Concept-based Instruction and Personal Response Systems (PRS) as an Assessment Method for Introductory Materials Science and Engineering

Maura Jenkins and Edward K. Goo
University of Southern California

Abstract

Personal response systems (PRS) are gaining in use as a method to engage students in large science and engineering lectures. Faculty pose questions to the class mid-lecture and receive immediate feedback via remote-control "clickers" as to whether students understand the underlying concepts necessary to solve problems on homework and exams. Thus, the pace of the lecture can be adjusted accordingly to focus on the most difficult concepts.

This method has been thoroughly developed for introductory chemistry and physics courses. Pioneers have developed ConcepTests, or multiple-choice questions that focus on conceptual understanding, rather than calculation. These questions encourage peer interaction, as instructors allow students to vote a second time after discussing their initial answer with classmates. Introductory Materials Science and Engineering shares many characteristics of the courses in which this method has been successful; lectures are often large, the course is required, and many students are non-majors.

In this paper, we share our experience in applying this method to an introductory materials science course. We will present data on student responses, test scores, demographics, and comparison to previous semesters without the response systems. Plans to develop a common bank of materials ConcepTests, building on existing concept inventories will also be discussed. Practical details about the equipment and software will be shared as well.

Introduction

Concept inventories, or multiple-choice exams focusing on 20-30 major concepts of a specific field, have recently experienced a surge in development as assessments independent of high-stakes testing. In recent years, concept inventories have been developed and tested for reliability in such fields as physics (mechanics)[1], statics[2], fluid mechanics[3], materials[4], and chemistry[5]. In developing these inventories, faculty focus on concepts and reasoning over computation, using varying degrees of rigor to distinguish between the two[2]. In many cases, the developers make use of open-ended responses from current students to develop distractors based on common misconceptions[5]. Reliability is tested by analyzing individual test items[3] or administering the entire assessment to multiple groups of students[4].

As a separate development, similar conceptual questions are being used in large lectures of introductory science classes such as chemistry and physics[6]. Faculty pose questions to the class mid-lecture and receive immediate feedback via remote-control "clickers" as to whether students understand the underlying concepts necessary to solve problems on homework and exams. Thus, the pace of the lecture can be adjusted accordingly to focus on the most difficult concepts. Faculty members observe improvements in student learning (grades or passing rate), attitude, and class participation. Few combine concept inventory questions with in-class activities[7], and even fewer publish analysis evaluating the effectiveness of this type of instruction[8].

In this paper, we present an example of how remote-control personal response systems (PRS) can be applied to an introductory Materials Science course using conceptual questions. Data analysis is provided to begin to evaluate the effectiveness of this method, and future work collecting more detailed data and combining concept inventories with lecture sessions is discussed.

Course-Specific Background and Motivation

Table 1 outlines the topics included in University of Southern California (USC) MASC 110, Introduction to Materials Science and Engineering. Unlike many introductory materials science courses, this course includes significant chemistry content and substitutes for the first-semester chemistry requirement for aerospace, mechanical, electrical and industrial engineering majors. A chemistry textbook is used[9], and materials science concepts are introduced through laboratory activities and lectures later in the semester.

Table 1. Topics included in Introduction to Materials Science and Engineering course discussed in this paper. The course substitutes for first-semester chemistry for aerospace, mechanical, electrical and industrial engineering majors. A chemistry textbook is used.

Syllabus (Lecture) Topics	Laboratories
Atoms and Molecules	Determination of Avogadro's Number
Atomic bonds	Atomic Spectroscopy
Gas, liquids and solids	Thermal Reduction of Copper Ore to Copper Metal
Chemical reactions, equilibrium and kinetics	Crystal Structures of Metals
Crystals, crystal defects and symmetry	Crystal Structures of Ionic Solids
Thermodynamics	Microstructure of Metals
Electrochemistry	Corrosion
Water	Hardness
Organic Compounds	Polymers
Polymers	Phase Equilibria
Minerals	
Metals	
Ceramics and Semiconductors	
Biopolymers	

This course shares many characteristics with courses in which personal response systems have been successful; lectures are large, the course is required, most students are non-majors, and the content is chemistry-focused. Although the graduate program provides teaching assistants, USC does not offer an undergraduate degree in materials science and engineering (a minor is offered). Therefore, none of the students in the course are majors.

