

---

# Classes That Click: Fast, Rich Feedback to Enhance Student Learning and Satisfaction

---

JOHN C. CHEN  
*California Polytechnic State University*

DEXTER C. WHITTINGHILL AND JENNIFER A. KADLOWEC  
*Rowan University*

## BACKGROUND

Our goal is to improve student learning in foundation engineering courses. These courses are prerequisite to many higher-level courses and are comprised of critically needed concepts and skills.

## PURPOSE (HYPOTHESIS)

We hypothesize that learning is improved by providing rapid feedback to students on their understanding of key concepts and skills. Such feedback also provides students with insight into their strategies for learning.

## DESIGN/METHOD

In two consecutive years, we conducted this study in two sections of a lower-level engineering mechanics course, Statics. One author taught both sections and a crossover design of experiment was used. In a crossover design, one section was randomly chosen to receive feedback with handheld computers (the "treatment" group) while the other received the "control," which was either a feedback system using flashcards (in year 1) or no feedback (year 2). After a certain period, the two sections swapped the

treatment and control. Student performance on a quiz at the end of each treatment period provided the data for comparison using an analysis of variance model with covariates.

## RESULTS

Findings from year 1 showed that there was no significant difference using either rapid-feedback method. In year 2 we found a significant and positive effect when students received feedback.

## CONCLUSIONS

This is a noteworthy finding, albeit within the constraints of the environment in which we conducted the study, that provides more evidence for the value of rapid feedback and the currently popular "clickers" that many professors are employing to promote classroom interaction and student engagement.

## KEYWORDS

clickers, peer discussion, rapid feedback

---

## I. 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 multi-year project, our hypothesis is that such learning is improved in an active, peer-assisted environment in which the students are provided frequent and rapid feedback of their state of learning.

## II. 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 provided through graded 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 (2001) 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 not knowing what students are actually learning; the inability to provide individualized feedback; address students' specific misconceptions; attend to diverse learning styles; and the struggle to engage students in learning.

Bransford, Brown, and Cocking (1999) 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 (2001) 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."

Previous research on feedback shows mixed results. In general, feedback is broadly defined as the provision of the correct answer, or whether an answer is right or wrong. Furthermore, there is no inference about the rapidity of the feedback, so it may be immediate or it may result from the return of graded assignments or examinations. Kluger and DeNisi (1996) note that feedback interventions have highly variable effects from improving performance to hindering it.

In a meta-analysis including 23,663 observations, these authors conclude that, on average, feedback improved performance but that in over one-third of the cases, it decreased performance. The authors theorize that feedback's effectiveness is decreased when it draws attention away from task learning and toward meta-task processes, such as resolving discrepancies between the feedback and self, attention to the self, and depletion of cognitive resources for task learning.

In a more recent review of the research on the use of feedback in education, Sims-Knight and Upchurch (2001) make five claims about feedback based on previous studies. First, the authors claim that informational feedback is effective when there is a clear right or wrong answer, and when the subject is tested immediately following the feedback provision. Second, they claim that when tested for retention and transfer, those who did not receive feedback fared better. Third, feedback can distract students from the learning task. Fourth, feedback is not a major variable in influencing student learning; other variables such as the classroom climate and the professor's organization and preparation have a much higher impact. Finally, the authors contend that teaching students to provide their own feedback is an effective alternative to professor-provided feedback.

Our project provides students with immediate, elaborated feedback and opportunities to improve learning designed to inform and motivate the students. 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.

### III. PROJECT DESIGN AND IMPLEMENTATION

#### A. Course Description and Project Implementation Outline

At Rowan University, Statics is a required course for sophomores in three of the four engineering disciplines (Civil and Environmental, Electrical and Computer, and Mechanical Engineering). The course content is similar to that of most engineering programs in the U.S., 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 and 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 (JCC) 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. In effect, we treated this semester 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 provide feedback; and developed most of the in-class exercises. In Fall 2004 and 2005, we repeated what was implemented in 2003 except that data were 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. The same author (JCC) always taught two sections of Statics for this study; a third section, not involved in this study, was available to the students. Having one instructor for the sections under study was done in order to minimize any differences in teaching style or content between the two sections of a given year. 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 (1997). 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. Concept questions deemphasize numerical calculations and instead focus more on conceptual understanding of a topic (such as drawing appropriate free-body diagrams). A skill quiz, on the other hand, is designed to check that students are able to do computations and apply specific skills (e.g., breaking a vector into components or doing a vector cross product). When the student responses from feedback 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, and another assessment exercise is presented to the students. When the responses show that a reasonable fraction of students understand (a distribution of answers, but at least about 30 percent with the correct answer), the professor directs the students to pair up with another student and explain the concept or skill to each other. (The students were not directed as to who should do the talking, but to merely have a two-way discussion about the problem posed.) Thereafter, the students are asked to either respond again to the same question, or to a different question on the same topic. The final scenario occurs when the student response shows a high percentage of correct answers, indicating that the students understand the topic. In this case, the professor confirms the correct answer and 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 examinations were given. Identical homework sets were assigned to the two sections. When a homework set was submitted by the students, a brief quiz (15–20 min) 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. Thus, these eight quizzes formed the whole of the quantitative data for this study. A different author (JAK) graded the quizzes for both sections of the course to eliminate any experimental bias by the instructor. No effort was made to blind the grading author from the treatment/control group, but this information was not directly or actively communicated either. In addition to this quantitative data, surveys were administered to the students during and at the end of the semester, as described later.

