Computer tutor versus solving problems by hand: a comparison in statics

Capt. Colin Engebretsen, US Air Force Academy

Captain Colin C. Engebretsen is an assistant professor of engineering mechanics at the US Air Force Academy. He is the Course Director for EM 220 Fundamentals of Mechanics and has also instructed EM 330 Mechanics of Deformable Bodies. He is the academic advisor for 13 cadets and works as the Assistant Advisor-in-Charge for the department of engineering mechanics.

Captain Engebretsen was commissioned in 2008 through the ROTC program at the University of North Dakota, obtaining a Bachelor's Degree in Mechanical Engineering. He obtained a master's degree in Aeronautical Engineering from the Air Force Institute of Technology while researching hysteretic damping in ceramic coated titanium. Additionally He has worked as a structural engineer on the KC-135 Stratotanker and Chief Engineer for Aircraft Battle Damage Repair Engineering at the Oklahoma City Air Logistics Center.

Dr. Paul S. Steif, Carnegie Mellon University

Paul S. Steif is a Professor of Mechanical Engineering at Carnegie Mellon University. He received a Sc.B. in engineering from Brown University (1979) and M.S. (1980) and Ph.D. (1982) degrees from Harvard University in applied mechanics. He has been active as a teacher and researcher in the field of engineering education and mechanics. His research has focused on student learning of mechanics concepts and developing new course materials and classroom approaches. Drawing upon methods of cognitive and learning sciences, he has led the development and psychometric validation of the Statics Concept Inventory – a test of statics conceptual knowledge. He is the co-author of Open Learning Initiative (OLI) Engineering Statics, and he is the author of a textbook Mechanics of Materials, published by Pearson.

COMPUTER TUTOR VERSUS SOLVING PROBLEMS BY HAND:
A COMPARISON IN STATICS

Introduction

Instructors increasingly seek to use computer-based technologies to enhance instruction in their classrooms. Recently, integration of video and web technologies has made it feasible for lecture type material to be provided readily to students outside of the classroom. It is more challenging to provide computer-based learning tools outside that enable students to practice using concepts and skills they have learned and to receive timely feedback, which is known to improve learning [1-4]. There have been examples of computer systems to allow students to practice and receive feedback in the realm of statics in particular, steps [5-7].

It is particularly difficult to give students computer-based means to practice solving problems that have complex arrangements of parts, and which have multiple pathways to solution. Recently, a computer tutor that enables students to practice truss problems was presented [8,9]. This tutor was inspired by the approach taken by cognitive tutors [10-12], with the tutor able to solve the problems on its own, and thus follow any solution path taken by a student. The tutor gives feedback on errors that enables the student to complete nearly all problems. Data collected from students using the tutor was analyzed and provided evidence that students were learning. Specifically, as students solved problems, which involve a variety of skills, errors using those skills decreased with practice.

Still the evidence of learning was internal to the tutor: errors using the tutor decrease with practice using the tutor. Thus, it remains to be seen whether use of such a tutor could still provide a learning experience that is equally effective as solving problems with pencil and paper. Summative assessment of student performance in courses is typically based on performance on written exams. While computer-based learning could provide beneficial instantaneous feedback, written homework has the advantage of giving students practice solving in the same mode as in the summative assessment. There have been some studies comparing the effect of computer-based instruction with traditional practice in physics. Some of these have found the computer-based instruction to lead to equivalent outcomes [13], and in some cases even improved outcomes [14]. In this paper we describe a controlled study testing how students who used truss tutor for homework performed on examinations in comparison with peers who did handwritten homework.

