2017 Zone IV Best Paper: Assessment of Long-term Effects of Technology Use in the Engineering Classroom

Dr. Sean St. Clair, Oregon Institute of Technology

Sean St.Claire is an Associate Professor and Chair of the Civil Engineering Department at Oregon Tech, where he teaches structural engineering courses and conducts research in engineering education. He is also a registered Professional Engineer.
Assessment of Long-Term Effects of Technology Use in the Engineering Classroom

Sean St.Clair
Oregon Institute of Technology

Abstract

This paper summarizes a longitudinal study of the effects of instructional technology on learning and knowledge retention. Two very different types of software were introduced to students in an undergraduate mechanics course and the effects of this technology use were assessed a number of different times over a twenty-five week period. The study took place in three sections of an undergraduate mechanics course. Students in one section used a tool-type software. In a second section, students used a content-type software. Students in the third section represented a control group and did not use any software. To ensure the study’s validity, demographic data were collected from and compared between each of the three sections and the study was replicated in a subsequent term. Students were assessed immediately following the intervention and five, ten, and twenty-five weeks later. The results of these assessments were compared among groups to determine the effect of the technology use on learning and retention. Results indicated that each of the three groups had very high levels of retention but that there were no differential levels of learning, retention, or long-term retention among the three groups. Students using technology, however, were able to solve a greater number of problems during the special session than those working by hand. It was thus concluded that instructional technology can make the educational process more efficient without hindering learning or long-term knowledge retention.

Introduction

The purpose of this research was to determine if a correlation exists between computer use and knowledge retention.

Computer use has become ubiquitous in all aspects of engineering education\(^1\). Students use spreadsheets to perform repetitive calculations or traffic simulation software to visualize and design an intersection. Students use electronic textbooks in lieu of printed volumes. Professors use courseware management systems to communicate effectively with their students. These are all examples of instructional technology.

Glennan and Melmed classify instructional technology into three categories: tool-type, content-type, and instructional management\(^2\). Tool-type technologies are those that have been developed for a purpose other than instruction but also have use in the classroom. An example of a tool-type technology is the use of commercial structural analysis software in a reinforced concrete design course. Alternatively, content-type technologies are those which have been developed specifically for use in education and are designed to deliver instructional content to student users. Electronic textbooks and intelligent tutoring systems are examples of content-type technologies. Instructional management technologies are those that are designed to assist in the administrative tasks of an instructor and include electronic grade books and courseware management systems.
Because instructional management tools are not manipulated by the students, these were not included in this study.

Knowledge retention, the recall or remembrance of previously learned material, is an essential part of education. Semb and Ellis point out that “the very existence of school rests on the assumption that people learn something of what is taught and later remember some part of it.” This recall is particularly true in engineering education, wherein many classes begin where previous classes have left off. Most upper division engineering courses have a number of prerequisites, which in turn have prerequisites of their own. This curriculum structure is based on the premise that students will remember, or retain, what was taught in earlier classes.

Research conducted in the area of knowledge retention, however, suggests that such retention is generally very poor. This may be linked to the fact that lectures remain the most popular teaching technique in higher education. This is an important point because it has been shown that lecturing, when compared to other forms of instruction, result in the lowest levels of retention. Studies on alternative instruction techniques, such as active learning, problem based learning, inductive reasoning, and discussion methods have shown that knowledge retention can be increased, sometimes significantly, when compared to lecture-based instruction.

The use of technology as an alternative instruction technique has led many, within engineering and other fields, to propose that the use of such technologies in the classroom may increase learning. Researchers point out that very few classroom-based studies exist, however, to define the relationship between instructional technology and retention. Many studies have been conducted in laboratory settings by cognitive scientists and psychologists in an attempt to establish a link between instructional technology and retention. Whether or not these laboratory studies—which often utilize simple learning activities and short retention spans—can be generalized to classroom learning is questioned by some.

