Extending the Journey toward a Virtual Introductory Statistics Course

Joseph Wisenbaker
University of Georgia
Department of Educational Psychology
320 A Aderhold Hall
Athens, GA 30602
USA
joe@coe.uga.edu

Drawing from web-based materials previously developed to supplement on-campus offerings of an introductory statistics course for graduate students in education, the author made an initial offering of an on-line virtual class in the fall of 2001. Poor student performance and dissatisfaction with this initial course organization led to a deeper reading of the literature on online teaching in general and the online teaching of statistics in particular. A much greater degree of instructor supplied organization, direction and interaction were incorporated into another offering of the course during the spring of 2003. Final examination scores and course evaluations improved markedly over the first on-line offering but remained somewhat lower and more variable than the results from on-campus offerings of the same course. Suggestions for improvements are offered based on instructor observations and student feedback.

Introduction

What follows represents an in-depth case study of one effort to implement a virtual version of an introductory statistics course offered to masters and doctoral students enrolled in the College of Education at the University of Georgia. It starts with a review of the development of the on-line material, some background on the circumstances that led to the initial offering of a virtual section of the course, and the changes that were incorporated into a subsequent offering. Details on the operation of both virtual sections and a review of their outcomes are also presented.

Questions addressed flow from a comparison of students enrolled in the virtual sections with students enrolled in a face-to-face section taught by the same instructor. Of particular interest were issues of their initial comparability, their performance on the final examination, and their evaluations of the course.

1. A Brief Review of Literature

The literature related to distance education is certainly extensive. A search of ERIC using the phrase ‘distance returned over 8,500 articles (adding 'online learning' and 'e-learning' to this yielded about 3.5% more hits). An on-line search using Google returned over 850,000 hits. Much of this literature consists of either straightforward position papers advocating greater use of distance education or ‘how to’ articles setting forth ways to use different distance education technologies. However, there are a large number of articles which include explicit comparisons of distance and face to face courses in terms of student attitudes or end of course performance. In a fairly massive review of actual studies contrasting distance education with face to face courses, Phipps & Merisotes (1999) stated that distance education courses have been found to have both favorable learning outcomes as well as high levels of student satisfaction. In an in-depth comparison of distance education and face to face

courses offered by Nova Southeastern University in the fall of 1999 (involving 138 different sections of 34 different courses), Fredda (2000) concluded that undergraduates' final grades did not differ due to context and that graduate students in internet courses did better than those involved in campus-based instruction. In a meta-analysis of studies examining the issue of student satisfaction under distance and in-class formats, Allen, Bourhis, Burrell and Mabry (2002) found a slight student preference for in-class formats and little difference in satisfaction. However, they did note that distance courses delivered with either video or audio content had lower levels of student satisfaction than did distance courses where the primary medium was written.

The literature explicitly addressing the teaching of introductory statistics courses via distance education and focusing on student outcomes is, of course, considerably more limited although containing some very well done studies. Martens, Portier and Valcke (1995) compared outcomes from first year college students taking an introductory course in statistics using interactive learning (computer mediated), independent learning and face-to-face classes. They found no differences in terms of student performance on subject oriented mastery tests and end of year exams. Dereshiwsky (1998) reports nearly identical (and highly positive) instructor and course evaluations of introductory statistics taught face-to-face verses on-line to graduate students in education. In looking at test outcomes associated with electronic mail, traditional and combined approaches to instructional delivery, Kennedy and McCallister (2000) found no differences in terms of student performance on end of course tests. Stephenson (2001) examined student performance and attitudes comparing in-class and tape-delayed video versions of an introductory statistics course offered to GM technical education students (mostly managers and engineers). In looking at outcomes over a 10 semester span from 1994 to 1999, he found no overall differences in terms of student grades or attitudes toward the course (although the earliest offerings were associated with better attitudes for the in-class students).

2. Development of On-line Material

For more than 20 years I have been teaching a variety of introductory statistics courses (basic statistics, analysis of variance, multiple regression analysis, etc.) to graduate students predominately in our College of Education. Since 1997, when the University of Georgia adopted WebCT as its supported package for implementing web-based instruction, I have been placing materials online. Rather than seeing these efforts as steps toward creating a virtual course, my primary interest was in making my class more accessible to students who could not always attend. I initially put the standard package of course syllabus, homework assignments and old, practice exams onto my web pages for each of my courses. In that form, the student access statistics in WebCT indicated that very few students made much use of these materials. At that point what I was doing hardly qualified as "on-line" even though it satisfied my university's request for courses having an "on-line presence."

