INFORMS Transactions on Education

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Evaluating the Integration of Technology into Business Statistics

Michael F. Gorman,

To cite this article:
Michael F. Gorman, (2008) Evaluating the Integration of Technology into Business Statistics. INFORMS Transactions on Education 9(1):10-19. https://doi.org/10.1287/ited.1080.0017

Vol. 9, No. 1, September 2008, pp. 10–19
ISSN 1532-0545 | 08 | 0901 | 0010

INFORMS Transactions on Education

DOI 10.1287/ited.1080.0017

Evaluating the Integration of Technology into Business Statistics

Michael F. Gorman
Department of Management Information Systems, Operations Management, and Decision Science, School of Business, University of Dayton, Dayton, Ohio 45469, michael.gorman@udayton.edu

This research explores the use of computer technology (specifically Microsoft Excel spreadsheets and the WebCT content delivery and student evaluation mechanism) as enablers of content delivery and as ways to improve pedagogy and learning. I evaluate the effect of these technologies on course delivery in Introductory Business Statistics, a sophomore-level required business course, but the lessons learned may apply to any analytical subject in the OR/MS field. I compare two different course offerings, a "technology light" version and a "technology centric" version, for both student satisfaction level (based on standard student evaluations) and level of learning achieved (based on student scores on a common course final), as well as provide an evaluation from the instructor perspective. The efficiency of instructor time and class time improves with technology, after initial startup costs are incurred. Technology components are generally well-received by students, and final test scores improve. I hypothesize that the increase in the students' outside-of-class workload and the assessment enabled by the technology may adversely affect student evaluations.

Key words: business education; business schools; computer-assisted instruction
History: Received: January 2008; accepted: July 2008. This paper was with the author 1 month for 1 revision.

1. Introduction

A recent issue of *INFORMS Transactions on Education* (Troxell 2006) points out the need for continued research into the effective use of technology in the classroom. The research described in this paper responds to this call by evaluating the effectiveness of spreadsheet and Web-based technologies in an Introductory Business Statistics course. To test their effectiveness, I compare a "technology light" (TL) course to a "technology centric" (TC) course in a number of dimensions.

The TC course was a paperless class; all assignments and tests took place using two technologies, Microsoft Excel spreadsheets (www.microsoft.com) and the WebCT (www.blackboard.com) content delivery and student testing and evaluation mechanism.[1] The TL course used limited Web-based course note retrieval (a Lotus QuickPlace website with static notes, study guides, problem solutions, and some in-class materials without testing or online discussions or grades) and Excel for graphing purposes only. The TL course required use of calculators and table lookups for in-class tests and homework assignments; the TC course strictly used Excel for such problems.

A technology-centric delivery carries some opportunities and some risks. Excel combines statistical theory with common practice, coupling the concepts of statistical theory with their implementation using a common business tool. Excel enables easy transition from small sample sizes to real-world, full-scale problems. However, Excel is another skill that must be learned as the students master statistical concepts; the additional skill acquisition can become an obstacle. Furthermore, the built-in functions of Excel may obfuscate the underlying statistical concepts from the student, reducing comprehension of these concepts.

WebCT enables ease of dissemination, collection, and update of information, as in the TL Web tool QuickPlace. In addition, WebCT functionalities allow for outside-of-class-time quizzing with immediate feedback and optional grading and free-form, less structured "discussion board" topics for student interaction. However, online content delivery and testing may have its own risks, including plagiarism, reduced attendance, and potentially a less personalized or diminished experience for the student (Haywood 2000, Yin et al 2002).

[1] The only exception was the final exam, whose questions were in hard copy to minimize the risk of file sharing; the answers were still tendered via WebCT.

Using student reactions and student test scores, this research evaluates student reception of the adoption of new technologies, as research by Martins and Kellermanns (2004) suggests is important, and whether these technologies enable or inhibit statistics learning. Various studies have proposed the use of spreadsheets to foster active learning of various subjects in the classroom (e.g., Johnson and Drougas 2004, type I and type II errors; Albritton and McMullen 2006, forecasting; Price and Zhang 2007, central limit theorem) but do not measure their effectiveness in terms of student learning or retention relative to a control group. Arbaugh (2005) evaluates the use of WebCT in distance education course design. Basile and D'Aquila (2002) survey accounting students on their acceptance of WebCT as a content delivery and assessment mechanism. Saadé and Kira (2004) find test scores improve in classes that use an interactive computer-aided tool for learning. Tsai and Wardell (2006a) evaluate the use of Visual Basic for enabling statistical learning and find that the tool has a statistically significant impact on student learning relative to a control group. This research evaluates the impact of extending the use of Excel and WebCT technologies in tandem from both student perception and pedagogical effectiveness perspectives to observe their effect on student learning.

