Evaluation of Computer-Based Assessment Methods for Engineering Courses

Neil L. Book and Oliver C. Sitton
Department of Chemical and Biological Engineering
University of Missouri-Rolla

Abstract

Several computer-based assessment methods have been used in chemical engineering classes at the University of Missouri-Rolla. This experience provides the basis for an evaluation of the technology and human factors involved in the application of these methods.

From the student's perspective, the major strength of computer-based assessments is the instantaneous feedback that is provided. From the instructor's perspective, the major strengths are twofold: 1) rapid grading and 2) individualized assessments. The major drawbacks for the students are the loss of partial credit and the insecurity associated with working in a virtual environment. Faculty find that the time required to construct effective computer-based assessments is much greater than that for paper-based assessments. Unfortunately, the increase in construction time can be greater than the reduction in grading time. With the current technology, it is extremely difficult to develop large, multi-step problems requiring problem analysis and solution synthesis skills without leading the student to the answer. This, of course, defeats the purpose of the assessment for this type of problem.

Introduction

The technology for implementing computer-based assessments is rapidly evolving. As technological shortcomings are identified, they are quickly eliminated. Thus, difficulties with human factors are more important for long-term implementation than technological shortcomings.

Students are very comfortable with the technology and adapt to its usage very rapidly. However, the technology is not perfectly reliable and this leads to insecurity for both the student and the instructor. Instructors seek technology that enhances learning. This enhancement is most useful if it comes with modest increases in time and effort and with technology that is user-friendly.

The benefits of active learning in various forms (active learning, collaborative learning, cooperative learning, and problem-based learning) are well-documented[6]. However, active learning exercises consume class time and the amount of time, depending on type of activity, can be dependent on class size[3]. Technology that provides the means to introduce active learning with minimal time consumption—especially transition time—is of great value.

Computer-Based Exams

Exams can be administered using computer-based testing systems and/or can be constructed to require the use of computers to solve problems.

Technology

Computerized classrooms[1] in which each student has a computer with access to all necessary software are required. Unfortunately, computers occasionally fail and all networks have downtime. Further, computer programs occasionally lockup and many have errors that can produce spurious results. Loss of files during an exam can be stressful and claims of computer difficulties can be used to cover a lack of preparation. Course policies, such as replacement of an exam score with the final exam score and/or giving exams when additional time can be granted to reproduce any lost work, are necessary. These policies alleviate some anxiety in the students.

Courseware (such as BlackBoard and WebCT) provides an environment for constructing and administering exams. Developing exams is only one capability of courseware. Special purpose programs (such as Respondus) are better environments for developing exams that can be uploaded into courseware systems. Multiple choice, multiple guess, matching, true/false, yes/no, and short answer types of questions can be machine graded. Effectively, there is no partial credit (part scores for matching and multiple guess types of questions are awarded) for machine graded problems. Essay questions may be posed but require manual grading off-line. Individualized exams are easily created. An exam can be constructed from several "pools" of questions. Each student's exam is composed of designated subsets of questions from the various pools. A random selection of the subset can be designated. Feedback can be provided for correct and incorrect responses upon exam submission or delayed.

Student Perspective

Student perspectives of computer-based exams vary widely. The absence of partial credit tends to be the source of much of the negative attitude. The insecurity associated with working in a virtual environment also creates negative perspectives. This insecurity tends to diminish with experience. Practice exams and/or assignments using the same computing environment are an effective means of limiting the insecurity[4]. Instantaneous grading and feedback is universally viewed as a positive. Comments regarding the use of "BlackBoard" exams from anonymous student evaluations of a senior-level, technical elective course in physical property estimation were 4 negative and 2 positive in a class of 21 and 4 negative and 1 positive in a second class of 17. Comments regarding the use of a suite of computer routines (MathCAD and Excel) to solve exam problems in the same courses were 1 negative and 2 positive and 1 negative and 1 positive, respectively. The majority of both classes accepted the exams without comment.

Instructor Perspective

Access to engineering software during an exam provides a mechanism for testing in a computing environment that is similar to the workplace. Grander, more realistic problems can be posed in limited testing time[4]. Unfortunately, students often treat software as a "black box" that produces results that are always correct and applicable. Posing exam questions that assess knowledge of the assumptions and processes within the software is difficult but important to do.

Individualized assessments minimize plagiarism. Machine grading with instantaneous feedback, especially of individualized exams, is a significant advantage of courseware systems. Unfortunately, the additional time required to construct exams that are amenable to machine grading, counteracts the time savings. Classes with ten or fewer students do not payout. The breakpoint is close to 30 students depending on course content.