In 1998 my college installed a video projector and camera for displaying written materials in my classroom. I rapidly abandoned the use of my classroom's dry erasable white board in favor of using a stack of copier paper on which I wrote the key terms, their definitions and in-class examples previously committed to a traditional chalk board. Not only did this help me avoid the problem of students asking questions about material just erased, but I could then allow students to copy the notes from lectures that they missed.

After a term of making those notes available, in 1999 I acquired an inexpensive scanner and began posting those lectures through WebCT thereby allowing students unlimited access.

Subsequently, I tried using those notes to lecture from hoping to create a collection of relatively stable materials that would eliminate the need for much in the way of subsequent scanning. Unfortunately, it proved impossible for me to adequately control the pace of my delivery (I tended to cover topics too quickly), and I soon abandoned that practice in favor of writing my definitions and examples live for each lecture.

The next term, in 2000, I started learning about RealAudio production and server software, the basic versions of which proved to be free of charge. Armed with an inexpensive tape recorder and a wireless microphone, I began creating an audio track of my daily lectures. An hour of so of replaying the tape through my desktop computer's audio input port allowed me to create streaming media files that were linked to my on-line lecture note pages. Soon after that I abandoned my trusty tape recorder in favor of running the digital encoder software from RealAudio on the in-class PC and capturing my lecture soundtrack directly. It was but a matter of a few post-class minutes to upload the resulting file to my own computer thereby eliminating the time needed to replay the lecture to create the streaming media files.

Additional materials have been added to my sites to enhance their value to my students. In 1998 a grant funded by the University System of Georgia made it possible to create a small collection of research and evaluation papers focusing on educational research in my own state illustrating the actual application of various statistical concepts and techniques. An invitation to do a guest lecture in a research methods course focusing on statistical power led to creating a web page with links to a variety of on-line materials from other universities and governmental agencies. Old exams and answers were also posted there.

Along the way, a handful of students opted to complete the requirements for several of my live courses using only the on-line materials. Located at some distance from the university, those students simply faxed me homework assignments and completed exams. While reporting that keeping to a schedule for working with the on-line materials required considerable self-discipline on their part, those students were just as successful in being able to respond to final exam questions as were students who regularly attend class meetings on campus.

3. Virtual Offering #1 (Fall 2001)

Responding to demands from one of our departments for an "extra" section of our introductory class to serve a cohort of students who could only physically appear on campus once or twice a month, we opted to create a virtual section of the class. It featured access to class materials through WebCT with the only lectures available being those which were already posted from the previous fall. Students had access to the instructors to get answers to their questions via e-mail or telephone and could post messages to the course discussion pages for broader interaction involving instructors and students.

Given the circumstances surrounding the course, the two "in-class" exams were reconfigured as open-book tests and the final, while held on-campus, permitted students access to their text and notes. In previous sections of the course none of the exams were open-book. Instead, students had been allowed access to self-generated "guides" in which they were encouraged to include formulas they deemed important, examples of worked problems, or other material they wanted to include. In those classes students were allowed access to their text only for tables they might need.

Prompted by student requests, a live review session for the course was held on a

Saturday afternoon about three weeks ahead of the final exam. Nearly all of the students attended with questions focusing on the circumstances of the final exam and specific problems with assigned homework exercises.

4. Virtual Offering #2 (Spring 2003)

The second on-line offering of the course (just over a year after the completion of the first) saw many changes in the structure of the class. While the primary 'lecture' material still consisted of the same prerecorded lectures that were made available to the first virtual section, those lectures were made available on a completely self-contained CD that required no use of WebCT for access. Interactions between student and instructor were much more highly structured. Once each week students were required to log into a one-hour plus, on-line discussion session operated through HorizonLive, a new system intended to support live on-line instruction (see http://www.horizonlive.com). Those sessions involved multi-way audio and an on-line whiteboard on which the instructor could write short answers to questions and work problems as requested by the students. The discussion sessions consisted primarily of question-and-answer sessions with a small amount of impromptu lecturing by the instructor. The instructor was available at other times to respond to telephone or e-mail requests from students.