A crossover design of experiment (Mason, Gunst, and Hess, 1989) was 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, or the time at which each section is held may affect student performance. Without the

crossover, what appears to be an effect due to the treatment may in fact have been due to the section in which the student is enrolled.

In our crossover design, one of two study groups (course sections in this case) was randomly chosen to receive instruction with the PDA-enabled feedback system while the other group acted as the control for a fixed period of time (or “treatment period”). For the next treatment period, the two sections were again randomly chosen for the roles of treatment and control, and this continued for the duration of the course. In this manner, each student acted as his or her own control to eliminate the non-correctible confounders. This design had the additional advantages of eliminating any bias that may be introduced by the professor in course delivery in the two sections, and eliminating any attitude bias that might result if students of either section received only the treatment or control for the entire course if swapping did not occur. The treatment periods generally lasted from two to five class meetings, as was determined logically based on the skills or topic being covered during the period. The treatment periods corresponded exactly with a homework set: when students submitted a homework set, the treatment period ended and a quiz was administered to measure students’ understanding of the material.

In Fall 2004 the control group used a flashcard system, similar to that described by Mehta (1995), to provide rapid feedback. In Fall 2005 the control group used no feedback as a comparison with the treatment group. Although the students in this latter group 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 traditional active-learning methods in these sessions (Felder, 1995; Felder and Brent, 2001; Smith et al., 2005). The students were instructed to work collaboratively 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 otherwise used identical teaching methods in both sections of the course, which included various active-learning techniques. In all cases, the students were provided with the correct solutions to the in-class problems and exercises. Table 1 shows the feedback method used for one section in each year of the study.

### B. Rapid Feedback Methods

Mehta (1995) developed the flashcard method for providing feedback to students. In short, students used double-sided and color-coded cards to display their answer to a multiple-choice question posed by the professor. Each card could display one of six possible responses. The cards provided a quick means for the professor to scan the class’s response and qualitatively determine the distribution of answers. The students were also able to see the class’s response by a quick visual scan. Because of this, the professor asked the students to respond simultaneously to the posed question so

that students who were uncertain of their answer could not vote with the majority response.

A fleet of PDAs was 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 models. All of the PDAs had wireless networking capabilities (802.11b or Wi-Fi) and communicated with the professor’s laptop computer using a peer-to-peer networking mode. The PDAs were available to the students only during the classes in which they were used. The software that was used to manage the inter-computer communications and to record and display student responses from the PDAs was a pre-beta version of OptionFinder VP, which was being developed by Option Technologies Interactive in 2004 but is now no longer available.

We developed this custom feedback system using wireless handheld PDAs since the now-popular “clickers,” also known by their various commercial trade names, were only available with infrared receivers at the time we started this project. This restriction was a barrier for our project for various reasons, and we were also able to get the needed software to use with the PDAs for free. Since the operation of our PDA-enabled feedback system was nearly identical to that of the clickers, we believe that our results could be duplicated in a clicker-enabled class.

Regardless of the feedback method used each time, the concept question or skill quiz was posed by the professor through his computer and was projected to the front of the classroom along with the possible solutions. The correct solution was embedded among incorrect answers, also known as “distracters,” which were derived from common student mistakes or misconceptions. Students were given time to reflect on the question posed and then asked to select from the possible solutions. The major differences between the two feedback methods were that the PDA/software-based method allowed for (1) quantitative and permanent recording of the student responses for future review and (2) a display of the tallied student responses that was projected up on the screen nearly instantaneously after the students responded. 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 common collaborative-learning techniques to encourage students to solve the problem and to share the solution with the class.

### C. Data Analysis

The goal of data 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 was the score on a quiz for the corresponding period of instruction where one section had the treatment and the other the control. This

	Treatment Period/Quiz Number							
	1	2	3	4	5	6	7	8
Fall 2004	na	F	F	P	P	P	F	P
Fall 2005	NF	NF	P	P	P	NF	P	NF

NOTE: Table lists feedback method for one of two study sections: the second section used the opposing feedback method for that year. “F” – Feedback with flashcard; “P” – Feedback with PDA; “NF” – No feedback; “na” – not applicable and not included in data set

*Table 1. Feedback method used for one section in each year of study.*

was 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 analysis of variance model with covariates using the Data Desk statistical software (Data Description, Inc., 2009):

$$y_{ijkl} = \mu + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \beta_3 x_{3ij} + \alpha_i + \gamma(\alpha)_{j(i)} + \delta_k + \tau_l + \varepsilon_{ijkl} \quad (1)$$

where

- $y$  = the score on the quiz,
- $\mu$  = the grand mean (average score with no factors taken in to account),
- $\beta$  = the absolute mean change in quiz score ( $y$ ) for each one-unit increase in the covariate
- $x_1$  = the student's Calculus I grade,
- $x_2$  = the student's Calculus II grade,
- $x_3$  = the student's Physics I grade,
- $\alpha$  = the Section in which the student is enrolled,
- $\gamma$  = the Student,
- $\delta$  = the Period (treatment period, or quiz),
- $\tau$  = the Treatment (PDA = "treatment" and flashcard/no feedback = "control"),
- $\varepsilon$  = normally distributed random error.

