Do Online Students Perform as Well as Lecture Students?

JOHN DUTTON  
College of Management  
North Carolina State University

MARILYN DUTTON  
School of Business  
North Carolina Central University

JO PERRY  
College of Engineering  
North Carolina State University

ABSTRACT

This paper reports research on whether online delivery performs as well as traditional lecture delivery for a computer science course at North Carolina State University. The comparisons made are for two large sections of the course for which almost the only difference was that one section attended on-campus lectures and the other did not. Where significant differences in outcomes appear for students who completed the course, they favor the online students. However, online students who started the course were less likely to complete it.

I. INTRODUCTION

The increasing access to the Internet and the World Wide Web has expanded the variety of media by which universities are able to offer distance learning opportunities. Universities have offered classes via television and video for decades. More recently courses and whole degree programs are being developed for delivery via the Internet.1,2

Online classes have many of the advantages of other forms of distance education. These classes make the university more accessible to mature students returning to school to update their current skills or acquire new ones.3 In addition, online classes provide greater flexibility to students who benefit from being able to control the time, pace and order in which they study the course materials.4,5

At the same time, online classes have at least as many disadvantages as other forms of distance education. Replacing on-campus face-to-face lectures with computer-based materials removes the structure provided by weekly class meetings, allowing less disciplined students to fall irretrievably behind in class work. Traditional lectures and class meetings also provide motivation, support and encouragement from the instructor and fellow students that online students may not receive.6 Because they are less likely to come to campus, online students may not get help and mentoring from faculty. Moreover, unlike televised or video formats, online classes typically do not allow students to see the instructor deliver the lecture and respond to students’ questions. Many people, especially faculty, worry that these disadvantages are likely have an adverse effect on student performance in online classes.

This paper addresses these concerns by comparing student outcomes in an online version and a face-to-face version of an introductory course in the computer language C++. First, we find that for those who completed the course, students in the online version did significantly better than students in the lecture version of the class. This is true even when we adjust for differences in effort and maturity. Second, the online class had a higher dropout rate. However, this difference is not statistically significant when we control for differences in effort and maturity level.

II. BACKGROUND

While the use of online courses in the university is relatively new, many studies have examined student performance in other forms of distance education. For the most part, the evidence indicates that students are not adversely affected by distance learning (see Russell for a large set of references).7,8 Recent studies by Hiltz,9 Davis,10 and Smeaton and Keogh11 are particularly relevant to this paper. Hiltz reports on student response and outcomes in classes at New Jersey Institute of Technology that employed one of three different modes of delivery: the traditional face-to-face or video lecture format with no online content, a face-to-face lecture with an online component, and a video lecture with an online component. Students generally responded favorably to the online component of the classes and evaluated the educational experience gained from online classes as equal to or better than that of traditional modes of delivery. A closer examination of student performance in the introductory computer programming class indicated that while completion rates tended to be lower in the online classes, students did better (more A’s and B’s) in the online version than in the traditional version of the class.

Davis examines student performance in Japanese language classes that used one of three modes of lecture delivery: on-campus class, audiographics, and satellite. Although the delivery modes differed, all students heard the same lecture at the same time and could participate in classroom discussion. Using examination scores as the measure of student performance, Davis finds a statistical difference between the audiographics group and the on-campus group in only two out of eight semesters. In one semester the on-campus group scored higher, in the other the audiographics group scored higher. During the five semesters that the satellite course was offered, there
was no statistical difference in scores between the on-campus group and the satellite group.

Smeaton and Keogh describe the application of technology to create an online version of a course in database management. They use scores in an end-of-course exam to evaluate student performance in each of the two modes of delivery. Comparing scores in the 1997 online version with those in a 1995 traditional lecture class, they find no significant difference between the two.

The Hirz and Davis studies, like most of their counterparts, looked at student performance in forms of distance education that allowed the student to view the lecture (via television or video) delivered in a classroom setting. Relatively little research has been published comparing distance education to traditional education when the distance mode is online web delivery. Online delivery is the subject of the Smeaton and Keogh study, but the examination of student performance is somewhat cursory and not the primary focus of the paper. Given the speed at which online courses are being prepared and offered at institutions across the nation, there is a strong need for further research on educational outcomes in those courses. Are the many doubts expressed about them truly warranted by the evidence?

This paper reports research on whether online delivery performs as well as traditional delivery for a computer science course at North Carolina State University. The comparisons made are for two large sections of the course for which almost the only difference was that one section attended lectures and the other did not.