The course is taught twice a year with approximately 75 students in fall and 50 in spring. The Fall 2004 class that was studied contained 69 students.

Method

InterWrite PRS™ "clickers" by GTCO CalComp (previously Educue) were made readily available to students by the textbook publisher[10]. Many large textbook publishers provide one or more types of these systems at a significant discount when bundled with student-purchased book. In this case, the PRS clickers would have cost up to $30 each if purchased by the university or department, but the publisher bundled them with the student textbook and lab manual for just $10 additional cost to students. Free receivers and software were also provided by the publisher for each 40 textbooks sold.

A bank of ConcepTest questions for chemistry available on a web site maintained by the University of Wisconsin-Madison[11] was used as a starting point. The instructor created other questions as needed. Sample questions used in class are listed in Table 2. The purpose of these questions was to test whether students understand concepts shortly after they are presented in lecture. As the emphasis is on conceptual understanding, the questions require minimal calculation. The response devices (detailed below) limit the format to multiple choice questions. Exam questions for this course test these same concepts, but are often in the format of numerical calculations. An example final exam question is also included in Table 2. Students have the opportunity to practice applying course concepts to problem-solving in homework assignments (problem sets), which are similar in format to the exams.

Table 2. Example conceptual questions posed to students in lecture and the response rates for each multiple-choice answer. Correct responses are in italics. A sample exam question is also included to illustrate the more computational nature.

In-Class Question or Prompt	**Multiple-Choice Answers and Student Responses (Correct responses in italics)**
Gold, silver, copper and platinum have positive reduction potentials.	*True 61%* False 39%
Of the following metals which will not show a ductile-brittle transition temperature: Al, Mg, Ti, Cu and Fe?	Al and Ti 51% Mg and Cu 7% *Al and Cu 23%* Cu and Ti 19%
Albite is a feldspar which is a network silicate with one fourth of the silicon replaced with aluminum. It also has Na. What is the empirical formula for Albite?	*$NaAlSi_3O_8$ 18%* $Na_5AlSi_3O_{10}$ 15% $Na_7AlSi_3O_{11}$ 22% $Na_9AlSi_3O_{12}$ 12% $NaAlSi_3O_6$ 33%
Related Final Exam Question: For the mineral $Ca_2Mg_4Fe(Si_4O_{11})_2(OH)_2$ give the oxidation state for each atom.	

The software used was TurningPoint™ (Turning Technologies)[12], which allows questions to be posed within PowerPoint. Once a question is posed, TurningPoint displays a countdown for the time allotted, and a small grid of numbers at the bottom or side of the screen lets students see if their response has been recorded. After the time is up, a histogram of responses is immediately generated. In many cases, students were encouraged to discuss their answer with a classmate before answering the same question

a second time. If a significant number of students answered incorrectly, the instructor led a discussion of the correct answer, including what was wrong with the incorrect answers.

Using a web site set up by university computing technology groups, students registered the serial number of their clickers to their student ID numbers. Attendance and participation in class, as measured by responses using the PRS clicker, counted toward 5% of students' final grades. Between 2 and 6 questions were posed during almost every 50-minute lecture.

The receivers are small, relatively mobile devices, so it is possible to store them on a portable cart along with an instructor computer. However, one receiver is required for every 25-40 students, so for large lectures it is far easier to mount them distributed across the front and sides of the room. For this course, we negotiated with an instructor from an unrelated department to teach in a physics classroom already equipped with receivers. The physics and materials science instructors brought their own laptops and plugged into a stationary instructor console linked to the receivers.