For the subscripts,  $i = 1, 2$ , for sections;  $j = 1, 2, \dots, n_j$ , for students within a section;  $k = 1, 2, \dots, 8$  for the period; and  $l = 1, 2$ , for the treatment or control.

Note that the model given in Equation (1) is the model we used to analyze data in Fall 2005. In 2004 the model also included a covariate for the student's cumulative GPA (Chen, Kadlowec, and Whittinghill, 2005, 2008). We found that this covariate was never significant and therefore excluded it from the model above. In addition, other covariates in the model, namely the students' grades in Calculus and Physics, are included in this covariate, making it redundant.

The 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. The Period (treatment 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, so no transformation of the data was needed (Mason, Gunst, and Hess, 1989).

#### IV. RESULTS

We have previously described in detail the results from Fall 2004 (Chen, Kadlowec, and Whittinghill, 2005, 2008), so only a summary is provided here. During that semester ( $N = 35$ ), we conducted a crossover experiment in which the two sections of students were provided rapid feedback and their performances on a series of quizzes were compared. The two rapid feedback methods used were the PDAs and the flashcards. The most important finding was that there was no statistically significant difference in student performance between these two groups. In other words, *it*

*did not matter how one provided rapid feedback*, as student performance was not dependent on this. Although we had thought that the "coolness" of the PDAs might affect a student's learning, it really would only affect their interest during the physical activity in class of reporting their answers. After all, both methods of feedback provided an active-learning activity and both introduced an element of competition in to the class lecture, albeit a very low-stakes one.

Student survey results from 2004 indicated 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 (>59 percent in the mid-course survey; 100 percent in the end-of-course survey). Interestingly, the students had a statistically significant stronger preference for the PDAs over the flashcards (93 percent versus 80 percent, respectively). 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 percent 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 latter is in contrast to the earlier result in the same survey whereby 100 percent of the students felt that having feedback was at least "somewhat helpful" to their learning, and it may simply reflect the students' self-confidence that they can perform at the same level regardless of the teaching method.

The rapid feedback also had impacts on the author (JCC) as instructor. Regardless of the feedback method, he 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. He 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. He observed that the students took the feedback questions in class quite seriously and tried hard to answer them correctly even though no grade was involved. This was an additional benefit because the students were forced to think about and apply the concepts now rather than later (or perhaps much later) when they sat down to do homework. Finally, the results of the rapid feedback questions allowed the instructor to note what concepts or skills were difficult for students and thus improve future instruction.

Table 2 presents results of our data analyses for the Fall 2005 cohort of Statics students ( $N = 44$ ). 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. The most noteworthy finding is that for all models examined *the treatment of having rapid feedback was statistically significant with a positive effect*. That is, student scores on the quizzes were higher when they were provided with rapid feedback, with effects of between 0.5 and 1.6 points, which corresponded to 5 percent to 16 percent increases in score (the exact effect size for each model is reported in the footnote to Table 2).

Some general observations can be made for all models that we examined in Fall 2005. First, none of the covariates included were significant (students' grade in Calculus I, Calculus II, and Physics I).

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 (Chen, Kadlowec, and Whittinghill, 2008). We hypothesized then that the students' grade in Calculus II was a reflection of their general abilities in mathematics, rather than specific concepts learned in that course, and that Physics I was significant because most of the concepts in Statics are derived directly from application of physics concepts. The fact that in Fall 2005 we found no significance in the covariate Calculus II may just be that the influence of that course on Statics is marginal (at Rowan, vector calculus is in Calculus III, which is taken concurrently with Statics). The fact that Physics I was not significant was puzzling but not worrisome since the more important finding—that the treatment was significant—was not dependent on this. Covariates were included in the analysis to account for variability in the quiz scores due to the possible effects of those courses. If they were not included in the model, their variability might have masked the effect of the factor under investigation.

The second general finding for Fall 2005 is that the Section in which the student belonged was not significant. This simply states that the two groups of students performed equivalently despite the fact that their classes were on different days and times. Third, the Student factor was always significant (at  $\alpha = 0.05$ ), 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  $\alpha = 0.001$ ), which 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.

The first row of results in Table 2 shows the basic model, which does not examine any interactions between the factors. (The remaining four rows show the results of the analysis if we add selected interactions. Because of the crossover, these are the only four possible combinations of interactions.) The treatment effect was significant at  $p = 0.0318$ . Subsequent models examined the two-way interactions between Section by Period, Student by Treatment, Period by Treatment, and finally the two two-way interactions of Section by Period and Student by Treatment together. In none of these models were any two-way interactions significant at  $\alpha = 0.05$ . The important result from these five models is that the Treatment effect was at least as significant in the models including an interac-

tion term as in the model without them (Equation 1). This gives us even more confidence in the statistical significance of our finding that the rapid feedback positively influenced the students' performance on the quizzes.