Alternatively, there are multitudinous examples of classroom-based research linking technology with learning and retention, but the vast majority of these lack any direct assessments and rely solely on indirect measures such as students’ opinions and self-evaluations. Other studies that quantitatively assess retention do so over a span of one month or less. The retention span was considered important in this study because students are generally expected to retain information well into a subsequent school term, often involving time frames exceeding one month.

In contrast to these many studies, this research was designed to take place in the classroom, while maintaining a high degree of scientific rigor both in terms of methodology and data analysis. Furthermore, the study was designed to determine a quantitative relationship between instructional technology and knowledge retention at intervals that more closely resembled semesters in length of time. Finally the study investigated tool-type technology use compared to content-type technology use.
Methodology

Research Questions

This study was designed to answer two basic research questions. First, do students learn and retain more information when using instructional technology as compared to traditional classroom techniques? Second, do students learn and retain more information when using a content-type technology as compared to using a tool-type technology?

Research Design and Control

To answer these questions, a nonequivalent control group research design was implemented wherein students who were already assigned to preexisting groups—sections of a course—were used as experimental and control groups. Three sections of a statics course were used in this study, one of which used the content-type software, another of which used the tool-type software, and a control group which did not use any software in the course.

Because students self-select into the research groups by enrolling in a particular section of statics, the research design is not strictly scientific and is open to a number of different biases. There are a few different methods that can be used to overcome these hurdles and increase the validity of the results. One such method used in this research was replication, meaning that the entire study was conducted twice with different groups of students. If the research yielded noticeable and significant differences in the results between the initial and replicated study, one could infer that one or more biases were interfering with the research. Another way to increase validity was to collect and compare demographic data between the three groups to determine if they differed significantly. If such demographic differences existed, they could be used as variables in the analysis of the results. Along with demographic data, preexisting contextual knowledge was collected via a pretest to determine if there were differences between groups. As another control, and because the sections were taught by different instructors, each of the instructors completed a teaching goals inventory to ensure that there were no significant differences. Each of the instructors also agreed to use identical homework assignments and exam questions for the topic of truss analysis. Finally, a guest lecturer taught all sections on the topic of trusses to reduce potential instructor-related bias. All of these controls were implemented to increase the validity of this research and to make as equal as possible the groups in this nonequivalent control group design.

Intervention

Three sections of a statics course taught during the same term were used as the three research groups. The intervention focused on one subject area within the statics course: truss analysis. Four lecture hours of each section were set aside for truss analysis. The researcher was the guest lecturer for each section during this topic, taught the same material to each section, and ensured that each section was working with the same knowledge base and skill set. Students were introduced to the basics of truss analysis during the first two lecture periods. The intervention took place during the third lecture period, wherein each of the three sections had a different experience.
One group used a tool-type software to complete a truss analysis exercise. Dr. Frame, the software used, is a structural analysis program used in both industry and education. It allows the user to quickly model, analyze, and manipulate two-dimensional structural models. Figure 1 is a screenshot of the software showing a truss model that has been built and loaded. The reactions are shown but the internal forces are not in this screenshot. Students used the software to answer both quantitative and qualitative questions about trusses. Figure A1 in the appendix contains a portion of the exercise that students were assigned to complete during this intervention session.

The second experimental group used content-type software titled Multimedia Engineering Statics (MES), which contained a module designed to teach truss analysis. The module included theory, examples, videos, and an easy-to-use simulation that allowed students to model, analyze, and manipulate a simple truss. Figure 2 shows one such model from MES, which students can manipulate and instantly see, in graphical form, the results of their manipulations. Students completed the same exercise as the tool-type group while using MES.

The comparison, or control, group used no software and simply completed an identical exercise by hand while the lecturer walked around the classroom answering questions.
Students completed the post-test during the fourth lecture period to measure their learning of the topic of truss analysis.

