J·T·L·A

The Journal of Technology, Learning, and Assessment

Volume 5, Number 5 · December 2006

Comprehensive Assessment of a Software Development Project for Engineering Instruction

Richard H. Hall, Timothy A. Philpot &
Nancy Hubing

www.jtla.org

A publication of the Technology and Assessment Study Collaborative
Caroline A. & Peter S. Lynch School of Education, Boston College

Volume 5, Number 5

Comprehensive Assessment of a Software Development Project for Engineering Instruction

Richard H. Hall, Timothy A. Philpot & Nancy Hubing

Editor: Michael Russell
russelmh@bc.edu
Technology and Assessment Study Collaborative
Lynch School of Education, Boston College
Chestnut Hill, MA 02467

Copy Editor: Kevon R. Tucker-Seeley
Design: Thomas Hoffmann
Layout: Aimee Levy

JTLA is a free on-line journal, published by the Technology and Assessment Study Collaborative, Caroline A. & Peter S. Lynch School of Education, Boston College.

Preferred citation:

Hall, R. H., Philpot, T. A., & Hubing, N. (2006). Comprehensive assessment of a software development project for engineering instruction. *Journal of Technology, Learning, and Assessment, 5*(5). Retrieved [date] from http://www.jtla.org.

Abstract:

This paper reviews a series of formative assessment studies that were conducted to inform and evaluate a large-scale instructional software development project at the University of Missouri – Rolla (UMR). The three-year project, entitled "Taking the Next Step in Engineering Education: Integrating Educational Software and Active Learning," was funded by the U.S. Department of Education Fund for the Improvement of Post Secondary Education (FIPSE). The assessment was carried out under the auspices of UMR's Laboratory for Information Technology Evaluation (LITE) and guided by the LITE model for evaluation of learning technologies. The fundamental premise of the model is that evaluation should consist of the triangulation of multiple research methodologies and measurement tools. Five representative evaluation studies, consisting of eight experiments, are presented here. The studies range from basic experimentation and usability testing to applied research conducted within the classroom as well as a multi-national cross-cultural applied dissemination survey conducted during the last semester of the project. This paper demonstrates that the LITE model can be an effective tool for guiding a comprehensive evaluation program. In addition, the research findings provide evidence that the instructional multimedia developed in this project can have a substantial positive impact in enhancing fundamental engineering classes.

Comprehensive Assessment of a Software Development Project for Engineering Instruction

Richard H. Hall
Timothy A. Philpot
Nancy Hubing
University of Missouri – Rolla

Introduction

The "Taking the Next Step in Engineering Education" Project

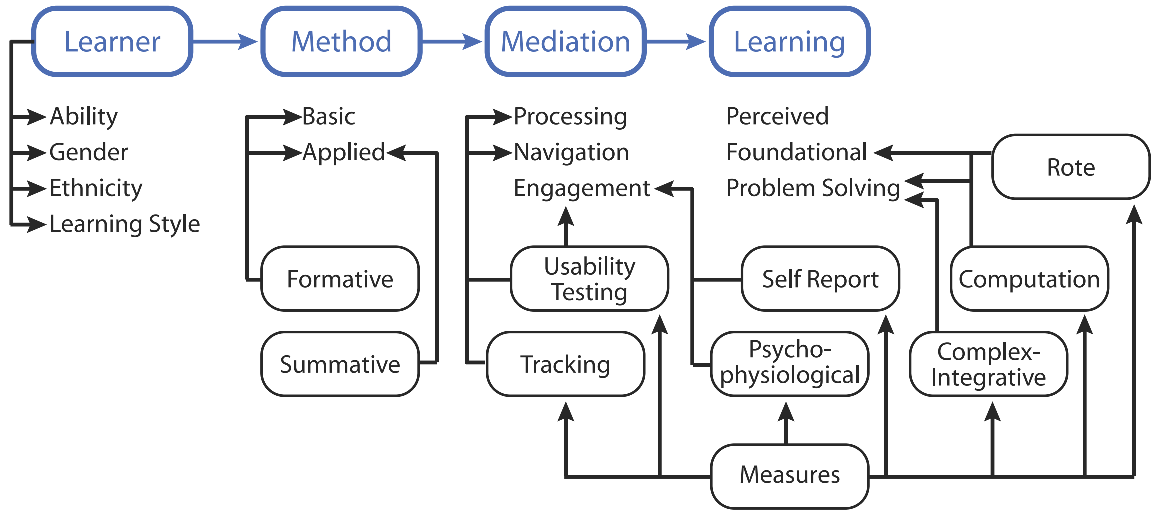
The University of Missouri – Rolla (UMR) recently completed a comprehensive three-year project sponsored by the U.S. Department of Education's FIPSE[1] program (Flori et al., 2002; Hall et al., 2002; Hubing et al., 2002; Philpot et al., 2002; Philpot et al., 2003a) entitled "Taking the Next Step in Engineering Education: Integrating Educational Software and Active Learning" (P116B000100). The project focused on enhancing learning in three core engineering courses through the development and implementation of a suite of animated and interactive courseware modules. The three core courses included Statics, Dynamics and Mechanics of Materials. More than 250 computer-based instructional examples, problems, games, and theory modules were developed. In addition, an extensive homework database administration system was created.[2] The project evaluation was carried out under the auspices of UMR's *Laboratory for Information Technology Evaluation* (LITE). Several assessment and evaluation studies have been carried out from the beginning of the project, representing diverse methodologies and measurement tools, based on an assessment framework described below.

LITE Assessment Model

The Laboratory for Information Technology Evaluation (LITE) at the University of Missouri – Rolla has developed a comprehensive framework for the assessment of learning-technology projects. This assessment model has evolved over the course of several projects that focused on the evaluation of software tools for engineering and science education (Eller, Hall, & Watkins, 2001; Hall et al., 2002; Hall, Watkins, Davis, Belarbi, & Chandrashekhara, 2001). A fundamental assumption of the model is that conclusions and recommendations should be based on the triangulation of information gleaned from multiple methodological and measurement tools.

Figure 1 (next page) is a graphical depiction of the LITE assessment model. As Figure 1 depicts, the model uses multiple data sources to examine the connection between the learner, the method used to develop learning, the tool(s) through which learning is mediated, and the resulting learning. During the early stages of an evaluation study, formative assessments of the method are conducted with the aim of improving the method. During later stages, assessment of the method focuses on summative effects. To document and inform the mediation process, learner-use is tracked and usability testing is conducted. In addition, the extent to which the mediation process engages the learner is examined through self-reports and by taking psycho-physiological measures such as tracking eye movements. During the initial stages of development, learning is assessed through self-reports. During later stages of development, more formal cognitive measures are collected. Depending upon the learning goals, these measures focus on complex-integrative processes, computation, and/or rote memory.

Figure 1: LITE Assessment Model



The model is meant a) to serve as a guide for large-scale assessment projects; and b) to provide some context for individual experiments. Over the course of a large-scale project, multiple studies will occur. Each study focuses on a subset of the aspects included in the model and is intended to refine the courseware modules or to examine the effects of the module on the learner. Consequently, any single study does not include all aspects of the model. However, over the course of an entire project, assessment activities will include many parts of the model. This paper reviews five studies that employed a total of eight experiments conducted over the course of the UMR project, *Taking the Next Step in Engineering Education*. These studies were selected to represent various methodological approaches. The experiments to be reviewed are presented in Table 1 (next page). Before presenting each of the five studies, we elaborate on the key elements of the LITE assessment model.

Table 1: Summary of Five Studies to be Reviewed

Methodology	Purpose	Conducted
Basic Experimentation (2 experiments)	Impact of interactive feedback	Fall 2001
Usability Testing	Impact of interactive feedback	Spring 2002
Applied: Class Context (3 experiments)	Using games to teach calculation procedures	Fall 2002 & Spring 2003
Applied: Class Context	Software in support of traditional lecture	Fall 2002
Applied: External Dissemination	Impact in multiple external contexts	Fall 2003

Learner Variables

Learner variables are included in many experiments. Ability, as measured by students' grades or more basic measures such as visual skills, can often serve as covariates. Visual skills include measures such as 3D rotation and 3D manipulation (Ekstrom, French, & Harman, 1979) Also, learner variables such as gender, ethnicity, and learning style often interact with factors under examination. For example, the variables of motivation and engagement explain a great deal of variance among the small number of student participants in the usability study discussed below. In the multi-national survey discussed at the end of the paper, the nationality of the school (U.S. vs. international) had a significant impact in mediating the effect of the software.