We administered three different surveys to the Fall 2005 cohort: one at the start of the course, one at the end of the course, and a single-question survey that was repeatedly embedded within each quiz that the students took. (This battery of surveys differed from that used for the 2004 cohort, in which only three surveys were administered in 2004: one each at the start, middle, and end of the course (Chen, Kadlowec, and Whittinghill, 2005, 2008).) The purpose of the survey at the start of the course was to discern any difference between the two sections with regard to their prior experience with rapid feedback as a teaching tool and their familiarity with the use of a PDA, on the assumption that any difference, if it existed, might explain one group's acceptance or rejection of the use of the PDAs. We asked the students whether they had ever had a class in which they were provided rapid feedback on a regular basis, whether or not they have previously used a PDA, and their level of expertise with the use of a PDA. Our statistical tests for homogeneity did not find any difference between the two sections, which support the notion that the two sections are in general equally novice with a PDA and with rapid feedback as a teaching technique.

A single survey question was embedded within each quiz administered at the end of each treatment period. The question asked, "Considering the period since the previous quiz, how useful have the classes been in helping you learn the course material?" The objective was to determine if the students felt that, from quiz to quiz, the classes were useful (with or without rapid feedback, as was the case). When we compared the two sections with each other and separately for each quiz, we saw no difference in the students' perception of usefulness. In general, they found the classes to be at least "somewhat useful" in nearly all cases. When we combined the two sections' responses to compare their responses from quiz to quiz, a chi-square test of independence found that the responses were significantly dependent on the quiz, meaning the topic during the treatment period ( $p = 0.0076$ ). That is, the students found that the usefulness of class was topic dependent, regardless of whether or not rapid feedback was employed to teach them the topic. One interpretation of this finding is that the use of rapid feedback as a teaching technique had no perceptible effect on the students during

Response	Covariates	Factors			Interactions (if any)
Score	1, 2, 3	Section	Student-in-Section*	Period***	Treatment* ( $p = 0.0318$ )
Score	1, 2, 3	Section	Student-in-Section*	Period***	Treatment** ( $p = 0.0062$ ) <u>Section/Period</u>
Score	1, 2, 3	Section	Student-in-Section*	Period***	Treatment* ( $p = 0.0144$ ) Student/Treatment
Score	1, 2, 3	Section	Student-in-Section*	Period***	Treatment* ( $p = 0.0204$ ) <u>Period/Treatment</u>
Score	1, 2, 3	Section	Student-in-Section*	Period***	Treatment** ( $p = 0.0033$ ) Section/Period & Student/Treatment

NOTE: For each model the factors marked with "\*" were significant at  $\alpha = 0.05$  (5 percent), with "\*\*\*" at  $\alpha = 0.01$  (1 percent), and with "\*\*\*\*\*" at  $\alpha = 0.001$  (0.1 percent). Underlined factors were significant at  $\alpha = 0.10$  (or 10 percent). Note that covariate 1 = Calculus I, 2 = Calculus II, and 3 = Physics I. The effect of having the rapid feedback (effect size) was an increase in the average quiz score (out of a maximum score of 10 points) of 0.5, 1.6, 0.5, 0.5, and 1.6, respectively, over the five models.

Table 2. Results of statistical analyses in various models of the data (fall 2005;  $N = 44$ ).

classes, which was still mainly driven by the topic or its perceived importance or interest.

The objective of the final survey was to assess the students' general perception of the use of rapid feedback in their overall learning experience during the semester. The 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 (Chen, Kadlowec, and Whittinghill, 2005, 2008). When asked how they rated the classes with rapid feedback, 40.5 percent answered that they were "very helpful" and another 54.8 percent answered that they were "somewhat helpful." When asked how they rated the classes without rapid feedback, 38.1 percent rated them as "very helpful" and 40.5 percent rated them as "somewhat helpful." Finally, a high percentage of the students (42.9 percent) believed that they would have done worse in a course taught by the same instructor without having rapid feedback, while 52.4 percent felt they would have performed at the same level. The remaining 4.7 percent believed they would have done better if rapid feedback was not used in instruction.

## V. DISCUSSION

The results beg the question "how much did the students learn in comparison with other instructors and instructional modes?" If the students learned very little of the course content, it matters little if rapid feedback improved on that. Furthermore, if the learning was minimal without rapid feedback, then it is likely that any intervention will result in learning gains. We can offer indirect evidence that the students in this study learned at least as much as students learned in a variety of classroom and institutional settings, and using a variety of teaching styles and techniques. The evidence for this comes from the students' performance on a well established Statics concept inventory developed at the Carnegie Mellon University (Steif and Dantzler, 2005). The Statics concept inventory, called the Concept Assessment Tool for Statics (CATS), is available online (Steif, 2009) and has undergone detailed psychometric analyses for validity and reliability. In addition, it has been examined in comparisons with classroom performance (Steif and Hansen, 2006). During the three-year span of this study, Rowan students took the CATS along with at least seven other institutions that spanned the variety of institutions in the U.S. (school size, public vs. private, graduate vs. undergraduate focus, etc.) and as a group performed somewhat better than the mean of all participating schools. (Note that the number of participants from each school varied greatly, but the Rowan cohort bettered the mean of the cohorts in each year.) This gives us confidence in stating that the Rowan cohort had at least as good of a conceptual understanding of Statics as other students taking the course elsewhere.