**Instruments and Assessments**

To answer the research questions, learning and knowledge retention were measured using a single instrument that was administered at three different points in time: 1) immediately after the use of technology to measure learning; 2) ten weeks after the intervention to measure retention; and 3) 25 weeks after the intervention to measure long-term retention. The instrument used in these measurements was developed with the input of four different professors who all taught statics. In a previous study, this instrument, as well as the pretest, were found to be valid and reliable\(^\text{19}\). Figure 3 illustrates the timeline of each of these assessments. At ten weeks, students were still enrolled in statics and so the retention posttest was administered in class. At 25 weeks, the civil engineering students were enrolled in a mechanics of materials class; instructors in different sections of that course allowed for the administration of long-term retention posttest in their classrooms.

**Replication**

As mentioned previously, the study was replicated in order to preserve the validity of the research. The replicated study was largely identical to the initial study. One difference was that one additional instructor was included in this study. The pretest, instruction, special session, posttest, exam questions, and retention test were all executed in exactly the same manner as with the initial implementation. The same instruments were used, the same lectures were delivered once again by the guest lecturer, and the same special session exercise was completed. The only other difference was in the exam questions, which were not reused because the exams from the previous term had been returned to the students and were thus in the public domain. The long-
term retention test was once again administered in the mechanics of materials course 15 weeks after the statics course ended.

![Figure 3: Timeline of Assessments](image)

**Participants**

Extremely rigorous analysis procedures were implemented on the participant data to ensure that the study was not influenced by the quasi-experimental nature of the naturalistic classroom setting.

**Instructors**

Three different instructors taught the six sections of statics that participated in this study. To reduce the possibility of an instructor-related bias, a guest lecturer was used. Furthermore, the three instructors each completed a widely-used Teaching Goals Inventory (TGI) to determine if there were significant differences between their approaches to teaching statics. Cluster scores for the TGI were compared using the non-parametric, bivariate correlation using Spearman’s rho that revealed significant correlations between all pairs of instructors: Instructors A and B ($\rho=.975$, $p<.01$), Instructors B and C ($\rho=.975$, $p<.01$), and Instructors C and A ($\rho=1.0$, $p<.01$). The results of the teaching goals inventory, combined with the use of a guest lecturer during the intervention, indicate sufficient attempts were made to ensure the results would not be tainted by instructor differences between the subject groups.

**Students**

Because students chose their own groups, rather than being randomly assigned to groups, participant data were analyzed to determine if the groups differed in any way. Demographic data were collected from all the students in the six sections of statics which participated in this study,
though not all students answered all questions. These data are summarized in Table 1. A chi-square comparison was made between the distributions of the groups for each demographic category. Two significant differences between distributions were found and are reported in Table 1. These differences were used, as needed, as factors in the final analysis. It can generally be inferred from the data that the groups within each of the two studies did not differ in any significant way.