As with the first virtual offering, the mid-term exams for this section were to be completed as take-home exams; however, students were required to complete and submit the same kinds of exam notes as in my live classes with the ground-rules for the tests being that students could access only their tables and those notes. The final exam, taken on-campus on a Saturday, was completed under the same restrictions.

5. Characteristics of Students

Those who enrolled in the first virtual section of the course were nearly all doctoral students in our Department of Adult Education in a program specifically targeting technical college administrators. While many degree seeking students in our college's graduate programs are full time teachers or school administrators, nearly all of the students in the virtual section held full time positions with considerable responsibility in their home institutions and were taking a full course load as part of their program requirements. While the students in the second virtual section were not predominantly drawn from the adult education program, most but not all were working full-time and lived at some distance from our campus.

The students in both virtual sections had to face the same rigorous admission standards applied to any other applicant to our graduate programs. In fact, they compared quite favorably to the students who had been enrolled in the previous live, face-to-face section. Table 1 contains statistics related to their performance on the Graduate Record Examination, the GRE (a major screening tool used my most selective graduate programs in the United States). While the students in the first virtual section had marginally higher verbal and quantitative scores than those in the live section, neither the means nor the standard deviations were statistically significantly different from one another (α=.05). The students in the second virtual section had marginally higher verbal and analytic scores than those in the in-class section, but, once again, neither the means nor the standard deviations were statistically different. This makes it somewhat reasonable to treat these students as reasonable similar in initial standing even though random assignment was not involved.

Because the students enrolled in the first virtual section had taken at least a year of coursework together, they all should have been used to the demands of doctoral courses and have a sufficient level of "community." Likewise, the fact that they had already taken several courses making extensive use of WebCT (though with a number of face-to-face half-day meetings per course), suggested that they had enough on-line experience so that they could successfully interact in that context. Unfortunately, nearly all of the use they made of e-mail and discussion features of WebCT revolved around issues entirely tangential to the course.

The students in the second virtual section were drawn from programs throughout the College of Education and, for the most part, did not have much experience in taking an on-line course. Likewise, most of those students were unacquainted with the other students enrolled. The only known exception to this involved a trio of students who worked together at a site at some distance from the campus.

Table 1: Student GRE Performance

		Analytic		Quantitative		Verbal	
Section	N	M	SD	M	SD	M	SD
Fall 2000 (Live)	22	547.27	94.27	536.82	72.79	491.82	92.36
Fall 2001 (Virtual #1)	27	527.04	101.82	540.74	103.33	537.41	73.93
Spring 2003 (Virtual #2)	15	553.33	141.57	531.67	118.39	519.17	120.71

6. Course Outcomes: Final Exam Performance

Because of major differences between the real and virtual versions of the course in the administration of "in-class" exams, performance on the final exam was the primary course outcome of interest. Since the exam questions used were unchanged from the ones administered to students in the face-to-face version of the course, performance for that class was compared to that from the virtual sections of the course.

The students in the on-campus section had a mean score (out of 100) of 84.4 with a standard deviation of 7.5. Those in the first virtual class had a mean of only 72.6 with a standard deviation of 20.9 while students in the second virtual class had a mean of 80.3 with a standard deviation of 13.1. Even with these small sample sizes, a one way ANOVA applied to this data yielded statistically significant results ($F(2,58) = 3.82$, $p = .03$) as did a test for homogeneity of variance (Levene Statistic=18.23, $p = .000$). When the Games-Howell procedure (which doesn't assume homogeneity of variance) was applied to look at post-hoc comparisons among the means, only the difference in exam scores was between the live section and the first virtual class was statistically significant ($p = .03$).

Looking at a breakdown of performance in greater detail, only 4.3% of the students in the face-to-face section obtained scores less than 70 while 42.9% of the students in the first virtual section did so as did 15.4% of the students in the second virtual section. At the same time, more students in the virtual sections had scores greater than 90 (25.0% in the first virtual section and 30.8% in the second) than did students in the live section (17.4%). Thus while a far greater proportion of the students in the virtual section performed quite poorly on the final

compared to what might otherwise be expected, a greater than expected number did quite well indeed (see Figure 1).