III. THE CLASS SETUP

The data for this study are taken from Jo Perry’s two class sections of CSC114, Introduction to Programming in C++, which were taught at North Carolina State University in the 1999 spring semester. CSC114 is required for computer science majors, many engineering majors (such as electrical engineering) and some majors with technical components (such as management information systems). It is also required for the “Certificate in Computer Programming,” a program for students with undergraduate degrees who wish to retrain for jobs as computer programmers. The course enrolls approximately 2000 students per academic year.

The on-campus lecture sections of CSC114 meet in two 50-minute lectures sessions plus a three-hour structured lab once per week. Classes are large, typically with 200 or more students, but the labs, which are taught mainly by undergraduates, are limited to 23 students each. Students’ grades are based on the following scheme:

- 10% lab average
- 5% small lab program average
- 5% homework
- 30% three programming projects
- 30% three tests
- 20% final examination.

CSC114 went online in the fall of 1997, and it has been online every semester since. It is the most popular online course at North Carolina State University, attracting 150–200 students each semester. The online version has been made progressively more accessible to off-campus students. Initially, it consisted of online lectures and course materials, but students were required to attend labs and take the exams (3 during the semester and a final exam) on campus. In the spring of 1998, the lab was put online and online students could choose to take the lab either online or on campus. Then in the spring of 1999, online students who could arrange for secure testing facilities were permitted to take their exams off campus as well, thus enabling students to take the entire course online.

The course website is complete and organized around the following pages:

1. General Information: contains the instructor’s name, office hours, email address, as well as required and recommended texts, the grading scheme, cheating policies, and expected workload.
2. Bulletin: contains daily announcements that the on-campus instructor would typically make at the beginning of class. The announcements include such items as availability of programming projects, information regarding the next test, corrections to assignments, and adjustments in schedules.
3. Lectures: contains links to all lecture material. Lectures are organized into 51 “Lessons,” each designed to take approximately 20 minutes to complete. Each lesson has an introduction, explanation of the significance of the subject matter, examples, and applications. Several sound clips (for overviews, class commentary) and self-test questions with answers accompany each of the lessons.
4. Regular and Online Labs: shows the schedule of lab work for the semester. The Online Lab pages also contain all of the lab materials and exercises. (Online labs are automatically graded by WebAssign® , a homework and testing facility at North Carolina State University.)
5. Assignments: contains links to all programming projects and short coding assignments. This page also instructs the students how to access their short answer homework, which is also graded by WebAssign®.
6. Calendar: displays a complete, day-by-day schedule for the semester, showing when lessons will be covered as well as a list of dates for tests, projects, homework, and labs.
7. Study Aids: contains links to tutorials and old tests. It also suggests some problem solving strategies for successfully completing assignments.

The only prerequisite for CSC114 is E115, a one-hour course in how to use popular applications on the campus workstations. This prerequisite is strictly enforced for students in the on-campus lab, but not for those in the online lab since online students are assumed to have access to personal computers and be able to use them effectively. In order to participate in the online lab, students must purchase and install MetroWerks CodeWarrior®, a multi-platform integrated C++ programming development environment. In addition to E115, there is also a calculus co-requisite, but it is only casually enforced.

In most respects, the online version and the lecture version of CSC114 are virtually identical. The primary difference between the two is the lab, which necessarily varies according to the computing platform being used (Unix for lecture students, PCs or Macintoshes for online). However, online lab materials parallel the lecture lab manual. In other respects, the two versions differ in only small details. The lecture students use the same website as the online students. They hear the same lecture as that presented in the website lessons. The two groups have the same homework and programming assignments and take the same tests; and both have access to a large bank of lab instructors who hold regular office hours. However, the online lab instructors typically interact with their students via email. Students in
the lecture section have the same access as the online students to the website but they cannot join the online listserv, which was established to facilitate communication among the online students. Online students are allowed to attend the on-campus lectures (which usually provide additional examples and commentary on the lecture notes), but they rarely take advantage of this extra benefit.

Programming assignments constitute a substantial part of student grades. Assignments must be submitted electronically and each has an electronically set due time which makes it impossible to submit “late” work without instructor intervention. Since the Internet connections are not totally reliable (with server and ISP failures, for example), the online lab instructors are more generous in accepting “late” work than the lecture lab instructors. It is the responsibility of the online lab instructor to decide whether a student’s failure to submit work on time was the fault of the student and act accordingly.

