EVALUATING COURSE WEB SITES AND STUDENT PERFORMANCE

Jesse M. Heines, Ed.D.
University of Massachusetts Lowell
Dept. of Computer Science
Lowell, Massachusetts 01854
1-978-934-3634
heines@cs.uml.edu

ABSTRACT

Erwin Boschmann [1] listed four major issues pertaining to "the infusion of technology in higher education" in a recent posting on *The Technology Source*. His third issue is that "there is little existing research on learning." This paper addresses that issue by analyzing a course web site now in its fifth semester of operation and focusing squarely on its effect on student learning. It discusses research methods for evaluating the web site's effect on student performance and statistical methods for identifying performance trends. Data on multiple undergraduate courses is presented, comparing student performance before and after the web site was introduced.

Keywords

Course Web sites, evaluating teaching methods, electronic class support.

1. CREATING AND EVALUATING COURSE WEB SITES

The creation of an effective course Web site is a great deal of work. Regardless of the tools used, an educator who creates a Web site and makes it an integral part of his or her course soon discovers that designing and producing the initial Web site are only small parts of the entire task. *Maintaining* that Web site throughout an entire semester—keeping its content current and its links up-to-date—is a much more time-consuming activity, by far.

David Jordan [5] attempted to "create a class Web site that would have a measurable effect on the students' education and/or morale" in a course entitled *Making of the Modern World*. Based on student responses to an instructor-designed survey and examination results in six semesters in which the course was offered (only the last of which used a course Web site), Jordan concluded that "the course seems to be better, but it is not knowable quite how much better, and the difference probably is not much." Given his finding that "creating and maintaining the site roughly doubles the work involved in teaching the course," Jordan was clearly disappointed that this effort did not have any "demonstrable educational effect."

ITiCSE '99 6/99 Cracow, Poland

Steve Gilbert [4] observes that many faculty, administrators, professional support staff, and students, though excited about new ideas and new uses of technology to improve teaching and learning, are also "frustrated and worn out with the amount of time and effort it takes to use technology responsibly and effectively." It is therefore critical, he concludes, to establish that such efforts bear fruit in measurable affective terms if not in measurable cognitive terms demonstrated by improved student performance.

Determining whether the use of technology "results in a greater understanding of the course content" [10] is also a great deal of work. William Trochim [9] *et al.* have investigated "a variety of methods for evaluating Web site development and use and the effects of the World Wide Web on the people who use it." Their findings, however, are less than conclusive.

Fitzelle and Trochim [3] conducted a student survey to measure students' *perceptions* of their own learning, but one cannot attach statistical significance to the results in terms of the Web site's effect on actual student performance. It is important to reiterate that Fitzelle and Trochim do not claim to have measured statistical significance. Rather, the fact that they did not attempt to do so may be an indication of the difficulty of doing so.

The task of evaluating whether the technology improves learning is so difficult, in fact, that some have taken a markedly pessimistic view toward the value of such research and the types of questions that it can answer. For example, in commenting on questions such as "what do computers teach best?", Stephen Ehrmann [2], who must certainly have seen his share of educational research as a program officer for the Fund for the Improvement of Postsecondary Education (FIPSE), has stated, "I don't think [those questions] *can* be answered in any reliable, valid way" [emphasis in original – JMH].

Others are more optimistic. In a widely quoted analysis of numerous studies on the effectiveness of computer-based instruction, Kulik and Kulik [6] found that this technology yields, on average, a 20% improvement in learning outcomes and speed. Even the more pessimistic Ehrmann believes that while the large global questions may never be answered, much is to be gained from evaluating "not so much what we can discover about the average truth for education at all institutions, but more what we can learn about our own degree programs and our own students"

[2, *op cit*, p. 27]. This is the more localized view taken in this study.

2. AN EXTENSIVE COURSE WEB SITE

The course home page begins with course-specific announcements. Further down is a list of all course assignments with links to their individual write-ups, a list of all lectures with links to detailed lecture notes for each class, and links to files they can download containing code discussed in class, code to be used for assignments, and useful utilities such as WinZip.

