Concepts Learning Using Technology for Rapid Feedback and Student Engagement

John C. Chen,[*] Jennifer A. Kadlowec,[*] and Dexter C. Whittinghill[†]
Departments of [*]Mechanical Engineering and [†]Mathematics
Rowan University, Glassboro, NJ

Abstract

In this project our goal is to improve student learning in the foundation mechanics course Statics. In this case improved learning is defined as knowledge retention (durability) and knowledge application in a different environment (transferability). We aim to do this by providing rapid feedback to students of their understanding of key concepts and skills being presented. The feedback system acts as the focal point and catalyst to encourage students to assist each other in correcting misconceptions or deepening each other's understanding of the topic or skill at hand. Furthermore, the system allows the professor to assess the students' level of comprehension (or misconception) in a just-in-time fashion, and thus guiding his or her pacing and coverage of the material. The rapid feedback is enabled through wireless-networked handheld personal digital assistants (PDAs) or flashcards. In the first two years of the study, we have implemented the system in two sections of Statics using a crossover design of experiment, where one section receives the rapid feedback 'treatment' (i.e., use of the PDAs) while the other (the 'control' group) receives rapid feedback on the exact same topics, but only through the use of flashcards. After a predetermined period, the sections swap their feedback treatment. Several swaps are achieved during the course, and in this manner each student acts as his or her own experimental control to assess the effectiveness of the treatment. This paper focuses on our experimental methods, the statistical analysis of data, and results of student learning and student satisfaction from the first implementation.

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

Copyright ©2005, American Society for Engineering Education

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[1] 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 demonstrate improved student learning 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. In that year, we collected some data to practice for what we might expect in the following year and focused on the details of implementing this project. Essentially, we treated the year as a 'trial run.' For example, we acquired all the 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 the most recent offering of the course in Fall 2004, we repeated what was implemented in the previous year except that data was taken for subsequent analysis. All of the reported results in this paper come from Fall 2004.

As mentioned previously, one of the authors taught two sections of Statics in Fall 2004. 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. At the start of any class, the students in each section are

Copyright ©2005, American Society for Engineering Education

provided with one of two means of receiving rapid feedback: a PDA or a flashcard. With the PDAs, students are paired up and share a single PDA, whereas with the flashcards, each student in the section is provided one. Details about the feedback methods are described later.

The in-class portion of this study is conducted in a similar manner to that described by Mazur.[4] 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 system (PDAs or flashcards) 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.

In addition to assigned homework sets, which were completed by students in two-person teams, quizzes and tests were used for student evaluation and data analyses for this study. 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 was given which covered some concept covered in the homework. Quizzes were designed to be similar, but not identical, between the two sections. The scores on the quizzes and examination questions were analyzed, as described later, to assess for differences between the two feedback methods.

A crossover design[3] of experiment is used in this study. 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 system (the 'treatment' group) while the other group will use the flashcard system (the 'control') for a fixed period of time. For the next 'treatment period,' the two sections simply swap the feedback method, 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.

Copyright ©2005, American Society for Engineering Education

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.

A fleet of 18 PDAs is used for the PDA-enabled feedback method. Half of the PDAs are Palm-OS-based and half are PocketPC-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 we use 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 (www.optiontechnologies.com).

Regardless of the feedback method, 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, which are derived from common student mistakes or misunderstanding. 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.

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, for the current year of study, a comparison between the two methods of providing rapid feedback 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 (the 'treatment') vs. flashcards (the 'control'), had an effect on the students' learning. The response variable tested is the score on a quiz or an examination question for the corresponding period of instruction where one section had the treatment and the other the control (or vice versa). 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 (PDA vs. flashcard) 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 package:

$$y_{mijkl} = \mu + \beta_1 x_{1,m} + \beta_2 x_{2,m} + \beta_3 x_{3,m} + \beta_4 x_{4,m} + \alpha_i + \gamma(\alpha)_{j(i)} + \delta_k + \tau_l + \varepsilon_m, \quad \text{where}$$

y = a transformed score on the quiz,

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

x_1 = the student's Freshman-year GPA (0.00 to 4.00, which includes x_2, x_3 and x_4.)