Since the inception of CSC114 online, Jo Perry has taught both the online section and a lecture section of CSC114 each semester. Her casual observations from the start were that the performance of online students on tests was significantly better than that of the lecture students.

IV. EMPIRICAL RESULTS OF STUDY

In order to evaluate the online delivery mode, we took the approach of comparing student success rates in the online and lecture versions of the class. We used conventional measures of success at the university level: grades and whether the student completed the class. We relied on two different indicators for grades. One was the score on the final examination. The other was a modified course grade (hereafter just “course grade”), which was computed using the relative weights defined in the syllabus for four of the course components: short labs, tests, final exam, and projects. These weights were the same for the two sections. The second measure of student success, completion of the course, was determined by whether or not the student took the final exam.

We began with a group of 345 students still registered when the drop period ended approximately one month after the first day of class. Of those, seven officially withdrew before the end of the semester and 14 were registered pass/fail. We removed both groups from the sample, the first because we had no information about why they withdrew and the second because their grades were not comparable to those of the normally graded students. We also removed the 12 auditors. This gave us a final sample of 312 students. Of these, 272 took the final examination and 40 did not. Failing to take the examination resulted in an F grade.

Table 1 shows the distribution of students in the two sections. More than 80% (258) of the sample were students officially enrolled in an undergraduate degree program at the university. The remaining group (lifelong students) contained students of all other classifications including non-degree and post baccalaureate students and students working toward the certificate in computer programming, as well as a few graduate students. Although they were typically discouraged from taking the class online, more than a third of the undergraduates opted for the online version. This compares with 87% of the lifelong students who took the class online.

To compare the two groups, we first look at student grades in the two sections. Our first hypothesis is that, for the 272 students who completed the course (took the final exam), online students did as well as lecture students. The hypothesis is well supported by the data. In fact, the data support a finding that online students did better than lecture students in some categories. Table 2 contains mean scores and mean differences in scores for the whole sample as well as a break down for undergraduate and lifelong students. The table also contains a p-value for each mean difference. The p-value indicates the likelihood of obtaining a difference as large as that observed if it occurred simply from randomness in the data. A low p-value implies that we would probably not observe such a large difference from purely random data and the difference must be the result of some systematic effect. By convention, we usually label any difference with a p-value of 0.05 or less as meaningful, that is, statistically significant.

First, consider the statistics for the class taken as a whole. For the whole sample, the online students scored 9.6 points higher on the Exam and 4.3 points higher on the Course Grade. The p-values for online-lecture Exam and Course Grade differences are 0.0000 (positive but less than 0.00005) and 0.0145. These are very low and well below the conventional cutoff point of 0.05, so the differences are statistically significant and we can state with a high level of confidence that online students did better than lecture students on both measures.

The results are slightly weaker when we break the sample of 272 into separate categories of undergraduate and lifelong students. Among undergraduates, the online students outperformed the lecture students by 7.5 points on the Exam and by 1.8 points on the Course Grade. The Exam difference had a p-value of 0.0024, well below 0.05, and so was statistically significant. The difference in undergraduates’ Course Grades in the online and lecture sections had a p-value of 0.3543. Because this p-value is so high, we cannot say with any assurance that the difference in Course Grades is meaningful rather than simply randomness in the data. Among undergraduates, then, it is clear that the online students clearly did better on the Exam but the difference in their Course Grade was too small to be statistically significant.
Among lifelong students, the online students scored 5.3 points below the lecture students on the Exam and 4.8 points below the lecture group on the Course Grade. In both cases the p-values were very high (well above 0.05), so neither difference was statistically significant. Consequently, we cannot say with any assurance which group did better among the lifelong students. We might ask why a difference of 4.3 was significant for course grade scores for the whole sample while the difference of 4.8 for the lifelong students was not significant. The reason is that the sample size for lifelong students alone was much smaller than for the total of all student and a given score difference is more significant if it is based on a larger sample.

Table 2 indicates a clear advantage for online students when we take the class as a whole, but the advantage is much less clear when we control for type of student by breaking the sample into two subgroups. Our next step was intended to gain a better understanding by considering other factors that might influence an individual student’s class performance.