Method

In terms of methodology, projects generally proceed from basic experimentation involving manipulation of specific variables to more comprehensive and applied experiments carried out in the context of ongoing classes. This paper presents a) two examples of basic research experiments where students were randomly assigned to groups within a controlled testing environment; b) a usability study in which a small group of students were examined in great detail as they used sample instructional software modules; c) two applied research studies conducted within the context of ongoing classes; and d) a survey that facilitated more comprehensive generalization across learning contexts, nationalities, and cultures.

Mediation

The mediation category refers to ways in which students use the technologies. Usability testing can be instrumental in this regard, in that it can provide important insight into the results of a larger scale experimental investigation. Tracking methodologies and psychophysiological measures are also valuable to capture aspects of the software utilization process in some projects. Though not used in the experiments reported here, these types of tools have been successfully used to assess the impact of learning environments in other projects. More specifically, we have used *galvinic skin response* (GSR) and *heart rate* – both indicative of autonomic/sympathetic nervous system arousal – to represent affect and learner fatigue (Hall et al., 2004b; Wilfred et al., 2004).

Learning

The UMR LITE model has been applied to a number of projects, primarily in science and engineering fields. Within this context, the focus of objective outcomes is mainly foundational knowledge and problem solving techniques. Attitudinal and perceived outcomes are often considered as well.

Measures

Many measurement approaches are included (as depicted in Figure 1 on page 6) and most are used throughout the course of a large-scale project such as the one described here.

Study 1: Effect of Interactive Feedback (Basic Experimentation)

Rationale

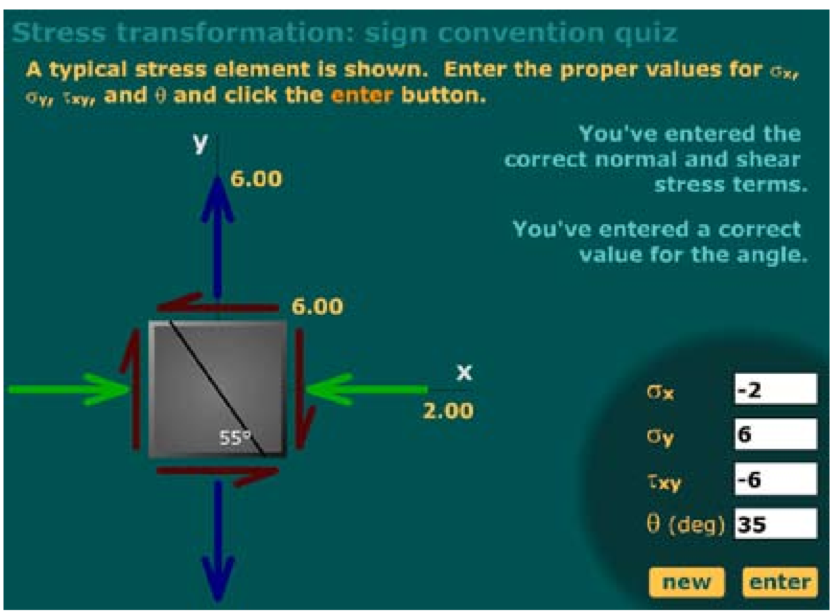
One of the first studies conducted within this project was an examination of the role of interactive feedback as a component of the software modules (Hall et al., 2002). There is evidence that providing learners with feedback as a component of instructional software systems can significantly enhance learning (Barron, 1996), particularly for novice learners (Dillon & Gabbard, 1998; Shin, Schallert, & Savenye, 1994). However, incorporating interactive feedback into instructional modules can often be resource and labor intensive since it requires the production of significantly more content in the form of problems and solutions. Further, developers are required to create a more complex and interactive learning environment. Therefore, it was important in this project to establish, from the outset, whether adding these feedback components would be effective enough to warrant the extra resources required.

Experimental Procedures

Two experiments were conducted using two courseware modules. In both of these experiments, students were recruited from ongoing Mechanics of Materials classes to study courseware modules pertaining to course concepts that had not yet been discussed in class. In both experiments, students completed these experiments in computer labs in controlled conditions, observed by experimenters. After studying the target modules for thirty minutes, all students completed brief tests on the subject matter and responded to subjective quantitative questions.

In both experiments, students were randomly assigned to one of two groups, either a *feedback* group or a *non-feedback* group. The study module consisted of animated movies that explained the relevant theory and specific aspects important in successfully applying the theory. For students in the feedback group, the study module also included a self-quiz component. An example of a feedback component is shown in Figure 2. Students in the non-feedback group were not given access to the self-quiz. The self-quiz component consisted of an interactive game feature in which the student was asked to respond to a series of questions, applying the information presented in the study module.

Figure 2: Typical Interactive Feedback of Software Feature



In experiment 1, students studied information pertaining to stress transformation equations and in experiment 2, they studied information on the Mohr's circle procedure for stress transformations. Thirty-five students participated in experiment 1 (15 in the feedback group and 10 in the non-feedback group). Twenty-eight students participated in experiment 2 (16 in the feedback group and 12 in the non-feedback group).[3]

At the completion of each experiment, all students completed a series of short-answer questions, which included completion, multiple-choice, and true/false (15 questions in study 1 and 10 in study 2). In terms of the LITE assessment model above, these questions were principally measures of foundational knowledge. In addition, they responded to the following four questions using a 9-point Likert scale (1 = *strongly disagree* and 9 = *strongly agree*).

a. I learned a great deal of information from the multimedia tutorials. *(learning)*

b. I found the multimedia tutorials to be very motivational. *(motivation)*

c. The web tutorials were effective in aiding me in recognizing how much I know and don't know about this topic. *(metacognition)*

d. I found the navigational scheme for the web tutorials to be logical and easy to navigate. *(usability-navigation)*

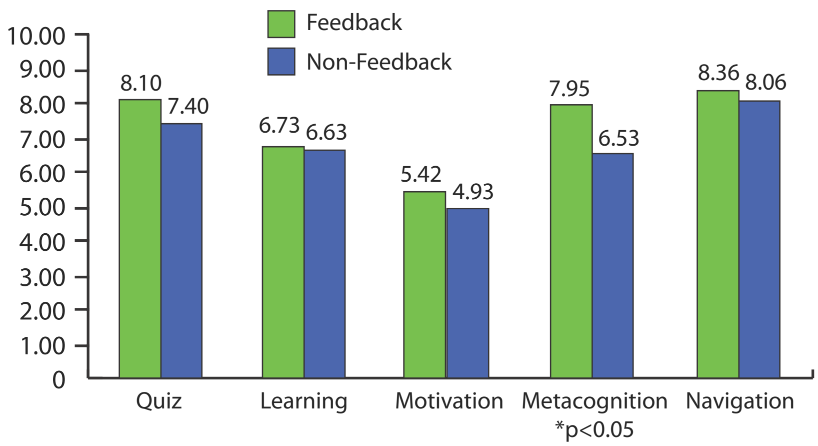
They were also asked to provide open-ended comments to support their responses to each of these items and to provide additional comments on the software effectiveness and/or suggestions for improvement.