Table 1: Summary and Comparison of Participant Data

<table>
<thead>
<tr>
<th>GROUP</th>
<th>INITIAL STUDY</th>
<th>REPLICATED STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Content</td>
<td>Tool</td>
</tr>
<tr>
<td>MAJOR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Civil Engineering</td>
<td>15 (37%)</td>
<td>13 (34%)</td>
</tr>
<tr>
<td>Industrial Engineering</td>
<td>19 (46%)</td>
<td>17 (45%)</td>
</tr>
<tr>
<td>Electrical Engineering</td>
<td>4 (10%)</td>
<td>6 (16%)</td>
</tr>
<tr>
<td>Other</td>
<td>3 (07%)</td>
<td>2 (05%)</td>
</tr>
<tr>
<td>GENDER</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>22 (54%)</td>
<td>19 (50%)</td>
</tr>
<tr>
<td>Female</td>
<td>19 (46%)</td>
<td>19 (50%)</td>
</tr>
<tr>
<td>RANK</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sophomore</td>
<td>12 (32%)</td>
<td>12 (32%)</td>
</tr>
<tr>
<td>Junior</td>
<td>19 (51%)</td>
<td>22 (59%)</td>
</tr>
<tr>
<td>Senior</td>
<td>6 (16%)</td>
<td>3 (08%)</td>
</tr>
<tr>
<td>GPA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.5 - 4.0</td>
<td>10 (27%)</td>
<td>6 (16%)</td>
</tr>
<tr>
<td>3.0 - 3.49</td>
<td>11 (30%)</td>
<td>12 (32%)</td>
</tr>
<tr>
<td>2.5 - 2.99</td>
<td>10 (27%)</td>
<td>12 (32%)</td>
</tr>
<tr>
<td>2.0 - 2.49</td>
<td>5 (14%)</td>
<td>5 (14%)</td>
</tr>
<tr>
<td>Below 2.0</td>
<td>1 (03%)</td>
<td>2 (05%)</td>
</tr>
<tr>
<td>STATUS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Part-time</td>
<td>2 (05%)</td>
<td>3 (08%)</td>
</tr>
<tr>
<td>Full-time</td>
<td>34 (92%)</td>
<td>33 (89%)</td>
</tr>
<tr>
<td>More than Full-time</td>
<td>1 (03%)</td>
<td>1 (03%)</td>
</tr>
<tr>
<td>ETHNICITY</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian, non-Hispanic</td>
<td>26 (70%)</td>
<td>28 (76%)</td>
</tr>
<tr>
<td>Asian/Pacific Islanders</td>
<td>6 (16%)</td>
<td>4 (11%)</td>
</tr>
<tr>
<td>African American</td>
<td>2 (05%)</td>
<td>1 (03%)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1 (03%)</td>
<td>2 (05%)</td>
</tr>
<tr>
<td>Indian</td>
<td>2 (05%)</td>
<td>2 (05%)</td>
</tr>
</tbody>
</table>

*Significant differences occurred in the distributions of ranks in the initial study: \(\chi^2(4)=30.096, p<.01\).

**Significant differences occurred in the distributions of majors in the replicated study: \(\chi^2(4)=12.863, p<.05\).

Students from each section completed a pretest to assess their preexisting contextual knowledge. The results from these pretests for both the initial and the replicated study are summarized in Figure 4, which reveals that the distributions of scores are fairly normal and do not differ greatly
between sections. An analysis of variance (ANOVA) test revealed no significant differences between the distributions of scores for both the initial (F(2,135)=2.799, p=.064) and the replicated studies (F(2,113)=.818, p=.444). These results further suggest that, despite being quasi-experimental due to the lack of random assignment, the different participant groups could be considered equal.

Figure 4: Pretest Results

Results

Extremely rigorous analysis procedures were used on the results as well. This section provides a snapshot of the results of these analyses.

Learning Posttest

A posttest was completed by students in the lecture that followed the special session. The results of the posttest were used to measure learning. These results are summarized in Figure 5. The distributions of these results are also fairly normal. An ANOVA performed on the results of the initial study revealed a significant difference (F(2,133)=5.958, p<.05) between the test scores, with the content group performing significantly better than the other two groups. Another analysis was performed using rank as a factor in the analysis because, as shown in Table 1, there was a significant difference in the ranks among these groups. With rank as a factor in the ANOVA, the effect of the treatment was no longer significant (F(2,127)=2.396, p=.095). There was no significant different between the groups on the replicated study (F(2,204)=.532, p=.589). It can thus be surmised that learning was equal among all groups.
Retention Test

Ten weeks later, students completed the test again, this time to measure their retention of what they had learned. Figure 6 contains the distributions of the results of their scores. While the results were much more variable, an ANOVA revealed no significant differences for the initial study ($F(2,134)=2.394, p=.095$) or the replicated study ($F(2,105)=.377, p=.687$), suggesting that technology use had little impact on knowledge retention.

Long-Term Retention Test

Fifteen weeks after the retention test, the civil engineering students were asked to complete the same exam to measure their long-term retention of truss analysis. The distributions of those scores are plotted in Figure 7. These data are also highly variable, but still fairly normally distributed and similar. Analyses revealed no significant differences in either the initial study ($F(2,32)=.227, p=.798$) or the replicated study ($F(2,22)=1.206, p=.318$). These data suggest that using technology in the classroom had neither a positive nor a negative effect on long-term retention.