Students indicated in a survey that the availability of detailed lecture notes for each class was the most valuable aspect of the course Web site. The maintenance of lecture notes proved to be a significant time for the professor. Notes were typically posted the day before each lecture, and students often came to class with a printout of the notes. Some claimed that this helped them better follow the lecture and others said that having their own copy made it easier to take notes right on the same sheet.

After the lecture, the professor updated the previous lecture's notes as he prepared the next lecture's notes to accurately reflect the exact material covered. This was seldom the same as what was originally planned and laid out in the notes due to shifts in response to student questions and other unforeseen factors. The notes then became a sort of journal of the dynamic flow of the course, and students reported that they found the revised on-line notes valuable in preparing for exams.

When the course Web site first "went public" in fall 1996, students didn't "hit" it as often as expected. The site had less functionality at that time, but it contained the same basic content—lecture notes and assignments—as it does today. The professor therefore tried to think of "carrots" that he could add to the site to encourage students to hit it. His first idea was an Anonymous Suggestion Box that allowed students to comment on the course without identifying themselves.

This feature was not heavily used. The professor received no more than a half dozen anonymous e-mail messages in any one semester. Perhaps students didn't feel that they needed to comment anonymously, but more likely that their main reason for sending e-mail was usually to get assistance, and the professor was unable to respond to them if they sent anonymous e-mail. Indeed, the professor received and replied to more than 400 personal e-mail messages from students over the course of the spring 1998 semester, and the Anonymous Suggestion Box ranked number 7 out of 8 on the survey in students' rankings of the value of various Web site components.

The professor's second "carrot" idea was to create a CGI program that allowed students to view their grades on all course assignments and exams so that they would know exactly where they stood in the course at all times. This feature turned out to be extremely popular with students, and they ranked its value as number 2 out of 8 on the survey, surpassed only by the lecture notes.

3. STUDENT GRADE ANALYSIS

3.1 Analysis of Final Student Grades

Beginning in 1994, the same professor had taught the course under study during four successive fall semesters. The course was taught without a course Web site in 1994 and 1995, and with a course Web site in 1996 and 1997. Except for the addition of the course Web site and the updating of some of the technical material, the course was essentially the same in each of the four semesters. A box-and-whisker quartile chart showing the distribution of final student grades in each of the four years is shown in Figure 1.

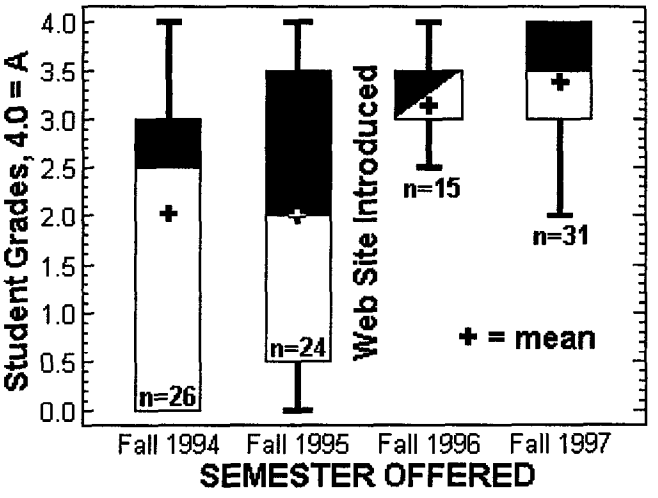

Figure 1. **Final Student Grades for the Same Course Over Four Years, Before and After the Introduction of the Course Web Site**

Analysis of variance results on the data in Figure 1 are presented in Table 1, and analyses of their ranges are shown in Table 2. These tables confirm what the professor had hoped: that there was a statistically significant difference in student performance after the Web site was introduced into the course.