Results

Quantitative Analysis

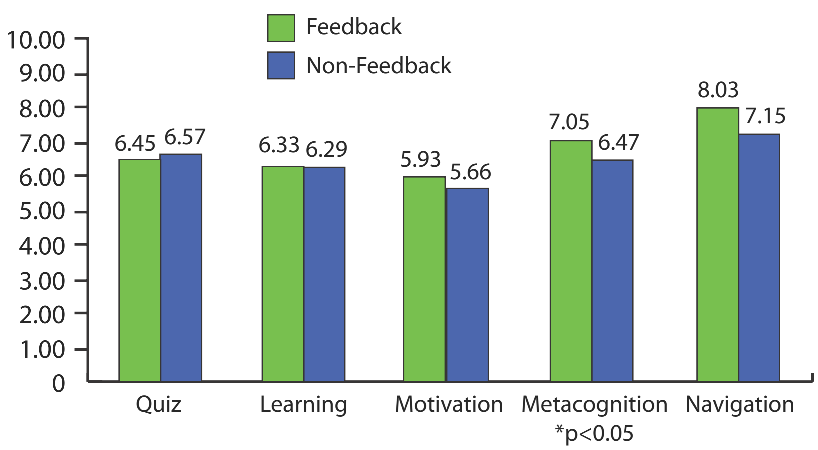
The quantitative analyses for each experiment consisted of a series of five between-subject *t*-tests with group (Feedback vs. Non-Feedback) serving as the independent variable. Test- and Likert-responses to each of the four subjective questions served as the dependent variables. The means for experiment 1 and 2 are depicted in Figures 3 and 4 (next page). In all cases, the mean for the feedback group was higher. The only statistically significant mean difference, however, was for the metacognition question in experiments 1 ($p<0.05$) and 2 ($p<0.05$).

Figure 3: Experiment 1 Mean Test and Subjective Rating Scores as a Function of Experimental Group



Note: y axis represents agree-disagree for subjective questions and percentage correct for test (e.g., 10.00 = 100%)

Figure 4: Experiment 2 Mean Test and Subjective Rating Scores as a Function of Experimental Group



Note: y axis represents agree-disagree for subjective questions and percentage correct for test (e.g., 10.00 = 100%)

Qualitative Analysis

Students' open-ended responses to each of the questions were examined with special emphasis on identifying differences between groups. Several themes emerged from the qualitative data. Three of the major themes and representative comments are provided. A more detailed description of this analysis is provided elsewhere (see Hall et al., 2002).

1. Students in the feedback groups were more positive about their perceived learning and especially about the degree to which it made them aware of their knowledge-level (metacognition).
   - *(feedback)*: “Made me realize I didn’t know anything.”
   - *(non-feedback)*: “I am not sure what I did right and what I did wrong.”
2. The modules could serve as a good supplement to class, especially for a student who was having difficulty.
   - “[I] have some previous knowledge of the material but the program offered some insightful hints that will prove very beneficial as supplemental material.”
   - “If I were unable to comprehend what was going on in class this would be a good place to seek info.”
3. The animations, graphics, and self-paced nature of the modules added a dimension that aided learning in a way that other methods could not accomplish.
   - “…it allowed you to go at your own pace.”
   - “Pictures and “movies” make it more interesting.”
   - “Nice to learn at your own pace and to go back and re-read.”

Conclusions

In the quantitative results, the feedback and non-feedback groups differed significantly on only one of the ten outcome measures. Those in the feedback group in experiment 1 rated their software significantly higher in the item referring to impact on metacognitive awareness. Despite this lack of significant effects, it is important to note that students in the feedback group scored higher on both quizzes and rated the software more positive on all subjective measures, so the means were consistently in a direction favoring the software with feedback. In addition, student comments for those in the feedback groups were more positive, and specifically pointed to advantages associated with the interactive components of the software. For this reason, it was concluded that interactive feedback should be an important part of subsequent software design. We also suspected, based on informal observation during the experiments, that students were using the software in very diverse ways, so that the impact of the feedback components may be strongly mediated by the way in which the software was used. This suspicion led to a more detailed analysis of students’ use of the software in the usability study that follows.

Study 2: Detailed Analysis of Software Use (Usability Testing)

Rationale

To gain further insight into how students were using the software, a usability study was conducted using the stress transformation module with feedback components (i.e., experiment 1). Complete details of the experimental procedures and results of this study are presented in Hall et al. (2005). The study focused primarily on a) student navigation patterns within the software including allocation of time spent on different features; and b) student engagement with the learning activities incorporated in the instructional software.

Synopsis

Students in the study could be grouped into two archetypes: the *engaged student* and the *non-engaged student*. The engaged student proceeded systematically, focused on aspects of the study module that involved the integration of graphics and equations, and completed the quiz at the end of the session. While the non-engaged student also proceeded sequentially, he or she spent very little time – often spending only seconds – on a scene. Although the interactive movies were more successful at holding the attention of the non-engaged user, they were largely unsuccessful at achieving the desired educational performance improvements because the non-engaged student skipped over the necessary theoretical background presented for the topic.

The usability study suggested some explanations for the results obtained in the basic research experiment. While the feedback group scored higher on all measures in the basic research experiment, these higher scores were not statistically significant (as might have been expected). One striking insight revealed by the usability study, however, was that most students did not reach those scenes that included the interactive feedback features. If one assumes that the students in the basic research study used the software in a similar manner, then a number of students counted in the feedback group did not actually use the feedback features. Consequently, positive effects of feedback might have been obscured or underestimated. The usability study also highlighted the dramatic variance of motivation and engagement among students. This variance may have created a large degree of within-group error in the experiment, again contributing to the non-significant effects.

Study 3: Using Games to Teach Statics (Applied: Class Context)

Rationale

Fundamental engineering courses such as Statics (one of the courses targeted in the project) seek to develop the student's ability to analyze basic engineering machines, mechanisms, and structures and to determine the information necessary to properly design these configurations. Fundamental calculations such as *centroids* and *area moments of inertia* are building blocks that students must employ to solve problems and develop designs in a variety of situations.

It is often assumed that repetition leads to proficiency; however, few students relish working dozens of problems on a particular topic. To make the learning process more enjoyable, repetition and drill on a specific topic can be encapsulated in a game context. Games have been found to be an effective method of increasing motivation, enjoyment and learning for many math and science topics that may otherwise seem boring to students (Amory, 2002; Jih, 2001; Smith, 1999; Westbrook & Braithwaite, 2001). There is evidence that such tools can be a particularly powerful for learning engineering concepts where visualization is important, such as engineering graphics (Crown, 1999). Through the challenge of the game, the student can receive the benefits of repetition without the sense of labor that they might feel otherwise. A game context provides students with a structure for learning and permits students to develop their skills at their own pace in a non-judgmental but competitive and often fun environment. Since the computer is a medium that is well suited for repetitive processes and for numeric calculations, computer-based games focused on specific calculation processes offer great potential as a new (or perhaps updated) type of learning tool for engineering mechanics courses. Therefore, a number of interactive games were developed as part of this project (Philpot et al., 2003b), and three experiments were conducted to examine the impact of this software within the context of ongoing classes. Two specific game modules, *The Centroids Game – Learning the Ropes* and *The Moment of Inertia Game – Starting from Square One*, were evaluated.

Experimental Procedures

Overview

In the 2002 academic year, the effectiveness of two games as teaching tools was assessed with two undergraduate Statics classes at UMR. In lieu of the customary lecture period, students were taken to a computer lab where a computer was available for each student. During the preceding class period, students had been introduced to the topic. At the start of the assessment class period, students were given a two-minute introduction to the relevant calculation procedure and then given approximately 40 minutes to play the game at their own pace. An instructor was present in the computer lab to answer questions about the topic and to clarify game procedures. Near the end of the class period, students were stopped and asked to complete a survey.

The activities of the spring 2003 class were more elaborate in that a second game was included and additional outcome measures were used, as summarized in Table 2 (next page). In the fall of 2002 the activity occurred during one class, and students used "The Centroids Game". In the spring of 2003 the activity occurred during two classes, one used "The Centroids Game" and the other used the "Moments of Inertia Game". In the fall the only outcome measure was a survey. In the spring of 2003, the outcome measure included this survey with some additional questions as indicated in Table 2. In addition, in spring 2003, a quiz over the material for both games was administered in a subsequent class period. These same quizzes were completed by students in "control" sections of the statics classes, where students got a traditional lecture covering the same topics that were covered in the games.

Table 2: Overview of Experiments for Using Games to Teach Statics

Class	Activity	Outcome Measures
Fall 2002	Played *Centroids Game* in lieu of lecture	Completed survey
Spring 2003	Played *Centroids Game* in lieu of lecture	Completed survey • included a question about textbook's presentation of topic Completed problem solving quiz after playing game. These results were compared with control sections.
Spring 2003	Played *Moments of Inertia Game* in lieu of lecture	Completed survey • included a question about textbook's presentation of topic Completed problem solving quiz after playing game. These results were compared with control sections.