7. Use of WebCT and its Relationship to Performance

One of the interesting features of WebCT is that it tracks student hits on course material and posts to the discussion page. Because of the way students were given access to lecture materials in the most recent on-line offering, it only makes sense to look at this for the students enrolled in the first virtual section. It was noted that while most students made considerable use of the web material (mean hits = 430), there was but a weak, but a non-significant, positive correlation between the number of hits and performance on the final (r=.20, p=.33). At the same time, there was a weak but negative relationship between posts to the discussion page and performance on the final (r=-.36, p=.07).

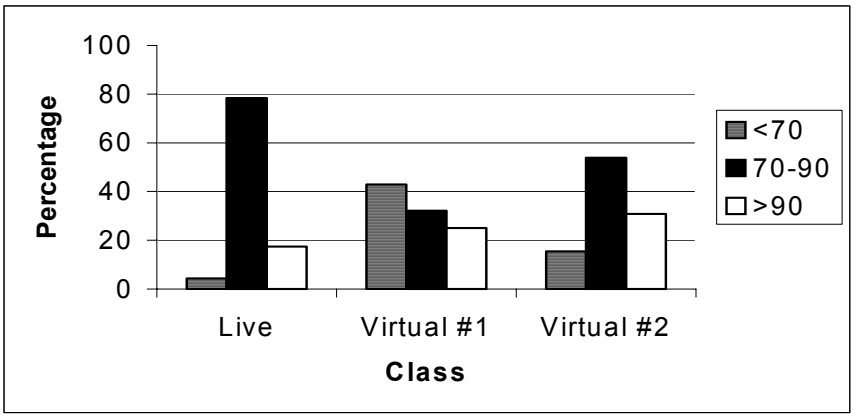

Figure 1: Performance on the Final Exam

The content of the discussion messages (of which there were 140) mostly revolved around social interchanges among the students (focusing on concerns for ill family members, sharing news about grandchildren, planning for a post-course social). Only 27 of the messages actually focused on the course itself (mostly dealing with anxiety about learning the material or dissatisfaction about the way in which the course was being offered) with but 2 of them actually relating to some particular assigned problem. Hence the use of the discussion facility seems to have had little use in the actual learning, but maybe important for maintaining group cohesion.

8. Course Evaluations

Student course evaluations are required in each section of a course. Those evaluations include ratings (on a 1-5 scale with higher values reflecting more favorable ratings) for both the course and the instructor. The overall course ratings averaged 4.4 under the face-to-face section, dropped precipitously to 2.8 in the first virtual section and bounced back somewhat to 3.1 in the second virtual section. At the same time, ratings of the quality of the instructor declined from a mean of 4.7 under the face-to-face section to a mean of 3.2 in the first virtual section with a rebound to 3.9 in the second virtual section (see Figure 2). Lest it be believed that this might merely reflect a general decline in the performance of the particular instructor, ratings from the same semesters in other courses either improved or stayed about the same as they had been previously.

At the end of each of the virtual sections, students were asked to provide specific

feedback on how the course might be improved. Of those who responded in the first virtual section (n=21), almost 62% indicated a desire for there to be a large number of on-campus class meetings with specific attention to working problems, illustrating the use of statistical software and responding to student questions about specific problems. The overall theme underlying their responses seemed to be a clear statement of their need for a greater degree of structure and support with regularly scheduled interactions via the internet if not in person.