1. Description of computer tutor

The tutor has been described previously [8]. As seen the screenshot in Figure 1, the user can define multiple subsystems, by selecting bars, partial bars, and pins.
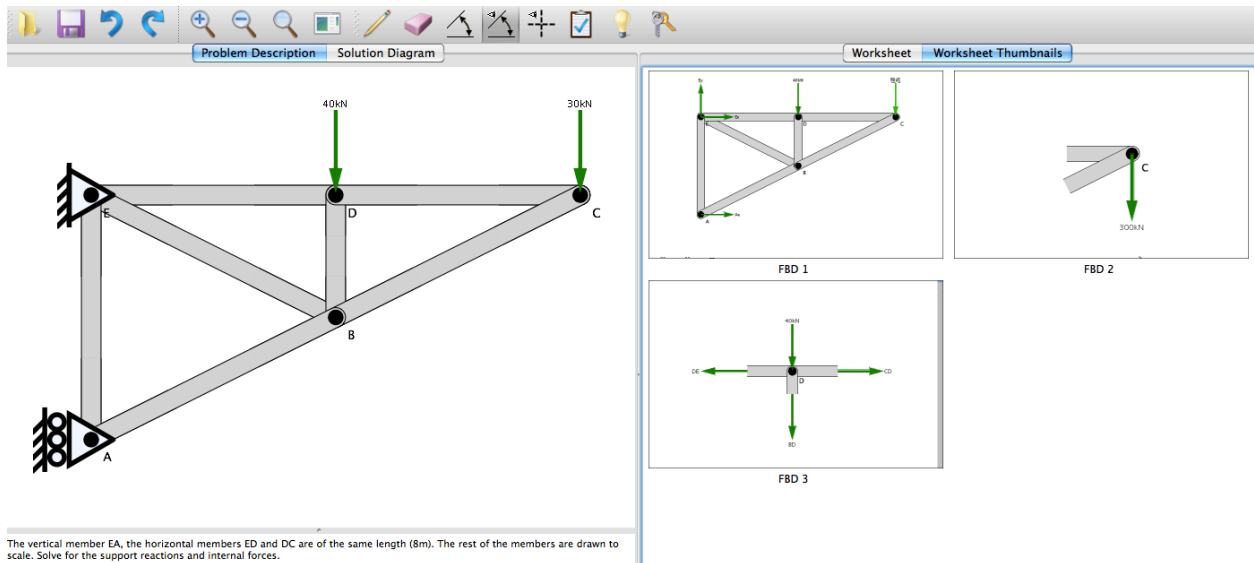
Figure 1. Screen shot of full display of tutor for trusses.

For any one of the subsystems, the user creates the free body diagram by drawing forces on selected points, categorizing each type of force (Figure 2). The user both categorizes the force being drawn as a known applied force, a support reaction (unknown or determined), or an internal force (unknown or determined), and then gives it a variable or number.
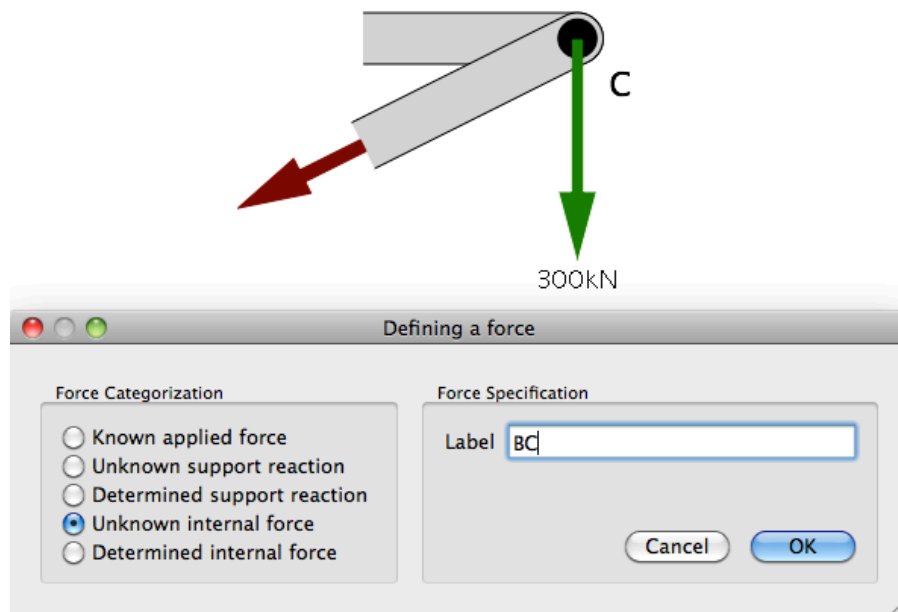

Figure 2. Screen shot of force being added to a free body diagram, showing force categorization.