2. Course Description and Comparison

The course considered in this study is the first statistics course in a series of two courses required for all undergraduate business students. The subject matter includes descriptive statistics, probability, and introductory inferential statistics. To maintain comparability, course attributes other than those under study were held constant. Both classes were offered in two-day-per-week, 75-minute sessions over 15 weeks in the fall semester. Both classes were taught by the same instructor. Both classes used Keller and Warrack (2003) as the textbook. Two variations on this course were offered—a TL and a TC course—to evaluate the impact of changing the course format by introducing extensive use of Excel and WebCT. The TL class was taught in 2002; the TC classes were taught in 2005. The TL course was given in a traditional classroom with instructor access to Excel, but without student computers. It used more traditional calculator-based and table-lookup (z-scores, t-values, binomials, etc.) methods for homework and test problems. Except for descriptive statistics' graphing homework problems (done in Excel), assignment calculations such as mean, variance, and correlation were completed manually after detailed, step-by-step in-class coverage of manual equation calculations. Although equivalent methods in Excel were demonstrated by the instructor and used in some homework assignments, Excel was not available for tests, so manual solution methods on small data sets were used out of necessity. The TC course used Excel built-in functions in place of calculators and table lookups, WebCT online testing for homework and tests, and working file depositories instead of paper. The course was administered in a computer lab, so Excel was used in all lectures, exercises, and testing.

In the TL class, tests were graded, most homework was graded, and occasional unscheduled quizzes were given on reading or the previous class material, for a total of 12 grading opportunities. The TC class utilized WebCT for quizzes on reading material and automated grading of take-home practice quizzes. Although the TC course had similar assignment requirements, all were graded, for a total of 30 grading opportunities during the semester.

3. New Technology Challenges

Introducing new technology into the classroom can present challenges with student acceptance (Martins and Kellermanns 2004). It was important that Excel didn't "get in the way" of learning statistics concepts; some students who may understand statistical concepts may be challenged to demonstrate them in the Excel environment. I had to provide sufficient Excel training for students to be comfortable learning statistics via Excel, but not so much that the course centered on spreadsheet skills. I offered in-class Excel basics and tips training to help students gain these valuable skills that are not core to statistics. This investment was 5%–10% of course time, but some of this time is recouped later in the semester when Excel functions are easier to master than statistical tables. For those students who were significantly deficient in Excel skills, an optional two-hour review session was offered (8% of students attended). These measures reduce but don't necessarily solve technology adoption problems. Students struggle with Excel, but from my experience perhaps no more than they struggle with traditional calculator and table-based methods.

Students showed little resistance to WebCT's online test-taking feature; most students in this section had had experience with online testing in prior classes. One drawback to using these tests is that automatic grading only applies to multiple-choice tests; thus, automatic grading allows for a high quantity of questions, but the preponderance of questions may be at a lower level of cognitive testing. Students provide no demonstration of solution method, and no partial credit can be given. Of course, a mix of questions could be used; not all need to be automatically graded, but the technology gives a strong incentive to

use multiple-choice-style questions, which were used in both the TL and TC course.

One risk posed by online testing is plagiarism. The homework component of the grade was 15% and pre-quizzes (reading quizzes) another 10%, making outside-of-class effort 25% of the grade. Information on problem solutions could easily be shared among class participants. More importantly, there is nothing to stop one student from logging on as another and completing the work; this is very difficult to detect. (This is also possible but less likely on WebCT tests taken within the classroom.) To mitigate this risk, a work file providing demonstration of the students' efforts was required to be turned in with each homework and test. Of course, the work files should also be compared for similar or identical work, but this can be exceedingly time-consuming. Inside class, occasional quizzes (10%), three tests (15% each), and the final (20%) made up 75% of the final grade to reduce the exposure of final grades to malfeasance occurring outside of class.