For this next step we used multiple regression analysis. Regression fits a linear function to the data and allows us to test the effects of several variables together on class performance. The variables that we believe affect performance are three—online versus lecture status, undergraduate versus lifelong status (for maturity level), and level of effort in the class.11 The equation to be estimated is as follows:

\[ \text{Grade} = C + \beta_1 \text{Online} + \beta_2 \text{Undergrad} + \beta_3 \text{Effort} \]

Here Grade can be either Exam or Course Grade, C is a constant term, while \( \beta_1, \beta_2, \) and \( \beta_3 \) are coefficients indicating the impact of each right hand side variable on Grade performance. Online is a dummy variable equal to 1 for online students and 0 for lecture students. The Online coefficient captures the score difference between online and lecture students, holding the other variables constant. Undergrad is another dummy variable, equal to 1 for undergraduates and 0 for lifelong students. This dummy variable coefficient should reflect a difference in score between undergraduates and lifelong students, holding the effects of effort and online status constant. (Undergrad is an indicator of maturity. We also tried age as a variable to indicate maturity but, surprisingly, found no significant effects of age on grade performance.) To indicate the degree of effort, we used a homework grade equal to the number of correct responses out of 239 possible over the course of the semester. Since a student could correct wrong answers and resubmit the homework multiple times, a persistent student could get 100 percent of the homework correct. We believe this homework grade reflects effort pretty accurately, and certainly much better than under traditional circumstances in which homework can be attempted only once.

Table 3 shows estimates of the ordinary least squares multiple regression coefficients for the models explaining the Exam score and the Course Grade using the sample of 272 students who completed the course. The two regressions indicate several things. First, according to the coefficient for Online, students in the online section outperformed lecture students by almost seven points on the Exam and one and a half points on the Course Grade. These differences between online and lecture students are both statistically significant, as shown by the low p-values associated with the Online coefficient values (0.0011 and less than 0.00005). These results generally support the findings in Table 2 of higher scores for online students, at least when the sample is considered as a whole. An advantage of using the regression is that we have controlled for two other factors that might have influenced performance, homework effort and undergraduate versus lifelong status. The grade advantage for online students stands up even when controlling for those two factors.

A second piece of information from the regressions is that undergraduates scored almost 8 points lower than lifelong students, holding the other explanatory variables constant. The low p-values associated with the Undergrad variable coefficients show that this effect is statistically significant and not simply the effect of data randomness.

The third result is for homework effort. We can see from the two regressions that correctly completing one more homework problem (out of a total of 239) led to a grade increase of 0.16 point. This is true for both the Exam and the Course Grade. In each case the p-value associated with the coefficient is very small (less than 0.00005), so we can be confident that this measure is statistically significant.

To summarize, the regressions demonstrate convincingly that online students scored higher, undergraduate students scored lower, and homework effort significantly raised performance on the Exam and Course Grade measures.

Note: we did consider a phenomenon that might have biased our results in favor of online students. Were there many students who effectively dropped the course (that is, quit doing the course work) but nonetheless took the final exam? And if so, were they predominantly in the lecture section? (Because they were predominantly on-campus, the cost to the lecture students of taking the exam might have been lower.) These effective drops who still took the exam might have dragged down the lecture average disproportionately. However, when we checked for students who had effectively dropped out but who nevertheless took the exam, we found only three clear cases, and one of these was an online student. Therefore, we decided that this phenomenon would have no important effect on the results.

Because we are interested in student success, we also examine another important measure of performance, completion rates for the course. If students drop out, they do not successfully complete the course. Table 4 indicates the percentage of students who
completed the course, as well as tests of differences in completion rates for the online and lecture sections. We disregarded students who left the class during the drop period. Students can drop a course freely during this period, and often do so because they have not decided to commit themselves to course completion. It seemed more appropriate to consider only students who had made the strong commitment of staying enrolled beyond the drop period. Of the 312 students still enrolled after the end of the drop period, 87.2% completed the course (took the final exam). The overall average masks a substantial difference between the online and lecture groups. For the online students the proportion completing the course was 79.4%, while for lecture students it was 93.6%. A test of whether the difference is significant gives the very low p-value of 0.0001, indicating that lecture students completed the course at a significantly higher rate than online students.

When we break the whole sample into undergraduate and lifelong sub-samples, we still find that for both groups online students were less likely to finish. However, among undergraduates the difference in percentages completing the course was smaller (only 9.9%) but still statistically significant (with a p-value of 0.0050). The difference for lifelong students was larger at 15.5% but not statistically significant (because of the smaller sample size). Its p-value was 0.1965. Note that neither of these group comparisons takes account of the degree of student effort.