Experience

Exams in which students have a suite of routines available that perform specific functions, such as computing activity coefficients using various models (UNIFAC, UNIQUAC, Wilson, or NRTL) or volumetric and thermodynamic properties from an equation of state (ideal gas, virial, Soave-Redlich-Kwong, Peng-Robinson, or Benedict-Webb-Rubin), allow for problems to be posed that have these calculations embedded in the solution. With routines like these available, students can be asked on exams to design compressors, make flash calculations, or compute reaction equilibria using nonideal mixtures. This would be impossible for students with calculators. Of course, it is important that each student understands the assumptions and processes contained in each routine. Class time must be devoted to training the students to use the routines and this requires special classroom facilities, however, the training exercises are an opportunity to switch to active learning mode. Development of routines across the curriculum is the most effective means of providing the suite. Both MathCAD and Matlab have been used to develop a suite of routines. In both cases, a considerable portion of a summer was devoted to adapting a number of existing routines and creating additional ones to form a suite of consistent, integrated routines.

Individualized exams are a necessity when students have networked computers. Also, the Synchroneyes Teacher program, which has a feature that allows the instructor to monitor each student computer during an exam, is recommended.

Constructing engineering exams in courseware systems is difficult and time consuming. Questions that assess problem analysis and solution synthesis skills are difficult to pose as multiple choice or short answer questions that can be machine graded. Multi-step problems typically are decomposed into a series of single-step problems that eliminates the necessity for the student to synthesize a solution. Ensuring that mistakes lead to incorrect answers for engineering problems is a challenge. This generally requires the identification of the common mistakes that will be made—something that is exceedingly difficult to do. These difficulties increase exam development time by a large factor.

Engineering problems frequently have numerical answers obtained from complex calculations. Rounding, significant digits, even the order in which the calculations are performed can produce slightly differing results that are correct. Unfortunately, many significant errors can produce results that are not far removed from the correct result. Machine graded exams have great difficulty identifying which of these two cases has occurred. Also, poorly-constructed multiple choice problems can, by the nature of the choices, identify the correct answer and/or eliminate certain choices from consideration. A good practice is to use the same set of potential answers for all problems. For example, the set of choices for systems that allow up to 20 could be: -500, -200, -100, -50, -20, -10, -5, -2, -1, 0, 1, 2, 5, 10, 20, 50, 100, 200, and 500, for each problem requesting a numerical answer. This provides potential answers that vary over orders of magnitude and provides a number of selections—even if the answer is known to be positive or negative. It is important that the question be stated to select the closest answer and that occasional problems be posed with answers that are larger than the largest choice (for example, 8000), smaller than the smallest choice, and close to zero (0.01 for example). The trick is to find conditions that ensure that mistakes will not result in selection of the correct answer. For example, if the correct answer is 2.53, mistakes must produce results that are greater than 3.5 or smaller than 1.5. Experience indicates that much time can be expended trying to anticipate mistakes and that many questions (or specified conditions for a question) are ultimately abandoned when a mistake is discovered that does not produce a result sufficiently different from the correct answer. Of course, all mistakes cannot be anticipated so students will make mistakes but still select the correct choice. These difficulties are not as significant for homework or remedial/review materials. Additionally, courseware systems allow these types of materials to be accessed outside the classroom (asynchronous learning). There is much potential in these areas—especially if large batteries of questions are available (as exist for engineering physics and calculus).

Reduced grading time is a function of class size while exam development time is not. For small classes, the increased development time is not recovered by the reduced grading time. Of course, an exam that is reused needs no development time.

Rapid Response Systems

Rapid response systems allow students to instantly respond to questions posed by the instructor.

Technology

Rapid response systems can be created through networked student computers or personal response devices (PRD). In either case, software on the instructor station provides a mechanism for the instructor to pose questions, receives the student responses, stores the responses in a database, and produces reports that analyze student responses. The instructor can synthesize questions off line (most systems allow this portion of the software to be loaded on faculty computers so that questions can be created at any time and outside of the classroom). These questions are downloaded to the instructor computer so that they can be electronically displayed during the class period. Typically, spur-of-the-moment questions can be handled where the instructor voices or writes the questions and possible responses. If the student computers are networked, the students provide responses using their mouse/keyboard. Personal response devices operate much like remote controls and send a signal (typically infrared) to receivers connected to the instructor station. Each student has a PRD that sends a unique identifier as well as the response to the receiver. When the response has been received at the instructor station, a visual signal is relayed to the student (frequently a matrix of identifiers displayed with the question that changes color or flashes to indicate receipt of the student response). Students can change their response by sending a new response. Timed responses (clock displayed) can be required. Recently developed systems allow questions to be posed on a slide in PowerPoint so that the instructor does not have to switch from PowerPoint to proprietary software to pose a question.

Student Perspective

Once students become comfortable with the technology, they respond favorably to the active learning environment that is created by the rapid response systems. The technology is user-friendly so the induction period is very short (2 or 3 class periods).