A second and final difference between Fall 2003 and Fall 2004 courses was virtual office hours conducted in a chat room within the Blackboard course management system[13]. Originally, office hours were scheduled for Sunday evenings online and Wednesday afternoons in person. Within the first few weeks, the instructor noted a sharp contrast in attendance, with far more students logging on than visiting the office. (Problem sets were due on Mondays, so the difference in attendance was not completely attributable to characteristics of the online chat environment.) As a result, office hours were adjusted to include set virtual hours with the option of making an appointment to meet in person.

Three major measures were used to assess the effectiveness of employing conceptual lecture questions and a PRS system in this course:

1. Comparison of PRS class participation and percent correct responses with exam and course grades
2. Development of a statistical model to predict course grade, including demographic variables, to see if PRS participation has greater effect than standardized test scores and similar measures
3. Survey (conducted using the PRS clickers) measuring student attitudes toward the system

As this was the first semester that PRS clickers were used in this course, overall exam and course grades were the only measures of student learning used to evaluate the new lecture format. A more valuable but time-consuming assessment would be test item analysis, which would allow individual student's understanding of specific concepts to be followed from lecture to exam. This would serve to address research questions such as, "Do students who incorrectly answer a question in class learn from their mistakes and answer correctly on the exam?" This level of analysis is tentatively planned for future semesters.

Statistical analysis was run using SPSS statistical software[14]. Pearson correlations (2-tailed significance) were calculated for all pairs of variables. Independent samples t-tests were employed to compare groups (In all cases students were split into just 2 categories, in part due to the relatively small sample size.). A stepwise linear multiple regression was also run to see if final course grade could be predicted, and if so, whether class participation with the PRS system was a factor. Unfortunately, the curve on which this course is graded precludes a valid direct comparison of the Fall 2004 with the Fall 2003 class.

Variables related to the course included summary measures for the whole semester as well as analysis of each midterm and the final exam as separate segments of the course. For example, participation and correct responses in the weeks between the first and second midterms were compared to the score on Midterm 2, as were total participation and correct responses for the entire semester compared to overall course grade. "Participation" is measured as total PRS clicker responses, as distinguished from correct responses, as a percentage of the total number of questions posed to the class. Midterm exams were administered during weeks 6 and 11 of the 15-week semester.

In comparing course attendance and participation to the final grade, it is valuable to see if standardized test scores and similar measures have better, comparable, or worse predictive value. The non-course variables included in this analysis were:

- SAT math, verbal, and composite scores
- high school GPA
- whether the student reported AP Calculus AB or BC or Chemistry exam scores
- university chemistry, math, and physics placement test scores
- minority status (underrepresented or not)
- gender
- whether the student was a freshman or upperclassman

Does this Method Clear up Student Misconceptions?

In order to assess whether the use of conceptual questions and PRS clickers helps to clear up student misconceptions, individual concepts would have to be tracked for each student from the initial lecture question to corresponding midterm and final exam questions. As individual test item data was not gathered, this level of analysis was not possible. However, in this case, most students answered the initial lecture question correctly (average 66% ± 11% for all students and questions), leaving few instances of students who have the opportunity to learn from their initial mistakes. It is likely that if this data were available, the sample size would not be sufficient to draw conclusions.

Are Conceptual Questions Better or Worse at Predicting Exam Performance than Admission Data?

The correlation between correct PRS clicker responses prior to Midterm 1 and score on Midterm 1 is only significant at the .05 level ($p < .05$). In this case, the chemistry and physics placement test scores and SAT math and composite scores are more significantly correlated ($p < .01$) with exam score than are correct PRS responses. In the case of Midterm 2 and the final exam, scores are correlated at the .01 level to correct responses in

the weeks leading up to each exam, along with chemistry placement test score and high school GPA. *Thus, later in the semester, correct responses using the PRS clickers is one of just a few measures that might be used to predict exam performance.*

Although correct PRS clicker responses are more highly correlated to exam scores later in the semester, an interesting result also emerges from data collected during the first few weeks of the semester. For the PRS questions leading up to Midterm 1 and the scores on that exam, most students fall into one of two different groups. Figure 1a is a scatter plot of Midterm 1 Score vs. Percent Correct PRS Clicker Responses. It is evident that most students fall into one of two different groups with distinctively different slopes. It is worth noting that the correct response and participation during the first portion of the course (weeks leading up to midterm 1) remain consistent with final counts for the entire semester, so this trend is not simply the result of elementary course material.