While we are confident that the gains demonstrated by students on quiz performance were due to the use of rapid feedback, it is necessary now to refine our use of that term. In the literature on learning sciences, feedback sometimes connotes simply the provision of the correct answer or whether the answer is right or wrong. Such "informational feedback" may go further by including an elaboration on the errors or the correct response (Sims-Knight and Upchurch, 2001). In addition, this feedback may not necessarily be transmitted rapidly; indeed, in most instances this feedback is provided through graded assignments or examinations. Clearly, our

use of the term "rapid feedback" does not fit this broad description. Not only did we always use elaborated informational feedback and provide that feedback immediately to the students, but the rapid-feedback activities included one-on-one student discourse and were followed by further practice when necessary. Thus, we not only provided feedback in the traditional sense, but also used that information to motivate students to improve their understanding of the problem at hand through repeated quizzing. This, we believe, is responsible for the positive gains we observed in student performance, especially in light of the research findings described earlier on feedback provided in the traditional meaning of the word.

Peer discourse was an integral part of the rapid feedback method and our study design, and may be at least partly responsible for the positive findings. A very recent study (Smith et al., 2009) highlighted the importance of peer discussion on gains in understanding in an undergraduate genetics course. The study had a similar implementation to ours, including the use of clickers and peer discussion, but it focused on conceptual questioning only and was based in a large class ( $N = 350$ ). The authors used pairs of conceptually similar questions that required application of the same principles or concepts. The students were first asked to answer one of the two questions individually (call this Q1). They were then invited to discuss the question with their peers (average group size of three) and asked to answer the same question again. Finally, students were asked to answer the second of the paired set of questions (Q2), again individually without peer input. The authors showed that, expectedly, the percentage of students who answered Q1 correctly after discussion was significantly higher. Furthermore, the percentage of students who answered Q2 correctly was also significantly higher, showing that they made gains in understanding from the peer discussion. The most significant finding from this study was that in peer-discussion groups in which no student knew the correct answer to Q1, some students still benefitted from the discussion and were able to answer Q2 correctly. The authors hypothesized that the participants in such groups (as well as students in groups in which the correct answer was known) were "arriving at conceptual understanding on their own, through the process of group discussion and debate." These findings, along with results from prior studies in science education (Crouch and Mazur, 2001; Mazur, 1997; Knight and Wood, 2005) showed that peer discussions almost always improved students' ability to solve conceptual problems, are consistent with our findings. Our results also demonstrated that students improved on their ability to solve applications problems as well as conceptual problems. Note that while peer discussion was required in the groups that received rapid feedback in our study (both the treatment and control groups in 2004 and the treatment group in 2005), it was not required in the no-feedback control group in 2005. Students in this control group were not prohibited or discouraged from discussing anything; the decision was at each group's discretion. Any peer or group discussions came about as a result of the other active-learning methods used, and our observation was that most students did not choose to have discussions at the level or intensity of the rapid-feedback groups.

Another possible explanation for the positive findings of this study may come from what is theorized about the psychology of learning. Atkinson and Shiffrin (1968, 1971) proposed a model of memory and learning that consisted of "sensory registers" that take in information from the environment through the various sensory

processes. The information then enters into the “short-term store” (STS), where it becomes under the control of the person. All information in the STS, which is limited in the total number of pieces of information that can be held, remains there so long as “control processes” are exercised on them to keep them there. Such processes include “rehearsal” (overt or covert repetition of information, such as repeating a phone number), “coding” (recasting of information by putting it into a context or mnemonic phrase), “imaging” (visualizing verbal information as images), along with others. Once information is lost from STS it cannot be recovered. For it to remain in memory, it must be moved from STS to “long-term store” (LTS), which is relatively permanent memory, though it may not always be retrievable or easily retrieved. Information entering the STS comes from a specific modality (visual, auditory, etc.), but associated information from the LTS is activated in all modalities to join it in the STS.

Rehearsal is one of the most important of the control processes. It either increases the momentary strength of information in the STS or otherwise delays its loss. It has also been shown through experiments to facilitate information transfer to the LTS. Furthermore, rehearsal techniques vary in their quality or efficiency in maintaining information in STS or moving them to LTS. We suggest that perhaps our method of providing rapid feedback to students is efficient rehearsal, and it not only keeps the information in STS but also facilitates its moving into LTS, where it helps the student in his learning when he is required to recall the material later in completing homework assignments or quizzes.

In their summary of the body of knowledge on memory phenomena applicable to learning, deWinstanley and Bjork (2002) present a similar interpretation on the empirical findings about effective learning. They claim that learning is an interpretive process whereby new information is stored by making associations and relationships to existing knowledge. Thus, what is important if students are to learn new material is that they be provided with opportunities to engage in processes that facilitate the encoding of the information for future retrieval. They further contend that the components of such processing are attention, interpretation, elaboration, generation, and retrieval practice. Focused attention to the learning task and the material is obviously important to learning. Learning also requires accurate interpretation and thorough elaboration of the new and unfamiliar information. Generation refers to the producing of new information from cues or partial information, which has been shown to be a powerful method of learning. Finally, retrieval practice is the intentional act of retrieving newly learned information from memory for the benefit of enhancing the likelihood of its future retrieval. It also serves as a tool to make the learner aware when he or she does not yet have the capability to retrieve this information. Again, it is possible our method of using rapid, elaborated feedback with student discourse is consistent with this interpretation of effective learning.

