworthy finding is that for all models examined *the treatment (of having rapid feedback) was statistically significant*. That is, student scores on the quizzes were positively influenced by the provision of rapid feedback. We will now elaborate more on the details of the statistical analyses.

Some general observations can be made for all models that we examined. First, none of the covariates included (students' grade in Calculus I, Calculus II and Physics I) were significant.

This is in contrast to our findings from Fall 2004, when we found that the quiz scores were dependent on the students' performance in Calculus II and Physics I.[6] We hypothesized then that the students' grade in Calculus II was a reflection of their general abilities in mathematics, rather than specific skills learned in that course, and the fact that Physics I was significant was because most of the concepts in Statics are derived directly from application of physics concepts. The fact that we find no significance in these covariates in Fall 2005 is somewhat puzzling, but not worrisome to us since the more important finding – that the treatment was significant – is not dependent on this.

Returning to Fall 2005, the second general finding is that the Section in which the student belonged was not significant. This simply states that there was no difference in performance between the two sections of the course despite the fact that they were offered on different days and times. Third, the Student factor was always significant (at α=0.05), which is not surprising since each student is expected to perform differently and somewhat consistently. Finally, we found that the treatment Period was highly significant (at α=0.001), which actually implies that the quizzes were inherently different in their degree of difficulty. Again, this is not surprising in that some topics in Statics are easier than others, and this finding simply reflects that fact.

Table 1: Results of statistical analyses in various models of the data.

Response	Covariates	← Factors →				Interactions (if any)
Score	1, 2, 3	Section	Student-in-Section*	Period***	Treatment* (p=.0318)	
Score	1, 2, 3	Section	Student-in-Section*	Period***	Treatment** (p=.0062)	<u>Section/Period</u>
Score	1, 2, 3	Section	Student-in-Section*	Period***	Treatment* (p=.0144)	Student/Treatment
Score	1, 2, 3	Section	Student-in-Section*	Period***	Treatment* (p=.0204)	<u>Period/Treatment</u>
Score	1, 2, 3	Section	Student-in-Section*	Period***	Treatment** (p=.0033)	Sec./Per., Stu./Treat.

For each model the factors marked with '*' were significant at α = 0.05 (5%), with '**' at α = 0.01 (1%), and with '***' at α = 0.001 (0.1%). Underlined factors were 'significant' at α = 0.10 (or 10%). Note that covariate 1 = Calculus I, 2 = Calculus II, and 3 = Physics I.

The first row of results in Table 1 shows the basic model, which does not examine any interactions between the factors. The treatment effect was significant at p=.0318. Subsequent models examined the two-way interactions between Section by Period, Student by Treatment, Period by Treatment, and finally a two-way interaction between Section by Period and Student by Treatment. In none of these models were any two-way interactions significant at α=0.05. The important result from these models is that the Treatment effect was at least as significant as in the model without them (Eq. (1)). This confirms the statistical significance of our finding that the rapid feedback positively influenced the students' performance on the quizzes.

We are currently at work completing the analysis of the student survey results from Fall 2005. Preliminary results show that students were positively receptive to the provision of rapid feedback in class and felt that it improved their learning in the course (as we had found in 2004). A high percentage of the students (42.9%) believed that they would have done worse in a course taught by the same instructor without having rapid feedback, while 52.3% felt they would have performed at the same level.

We are also analyzing results of student performance in a follow-on course, Solid Mechanics, in the current semester (Spring 2006). For this part of the study, we will be able to compare student performance on in-class concept quizzes (delivered through the PDAs) on four student populations: those who had rapid feedback when learning this concept, those who did not but had the same instructor and at the same time as the previous group, those who learned the concept from a different instructor but during the same semester, and those who learned the concept at a different time or place. We will describe these results at the conference and in future publications.

Conclusions

Our most noteworthy finding to date is that rapid feedback has a significant and positive effect on student performance. This confirms the value of providing frequent and rapid feedback to students. We theorize that this provides the students with knowledge of their state of learning, allows them to make adjustments in their strategies for learning, and encourages immediate reflection on and practice in the concept or skill at hand. Although we did not utilize the currently popular feedback devices known as 'personal response systems' or 'clickers' that are offered by several commercial vendors, our method of using handheld wireless computers no doubt is analogous to these devices, which are gaining in popularity in higher education. Our findings provide strong evidence for the usefulness of these feedback devices.

Acknowledgements

The authors gratefully acknowledge the support of the National Science Foundation through grants DUE-0243227 and EIA-0312868.

References

1. Freeman, M. and J. McKenzie, "Aligning Peer Assessment with Peer Learning for Large Classes: the Case for an Online Self and Peer Assessment System," in *Per Learning in Higher Education*, D. Boud, R. Cohen, and J. Sampson, eds, Kogan Page Limited, London, 2001, pp.156-169.
2. Bransford, J.D., Brown, A.L., and Cocking, R.R., eds., *How People Learn: Brain, Mind, Experience, and School*, National Academy Press, Washington, DC, 1999.
3. Mazur, E, *Peer Instruction: A User's Manual*, Prentice Hall, Upper Saddle River, NJ, 1997.
4. Mason, R.L., Gunst, R.F. and Hess, J.L. *Statistical Design and Analysis of Experiments, with Applications to Engineering and Science*, Wiley Series in Probability and Mathematical Statistics, New York, 1989.
5. Mehta, S.I., "A Method for Instant Assessment and Active Learning," *J. of Engr. Educ.*, 84:295, 1995.
6. Chen, J., Kadlowec, J., and Whittinghill, D., "Using Technology for Concepts Learning and Rapid Feedback in Statics," *Proceedings of the 2005 ASEE Annual Conference*, Session 1668, Portland, OR, USA, 2005. See also, Kadlowec, J., Chen, J., and Whittinghill, D.C., "Using Rapid Feedback to Enhance Student Learning in Mechanics," *Proceedings of the 2005 Frontiers in Education Conference*, Indianapolis, IN, USA.