Once a free body diagram is completed and is correct, the user types in equations of equilibrium as shown in Figure 3. If the user has written down an equation with one variable, the built-in calculating facility can solve the equation.
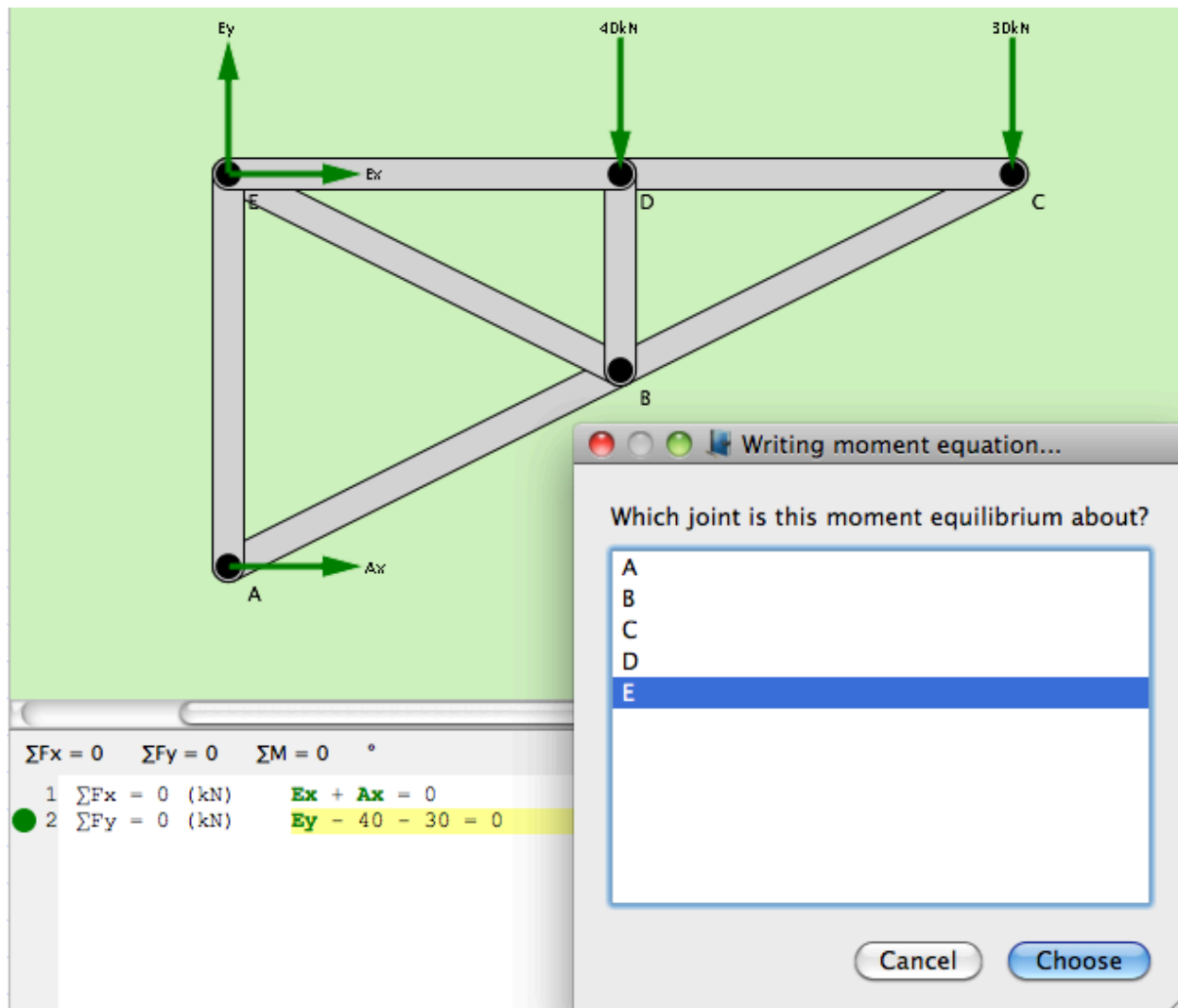
Figure 3. Screen shot of writing equations, and choosing moment center.