Atkinson and Shiffrin (1968, 1971) and deWinstanley and Bjork (2002) emphasize the importance of encoding or rehearsing for the learning of new material. This is in line with the traditional view of learning as being an act of studying (encoding or rehearsing) in order to learn new material, and testing (retrieval) as being a learning-neutral act that serves only to assess one’s learned knowledge but does little to enhance the learning or knowledge retention process. Recent research (Karpicke and Roediger, 2008) has changed this view. In a study of college students’ learning of paired

Swahili-English vocabulary words, Karpicke and Roediger showed that retrieval of learned information is critically important for long-term retention of that information, and is much more effective than repeated encoding (what the authors called “studying”) of that information. Students in this research were asked to learn 40 vocabulary words through repeated study-test conditions. In one condition (C1), upon the successful learning of a new word (ability to correctly define it), that word pair was dropped from future study but was retained on all future tests. In a second condition (C2) all word pairs were repeatedly studied but only pairs that were not yet learned were included on future tests. The third condition (C3) dropped all learned word pairs from both future study and tests. After eight such study-test periods, nearly all 40 words were learned by the subjects at essentially the same rate. The subjects were then tested one week later for their knowledge retention. The results showed that students in the C1 condition retained 80 percent of the vocabulary words compared to C2 and C3, which both showed retention of about 35 percent (this was a four standard-deviations increase of retrieval over non-retrieval conditions). This demonstrates that repeated studying has little benefit (comparing C2 to C3) and highlights the critical importance of retrieval practice for knowledge retention over additional encoding (comparing C1 to C2). While the cited study was on language learning and its applicability to our case is unknown, it does offer an intriguing possible explanation for the results that we observed.

Our results suggest that the role of using clickers or clicker-like devices to be at least partly responsible for the findings, especially given the recent rise in popularity and claims of effectiveness for this classroom communication tool. Typical clicker use in higher education almost always includes some type of rapid feedback, though it may not always be elaborated. Thus, findings from prior research on clickers are relevant to our study, though the two effects of clicker use and rapid feedback (including the elaboration of students’ gains through peer discussions) cannot be separately assessed, as with our case. The literature concerning clickers in the classroom is vast (see for example the compilation at Vanderbilt University’s Center for Teaching (2009)), but reviews of this literature generally agree that most studies are anecdotal in nature and not systematic enough to draw clear conclusions or to determine the magnitude of the effect (Caldwell, 2007; Roschelle, Penuel, and Abrahamson, 2004; Fies and Marshall, 2006). This discussion, therefore, will focus on the few reviews of the clicker literature (Caldwell, 2007; Roschelle Penuel, and Abrahamson, 2004; Fies and Marshall, 2006) and on those quantitative studies based in engineering education.

Caldwell (2007) reported that the vast literature across a variety of disciplines generally finds that clicker use improved student outcomes including higher exam scores, improved passing rates, and student comprehension, and that students viewed clickers highly favorably, though their ratings were less consistent when asked if the clickers helped them learn. The literature also showed general agreement that clickers “tend to change the atmosphere of lectures,” by encouraging students to be actively involved and becoming “emotionally” or “psychologically” invested in the answer they commit to in response to a question. Caldwell noted also that instructors who use clickers rate them favorably because students are more “active and attentive,” and are thus “more pleasant to teach.”

Roschelle, Penuel, and Abrahamson (2004) surveyed a wide range of clicker implementations from K-12 to higher education and found consistent results. The university settings included

classrooms and lecture halls and covered mathematics, physics, chemistry, biology, premedical education, business, and computer science. K-12 results were gathered from middle- and high-school mathematics, physics, and chemistry, as well as reading from all grade levels. Among the most commonly reported outcomes from these studies were increases in student engagement, understanding of subject matter, and enjoyment of class. Again, the authors warned of a lack of scientific rigor among the great majority of the studies, but suggested that the collective and consistent findings represent a real phenomenon associated with clicker use.

Fies and Marshall (2006) described similar findings of the benefits of clicker use, but also cited improved awareness of the students' understanding by both instructors and students, which presumably led to more responsive instruction. The authors also described less frequently cited findings, such as the students' preference for small-group discussions after a clicker question as opposed to whole-class discussion. Also noted was the benefit of anonymity that clickers provided, which allowed more students to participate without the fear of public humiliation or domination by a few vocal classmates.

Boyle and Nicol (2003) and Nicol and Boyle (2003) implemented both clickers and peer discussion in a large engineering mechanics course and administered questionnaires, surveys, and focus groups (no comparison of measurable outcomes was included) to study the impacts on student learning. In general, the study found that students felt overwhelmingly positive toward using the clickers in a large class (117 students). They self-reported being more motivated and engaged, and feeling that they learned more in comparison with a lecture-only class. Our results from student surveys, from a similar course but in small classes, generally agree with these findings.

Paschal (2002) conducted a study in a physiology for biomedical engineers course and found no statistically significant difference in students' test scores when comparing clicker-enabled classes in one year with traditional lectures from a prior year. The study period covered only the first one-third of a semester-long course and the instructional method thereafter reverted back to traditional lecture for the test group. Another important difference between the two study groups was the elimination of homework assignments in the clicker class and their replacement with in-class quizzes based on the reading assignments. Survey results indicated that students in the test group strongly preferred the clicker-enabled classes as "optimal for [their] learning and [their] time management." A significant and likely confounding factor for these results, as the author noted, were the September 11 terrorist attacks that occurred during the study period of the clicker-enabled classes. Also confounding the results were the changes in homework policy and quizzing between the two study groups.