Source	*Sum of Squares*	*Deg. of Freedom*	*Mean Square*	*F Ratio*	*P Value*
Between groups	40.70	3	13.56	10.26	1.0E-4
Within groups	121.65	92	1.32		
Total	162.35	95			

Table 1. **ANOVA Results of Final Student Grade Data Presented in Figure 1.**

Statlets Interpretation: Since the P-value of the F-test is less than 0.05, there is a statistically significant difference between the means of the various groups at the 5.0% significance level.

Contrast	*Difference*	*+/- Limits*	*Significant*
Fall 1994 - Fall 1995	0.038	0.646	
Fall 1994 - Fall 1996	-1.128	0.740	*
Fall 1994 - Fall 1997	-1.348	0.607	*
Fall 1995 - Fall 1996	-1.166	0.751	*
Fall 1995 - Fall 1997	-1.387	0.620	*
Fall 1996 - Fall 1997	-0.220	0.718	

Table 2. **Analysis of Ranges of Final Student Grade Data Presented in Figure 1.**

Statlets Interpretation: * = statistically significant difference at the 5.0% significance level using Fisher's least significant difference procedure.

3.2 Analysis of the Probability of Success

On April 16, 1997, the author had the opportunity to discuss this research with Dr. Ben Shneiderman, the University of Maryland's renowned Human Factors researcher. While lending his insights to the topic and his advice on data analysis, Shneiderman suggested that the author research analytical techniques for focusing on the performance of the lower half of the class, that is, the poorer students. Shneiderman's reasoning was that [8, paraphrased by author] "good students will learn regardless of the instructional techniques employed. Poor students, however, may show marked differences in performance when something like a course Web site is made available as an enhancement to traditional classroom instruction."

Given this direction, the author enlisted the help of Dr. Shelley Rasmussen to devise such an analysis. Rasmussen [7] suggested computing an estimated "probability of success," defined as the ratio of the number of "successful" students to the total number of students. Assuming random sampling, the standard deviation of this ratio could be approximated by its standard error. Thus, if the estimated probability of success is

$$\hat{p} = \frac{number\ of\ successes}{total\ number\ of\ cases}$$

then the standard error of $\hat{p}$ is

$$SE(\hat{p}) = \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}}$$

A traditional rule-of-thumb cautions that these formulae are "reasonable to use if the numbers of observed successes and failures is at least five." The problem, then, is to define what constitutes "success." Unfortunately, there was no single final letter grade at which the data could be cut so that at least five students fell into the "success" and "failure" categories in all four years. Setting the success criterion as "BC or Better" resulted in only 1 student in the "C or Worse" category for 1997. Setting the criterion as "B or better" resulted in only 3 students in the "BC or Worse" category for 1997. And setting the criterion as "AB or Better" resulted in only 3 students in the "AB or Worse" category for 1994. Thus, statistical significance on this measure cannot be established on this dimension for the data used in this study. Nonetheless, Figure 2 shows these data presented using a success criterion of "B or Better" to provide a feel for the type of analysis that Shneiderman suggested, and one that would be reasonable to use on larger samples.

4. CONCLUSIONS AND DIRECTIONS FOR FURTHER STUDY

Evaluation of a course Web site is clearly an inexact science, but it can be seen that analysis is possible within limited parameters and when the scope of interpretation of results is well-defined. Work such as that by Trochim [9, *op cit*] may shed light on which analysis techniques are better than others for specific purposes, but it will remain difficult to control all variables that can influence student performance, particularly in university core courses.

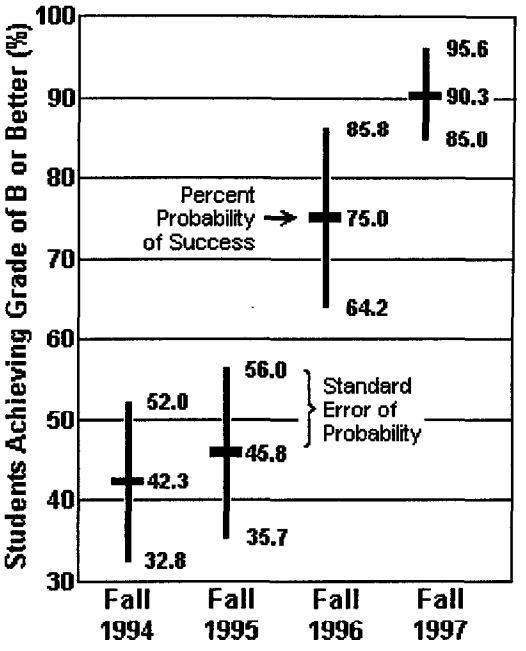

Figure 2. **Probability of B or Better Grade Success (same course over four years)**