Copyright ©2005, American Society for Engineering Education

x_2 = the student's Calculus I grade (0 to 40),
x_3 = the student's Calculus II grade (0 to 40),
x_4 = the student's Physics I grade (0 to 40),
α = the Class (Section 01 (8 am class meeting) or Section 02 (10:50 am meeting)),
γ = the student nested in class, or Student-in-class,
δ = the Quiz (or period or topic),
τ = the Treatment (PDA = 'treatment' and flashcard = 'control'),
ε = random error.

The quiz scores were transformed because they were skewed (y = the 'score' squared). The Freshman-year GPA and the Calculus I, Calculus II and Physics I grades were treated as continuous covariates. The Class factor was discrete, and the Student factor was discrete, and nested in class (student 1 in Section 01 is not the same as student 1 Section 02). The Quiz factor was discrete and included because some quiz topics may be intrinsically more difficult than others. The Treatment factor was discrete as well.

Results

For various reasons, only five of the eight quizzes were included in the analysis (questions from the examination showed no difference between the classes and/or treatment). When the model above was analyzed, Calculus II, Physics I and Student-in-class factors were significant at α = .05. We will address the terms in their order in the model. Recall that factors other than the Treatment are in the model to account for likely sources of variability in the quiz scores. That way, any variability due to the Treatment is not masked by the other factors and we can detect the Treatment effect.

Because Freshman GPA is based on the grades for Calculus I, Calculus II and Physics I, it is not surprising that with these included in the model, and the latter two significant, Freshman GPA is not significant. It was not surprising that Calculus II and Physics I were significant (p-values = .0275 and .0018, respectively), because each course contains skills and concepts important to Statics. That Calculus I was not significant might be because the most important calculus techniques used in Statics come from Calculus II, though we cannot be certain of this reasoning. Although Class was not significant (p-value = .0752), which reinforced preliminary results from Fall 2003, that it was only marginally so justifies our having it in the model. Student-in-class was significant (p-value = .0009), which should be expected, as scores should always depend on the individual student. That Quiz was not significant may or may not be surprising. The fact that the scores for different quizzes were essentially the same indicates that the quizzes inherently adjusted for the difficulty of the material, or that the periods of instruction were constructed so that no period or topic was inherently more difficult (the authors wish we could claim responsibility for this). Finally, and perhaps most importantly, the Treatment (PDA or flashcard use) was not significant (p-value = .0947). This result suggests that using PDAs or the flashcards to provide feedback to the students had little effect on their score. 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 would be affected by outside work (such as studying!) and inherent interest or

Copyright ©2005, American Society for Engineering Education

motivation in the material, neither of which would be greatly influenced by the fact that a PDA was used in class.

Although we had a large number of observations, the nature of the crossover design and the fact that the students were nested in the classes meant that there were only four two-way interactions that could be added to the model, and these were generally one-at-a-time. When the Section-by-Quiz interaction was added, it was not significant, but the Treatment was significant at $\alpha = .05$ (p-value = .0394). All of the other factors were significant or not, as before. It is not obvious to us why this is the case, except that the three factors – Class, Quiz and Treatment – are involved in the crossover, and adding this interaction somehow made the effect of the treatment stand out. Because the Treatment factor was not significant in the first model or in any subsequent models with one or two interactions, we feel that our discussion in the previous paragraph still applies.

When we added the Student-in-class-by-Quiz interaction, nothing was significant. Essentially the 'team' of factors were sharing the work of explaining the quiz scores, with no factor standing out. When we added the Student-in-class-by-Treatment or the Quiz-by-Treatment interactions (one at a time), they were not significant, and only the Calculus II, Physics I and Student-in-class factors were significant at $\alpha = .05$, which is the same result as described earlier. When we added the pair of two-way interactions, Section-by-Period and Student-in-Section-by-Treatment, Calculus II, Physics I, Section, and Student-in-section factors were significant at $\alpha = .05$. That Section became significant ($p = 0.0498$) for the first time was intriguing but not considered important.

A second set of results we obtained was from two surveys administered to the cohort, one approximately halfway through the course, and the other on the final day of the course. Each survey, along with the responses, is shown in Table 1. Note that the responses from both sections of students were combined in Table 1 since chi-square tests of the results showed no differences between the sections.