Survey Items

The survey consisted of a number of 11 to 12 items in which students responded to Likert-type statements using a 9-point scale where 1 = *strongly disagree* and 9 = *strongly agree*. The questionnaire used in assessing *The Centroids Game* is shown below, and a similar questionnaire was used for *The Moment of Inertia Game*.

1. After using *The Centroids Game*, I felt confident in my ability to calculate centroids for composite bodies.
2. After using *The Centroids Game*, I was able to visualize the procedure for calculating centroids.
3. After using *The Centroids Game*, I understood which cross-sectional dimensions to include in my calculations when working a centroids problem.
4. *The Centroids Game* helped me to recognize how much I know and don't know about the procedure for calculating centroids.
5. I found *The Centroids Game* to be motivational concerning the procedure for calculating centroids.
6. I liked playing a game to help me get better at calculating centroids.
7. I learned a great deal about the procedure for calculating centroids from *The Centroids Game*.
8. I learned a great deal about the procedure for calculating centroids from my Statics textbook *(spring 2003 only)*.
9. I thought the time spent playing *The Centroids Game* was a worthwhile use of my study time.
10. The procedure for playing *The Centroids Game* was easy to understand.
11. The number of questions and the number of rounds used in *The Centroids Game* seemed about right to me.
12. Give your overall evaluation of *The Centroids Game* on the procedure for calculating Centroids, using the 1…9 scale, with 1 being *very poor* and 9 being *outstanding*.

Results from Game Assessment

Survey

For both game assessments, the survey results are summarized in Table 3 for both fall and spring Statics classes.

Table 3: Mean Responses To Game Questionnaire Items

(Scale: 1 = *strongly disagree*, 9 = *strongly agree*)	***Centroids Game***		***Moments of Inertia Game***
Survey Statements	**Fall 2002 class** (n=27)	**Spring 2003 class** (n=23)	**Spring 2003 class** (n=23)
1. Confidence in ability to perform the calculation	8.00	8.61	8.17
2. Visualization of calculation procedure	8.50	8.35	8.17
3. Understanding necessary cross-sectional dimensions	8.38	8.13	8.35
4. Recognize how much I know and don't know	7.96	7.30	7.83
5. Motivation	7.75	7.39	7.52
6. I liked playing the game	8.21	8.04	8.09
7. I learned a great deal about the procedure from game	7.75	7.35	7.87
8. I learned a great deal about the procedure from textbook	N/A	3.17	2.17
9. Worthwhile use of study time	7.83	7.52	7.78
10. Game procedure was easy to understand	8.58	8.70	8.30
11. Number of game questions about right	7.92	7.70	7.96
12. Overall evaluation of game	8.38	8.04	8.13

Students were also asked to comment on their overall evaluation of the games. Their comments were consistently positive, as characterized by representative comments such as:

- "I think it's a much easier way to do homework and I did 10 times as many problems as I normally do. I have this concept down very well."
- "Easy to understand. Helps to teach by progression… easy-to-hard."
- "It showed me everything I didn't know and allowed me to learn."
- "Most fun I've had while learning in a long time."

Two open-ended questions were included in *The Moment of Inertia Game* questionnaire to explore students' perceptions of instructional software in general, particularly after having just had an experience with the game.

1. Are there things you really dislike about instructional software? Do you think software is a waste of time or just no-good? What really bugs you about this stuff?
2. Are there things that you really like about instructional software? Have you tried instructional software? Are there any programs that you think are really good?

Students' responses to these open-ended questions were combined and categorized according to themes. Two themes that emerged from students' comments and some representative student comments are presented below:

Theme 1: Students felt very positive about instructional technology in general and *The Moment of Inertia Game* in particular. The principle advantages cited were a) immediate feedback; b) aid in visualization; and c) increase in motivation and enjoyment.

a. *Immediate Feedback*
   - "It's a great way to do homework and it gives you the correct answers right away – that way I KNOW I'm doing it right, every time."

b. *Aid in Visualization*
   - "If it is good visually and outlines steps, it can be very helpful."

c. *Increase motivation and enjoyment*
   - "I enjoyed it thoroughly. I like the competitive view, try to get the better score."

Theme 2: It is important that instructional software is integrated with the class and instructor.

- "I like it in class if the prof [sic] is walking around helping."
- "I think it [instructional technology] is a good idea, but must be assigned in class."

Comparison with Textbook

To compare student ratings of the game with their textbook, survey statement 8 was added to the spring 2003 questionnaire. The responses to statement 7 were compared with the responses to statement 8, using a within-subjects t-test. In both *The Centroids Game* assessment [$t(22)=10.098$, $p<0.001$] and *The Moment of Inertia Game* assessment [$t(22)=6.86$, $p<0.001$], this test was statistically significant. On a scale of 9, students' agreement with the statement that they learned a great deal from the game was more than two to four times higher than their rating of the same statement for the textbook.

Impact of Game on Learning

In the spring 2003 assessment experiment, a single-problem quiz was administered to students at the end of the class period following completion of *The Centroids Game* exercise. To serve as a control group, students in four additional Statics sections were also given the same quiz. None of the students in the control groups had exposure to *The Centroids Game*. Students in the control group took the quiz either one or two class periods after the topic of centroids of composite areas had been discussed in lecture. Students in the control group, therefore, had some opportunity to review notes and work assigned homework problems in the days following their in-class exposure to this topic. Students in both the experimental and control groups, however, were not told about the quiz before the class period in which it was administered.

Students were asked to compute the vertical location of the centroid for a double-tee shape (Figure 5). For the purpose of this experiment, quizzes were marked either *100% correct* or *incorrect* (for any type of error). The results of the quiz are shown in Table 4 .

Figure 5: Double-Tee Shape Used for the Centroid Calculation Quiz

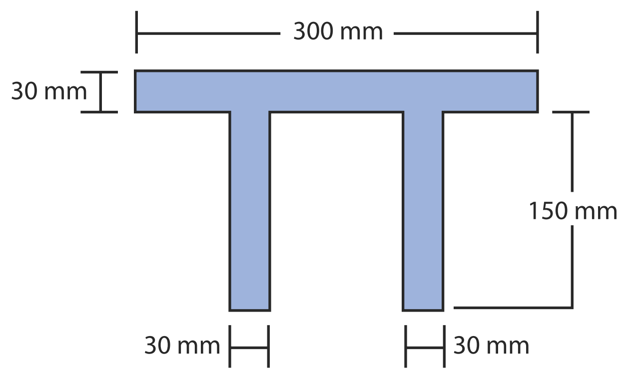


Table 4: Quiz Results from *The Centroids Game*

***The Centroids Game* Quiz Results**	**Total Number of Students**	**Correct Responses**	**Incorrect Responses**
Students who played *The Centroids Game*	23	23	0
Students in control group	91	55	36

$x^2(1) = 10.50, p < 0.01$

Note: disparity in control vs. experimental group size is due to the fact that there was only one experimental section of the class and multiple control sections.

An analysis was conducted to compare problem scores for students in the test group with those in the control group. Since these data consisted of dichotomous data, a Pearson Chi-Square was computed to test for significant differences in the distribution of correct and incorrect responses between the groups (test vs. control). This test was statistically significant, indicating that those in *The Centroids Game* group performed significantly better on the quiz problem than those in the control group.

To compare students who used *The Moment of Inertia Game* to those who learned in a traditional lecture, the test class was compared with a control group of three Statics classes that had not used the game. Students in both the test group and the control group were given a brief quiz at the beginning of the class period after moments of inertia for composite areas had been presented, either by the game or in a lecture. Students in both groups, therefore, had some opportunity to review notes and work assigned homework problems in the two days following their in-class exposure to this topic. This assessment differed from *The Centroids Game* assessment in that students in both the test and control groups were told in advance about the upcoming quiz. Students were asked to compute the area moments of inertia I_x and I_y for a tee-shape about both the horizontal and vertical centroidal axes. The vertical location of the centroid was explicitly given. Quizzes were graded and grouped into three categories: *100% correct* if the student correctly determined both I_x and I_y, *partially correct* if the student correctly determined either I_x or I_y or if they simply made a calculation error while performing the correct procedure, or *100% incorrect* if the student did not demonstrate understanding of the proper calculation procedure. The results of the quiz are shown in Table 5.