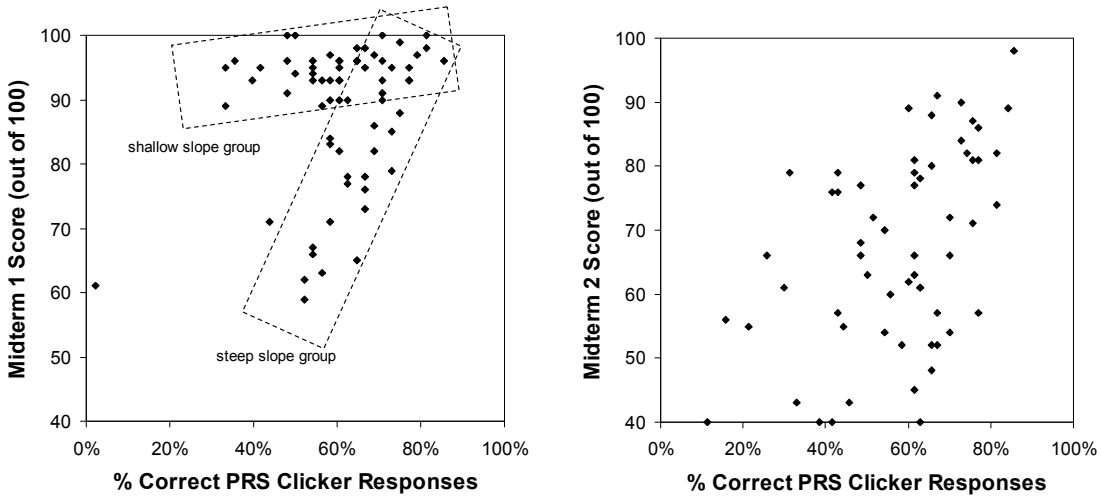

Figure 1. Scatter plots of Midterm Exam Score vs. % Correct PRS Clicker Responses for the Midterm 1 and Midterm 2. During the first few weeks of the semester (Midterm 1, Fig. 1a), most students fall into one of two different groups with distinctively different slopes (indicated by dashed-line boxes). By Midterm 2 (Fig. 1b), the scatter plot has a similar shape, but the two groups are no longer distinct.

To determine the nature of this difference, students were separated into a steep slope or shallow slope group, as indicated by the dashed-line boxes in Figure 1a. (Students with correct responses of greater than 70% were not included in this analysis because they were too close to the intersection of the trend lines.) An independent samples t-test was run to determine what factors, if any differed statistically significantly between the two groups. Significant at the .05 level were three variables (listed in order of decreasing significance):

1. whether the student reported an AP Calculus BC exam score
2. SAT composite score
3. SAT math score

In each case, the test score is higher for the group with the shallower slope, who also had higher Midterm 1 exam scores. Although the exam questions were more mathematical than the conceptual questions posed in lecture, calculus was not necessary to solve them. (Refer to the sample exam question in Table 2 above.) It is more likely that problem-solving skills, rather than math background, is an underlying factor. What the above result suggests about PRS clicker use at the beginning of the semester (which is necessarily review for some students) is that for a group of less well-prepared students, correct PRS responses have a very strong relationship with good grades. It would have been a stronger argument for PRS systems to show that participation, rather than correct responses, is correlated to high exam grades (which was not true for this group in the first few weeks of the semester); however, the above result illustrates that short quizzes embedded in lecture are as valid an assessment of learning as exams.