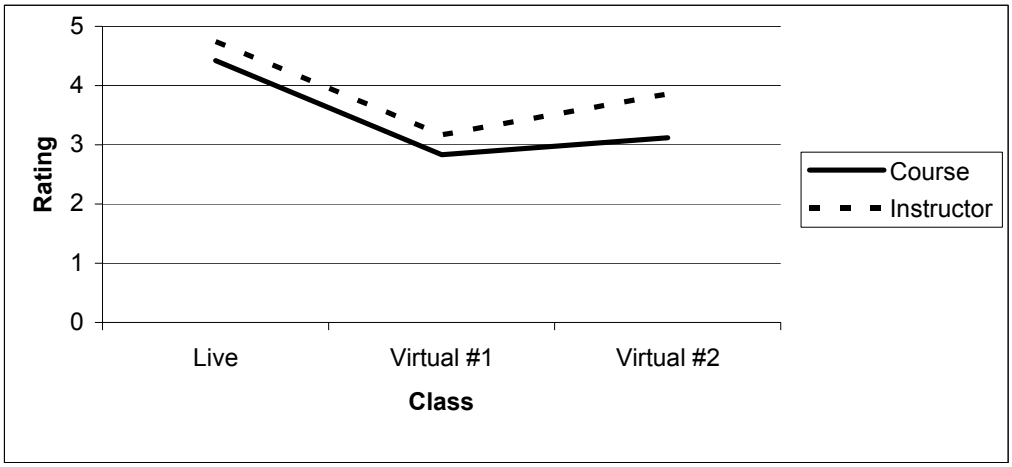

Figure 2: Evaluations of Course and Instructor

The responses from the students in the second virtual section were quite different from those in the first. Generally, there was more diversity in their responses and less commonality. When responding to the question about their biggest challenges in taking the course in an on-line format 3 mentioned time management/pacing issues; 3 wrote about the difficulties posed by a lack of consistency in the use of symbols between the instructor and the text; 2 voiced concerns about technical problems with HorizonLive; and 2 mentioned the need for more of a video type format for the recorded lectures. The balance of the challenges noted were mentioned by only one person apiece and ranged from 'needing more sequential assistance than provided by the discussions' to 'isolation'.

The kinds of changes they recommended were also very diverse. Four mentioned the need for even greater organization and explicit pacing through the homework problems; 3 wanted greater facilitation in the formation of study/work groups; 2 wanted the homework problems to be more like the exam questions with greater emphasis on determining what procedures would be appropriate; 2 wanted more explicit instruction in the use of their calculators; 2 wanted some face-to-face meetings; 2 wanted more time for the course; 2 wanted greater consistency in the use of symbols; and 2 wanted more differentiated courses to provide greater support for students needing it. The rest of the suggestions were mentioned by one person each.

9. Discussion

In spite of promising developments along the way and a population of students who would seem ideal for working in such an environment, the first virtual section of the introductory statistics course fell far short of the mark in many respects. While some students did quite well on the final exam, a large portion did poorly indeed. Overall performance was substantially below what had been observed in the live version of the same course taught by the same instructor. Students were clearly much more dissatisfied with the course and the

instructor.

What then might account for these unhappy results so at variance from what might be expected given what has been reported in the general literature about distance education and the more limited literature about teaching introductory statistics via distance education? A more careful reading of those reports and articles suggests several possibilities. First, the circumstances considered in everything I reviewed seem to strongly suggest that students participating in distance settings were very much self-selected. The students in my first virtual section had very little choice in the matter, and, despite my use of an on-line environment with which the students were all familiar, very much wanted a real in-class component in addition to what they were given.

The kind of shifts in student attitudes across multiple offerings of the same statistics course noted by Stephenson (2001) may very much relate to successive waves of students being better able to make informed decisions about participating in virtual courses. The relatively small number of students enrolled in the second virtual section who expressed a desire for more in-class meetings may be seen as consistent with that idea.

Second, apart from the timing of exams, there was really very little in the way of structure imposed on the students in the first virtual section and, from my examination of how they used the on-line discussion forum, it would seem that most of their own interactions had little to do with the content of the course. Their responses to the end of course questionnaire clearly indicated a strong desire for greater structure. This, to a considerable extent, tends to reinforce the notion that many of the students thrown into that section of the virtual class may not have been sufficiently self-directed to do well in such an environment.

The greater structure imposed on the students in the second virtual section would seem to have addressed a portion of those concerns. Certainly their performance was better overall although still quite variable. Likewise, their attitudes about the course seemed a little better but with several still indicating the need for greater structure and something more lecture-like in the weekly on-line sessions.

The shift in test taking conditions requiring that students create their own notes for use with exams and that they not be open-book may also have led to their better scores on the final. That approach has long been a staple in the statistics courses I offer on-campus under the belief that it promotes students to have greater engagement with the material than would be the case with open-book tests. Requiring that of the students in the second virtual section only necessitated my insisting that they turn in copies of those study notes along with their exams.