Once a support reaction or internal force is determined, the user can declare it as solved and place it in a diagram where the solution is displayed. Here interpretation of the computed result is required: a support force must be drawn in its actual sense and given a positive magnitude, and an internal force of a bar is given a magnitude and described as tensile or compressive.

Results from analyses of student data[8,9] support the hypothesis that practice with the tutor leads to reductions in errors using the tutor.

2. Setting for Study

The study took place at the Air Force Academy, at which all students share a broad core curriculum, including a semester-long statics and mechanics course. Class sizes range from 23 to 28 students, consisting of approximately 94% sophomores, and 6% juniors.

The course was graded out of 1000 points and consisted of three 125-point exams and a 250-point final exam. The second regular semester exam and the final exam each had a "work-from-scratch" truss problem. Trusses constituted a combined 70 points on these exams; when added to the approximately 10 points for the truss homework, the topic of trusses accounts for 8% of the students' grade in the course.

The truss problem from the second regular semester exam is shown in Figure 4, and the truss problem from the final exam is shown in Figure 5.
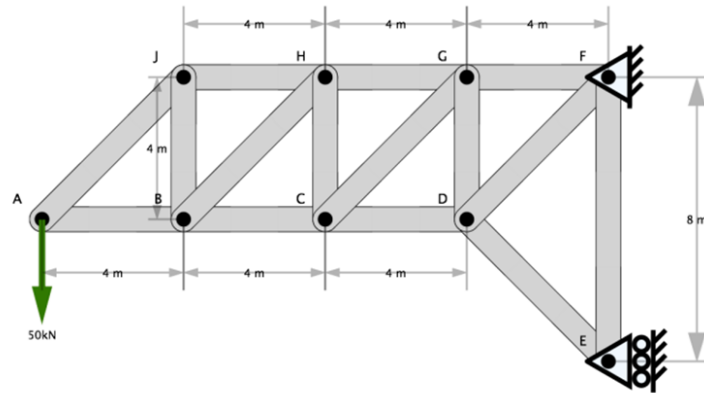

Figure 4. Problem from exam 2, in which students are asked to solve for the forces in members JH, BH, BC.
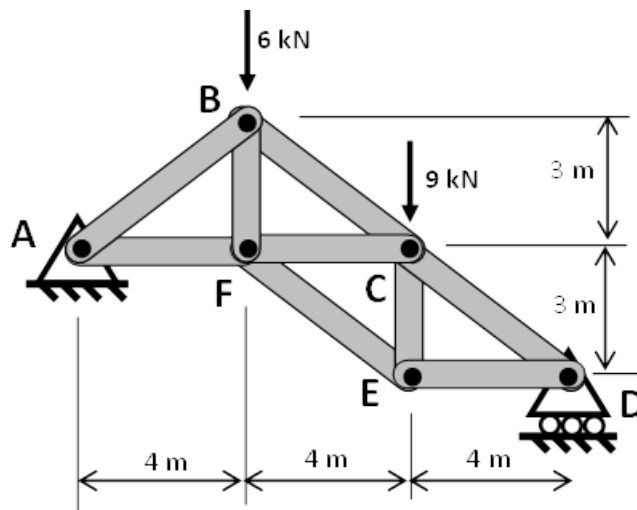

Figure 5. Problem from the final exam, in which students are asked to solve for the forces in members BC, FC, and FE.