Instructor Perspective

There are many positive aspects to rapid response systems[2]. Each class can be initiated with a question or two that covers the assigned reading materials/homework. In addition to assessing the level of preparation, these questions can be used to poll attendance/tardiness. Questions interspersed throughout the class period assess level of understanding and provide insight into the effectiveness of the presentation. Responses are rapid so that each question can be posed, polled, and responses displayed in a few minutes (especially if the instructor does not have to change software). Immediate feedback is provided to the student.

The construction of effective questions generally requires additional class development time. If they are embedded in a PowerPoint presentation, then development time is required for both the questions and the presentation. Questions spaced throughout the class period require the student to switch from passive to active learning. This helps to avoid many of the pitfalls to lectures based around PowerPoint presentations[5]. Class periods tend to have more structure (so courses being taught for the first time are more difficult to employ) although the ability to construct spur-of-the-moment questions does provide flexibility.

Quizzes cannot be individualized using personal response devices. In a large, crowded classroom where students must make line-of-sight connection between their personal response device and the receiver, it is easy for students to share their responses with neighbors. This can be used to advantage by allowing students to discuss the question before making responses.

Experience

Attempts to use courseware systems (for example, BlackBoard or WebCT) for rapid responses were unsatisfactory. These systems require networked computers and internet login times greatly exceeded the class time allotted for rapid response quizzes. Additionally, occasional problems with the network restricted or delayed student access so that developed quizzes could not be administered. Paper-based quizzes are dependable. However, the class time required to administer them is a function of class size, they have to be graded outside of class, and feedback to the student is delayed. The courseware systems do allow immediate feedback to students and quizzes can be individualized but do not allow the instructor to immediately display responses. The courseware quizzes have been used effectively to identify students at risk. For example, scores on a quiz requiring the identification of the points, curves, and regions on PT and PV diagrams for a pure substance generally serve as good predictors of performance in thermodynamics and physical property estimation courses.

Experience with personal response devices, especially those that allow questions to be posed in PowerPoint, is encouraging. The devices that have been tested allow responses only to multiple choice and true/false types of questions (typically 6 to 10 possible responses). The class time required for administering quizzes is independent of class size (so this is very desirable for large classes). Colleagues in the Chemistry Department obtained a 10 point improvement (from roughly 70 to 80) in the average scores on exams (100 points maximum) in the freshman chemistry course (approximately 150 students in each class) from one year (no PRD) to the next (PRD with companion textbook). Systems that allow more complex responses (numerical or graphical) are desirable for engineering courses. A system that allows textual responses (comments, short answers, and/or numbers) is currently being tested in courses on chemical engineering economics and chemical process safety. Grading of the responses must be done off-line and it is more difficult to provide rapid feedback.

Summary

Personal response devices create an active learning environment in the classroom. The rapid response allowed by these systems minimizes class time and maximizes feedback to both instructor and student.

Exams during class time that are administered using courseware systems require large class sizes and/or multiple uses to be effective. Uses outside the classroom (asynchronous learning) have the greatest potential.

Exams where students have a suite of routines at their disposal allow for test problems that are grander and closer to the scale that will be encountered in the workplace. The routines in the suite can be developed across the curriculum integrating and reinforcing content from courses.

Bibliography

1. Book, N. L., D. K. Ludlow, and O. C. Sitton, “Development and Implementation of a Computer-Based Learning System in Chemical Engineering,” Proceedings of the 2001 American Society for Engineering Education Annual Conference & Exposition, ASEE, (2001).

2. Falconer, J. L., “Use of ConcepTests and Instant Feedback in Thermodynamics,” Chemical Engineering Education, 38, 64, (2004).

3. Felder, R. M., and R. Brent, “FAQS. II: Active Learning versus Covering the Syllabus and Dealing with Large Classes,” Chemical Engineering Education, 33, 276, (1999).

4. Felder, R. M., and R. Brent, “FAQS. V: Designing Fair Tests,” Chemical Engineering Education, 36, 204, (2002).

5. Felder, R. M., and R. Brent, “Death by PowerPoint,” Chemical Engineering Education, 39, 28, (2005).

6. Prince, M., “Does Active Learning Work? A Review of the Research,” J. of Engineering Education, 93, 223, (2004).

NEIL BOOK
Dr. Neil L. Book is an Associate Professor in the Chemical and Biological Engineering department at the University of Missouri-Rolla. He earned a BSChE from the University of Missouri-Rolla in 1972, an MSChE from the University of Colorado in 1974 and a PhD in Chemical Engineering from the University of Colorado in 1976. Professor Book teaches a wide variety of courses including material balances, thermodynamics and chemical process safety.

OLIVER SITTON
Dr. Oliver C. Sitton is an Associate Professor in the Chemical and Biological Engineering department at the University of Missouri-Rolla. He earned a BSChE (1975), MSChE (1976) and PhD (1978) degrees from the University of Missouri - Rolla. He was a Senior Research Engineer for the Monsanto Company when he joined the faculty at the University of Missouri–Rolla. Professor Sitton currently teaches courses in thermodynamics, separations and biotechnology.