Table 5: Quiz Results from *The Moment of Inertia Game*

***Moment of Inertia* Quiz Results**	**Total Number of Students**	**100% Correct Responses**	**Partially Correct Responses**	**100% Incorrect Responses**
Students who played *The Moment of Inertia Game*	23	20 (87%)	2 (9%)	1 (4%)
Students in control group	55	26 (47%)	14 (25%)	15 (27%)

$x^2 = 10.71$, $p < 0.01$

Note: disparity in control vs. experimental group size is due to the fact that there was only one experimental section of the class and multiple control sections.

A Pearson Chi-square analysis was again used to test for statistical significance between the distributions of scores for those in the test group versus those in the control group. The Chi-square test was statistically significant, $X^2(2)=10.71$, $p<0.01$. The frequencies displayed in Table 5 indicate that the significant Chi-square was due to the fact that virtually all of the students in the test group scored 100% correct on the quiz while over half of the students in the control group received *partially correct* or *100% incorrect*.

Conclusions

This assessment was one of the strongest and most positive with respect to indicating the software's effectiveness. Students' responses to all questionnaire items were very positive in comparison to the scale midpoint of 4.5 (Table 3 on page 18). In addition, students rated the game as significantly and overwhelmingly more effective than the textbook. Finally, student scores on objective measures of their learning were significantly higher for those who used the games in two different experiments.

The consistent and positive findings in this study are also indicative of improvements in the software design that were informed by initial research, and reflect the improving nature of the learning technologies informed by iterative evaluation.

Study 4: Instructional Multimedia as Support for a Traditional Lecture in Statics (Applied: Class Context)

Rationale

Studies conducted for the project such as the games experiments presented above provide strong evidence that the software can serve as an effective alternative to traditional lecture. This is consistent with other research (Merino & Abel, 2003; Steif & Naples, 2003), which indicated that these tools can be effective substitutes, so long as they are well designed and used in the proper context (Babu, Suni, & Rasmussen, 1998; Kadiyala & Crynes, 2000). However, very few investigations have examined these tools as enhancements for a traditional lecture. In the classroom, software instructional aids can potentially free the lecturer from writing out individual problems step-by-step on the board, which allows time for higher-level interactive discussion. In addition, these types of modules have the potential to allow for elaborate multi-dimensional dynamic visualizations, which go far beyond a two-dimensional diagram or drawing, providing the student with a much more realistic and concrete representation of spatial concepts (Crown, 2001; Sorby & Baartmans, 2000). Therefore, an assessment experiment was conducted which focused on the use of a multimedia module as support for a "traditional" classroom lecture.

Experimental Procedures

In the fall of 2002, one Statics instructor gave a lecture to two Statics classes on the topic of *truss analysis* using the method of sections. In the control class, the instructor used "traditional" tools (chalkboard) to support the lecture and presented the theory followed by example problems. The experimental class also included a discussion of theory followed by example problems, but the examples were provided in the form of multimedia modules via a computer and projector. The multimedia modules allowed the instructor time to cover the examples and to cover multiple choice questions included in the module in an interactive format with the class. The class, as a whole, determined the answers to the multiple-choice questions through discussion lead by the instructor. In both classes, the same total amount of class time was devoted to the topic. There were 24 students in the control class and 27 in the experimental class.

Students completed a quiz over the material presented in the lecture, following the lecture. This quiz consisted of three multiple-choice questions, each worth three points. After completing the quiz, students responded to the following seven questions, using a 10-point Likert scale (1 = *strongly disagree* and 10 = *strongly agree*).

1. I learned a great deal of information from today's lecture/presentation on the method of sections for analyzing trusses. *(learning)*
2. I found today's lecture/presentation helped me to better visualize how to apply the method of sections to trusses. *(visualization)*
3. I found today's lecture/presentation helped me to better understand which equation to choose when working a truss problem with the method of sections. *(understanding)*
4. Today's lecture/presentation on the method of sections helped me to recognize how much I know and don't know about this topic. *(metacognition)*
5. I found today's lecture/presentation on the method of sections for analyzing trusses to be motivational. *(motivation)*
6. Today's lecture/presentation on the method of sections helped me to see how this technique has "real world" engineering applications. *(application)*
7. Give your overall evaluation of today's lecture/presentation on the method of sections, using the 1...10 scale, with 1 being *very poor* and 10 being *outstanding*. *(overall)*

Students were also asked to provide open-ended explanations for their ratings to each of these questions.

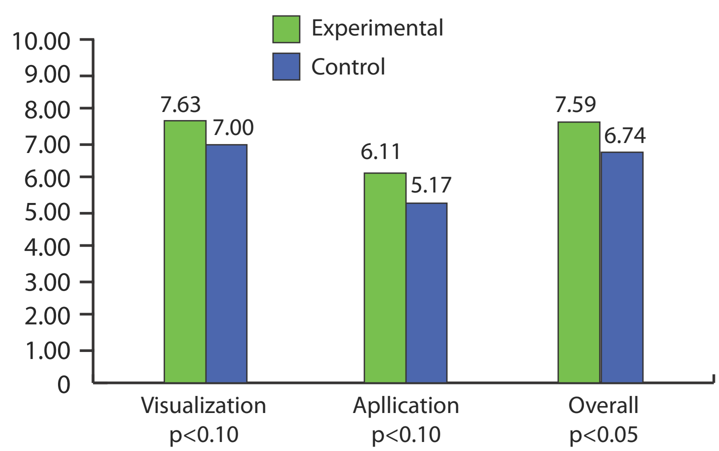
Results

Quantitative Analysis

An examination of the distribution of quiz scores indicated that there was very little variance, with the majority of students (40) receiving a perfect score (9) and the other 11 students receiving a score of 6. Given the strong ceiling effect, it was not surprising that further analysis revealed no statistical difference between the experimental and control classes.

A series of seven between-subjects *t*-tests were computed to compare the groups on questionnaire ratings. Group (experimental vs. control) served as the independent variable in each *t*-test while one of the seven questionnaire items served as the dependent variable. The means for these three items are displayed in Figure 6. Due to the formative nature of these analyses a significance level of $p<.10$ was selected.

Figure 6: Mean Ratings on Questionnaire Responses as a Function of Experimental Condition



The groups differed significantly in their ratings for Question 7 [$t(48)=2.51$, $p<0.05$], which was the rating of overall effectiveness. In addition, the differences were significant on question 2 [$t(48)=1.65$, $p<0.10$], which asked students to rate the effectiveness of the lecture to help them visualize, and question 6 [$t(48)=1.79$, $p<0.10$], which asked students to rate the applicability of the lecture to "real world" engineering. In each case, the experimental group's mean was higher.

Qualitative Analysis

Students' open-ended responses were examined, labeled, and categorized. Categories that emerged with a substantial number of comments were considered themes. Representative comments from students in the each group are presented below.

Theme 1: Those in the experimental group were more positive, overall, about the lecture.

a. *Experimental Group*

- "I learned a lot today.
- The lecture technique for me is very beneficial.
- I understood the lecture very well and it helped me understand the method of sections."

b. *Control Group*

- "[I] felt rushed through it."
- "I have already done the homework for this week."

Theme 2: Those in the experimental group found that the lecture to aid in visualization of the concepts.

a. *Experimental Group*

- "The computer animation helped me see how to section a truss…"
- "Computer diagrams are neater and easier to see what is going on."
- "The animation made it easy to see how the truss is broken apart and analyzed."

b. *Control Group* (no relevant comments on visualization)

Theme 3: Those in the experimental group found that the interactive nature of the lecture encouraged them to remain engaged.

a *Experimental Group*

- "Making us take a quiz at the end kept me awake despite how tired I am."
- "I think I may have found other lectures more helpful if I had been forced to pay attention."
- "It is more entertaining, easier to stay awake, and just as informative."

b. *Control Group*

- "To be fair, I'm not generally motivated/ excited by lectures."
- "Motivate me for what? I mean I like Statics but nothing can motivate me to want to do homework."