Ten class sections were chosen for the study specifically from instructors who wished to participate and had back-to-back class periods to serve as an experimental and a control group. Each instructor taught the same truss lesson to both class periods and assigned the same homework to all students. The homework assignment consisted of eight truss problems, three requiring the method of joints, and 5 requiring the method of sections. Homework problems and exam problems were of similar style and comparable difficulty. The control groups completed the homework in a traditional pencil-and-paper format with the solution available to them if they got stuck and needed assistance. The experimental group completed the same set of problems using the computer tutor.

Instructors were also allowed to assign additional, non-homework truss practice or assessments at their discretion as they would in a normal class setting. This additional work typically took the form of quiz problems or exam review sheets, which were done in traditional pencil-and-paper format. Every instructor used some form of additional practice, and the same practice was assigned to both the control and experimental groups equally. Many additional resources such as practice problems or past examples were available to all students to reference at their discretion. The experimental group also had the option to complete additional problems using the computer tutor for further practice.

Across all sections, complete data were obtained for 118 students in the control condition and 111 in the experimental condition. Scores of students on all exams, and on the two exam problems involving trusses in particular, were recorded. For each instructor we designate Section A as the section in which students solved truss problems by hand (control) and Section B as the section in which students solved truss problems using truss tutor (experimental).

3. Results of study

To gain insights into whether there are significant differences in students in the two sections taught by a given instructor, we computed the total score on all exam problems not involving trusses; these results are shown in Table 1. For these non-truss problems, there should be no effect of the homework condition explored here. It can be seen that the means in the two sections are nearly the same. Indeed, there are no significant differences from one instructor to the other. Thus, as judged by exam scores, all instructors are equally effective. This comparison across sections suggests inherent differences in the students in the two sections would not affect the subsequent results related to the effects of instruction.

Table 1: Comparison of students in each pair of sections (A and B) on all non-truss exam problems; shows no differences across sections and from one instructor to the next.

Instructor	Section A		Section B	
	Mean	Standard Deviation	Mean	Standard Deviation
1	78.18	10.63	75.70	11.36
2	77.25	10.21	76.42	13.21
3	80.13	8.58	76.98	11.11
4	75.77	7.57	76.55	9.58
5	78.72	8.63	79.21	9.68

Means and standard deviations for scores on the truss problem that appeared in exam 2 are shown in Table 2.

Table 2: Comparison of performance of students in control and experiment conditions on truss problem in exam 2

Instructor	Section A: Control (Handwritten)		Section B: Experimental (Computer Tutor)	
	Mean	Standard Deviation	Mean	Standard Deviation
1	94.64	11.78	91.63	17.43
2	93.85	10.63	84.27	25.12
3	92.74	13.04	90.98	12.72
4	91.70	16.98	89.77	19.76
5	97.20	5.97	96.63	4.86

It can be seen that there are minor differences between the performances by students in the control (handwritten) and experimental (computer tutor) conditions. None of these differences are statistically significant: even for the sections of Instructor 2, where the difference in the means for the two conditions is largest, a t-test reveals $t = 1.72$, $p = 0.092$. Even this test is likely of dubious value. Of 238 students, 137 had scores of 100 (making the median score 100). With such a high mean, and such a skewed distribution, this set of scores may not be very useful in sorting out differences among instructors or homework conditions.

Scores for the truss problem on the final exam provide a more useful measure with which to compare conditions, with an overall mean of 75.6 and standard deviation of 22.1. The comparison of performance on the truss problem of the final exam for the two conditions is shown in Table 3.

Table 3: Comparison of performance of students in control and experiment conditions on truss problem in final exam

Instructor	Section A: Control (Handwritten)		Section B: Experimental (Computer Tutor)	
	Mean	Standard Deviation	Mean	Standard Deviation
1	76.91	22.01	73.20	26.31
2	81.94	20.48	82.38	23.48
3	70.48	18.02	66.81	25.55
4	80.00	19.63	74.67	20.16
5	73.06	25.21	76.36	14.87

Again, there are minor differences, but the differences are not systematic: for three sections the control section performed better and for two sections the experimental section performed better. None of the differences is statistically significant: even for Instructor 4 with the largest difference between control and experimental means, the result is not significant ($t = 0.87$, $p = 0.391$.) Thus, it appears that using the computer tutor as the prime problem solving experience leads to essentially the same outcomes as handwritten homework.