Finally, I never effectively established rapport with the students in the first virtual section, either in person or on-line. If the very strong recommendations about this noted by Dereshiwsky (1998) are to be taken seriously, success in a virtual class and ultimately positive attitudes on the part of the students are critically dependent on this. Thinking about my own experiences in face-to-face classes, I believe that a high level of instructor engagement may go a long way in keeping students working with the material, progressing at a reasonable pace, and not discouraged by the common struggles with material that so many students find to be difficult. In fact, the primary motivator for offering the virtual section was the lack of available time for me to work directly with the students. If anything, much more time devoted to interacting with students on-line may be necessary for a successful virtual course.

The inclusion of weekly on-line discussion/question-and-answer sessions in the second virtual section would seem to have helped create a higher degree of rapport between students and instructor. Their evaluations of the quality of the instructor were certainly much more positive than were those from the students in the first virtual section. In addition, those sessions gave me much more of a chance to address student uncertainties than the simple provision of access to the instructor via telephone or e-mail in the first virtual session.

While the changes I implemented in the second virtual section of my course led to some clear improvements in outcome, there is clearly considerable room for further improvement. A stricter week-to-week schedule in terms of course content coverage seems like a very reasonable request from the students. Likewise, setting up a mechanism that would encourage students to work together on specified homework problems might naturally provide a way to reduce the isolation several students most recently reported. While my expectation that such productive collaboration might arise spontaneously was clearly unreasonable in the first virtual section, failing to try to promote it in the second virtual section was not a productive approach.

My intent in documenting my efforts and the problems that students experienced is to offer some cautionary advice to other instructors who may try something similar. Rather than concluding that such efforts are likely to result in similarly poor outcomes, my experiences and those of other, more successful instructors who have shared their own experiences serve to illuminate the kind of instructor engagement that may be one of several necessary components to ensure an acceptable range of outcomes.

Likewise, successive efforts to arrive at an optimal organization of an on-line class may be something that instructors venturing into this area should expect. Instructors differ among themselves in teaching styles and philosophy of instruction as much, if not more than students vary in ability and orientation to learning statistics. Our learning from missteps in trying to take advantage of this new environment for teaching is just as important as students learning from their mistakes in trying to master the material.

Ultimately, the most effective technology for offering introductory statistics courses over the internet is probably that which helps keep students engaged with their own efforts to learn and instructors engaged with them as they try to do so. Good teaching promotes good learning through the relationship forged between teacher and student, whether that relationship is face-to-face or at some distance.

10. References

Allen, M., Bourhis, J., Burrell, N., & Mabry, E. (2002). Comparing student satisfaction with distance education to traditional classrooms in higher education: A meta-analysis. *The American Journal of Distance Education, 16*(2), 83-97.

Dereshiwsky, M. (1998). *"Go Figure": The Surprising Success of Teaching Statistics via Internet.* Paper presented at the meeting of the Teaching in the Community Colleges Online Conference "Online Instruction: Trends and Issues II", Kapiolani Community College. (ERIC Document Reproduction Service No. 449 975)

Fredda, J. V. (2000). *Comparison of selected student outcomes for internet versus campus based instruction* (Report 00-08). Ft. Lauderdale, FL: Nova Southeastern University Research and Planning. (ERIC Document Reproduction Service No. 443 374)

Kennedy, R. L., & McCallister, C. J. (2000). *Basic statistics via the internet.* Paper presented at the meeting of the Annual Meeting of the Mid-South ERA, Bowling Green, KY.

Martens, R., Portier, S., & Valcke, M. (1995). *Comparing study outcomes from different learning environments for statistics* (Vol. 65). Heerlen, Netherlands: Open University Center for Educational Technological Innovation. (ERIC Document Reproduction Service No. 383 745)

Phipps, R., & Merisotis, J. (1999). *What's the difference: A review of contemporary research on the effectiveness of distance learning in higher education.* The Institute of Higher Education Policy. Retrieved October 2. 2002, from http://www.ihep.com/Pubs/PDF/Difference.pdf

Stephenson, W. (2001). Statistics at a distance. *Journal of Statistics Education, 9*(3). Retrieved October 2, 2002, from http://www.amstat.org/publications/jse/v9n3/stephenson.html