Theme 4: Those in the experimental group felt the lecture helped them to recognize the problem more multi-dimensionally and globally.

a. *Experimental Group*

- "On the computer it was clearer because it showed several wrong ways also."
- "Discussing the different choices made it easier to see why some are better than others."
- "This lecture helped me see which areas and clearly understand those that are harder."

b. *Control Group* (no relevant comments regarding this theme)

Conclusion

These results support the contention that the multimedia learning tools discussed here can enhance a "traditional" classroom lecture. Although the groups did not significantly differ in quiz scores, those in the experimental group gave significantly higher ratings in their overall assessment of the lecture. These results were also consistent with the qualitative analysis of students comments, in which students in the control group had a greater number of positive comments about the lecture overall, as compared to those in the control group. These results extend this previous research by indicating that the software developed as part of this project can also significantly enhance a traditional classroom lecture.

A closer examination of the data provide us with some insights into the aspects of the multimedia lecture that lead to the more positive views of the students. It appears that two basic factors lead to the effectiveness of the multimedia-enhanced lecture. First, students found the tools as important aids in visualization and second, the tools supported a more engaging and interactive lecture style. It is not surprising that the dynamic three-dimensional representations of the content allowed students to more readily visualize some of the concepts in Statics, which involve complex multi-dimensional relations. A quick examination of the modules[4] developed as part of this project provides ample evidence that they are able to provide a level of visualization not possible via traditional teaching methods (e.g., an instructor's two-dimensional drawing or verbal description). Combined with the positive subjective and qualitative findings from this applied experiment, this provides support for the contention that such tools can be particularly powerful when combined with a professors' commentary within the lecture format.

A subtle, yet more fundamental, characteristic of the multimedia-enhanced lecture was the transformational effect that it had on the instructors' lecture. A number of researchers have pointed out that learning technologies can only be effective if they are used in conjunction with good pedagogical practice, such as increasing student interactivity and engagement (Kadiyala & Crynes, 2000). Evidence from this study appears to support this contention. Since the instructor was not required to work through the examples on the board step-by-step, she was free to spend more time discussing different aspects of the problems. As a result, the students in the multimedia-enhanced class were able to view problems in a more global fashion. Further, multiple-choice questions were built into the multimedia modules, and the time saved by displaying examples contained within the instructional modules allowed time to cover these questions, which the instructor did, in an interactive discussion format. Further, the time saved by displaying examples contained within the instructional modules allowed more class time for an interactive discussion of the questions built into the multimedia modules.

Study 5: Evaluation of Software in Multiple Instructional Environments (Applied: Dissemination Evaluation)

Rationale

The purpose of this study was to extend the project evaluation by examining the effect of the software in a number of instructional settings within different types of universities and diverse cultural contexts. Professors from a number of U.S. and international colleges and universities were contacted during the fall of 2003 and offered an opportunity to access four different interactive instructional modules. The modules covered four subjects: *Centroids and Moments of Inertia, Stress Transformations, Mohr's Circle Stress Transformations*, and *Structural Analysis (Trusses and Frames)*. The professors were told that they could integrate the modules into their classes in any way they chose and were also requested to encourage students to complete Internet-based surveys over the modules used. Students of nineteen different faculty responded to the survey, representing sixteen different schools. Ten of the schools were located in the U.S. and six were located outside of the U.S.

Results from the Centroids and Moments of Inertia (CMI) survey will be discussed here. The conclusions from this survey are representative of the three other studies. For the CMI study, students from four U.S. schools and three schools from outside of the U.S. participated in the survey (Table 6, next page).

Table 6: CMI Study Participating Schools

School	Location	Number of Student Participants
Ohio State University	Columbus, Ohio	87
Texas Tech University	Lubbock, Texas,	3
Virginia Western Community College	Roanoke, Virginia	3
Monroe County Community College	Monroe, Michigan	8
Tec de Monterrey	Monterrey, Mexico	4
Instituto da Technologia da Amazonia	Amazonia, Brazil	29
University of Sarajevo	Zenica, Bosnia & Herzegovina	8
		Total: 142

Experimental Procedures

In addition to a number of general survey statements, the CMI survey included a series of six questions that asked students to rate their knowledge on an important aspect of the topic covered in the module *before* and *after* they completed the module. On each of these questions students were asked to rate their degree of agreement from 1 to 9, with 1 representing *strongly disagree* and 9 representing *strongly agree*. The six before/after questions for the CMI module were:

1 & 2 *Before/After* using the centroid and moment of inertia review, I was confident in my ability to determine the centroid location for composite shapes…

3 & 4 *Before/After* using the centroid and moment of inertia review, I was confident that I could correctly determine the moments of inertia (about both the horizontal and vertical centroidal axes) for composite shapes…

5 & 6 *Before/After* using the centroid and moment of inertia review, I was confident that I could correctly determine the moments of inertia (about both the horizontal and vertical centroidal axes) for shapes consisting of standard steel shapes.

A “before” and “after” composite score was created for the survey by averaging the responses to the three *before* and three *after* questions, respectively.

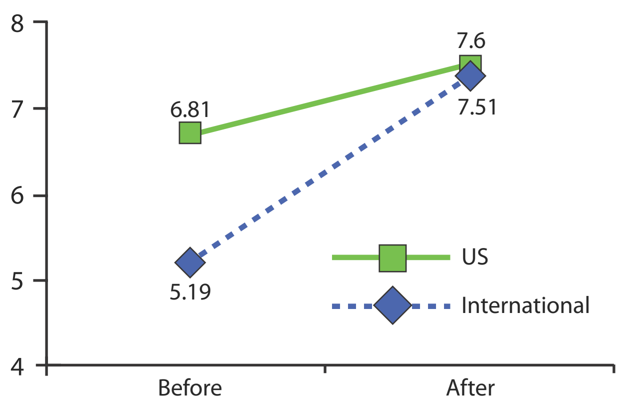
Comparison of Pre vs. Post Knowledge Ratings

Analysis

In order to examine the effect of location (U.S. vs. International), a location variable was created with students from U.S. schools classified as *U.S.* and students from schools outside the U.S. classified as *International*. A two-way mixed analysis of variance (ANOVA) was computed with *gain* (before vs. after) serving as a within-subject independent variable, *location* (U.S. vs. International) serving as a between-subject independent variable, and *ratings* serving as the dependent variable.

In the CMI analysis, significant effects were found for Gain $F(1,132)=90.17$, $p<0.001$; Location $F(1,132)=12.36$, $p<0.01$; and Gain × Location $F(1,132)=21.52$, $p<0.001$. Figure 7 displays the means associated with this ANOVA.

Figure 7: Ratings as a Function of Gain and Location for the CMI Questionnaire



Interpretation

The significant main effects for Gain and Location indicate that, overall: 1) students rated their knowledge after using the software as significantly higher than before; and 2) the international students rated their knowledge of the topics higher. However, both of these main effects are better explained by the significant interaction. In both cases, the international students rated their knowledge before as being substantially greater than the U.S. students, but this gap was, for the most part, closed in the after-software rating. As a consequence, the large pre-to-post rating gain was mainly the result of the U.S. students’ increase, while the increase was not so dramatic for the international students.

This effect may be interpreted in a number of ways. The most obvious, is that the U.S. students benefited more from the software. It is also possible that the high initial ratings for the international students simply left little room for improvement on the post ratings. Another possibility is that the U.S. students are simply less confident initially and become more confident as a result of their experience with the software. However, this interpretation is inconsistent with the 2003 Trends in International Math and Science Study (TIMSS), which indicated that U.S. students tend to rate their math skills very high relative to students in other countries; in fact, much higher than would be warranted by their actual scores (Mullis, Martin, Gonzalez, & Chrostowski, 2003). In summary, the ratings increased significantly across groups, but the effect was exhibited much more dramatically with the U.S. students.