The finding that there is essentially no difference between using the computer tutor and written homework may be the result of compensating factors. Students using the tutor have the benefits of receiving immediate feedback and sufficient guidance and hints in the cases of errors that they

usually complete all problems. On the other hand, students using the tutor have the disadvantage of not gaining comfort in the modality they are tested on. However, the study by no means reflects the full extent of student practice. Indeed, as pointed out above, all instructors assigned additional practice problems to students, which were carried out with pencil and paper. Students in the experimental group may have later solved additional problems with the computer tutor.

The only difference for the final exam truss problem that may be significant is not between homework conditions, but between instructors. We grouped the control and experimental sections together for each instructor and used the Tukey Simultaneous Test for Differences in the Means to compare each pair of sections (there are 10 such comparisons). The only difference that was significant at the 95% confidence level was between Instructor 2 and Instructor 3; the probability that this difference is due to random chance is $p = 0.03$. It is not clear if sections taught by Instructor 3 received less additional practice than those of other instructors, but we know these students otherwise performed comparably well in the course overall.

The role of the homework condition may be a modest contributor in any event. Student grade point average (GPA) upon entering statics was available. A scatter plot of average score on all non-truss problems versus GPA is shown in Figure 6. The Pearson correlation between GPA and exam average is $r = 0.73$, which is rather high in educational settings. Thus, whatever contributes to obtaining good grades (e.g., intelligence, work patterns, motivation, etc.) explains over 50% (r^2) of the variability in student performance on course exams overall. It may be difficult for a narrowly targeted instructional intervention to overwhelm this effect.
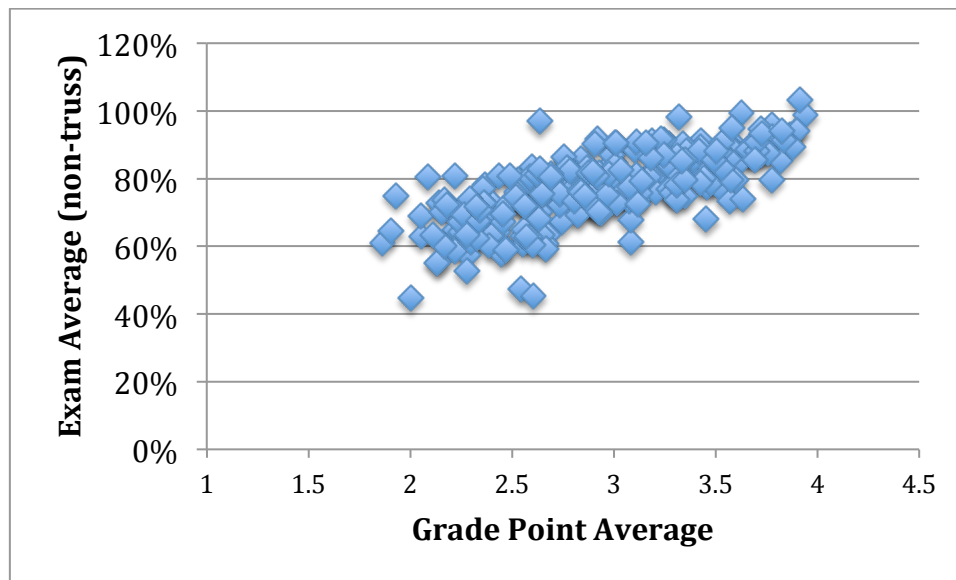

Figure 6. Scatterplot of exam average versus grade point average.