Roselli and Brophy (2006) studied the effectiveness of clickers in concert with several other educational innovations and compared the outcomes to traditional teaching of a biomechanics course. The authors found that, not surprisingly, the clicker classes included many more instances of formative assessment of student understanding and thus offered more opportunities for the instructor to adjust to the students' needs. Exit survey results showed that students liked the anonymity of the clickers, felt it was a good use of class time, and thought it helped them focus in class.

In hindsight our findings are perhaps not surprising, but rather expected. After all, our in-class activities when using rapid feedback are simply a collection of well-documented active-learning tech-

niques (Mazur, 1997; Mehta, 1995; Lochhead and Whimbey, 1987; Barkley, Cross, and Major, 2004) that were implemented through a relatively new technology. Active learning, in its various forms, has been conclusively shown to improve student learning (Prince, 2004; Hake, 1998). Furthermore, using the clickers as we did directly addresses six of the "Seven Principles for Good Practice in Undergraduate Education," as described by Chickering and Gamson (1987, 1996). Perhaps the one surprising result is that the use of rapid feedback via the PDAs and peer discussion was an improvement over other active-learning techniques without the prompt feedback (in the 2005 control group). This improvement manifested in higher quiz scores, albeit with small effect sizes. The question of whether this finding is specific to this study or is more generalizable to other classroom environments is unanswerable at this point.

## VI. CONCLUSIONS

Our most noteworthy findings are that PDA-enabled rapid feedback has a significant and positive effect on student performance when compared to no rapid feedback and, from 2004, that the rapid feedback provided through the use of flashcards appears to be just as useful. If we were to summarize our research findings from the two years, it would be that "it does not matter how one provides rapid feedback to students, as long as it is provided." We would also emphasize that the rapid feedback should be elaborated and accompanied by peer discussion, and used to motivate students through repeated quizzing to examine their current state of understanding, especially at the conceptual level. Finally, it is clear from our survey data that student satisfaction in classes in which rapid feedback was provided was extremely high, which further increases the appeal of this mode of instruction.

Our findings confirm 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 collectively as classroom 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 in enhancing student learning and satisfaction.

Although we are confident in making the above conclusions based on our results, it should be noted that our study was limited to a single engineering course, and in a small-class environment (<25 students per section). Thus, we can make no generalizations about the effectiveness of frequent and rapid feedback in other disciplines or learning environments. It does suggest that further studies in such environments might be worthwhile.