Nonetheless, the author is encouraged not only by his students' positive feedback on the value of course Web sites, but also by the positive effects on their performance indicated by the results of statistical analysis. While one cannot generalize these results to all courses taught by all professors, one can surely state that contrary to the conclusion drawn by Jordan [5, *op cit*], creation and maintenance of the course Web site is indeed worth the extra time and effort that it entails.

Analyses on data from additional courses taught by the professor will be conducted in the coming months. In addition, historical data on student performance is available for course 91.101, Computing I, which has been taken by over 200 students per semester over the last few years and has historically high failure and drop-out rates. This course has been enhanced by the availability of a course Web site for the first time in the current fall 1998 semester, and analyses of student performance results will be available in time for reporting at the conference in June 1999.

5. ACKNOWLEDGMENTS

All data analysis was performed using Statlets, a statistical package available from NWP Associates, Inc. of Princeton, NJ, and provided free of charge to academic institutions. Please see http://www.statlets.com/.

Dr. Shelley Rasmussen, Dept. of Mathematics, University of Massachusetts Lowell, provided invaluable guidance on the statistical analyses used in this study.

6. REFERENCES

[1] Boschmann, Erwin. Moving toward a more inclusive reward structure. The Technology Source, October 1998. Paper posted at URL http://horizon.unc.edu/TS/development/-1998-10.asp and http://www.microsoft.com/education/hed/articles/facnov98.htm.

[2] Ehrmann, Stephen C., 1995. Asking the right questions: What does research tell us about technology and higher learning? Change 27(2):20-27, March/April 1995.

[3] Fitzelle, George T., Jr. and William M. K. Trochim. Survey Evaluation of Web Site Instructional Technology: Does It Increase Student Learning? Annual Conference of the American Evaluation Association, Atlanta, GA, November 1996. Paper posted at URL http://trochim.human.cornell.-edu/webeval/webques/webques.htm.

[4] Gilbert, Steven W. Lazy students & faculty? Posting to the Teaching, Learning, and Technology list processor of the American Association for Higher Education (aahesgit@-list.cren.net) dated February 5, 1998, message number AA-HESGIT 2.

[5] Jordan, David. Evaluation of a web site as a teaching tool in MMW-1 (Track C), Autumn, 1997. Paper posted at URL http://www.uscd.edu/~dkjordan/mmw/mmweval/mooeval97.-html.

[6] Kulik, Chen-Lin C., and James A. Kulik. Effectiveness of computer-based instruction: An updated analysis. Computers in Human Behavior 7(1-2):75-94, 1991.

[7] Rasmussen, Shelley. An Introduction to Statistics with Data Analysis. Brooks/Cole Publishing Company, Pacific Grove, CA, 1992.

[8] Shneiderman, Ben. Personal conversation. October 16, 1997. See http://www.cs.umd.edu/~ben/ for information on Shneiderman's work.

[9] Trochim, William,. Evaluating Websites: Introduction. Annual Conference of the American Evaluation Association, Atlanta, GA, November 1996. Paper posted at URL http://trochim.human.cornell.edu/webeval/intro.htm.

[10] Williams Glaser, Candyce. Creating new standards for higher education: effecting pedagogical change in the undergraduate curriculum through the integration of technology. Unpublished Doctoral Dissertation, Peabody College (of Education and Human Development), Vanderbilt University. Quote from a posting to the Teaching, Learning, and Technology list processor of the American Association for Higher Education (aahesgit@list.cren.net) dated February 25, 1998, message number AAHESGIT 20.