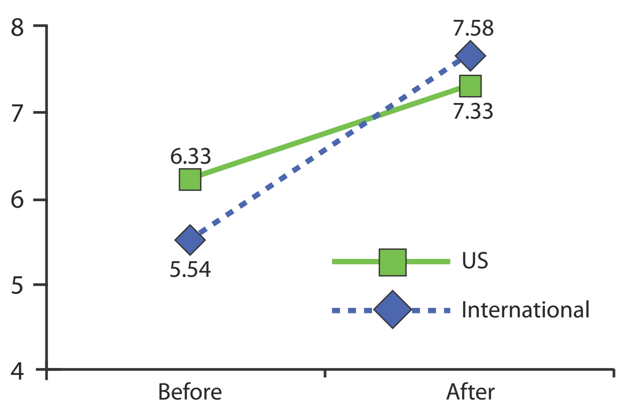
Males vs. Females

Analysis

Change in pre- to post-test rating was also examined as a function of gender in a mixed analysis of variance (ANOVA) for the CMI survey. *Gain* (before vs. after) again served as a within-subject independent variable, *gender* (male vs. female) served as a between-subject independent variable, and *ratings* served as the dependent variable. In reporting the results below we will not discuss the main effect for Gain since that effect is redundant with respect to the previous ANOVA.

In the CMI analysis, there was no significant main effect for *gender*, but there was a significant *gender* × *gain* interaction, $F(1,132)=6.26$, $p< 0.05$. The means associated with this interaction are displayed in Figure 8.

Figure 8: Ratings as a Function of Gain and Gender for the CMI Questionnaire



Interpretation

The significant interaction that was found in the CMI analysis indicates that, while both males and females increased in their rating, the increase was greater for the males. They initially rated their knowledge lower, and subsequently rated their knowledge slightly higher, following their experience with the software. This indicates that, to the extent that the positive impact of the software differed as a function of gender, this impact was greater for males. However, it is important to note that ratings increased for both males and females.

Conclusions

The survey results provide further support for the effectiveness of this instructional multimedia. This support was almost as strong as it could be given the subjective questionnaires that were used. Students rated their knowledge as greater after using the software than before and demonstrated strong agreement with a number of other statements indicating the effectiveness of the software.

There are, however, two important qualifications. The positive impact was manifested more for students from the U.S. and for males, as opposed to non-U.S. institutions and females. Before addressing potential explanations for these effects, it is important to note that ratings increased for both nationality and gender groups. The effect was simply more pronounced within the U.S. and male groups.

With respect to the differential effect of nationality, there are a number of reasons why such an effect would occur, most of which center around the fact that the software was created at the University of Missouri – Rolla, a U.S. university. In fact, professors who were born and raised in the United States provided all of the content and most of the software design. All of the written information in the software was written by these professors in English, and for many of the students in the international universities, this was most likely not their first language. This may have been particularly relevant in comparing the software and textbooks, where the text may not have been written in English. Beyond language, there may also be cultural difference in the way that students in U.S. vs. non-U.S. schools view – and most effectively learn – engineering concepts.

The significant interaction involving gender is more difficulty to explain. One potential explanation is that males and females may differ to some degree in the ways in which they process spatial displays, or more accurately in their *preference* for such displays (Hall & Hickman, 1999). Another, perhaps more important, explanation is that this effect was somewhat of an anomaly in that it does not occur consistently in these types of context.

In fact, a more detailed analysis of these surveys including all modules was carried out and published elsewhere (Hall, Hubing, Philpot, Flori, & Yellamraju, 2004a), and results indicated that gender did not interact with pre vs. post test when the other three modules were examined (though the nationality interaction effect was consistent for all modules).

Overall Conclusions

Project

In terms of the project, this set of studies provide strong evidence that this instructional software can be effective across a variety of contexts in comparison to different types of criteria (e.g., pre-knowledge, class textbooks, control groups) based on both objective and subjective phenomenon. The progression of the research also indicates that initial systematic research on software components lead to more effectively designed software.

The findings of these assessment studies have been used to improve the computer-based instructional media in several ways. From the assessment results, it was clear that individual animated movies must be kept to a modest length. For example, one of the movies used in the assessment studies was a comprehensive movie containing about 25 scenes that presented four aspects of the theory. This movie was revised into four separate movies, with each discussing only one aspect of theory. The assessment studies revealed strong negative student reaction to animations detailing mathematical operations on equations but strong, positive student reaction to animations that helped students visualize the internal response of solid objects to external loads. Based on these findings, subsequent movies have focused more attention on visualization and less attention on algebraic equation details. The assessment studies also suggested that it was worthwhile to continue including interactive feedback components in the movies. Additional feedback styles have been developed, and these feedback features have proven to be both popular with students and effective as instructional tools. Tim writes stuff here, about changes that resulted based on the assessment

Model

In order to explore the efficacy of the model that served as a framework for this project, we'll consider each of the main model components within the context of the studies discussed.

Learner variables were not identified as playing an important role in the majority of studies, but they did prove to be very important in some cases. The most dramatic case occurred in the usability testing of the interactive feedback modules. It was apparent that there were substantial differences between users in terms of engagement, and this, subsequently, lead to substantial differences in their strategies for working with the software. This provided us with useful insight into, and possible explanations for, some of the non-significant effects in the first two experiments. In terms of modification and evolution of our assessment model this indicates that we should also consider engagement and/or motivation as an important individual difference variable and work to identify tools for its systematic measurement. Learner variables also proved to be important in the final study, which focused on evaluation of the software across different international contexts. The learner's nationality and gender proved to play a significant role in perceived effectiveness of the software. In terms of the model, it may be useful to add culture and/or nationality as an important learner variable to consider.

The experimental methodologies employed are listed in Table 1 (page 7) above so we won't address them further here except to note that the mix of methodologies clearly allowed for a more holistic view of important issues and contextual factors with respect to the use of these technological tools. As for the mediation component of the model, it was only explicitly addressed in the usability study, where learners were observed directly as they used the software. The results of this study were largely unexpected, in that some subset of users clearly did not use a rationale strategy as we had anticipated they would. This, of course, could account for significant and unanticipated outcomes and, indeed, provided important insights into the quantitative experiments on the interactive software that preceded the usability analysis.

With respect to learning outcomes, the most common measures applied were quantitative measures of perceived learning, which included not only the degree to which student's learned, but also related concepts such as motivation to learn, metacognition, and degree to which the material related to "real world" applications. In addition, measures of foundational learning, in the form of multiple choice tests addressing basic concepts, were applied in most studies. Problem solving measures were used in two of the experiments covering games used for calculation, in the form of problems students were required to complete. In the majority of experiments described here, both perceived and objective (foundational or problem solving) outcomes were used. Although the perceived outcomes appeared more sensitive to differences between groups, the different measures were generally consistent in terms of the direction of the means.

Finally, it is important to note that both quantitative and qualitative measures were used in the majority of experiments. The qualitative measures generally consisted of questionnaires with open-ended questions to which students responded. These were then categorized and labeled based on themes that emerged in the data. For the most part, the qualitative results supported the quantitative results, while providing additional insights. For example, in the interactive software research and the applied classroom experiments, a number of themes were identified that helped to explain why those who were exposed to the software perceived the software to be more effective for learning. These qualitative analyses provided guides for software modifications, as well as for indicating the most effective ways to integrate software into instruction.

In summary, a review of the model within the context of these studies indicates that the majority of the general categories and measures were utilized and the combination of methodologies and measures served to provide additional insights into processes and outcomes that they may not have provided had they been applied in isolation. Overall, this project provides an example of how the triangulation of diverse measurement tools and research methodologies can be combined to provide a rich and integrative picture of a large-scale project that, in turn, informs – and allows for the evaluation of – instructional software development.

Endnotes

1. Fund for the Improvement of Postsecondary Education
2. Examples of many of these modules are available at http://www.umr.edu/~bestmech.
3. The experiment directions and the modules for both experiments and both groups are available at http://campus.umr.edu/lite/feedback.
4. Go to http://web.umr.edu/~bestmech for examples.

References

Amory, A. (2002). Building and Educational Adventure Game: Theory, Design, and Lessons. *Journal of Interactive Learning Research, 12*, 249–263.

Babu, S. V., Suni, I. I., & Rasmussen, D. H. (1998). Development of a CD-ROM in Thin Film Technologies: Design, Usability Assessment, and Challenges. *Journal of Engineering Education, 87*, 583–589.

Barron, A. E. (1996). Design Guidelines for the World Wide Web. *Journal of Interactive Instruction Development, 8*(3), 13–17.