## ACKNOWLEDGMENT

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



## REFERENCES

- Atkinson, R.C., and R.M. Shiffrin. 1968. Human memory: A proposed system and its control processes. In *The psychology of learning and motivation*, Vol. 2, eds. K.W. Spence and J.T. Spence, 89–195. New York: Academic Press.
- Atkinson, R.C., and R.M. Shiffrin. 1971. The control of short-term memory. *Scientific American* 225 (2): 82–90.
- Barkley, E., K.P. Cross, and C.H. Major. 2004. *Collaborative learning: A handbook for college faculty*. San Francisco, CA: Jossey-Bass.
- Boyle, J.T., and D.J. Nicol. 2003. Using classroom communication systems to support interaction and discussion in large class settings. *Association for Learning Technology Journal* 11 (3): 43–57.
- Bransford, J.D., A.L. Brown, and R.R. Cocking, eds. 1999. *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academies Press.
- Caldwell, J.E.. 2007. Clickers in the large classroom: Current research and best-practice tips. *Life Sciences Education* 6 (1): 9–20.
- Chen, J.C., J.A. Kadlowec, and D. Whittinghill. 2005. Using technology for concepts learning and rapid feedback in statics. In *Proceedings of the 2005 American Society for Engineering Education Annual Conference and Exposition*. Portland, OR.
- Chen, J.C., J.A. Kadlowec, and D. Whittinghill. 2008. Using handheld computers for instantaneous feedback to enhance student learning and promote interaction. *International Journal of Engineering Education* 24 (3): 616–24.
- Chickering A.W., and Z.F. Gamson. 1987. Seven principles for good practice in undergraduate education. *AAHE Bulletin* 39 (7): 3–7.
- Chickering, A.W., and S.C. Ehrmann. 1996. Implementing the seven principles: Technology as lever. *AAHE Bulletin* 49 (2): 3–6.
- Crouch, C.H., and E. Mazur. 2001. Peer instruction: Ten years of experience and results. *American Journal of Physics* 69 (9): 970–77.
- Data Description, Inc. <http://www.datadesk.com> (last accessed, September 2009).
- deWinstanley, P.A., and R.A. Bjork. 2002. Successful lecturing: Presenting information in ways that engage effective processing. In *New direction for teaching and learning; special issue: Applying the science of learning to university teaching and beyond*, no. 89, eds. D.F. Halpern and M.D. Hakel, 19–31. San Francisco, CA: Jossey-Bass.
- Felder, R.M. 1995. A longitudinal study of engineering student performance and retention. IV. Instructional methods and student responses to them. *Journal of Engineering Education* 84 (4): 361–67.
- Felder, R.M., and R. Brent. 2001. Effective strategies for cooperative learning. *Journal of Cooperation and Collaboration in College Teaching* 10 (2): 63–69.
- Fies, C., and J. Marshall. 2006. Classroom response systems: A review of the literature. *Journal of Science Education and Technology* 15 (1): 101–09.
- Freeman, M., and J. McKenzie. 2001. 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*, eds. D. Boud, R. Cohen, and J. Sampson, 156–169. London: Kogan Page Limited.
- Hake, R.R. 1998. Interactive-engagement vs. traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics* 66 (1): 64–74.
- Karpicke, J.D., and H.L. Roediger. 2008. The critical importance of retrieval for learning. *Science* 319 (5865): 966–68.
- Kluger, A.N., and A. DeNisi. 1996. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin* 119 (2): 254–84.
- Knight, J.K., and W.B. Wood. 2005. Teaching more by lecturing less. *Cell Biology Education* 4 (4): 298–310.
- Lochhead, J., and A. Whimbey. 1987. Teaching analytical reasoning through thinking aloud pair problem solving. *New Directions for Teaching and Learning* 30: 73–92.
- Mason, R.L., R.F. Gunst, and J.L. Hess. 1989. *Statistical design and analysis of experiments, with applications to engineering and science*. New York: Wiley Series in Probability and Mathematical Statistics.
- Mazur, E. 1997. *Peer instruction: A user's manual*. Upper Saddle River, NJ: Prentice Hall.
- Mehta, S.I. 1995. A method for instant assessment and active learning. *Journal of Engineering Education* 84 (3): 295–98.
- Nicol, D.J., and J.T. Boyle. 2003. Peer instruction versus class-wide discussion in large classes: A comparison of two interaction methods in the wired classroom. *Studies in Higher Education* 28 (4): 457–73.
- Paschal, C.B. 2002. Formative assessment in physiology teaching using a wireless classroom communication system. *Advances in Physiology Education* 26 (4): 299–308.
- Prince, M. 2004. Does active learning work? A review of the research. *Journal of Engineering Education* 93 (3): 223–31.
- Roschelle, J., W.R. Penuel, and L. Abrahamson. 2004. Classroom response and communication systems: Research review and theory. In *Proceedings of the Annual Meeting of the American Educational Research Association*. San Diego, CA. [http://ubiqcomputing.org/CATAALYST\\_AERA\\_Proposal.pdf](http://ubiqcomputing.org/CATAALYST_AERA_Proposal.pdf) (last accessed September 2009).
- Roselli, R.J., and S.P. Brophy. 2006. Experiences with formative assessment in engineering classrooms. *Journal of Engineering Education* 95 (4): 311–24.
- Sims-Knight, J., and R.L. Upchurch. 2001. What's wrong with giving students feedback. In *Proceedings of the 2001 American Society for Engineering Education Annual Conference and Exposition*. Albuquerque, NM.
- Smith, K.A., S.D. Sheppard, D.W. Johnson, and R.T. Johnson. 2005. Pedagogies of engagement: classroom-based practices. *Journal of Engineering Education* 94 (1): 87–101.
- Smith, M.K., W.B. Wood, W.K. Adams, C. Wieman, J.K. Knight, N. Guild, and T.T. Su. 2009. Why peer discussion improves student performance on in-class concept questions. *Science* 323 (5910): 122–24.
- Steif, P.S., and J.A. Dantzer. 2005. A statics concept inventory: Development and psychometric analysis. *Journal of Engineering Education* 94 (4): 723–29.
- Steif, P.S., and M. Hansen. 2006. Comparisons between performances in a statics concept inventory and course examinations. *International Journal of Engineering Education* 22 (5): 1070–76.
- Steif, P.S. <http://engineering-education.com/CATS> (last accessed September 2009).
- Vanderbilt University Center for Teaching. [http://www.vanderbilt.edu/cft/resources/teaching\\_resources/technology/crs\\_biblio.htm](http://www.vanderbilt.edu/cft/resources/teaching_resources/technology/crs_biblio.htm) (last accessed September 2009).

## AUTHORS' BIOGRAPHIES

John Chen is an associate professor of Mechanical Engineering. He received his B.S. in mechanical engineering from the University

of Virginia in 1985 and his M.S. and Ph.D. in mechanical engineering from Stanford University in 1991. He has been a faculty member and active member of ASEE 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 was on the faculty at Rowan University from 1998 to 2008 and moved to California Polytechnic State University in 2008.

*Address:* California Polytechnic State University, Building 13, Room 226, San Luis Obispo, CA 93407; telephone: (+1) 805.756.2144; fax: (+1) 805.756.1137; e-mail: jchen24@calpoly.edu.

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 American Society for Engineering Education since 1998 and is actively involved with the Mechanics Division.

*Address:* Rowan University, Department of Mechanical Engineering, 201 Mullica Hill Road, Glassboro, NJ 08028-1701; telephone: (+1) 856.256.5344; fax: (+1) 856.256.5241; e-mail: kadlowec@rowan.edu.

Dexter Whittinghill is an associate professor in the Mathematics Department at Rowan University. He is in his fourteenth year at Rowan, and has been a professor since 1984 upon receiving 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 and is currently active within statistical education groups within the statistics and mathematics communities.

*Address:* Rowan University, Department of Mathematics, 201 Mullica Hill Road, Glassboro, NJ 08028-1700; telephone: (+1) 856.256.4500 ext. 3879; fax: (+1) 856.256.4816; e-mail: whittinghill@rowan.edu.