A similar scatter plot of Midterm 2 Grade vs. Percent Correct PRS Clicker Responses prior to that exam is shown in Figure 1b. The trend of two groups (slopes) of students has disappeared, although the triangular shape remains. Again, for this portion of the semester, percent correct PRS clicker responses is a stronger indicator of exam performance than most standardized test scores or other pre-course preparation. Also, there are no statistical differences between students in the two initial slope groups in later exams or the final course grade.

To determine whether final course grade could be predicted, a stepwise linear regression model was generated. The first iteration (model) accounted for 43% of the variation in course grade using percent correct PRS clicker responses alone. The second and final iteration produced a model accounting for 53% of the variation in final course grade using percent correct clicker responses and high school GPA.

Can the Effects be Explained by Class Participation Alone?
No attendance data exist for previous offerings of this course, but it is the instructor's impression that attendance has increased dramatically since the introduction of PRS clickers. It is possible, then, that any increase in student learning could be a result of increased attendance rather than the conceptual questions posed during lecture.

Over the course of the entire semester, both percent correct responses and attendance measured by response rate using the PRS clickers are highly significantly correlated with both final exam grade and course grade ($p < .001$). It is difficult to separate participation from correct responses, since most students answered correctly when they did answer. The average and standard deviation of percent correct responses for the entire class was $66\% \pm 11\%$, while the corresponding value for participation was $87\% \pm 18\%$. Thus, for this dataset it is impossible to separate the effects of attendance from those of the conceptual questions themselves.

Does This Method Have Different Effects on Students from Different Groups?
To determine if there are any significant differences between various groups of interest, independent samples t-tests were run based on

- whether or not the student reported an AP Calculus BC exam score

- whether the student was a freshman or upper-class student
- whether or not the student was an underrepresented minority
- whether the student was female or male

The only statistically significant difference between any of the groups listed above was in the response rate (questions answered, regardless of whether they were correct). Freshmen were more likely than upperclassmen to respond using their PRS clickers: 91% vs. 74% response rates.

Other results related to response rate are also of interest. Both women and underrepresented minorities answered questions using the PRS clickers more often than other groups. In the class of 69 students, there were 12 females and 9 minority students. In the case of women, the difference was 94% vs. 85% response rate, and for minorities, it was 93% vs. 86%. Although these numbers are encouraging, the difference is not statistically significant (The standard deviation of the majority group is ~19%.).

How Do Students Feel about the Clickers?

The final assessment of PRS clicker use in this introductory materials science class is a student survey conducted in class using the devices. The results are listed in Table 3. Students perceived the clicker exercises to be very helpful in understanding concepts, providing immediate feedback, and helping them feel more comfortable answering when unsure. Additionally, 60% of students routinely discuss their answers with classmates, as they were encouraged to do. These are very positive responses corresponding to the major goals and benefits of PRS clicker use. Most students did not find the questions helpful in preparing for problem sets or increasing confidence for exams. This is perhaps an indication to reevaluate the balance of computational and conceptual problems in these areas of the course. Finally, 54% of students agreed that they would prefer to take a course using PRS clickers if other aspects of the course were equal.

Future Work

The survey results indicate that students perceive a disconnect between the types of questions posed in class and the types included in homework assignments and on exams. Next steps include integrating existing chemistry and materials science concept inventories more systematically into the course, perhaps by adding more conceptual questions to exams and homework assignments. It stands to reason that a method aimed at augmenting conceptual understanding might not be properly evaluated using computational exams. Thus, concept inventories would probably serve as better assessments of both student learning and of PRS clickers than exams, and the fact that the concept inventories are standardized would support a more rigorous evaluation of the methods' value. More detailed data collection from non-PRS sources is also necessary, e.g. individual students' scores on individual test and concept inventory items.

Plans are also underway to share these results as well as classroom facilities on our campus so that other faculty will find it easy to incorporate this and similar methods into their teaching.

Table 3. PRS Clicker Use Student Survey Results. 51% of students almost always discussed questions in small groups with peers, and 54% agreed that they would prefer a course using PRS clickers.