3. Summary and conclusions

A computer tutor had been developed previously to give students practice solving truss analysis problems. The tutor offers feedback to students on their errors, and guides them to completing most problems correctly. Data obtained from student use provided evidence that the feedback enabled to make fewer errors as they practiced. However, it remained to be seen whether use of the computer tutor enables students to learn material as well as the conventional solving of such problems by hand with pencil and paper.

This paper presents a study in which students solving problems with a computer tutor are compared with students solving the same problems by hand. Five instructors of statics, each teaching two sections, participated in the study. The two sections were taught in the same way, except during the week in which students solved truss problems on homework, students in one section (control) solved truss problems in the conventional way by hand, and students the other section (experimental) solved truss problems with the computer tutor. Students in different sections were compared both for how the performed on truss problems appearing on exams and on all other problems appearing on exams.

By looking at scores on all exam problems not dealing with trusses, it could be seen that there were no statistical differences between the students in the different sections. Further, we found no statistically significant differences between the experimental and the control groups with respect to performance on the exam problems dealing with trusses. Thus, it appears that the mode in which homework is initially carried out does not play a significant role in affecting the learning outcome. There may have been benefit of an unknown magnitude associated with additional practice problems, benefit that accrued equally to both groups. Still, even in a distance-learning environment in which students carry out homework by computer, for example, students may independently undertake additional practice problems, which can have benefits.

Acknowledgements

Bibliographic Information

1. Black, P. and William, D., *Assessment and Classroom Learning*, Assessment in Education, Vol. 5(1), 1998, pp. 7-73.

2. Lewis, M. W. and Anderson, J. R. (1985). Discrimination of operator schemata in problem solving: Learning from examples. *Cognitive Psychology, 17,* 26-65.

3. Bangert-Drowns, R. L., Kulik, C.-L., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61, 213-238.

4. Corbett, A.T. and Anderson, J.R. (2001). Locus of feedback control in computer-based tutoring: Impact on learning rate, achievement and attitudes. *Proceedings of ACM CHI'2001 Conference on Human Factors in Computing Systems*, 245-252.

5. Roselli, RJ, Howard, L and Brophy, S. (2006). A computer-based free body diagram assistant. Computer *Applications in Engineering Education* 14: 281-290

6. Dannenhoffer, J. and Dannenhoffer, J. (2009). An online system to help students successfully solve statics problems. Proceedings of the 2009 American Society for Engineering Education Annual Conference and Exposition, Austin, Texas, June 2009.

7. Dollár and P. S. Steif, (2008). An Interactive, Cognitively Informed, Web-Based Statics Course, *International Journal of Engineering Education*, Vol. 24, No. 6, pp. 1229-1241.

8. P. S. Steif, L. Fu, L.B. Kara, "Development of a cognitive tutor for learning truss analysis", Proceedings of the American Society for Engineering Education Annual Conference & Exposition, Indianapolis, IN, June 2014

9. P. S. Steif, L. Fu, L.B. Kara, "The Potential for Computer Tutors to Assist Students Learning to Solve Complex Problems", Proceedings of the American Society for Engineering Education Annual Conference & Exposition, Indianapolis, IN, June 2014

10. Anderson, J. R., Boyle, C. F., and Reiser, B. J. (1985). Intelligent tutoring systems. *Science* 228, 456-468.

11. Koedinger, K. R., Anderson, J. R., Hadley, W. H., and Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education* 8, 30-43.

12. Koedinger, K. R. (2002). Toward evidence for instructional design principles: Examples from Cognitive Tutor Math 6. Presented at Proceedings of PME-NA XXXIII, Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education 2002.

13. Bonham, S. W., Deardorff, D. L., & Beichner, R. J. (2003). Comparison of student performance using Web- and paper-based homework in college-level physics. *Journal of Research in Science Teaching*, 40(10), 1050–1071.

14. VanLehn, K., Lynch, C., Schulze, K. Shapiro, J. A., Shelby, R. H., Taylor, L.,Treacy, D. J., Weinstein, A., & Wintersgill, M. C. (2005). The Andes physics tutoring system: Lessons Learned. *International Journal of Artificial Intelligence and Education*, 15(3) 1–47.