Crown, S. (1999). The Development of a Multimedia Instructional CD-ROM/Web Page for Engineering Graphics. *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 1026–1031.

Crown, S. W. (2001). Improving Visualization Skills of Engineering Graphics Students Using Simple JavaScript Web Based Games. *Journal of Engineering Education, 90*, 347–355.

Dillon, A., & Gabbard, R. (1998). Hypermedia as an educational technology: A review of the quantitative research literature on learner comprehension, control, and style. *Review of Educational Research, 68*, 322–349.

Ekstrom, R. B., French, J. W., & Harman, H. H. (1979). Cognitive factors: Their identification and replication. *Multivariate Behavioral Research Monographs, 79*.

Eller, V. E., Hall, R. H., & Watkins, S. E. (2001). Multimedia web-based resources for engineering education: The media design and assessment laboratory at UMR. *Proceedings of the American Society for Engineering Education*, http://litelab.org/publications/asee_2001_eller_multimedia.pdf.

Flori, R. E., Oglesby, D. B., Philpot, T. A., Hubing, N., Hall, R. H., & Yellamraju, V. (2002). Incorporating Web-Based Homework Problems in Engineering Dynamics. *Proceedings of American Society of Engineering Education Conference*, http://litelab.org/publications/asee_2002_flori-incorp.pdf.

Hall, R. H., & Hickman, L. L. (1999). Imagic and textual components of web page design: The role of gender in subjective ratings. *Virtual University Journal, 2*, 50–53.

Hall, R. H., Hubing, N., Philpot, T. A., Flori, R. E., & Yellamraju, V. (2004a). Assessment of Engineering Mechanics Instructional Software in a Variety of Instructional Settings. *Proceedings of the American Society for Engineering Education*, http://www.asee.org/acPapers/2004-1663_Final.pdf.

Hall, R. H., Philpot, T. A., Oglesby, D. B., Flori, R., Hubing, N., Watkins, S. E., & Yellamraju, V. (2002). A model for the evaluation of innovative engineering courseware: Engineering an assessment program. *Proceedings of the American Society for Engineering Education*, http://litelab.org/publications/asee_hall_model.pdf.

Hall, R. H., Watkins, S. E., Davis, R. L., Belarbi, A., & Chandrashekhara, K. (2001). Design and Assessment of Web-Based Learning Environments: The Smart Engineering Project and the Instructional Software Development Center at U.M.R. In L. R. Vandervert & L. V. Shavinina (Eds.), *Cybereducation: The Future of Long Distance Education* (pp. 137–156). New York: Mary Ann Liebert, Inc.

Hall, R. H., Wilfred, L., Hilgers, M. G., Leu, M. C., Walker, C. P., & Hortenstine, J. M. (2004b). Virtual Terrorist Attack on the Computer Science Building: Design and Evaluation of a Research Methodology. *Presence-Connect, 4*(4), http://www.presence-connect.com.

Hubing, N., Flori, R., Hall, R. H., Oglesby, D. B., Philpot, T. A., & Yellamraju, V. (2002). Interactive Learning Tools: Animating Statics. *Proceedings of the American Society for Engineering Education*, http://litelab.org/publications/asee_2002_hubing_interact.pdf.

Jih, H. (2001). Promoting Interactive Learning through Contextual Interfaces on a Web-Based Guided Discovery CAL. *Journal of Computers in Mathematics and Science Teaching, 20*(4), 367–376.

Kadiyala, M., & Crynes, B. (2000). A Review of Literature on Effectiveness of Use of Information Technology in Education. *Journal of Engineering Education, 89*, 177–189.

Merino, D. N., & Abel, K. D. (2003). Evaluating the Effectiveness of Computer Tutorials versus Traditional Lecturing in Accounting Topics. *Journal of Engineering Education, 92*, 189–194.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., & Chrostowski, S. J. (2003). *TIMSS 2003 International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*: International Association for the Evaluation of Educational Achievement.

Philpot, T. A., Flori, R., Hall, R. H., Hubing, N., Oglesby, D. B., & Yellamraju, V. (2002). Interactive learning Tools: Animating mechanics of materials. *Proceedings of American Society for Engineering Education Conference*, http://litelab.org/publications/asee_2002_philpot_interact.pdf.

Philpot, T. A., Hall, R. H., Hubing, N., Flori, R., Oglesby, D. B., & Yellamraju, V. (2003a). Animated instructional media for stress transformations in a mechanics of materials course. *Computer Applications in Engineering Education, 11*, 40–51.

Philpot, T. A., Hubing, N., Hall, R. H., Flori, R., Oglesby, D. B., & Yellamraju, V. (2003b). Games as Teaching Tools in Engineering Mechanics Courses. *Proceedings of the American Society of Engineering Education*, http://litelab.org/publications/asee_2003_philpot_games.pdf.

Shin, E., Schallert, D., & Savenye, C. (1994). Effects of learner control, advisement, and prior knowledge on young students' learning in a hypertext environment. *Educational Technology Research and Development, 42*, 33–46.

Smith, S. (1999). An Interactive Tool for Math and Science Learning. *Proceedings of International Conference on Mathematics/Science and Technology*, 166–170.

Sorby, S. A., & Baartmans, B. J. (2000). The Development and Assessment of a Course for Enhancing the 3-D Spatial Visualization Skills of First Year Engineering Students. *Journal of Engineering Education, 89*, 301–307.

Steif, P. S., & Naples, L. M. (2003). Design and Evaluation of Problem Solving Courseware Modules for Mechanics of Materials. *Journal of Engineering Education, 92*, 239–247.

Westbrook, J., & Braithwaite, J. (2001). The Health Care Game: An Evaluation of a Heuristic, Web-Based Simulation. *Journal of Interactive Learning Research*, 12, 89–104.

Wilfred, L. M., Hall, R. H., Hilgers, M. G., Leu, M. C., Hortenstine, J. M., Walker, C. P., & Reddy, M. (2004). Training in Affectively Intense Learning Environments. *Proceedings of the AACE E-Learn Conference*, 2233–2240, http://dl.aace.org/16761.

Acknowledgement

This work was supported in part by a grant from the United States Department of Education Fund for the Improvement of Post-Secondary Education (FIPSE #P116B000100).

Author Biographies

Richard H. Hall is Professor of Information Science and Technology at the University of Missouri – Rolla. He received his BS degree in Psychology from the University of North Texas, and Ph.D. degree in Experimental Psychology from Texas Christian University. He is the co-director of UMR's Laboratory for Information Technology Evaluation, and his research focuses on web design, and evaluation with a focus on web based learning technologies.

Mailing address:
Dept of Information Science and Technology
University of Missouri – Rolla
Fulton Hall 101
1870 Miner Circle
Rolla, MO 65409-0320

Email:
rhall@umr.edu

Timothy A. Philpot is an Associate Professor in the Interdisciplinary Engineering Department at the University of Missouri – Rolla. He received a B.S. degree from the University of Kentucky, a M.Eng. degree from Cornell University, and a Ph.D. degree from Purdue University. All three degrees were in Civil Engineering. His research focuses on instructional software for engineering, and he is the author of two award-winning educational engineering software packages: MDSolids and MecMovies.

Mailing address:
Dept. of Interdisciplinary Engineering
University of Missouri – Rolla
1870 Miner Circle
Rolla, MO 65409-0210

Email:
philpott@umr.edu

Dr. Nancy Hubing received the BSEE degree from Purdue University in 1982, and the MS and PhD degrees in Electrical and Computer Engineering from North Carolina State University in 1984 and 1989, respectively. She served on the faculty of the University of Missouri – Rolla from 1989 to 1999 in the Electrical and Computer Engineering Department, and from 2000 to 2006 in the Basic Engineering Department. She taught physics at Rolla High School during the academic year 1999 – 2000, and served as director of UMR's Freshman Engineering Program 2005 – 2006. Dr. Hubing is particularly interested in engineering education and K–12 outreach.

J·T·L·A

The Journal of Technology, Learning, and Assessment

www.jtla.org

Technology and Assessment Study Collaborative
Caroline A. & Peter S. Lynch School of Education, Boston College