Data Analysis Summary

A small snapshot of all the data that were collected and analyzed is presented here. Much more data were gathered and analyzed. Completion times, response rates, and differential scores—the numerical difference between retention test scores and posttest scores—were all analyzed and revealed no significant groups. Further analyses included information about question types (factual, qualitative, and quantitative), comparisons across implementations, and comparisons across demographics. While the complete set of data and analysis is outside the scope of this
paper, such information is available from the author upon request. The results, however, of all the data analyses are the same: the use of technology in this study did not significantly affect student performance on any of the assessments conducted.

**Figure 6: Retention Test Results**

**Figure 7: Long-Term Retention Test Results**
Findings, Conclusions, and Future Work

In this study, all three groups performed equally well on all assessments during both implementations of the study. Furthermore, all three groups retained a great deal of information because average scores on each subsequent assessment remained at least equal. These findings are in contrast to what was expected. The technologies chosen for this study were very interactive and it was assumed that the active nature of the technology use would increase both learning and retention when compared to those that did not use technology. The reason for this result is likely due to what was done by the control group. During the special session the control group did not listen to lectures but were working on truss problems with other students. It is possible that this team-based exercise was engaging enough to be considered active learning and was equally active as using the instructional technology. This finding has a few implications.

The first conclusion based on this finding is that instructional technology should not be implemented solely as a means of increasing retention. Instructional technology can be expensive and time-consuming to implement and does not result in greater learning and retention than simply using in-class, team-based activities. It was concluded that technology should be implemented for a specific reason other than to increase knowledge retention. One possible reason for implementation became apparent in this study: students were able to complete far more problems when facilitated by technology than they could complete by hand. As such, technology was an efficient tool that exposed students to many more problems, solutions, and situations than they would have seen otherwise. Instructors should feel confident, however, that when technology is implemented for a specific and useful reason, student learning and knowledge retention will not be adversely affected.

The second implication of these findings is the benefit that was realized from using in-class activities. In this study, simply having students work together on truss problems in class was as effective as implementing technology. Such in-class activities are easy to do, require little preparation, and cost nothing. The results of this study suggest that instructors who are concerned about knowledge retention should be implementing in-class activities.

As with any research project, the findings and conclusions resulting from this study suggest that there is more work to be done in the areas of knowledge retention and instructional technology in engineering education. Application of the framework used in this study to a different content domain could lead to more firm conclusions about the long-term effects of technology in the classroom. For instance, control groups may have performed as well as the experimental groups because truss analysis is an active process that requires visualization, calculation, and a systematic thought process. If this study was replicated with a topic that was not so inherently engaging, the results would further define the relationship between technology and knowledge retention.

Another worthy area of future investigation is one that statics instructors perceived as important: the qualitative understanding of truss behavior. It was assumed that technology could help students view the truss as a whole rather than as a system of discrete parts and that this would help students determine conceptually how an entire truss would behave under certain conditions. More extensive and personal research, such as interviews and focus groups, may reveal the
students’ thought processes and determine if technology could help students in such a manner. Ultimately, additional naturalistic studies of technology use in education are required to understand more fully its potential influence on learning.

References


17. Dr. Software LLC. *Dr. Frame*, software, 2001.


Appendix A
Special Session Instrument

For the following exercises, consider the truss in the figure below, which is supported by a pin at H and a roller at D.

If $\theta = 45^\circ$ and a 9kip force is applied vertically at A in the downward direction, determine the forces in GH, CH, and AF and indicate whether the member is in tension or compression.

What is the relationship between $\theta$ and the force in members CH and GH?

Do any of the member forces change if the 9kip force is applied at F rather than A; if so, how?

How do the member forces change if the 9kip load is distributed across the bottom cord: 3kip loading at each of the joints G, F, and E?

Figure A1: Portion of Special Session Exercise