In general, the survey results show that students have relatively little familiarity with PDAs (based on the mid-course survey), but still an overwhelming majority in either survey found that the PDAs (and the associated rapid feedback method) enhanced their learning experience (73% and 94%, respectively). In both surveys, a majority of students found that rapid feedback with either the flashcards or the PDAs was at least 'somewhat helpful' to their learning, with a stronger preference in both surveys for the PDAs. In fact, in the second survey, our chi-square test of independence found that the PDAs were considered by the students to be significantly more helpful than the flashcards ($p = 0.0089$). We attribute this finding to the immediate availability of the tallied responses that was provided to the students using the computer and software. Finally, in comparing the results between the two surveys, it is obvious that as the course progressed, the students' acceptance of rapid feedback using either method increased. This is seen in the results showing that the percentage of students who found either method to be at least 'somewhat useful' increased from 59% to 80% for the flashcards and from 73% to 94% for the PDAs.

Copyright ©2005, American Society for Engineering Education

Mid-Course Survey

Question	Response	Response Count	Percentage
Prior to this class, have you used a PDA?	Yes, I have one.	3	8.8
	Yes, but it was someone else's.	10	29.4
	No.	21	61.8
Rate your familiarity with PDAs	No experience.	17	50.0
	Beginner.	10	29.4
	Somewhat familiar.	5	14.7
	Expert.	2	5.9
How useful were the flashcards for your learning?	Very helpful.	5	14.7
	Somewhat helpful.	15	44.1
	No difference.	9	26.5
	Somewhat hindered.	4	11.8
	Very hindered.	1	2.9
How useful were the PDAs for your learning?	Very helpful.	11	32.4
	Somewhat helpful.	14	41.2
	No difference.	5	14.7
	Somewhat hindered.	3	8.8
	Very hindered.	1	2.9

Final Survey

Question	Response	Response Count	Percentage
How helpful was rapid feedback (either method) to your learning?	Very helpful.	16	45.7
	Somewhat helpful.	19	54.3
	No difference.	0	0.0
	Somewhat hindered.	0	0.0
	Very hindered.	0	0.0
How helpful was using the FLASHCARDS to your learning?	Very helpful.	5	14.3
	Somewhat helpful.	23	65.7
	No difference.	5	14.3
	Somewhat hindered.	2	5.7
	Very hindered.	0	0.0
How helpful was using the PDAs to your learning?	Very helpful.	16	45.7
	Somewhat helpful.	17	48.6
	No difference.	1	2.9
	Somewhat hindered.	1	2.9
	Very hindered.	0	0.0
Do you think you would have done better or worse if this course were taught by the same professor, but in a more traditional method of teaching?	Much better.	0	0.0
	A little better.	0	0.0
	No difference.	12	35.3
	A little worse.	21	61.8
	Much worse.	1	2.9

Table 1: Results from the mid-course and final surveys administered to the Fall 2004 cohort.

Copyright ©2005, American Society for Engineering Education

Conclusions

Based on our results from the second year of our study, we can conclude that student scores in a Statics course were significantly associated with their prior performance in Calculus II and Physics I (both from the second semester of the freshman year). Most importantly, we found no difference between the scores when the students were provided with rapid feedback facilitated by the use of either flashcards or PDAs and software, which we found to be mildly surprising. Next year, our project will continue with a crossover design between two sections in which the treatment (using the PDAs) will be contrasted with the control (using no feedback method), while maintaining the remainder of this past year's project design.

The survey results indicate that students unanimously felt that having rapid feedback of their state of learning was somewhat or very helpful to them. Furthermore, a 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. Although our analysis found the use of PDAs versus flashcards did not affect the students' learning (as measured by the quiz scores and exam questions), the PDAs were *perceived* by students as being more helpful. 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.

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 *Peer 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. 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.
4. Mazur, E. *Peer Instruction: A User's Manual*, Prentice Hall, Upper Saddle River, NJ, 1997.
5. Mehta, S.I., "A Method for Instant Assessment and Active Learning," *J. of Engr. Educ.*, 84:295 (1995).

Biographic Information

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.

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. Currently she is the Program Chair of the Mechanics Division.

DEXTER C. 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.

Copyright ©2005, American Society for Engineering Education