Survey Question	Responses via Clickers
How helpful or unhelpful were the clicker exercises in helping you to understand the concepts of this course?	20% Not very helpful 14% Somewhat helpful 8% Neither helpful nor unhelpful **39% Somewhat helpful** 18% Very helpful
How helpful or unhelpful were the clicker exercises in preparing you to complete your assignments?	**37% Not very helpful** 16% Somewhat helpful 20% Neither helpful nor unhelpful 22% Somewhat helpful 4% Very helpful
How helpful or unhelpful were the clicker exercises in increasing your confidence in your ability to take course exams?	**44% Not very helpful** 28% Somewhat helpful 20% Neither helpful nor unhelpful 4% Somewhat helpful 4% Very helpful
How frequently or infrequently did you work in small groups to discuss a question as part of a clicker exercise?	20% Never or almost never 11% Infrequently 9% About half the time 9% Frequently **51% Always or almost always**
How helpful or unhelpful were the clicker exercises in providing you with immediate feedback on your understanding of a course concept?	8% Not very helpful 20% Somewhat helpful 10% Neither helpful nor unhelpful **40% Somewhat helpful** 22% Very helpful
How helpful or unhelpful were the clickers in making you comfortable in providing an answer in cases when you were unsure or didn't know the answer?	15% Not very helpful 2% Somewhat helpful 21% Neither helpful nor unhelpful 19% Somewhat helpful **43% Very helpful**
All other things being equal, I would prefer taking a class that uses clickers over one that does not use clickers.	22% Strongly disagree 6% Somewhat disagree 18% Neither agree nor disagree **34% Somewhat agree** 20% Strongly agree

Acknowledgements

The authors would like to thank USC's NICE-PRS discussion group members, including Chris Gould; USC Center for Scholarly Technology staff members Rick Lacy, Sue Gautsch, Otto Khera and others; and Brooks-Cole sales rep Sylvia Sotomayor.

References

[1]Hestenes, et al. "Force Concept Inventory." *The Physics Teacher* **30** (3), 141, (1992).

[2]Danielson, S. "Developing Statics Knowledge Inventories." *2004 Frontiers in Education Conference.*

[3]Martin, J., J. Mitchell and T. Newell. "Development of a Concept Inventory for Fluid Mechanics." *2003 Frontiers in Education Conference.*

[4]Krause, S., J. Decker and R. Griffin. "Using a Materials Concept Inventory to Assess Conceptual Gain in Introductory Materials Engineering Courses." *2004 Frontiers in Education Conference*.

[5]Krause, S., J. Birk, R. Bauer, B. Jenkins, and M. J. Pavelich. "Development, Testing, and Application of a Chemistry Concept Inventory." *2004 Frontiers in Education Conference.*

[6]Mazur, E. *Peer Instruction: A User's Manual.* (Prentice Hall, Upper Saddle River, NJ, 1997)

[7]Mitchell, J. and J. Martin. "Interactive Session: Introduction to Concept-Based Instruction." *2004 Frontiers in Education Conference.*

[8]Crouch, C.H. and E. Mazur. "Peer Instruction: Ten Years of Research and Results." *American Journal of Physics* **69**, 970-977 (2001).

[9]Oxtoby, D.W., H.P.Gillis, and N.H. Nachtrieb. *Principles of Modern Chemistry.* 5th ed. (Brooks-Cole, 2002).

[10]InterWrite PRS™ by GTCO CalComp at http://www.gtcocalcomp.com/interwriteprs.htm

[11]Ellis, A. B., A. Cappellari, G.C. Lisensky, J.K. Lorenz, K. Meeker, D. Moore , K. Campbell, J. Billmann, and K. Rickert. http://www.jce.divched.org/JCEDLib/QBank/collection/ConcepTests/

[12]TurningPoint™ by Turning Technologies at http://www.turningtechnologies.com/downloads.htm

[13]Blackboard™ course management system by Blackboard Inc. at http://www.blackboard.com

[14]SPSS™ statistical analysis software by SPSS Inc. at http://www.spss.com