We can only speculate as to why fewer online students completed the course. Many of the lifelong students are non-degree and part-time students, and the consequences for them of not completing a class may be lower than for undergraduates enrolled in a degree program. This difference is reflected in the lower completion rate for the lifelong group, 72.2% compared to 90.3% for undergraduates, a statistically significant difference. The difference in completion rates for these two groups may have affected the comparison of the online and lecture completion rates. Lifelong students made up a third of the online group but less than 5% of the lecture section. Their lower completion rate would much more heavily affect the results of the online section.

To delve further into the determinants of course completion, we used logit regression to estimate the effects of online status, undergraduate versus lifelong status and effort together on the probability of completing the course. The coefficients from a logit regression indicate the effects of the explanatory variables on the log of the odds (e.g. one-to-four or three-to-two) of completing the course. From these coefficient values we can also derive the effects of the explanatory variables on the probability (e.g. a 10 percent chance or 80 percent chance) of completing the course. Even though these probability effects must be calculated in an extra step, we focus on them because they are intuitively easier to interpret. The probability effects are not constant since they depend on the values of the explanatory variables. We follow the convention of using the mean values of the explanatory variables in the computations.

There are three sets of values reported in Table 5: each variable’s logit coefficient, the p-value associated with that coefficient, and the variable’s effect on the probability of completing the course. For example, the values in column one indicate that the odds of completing the course are higher for undergraduates than for lifelong students, and the low p-value means that the difference is statistically significant. The 0.0828 probability effect implies that the probability of finishing is 8.28% higher for undergraduates than for lifelong students.

Online status in Table 5 has negative coefficients and negative probability effects in all three regressions. The negative probability effects imply that online students have a lower likelihood of finishing, regardless of whether we look at the whole group or separate them by class status. However, the negative effect of online status is not statistically significant, as indicated by p-values of 0.1282, 0.1407, and 0.5597, all well above 0.05.

The nonsignificance of online status differs from the results of Table 4 because in the logit regression we have controlled for the effect of student effort on the likelihood of completing the course. The effect of the effort variable is positive in all three regressions. For lifelong students, correctly working each additional homework problem increases the probability of course completion by 0.70%. For undergraduate students the effect is 0.11% per problem. For the whole group the effect is 0.12% per problem. The effort coefficients have very low p-values in all three regressions, so effort is definitely a significant variable in predicting the probability of course completion.

The logit regressions support the interpretation that completion rates are influenced positively by effort in the course and negatively by lifelong status. No clear effect of online status emerges from these results.

V. Conclusion

A summary of our findings, then, is that, as a group, online students who completed the course generally did significantly better than lecture students. This result holds up when we consider undergraduate students only. For lifelong students the result is reversed, but the difference is not statistically significant. When we perform a regression estimation that controls for undergraduate versus lifelong status and for effort, we find that the online grade advantage holds up. We also find that, among those who completed the course, undergraduates did less well than lifelong students and homework effort had a strong positive effect on Exam and Course Grade.

Online students had a smaller likelihood of completing the course than did students in the traditional lecture section. This difference is significant if we consider only differences in proportions without controlling for differences in effort or undergraduate versus
lifelong status. When we use a logit equation that controls for these other factors, the negative effect of online status on probability of completion is no longer significant. It turns out instead that the low completion rate is associated with lifelong status and with low homework effort. The apparent association of low completion with online status came about because lifelong students take their courses disproportionately in online form. When we look at undergraduate students alone and control for effort, online status is not a significant determinant of completion rate.

Our results demonstrate that online students can perform at least as well as students in the traditional lecture setting. They do not show that any student, randomly selected, can do as well in an online class. Undoubtedly self-selection means that students with greater computer skills and/or greater maturity are more likely to opt for an online course. However, we have reduced the effects of those factors in this study. The subject matter of the course reduces variability in computer skills. And in the empirical work we have controlled for maturity differences to the extent we could. So our results provide support for further development of online course delivery. Our results also show that the completion rate is an important issue in providing online education, but that dropping out is a problem associated with lifelong status rather than online status.

An important caveat for all our findings is that we have studied only one course in one specialized topic. An important area for additional research is comparison of student outcomes in other types of courses to see how far our results extend.

REFERENCES


12. The website URL is: http://courses.ncsu.edu/CSC114/CSC114.html.