WebCT poses some risk for cheating during tests as well. Students and all class materials are online during tests; thus, course notes and example problems are highly accessible during testing. As I had in the TL class section, in the TC course I allowed a "crib sheet" of key course notes to each student, so the benefit of browsing online class materials was reduced. The crib sheet also allowed for creating more challenging, thought-provoking, and analytical tests rather than asking rote memorization questions. To increase the cost of cheating, I used a product called SynchronEyes™ (www.synchroneyes.com) to monitor students during the test to reduce any form of online browsing, communication, or file sharing. (Incidentally, this product also provided me great insight into the students' thinking processes during the test. In effect, I could look over the student's shoulder as he or she worked through problems.) I looked for wrongdoing such as email being used or the course notes website being open. Perhaps partially because of these preventative measures, I observed no evidence of online malfeasance.

Electronic tests are more easily leaked between sections of the class than are hard-copy versions (via emailing working files or cut and paste of test questions). Scheduling classes in consecutive time slots reduced the potential for any such leak, but as an added precaution I used different test questions, input data, random question ordering, and random option ordering to reduce the potential for any malfeasance. Because the final is the most heavily weighted component of the grade and is the key metric of learning for this study, I used hard copies of questions so that there was no electronic source for them available.

Table 1 Demographics of Subject Groups

(a) Class demographics

	Tech light	Tech centric
Count	35	66
Males (%)	60	56
Females (%)	40	44
Avg. age	19.8	19.8

(b) SAT and ACT math score comparison

	TL test scores			TC test scores			Mean diff
	Count	Mean score	Std. dev.	Count	Mean	Std. dev.	*P*-value
SAT math	24	559.17	15.20	41	570.73	9.67	0.262
ACT math	32	24.22	0.60	58	24.74	0.37	0.229

4. Student Sampling

My sample for this study was the students who enrolled in the one section of the TL course (35 students) and the students who enrolled in the two sections of the TC course (65 students). Table 1(a) shows the general demographics of the two groups, which are similar in gender mix and average age. Table 1(b) shows descriptive statistics for SAT and ACT math scores. Although we could not impose random sampling techniques on the enrollees for each section (students self-selected), *t*-tests show that on a univariate basis there were no statistically significant differences between SAT and ACT math scores between the samples. These demographic statistics and test scores fairly reflect the general student population taking business statistics.

5. Course Effectiveness Assessment

This section presents quantitative and qualitative assessments of the two course offerings from both the student and the instructor perspectives. Section 5.1, Technology Component Assessment, evaluates the student ratings of various technology components of the TC course through responses to student surveys offered in these courses. Sections 5.2 and 5.3, Course Effectiveness—Qualitative Student Assessment and Course Effectiveness—Quantitative Student Assessment, compare the student assessment of the two different course offerings overall both qualitatively and quantitatively. Section 5.4, Course Effectiveness—Student Learning, evaluates course effectiveness by comparing student scores on identical finals given to all classes. The test score analysis is presented by subject matter area (descriptive statistics, probability, and inferential statistics) and by question type (equation-based problems or conceptual understanding questions). Finally, §5.5, Course Effectiveness—Instructor Assessment, provides insights from the instructor's perspective for classroom efficiency, course workload, and general satisfaction level.

Table 2 Student Rating of Technology Components of Course

Question topic	Average	Median	Mode	Std dev	*N*
In-class class Excel files	1.58	1	1	0.73	57
WebCT as an info depository	1.75	2	1	0.74	57
Online gradebook	1.79	2	1	0.80	57
Online syllabus	1.91	2	2	0.89	57
Sample tests	2.02	2	1	1.08	57
Multiple grading opportunities	2.09	2	2	0.93	57
Posted HW solutions	2.14	2	2	0.88	57
Section review notes	2.18	2	2	0.71	57
Real-world examples	2.21	2	3	1.01	57
Email/bulletin board	2.47	2	2	1.02	57
Instructor's organization	2.51	3	3	0.87	57
Homework assignments	2.58	3	3	0.78	57
Excel-based test questions	2.65	3	3	0.86	57
Instructor availability	2.68	3	3	0.93	57
Instructor's teaching style	2.95	3	3	0.95	57
In-class tests	3.05	3	3	0.87	57
Online discussion topics	3.19	3	3	0.93	57
Prequizzes on reading	3.21	3	3	1.00	57
Course notes (slides)	3.28	3	4	1.03	57
Tests	3.35	3	3	0.94	57
Textbook	4.12	4	5	0.91	57

Rate how helpful you found the following course components:

Question coding:	Extremely helpful	Very helpful	Helpful	A little helpful	Not helpful
	1	2	3	4	5

5.1. Course Effectiveness—Technology Component Assessment

Table 2 presents a comparison of the two TC course sections for students' opinion on various components of the course. Students respond to the question, "How helpful were the following elements of the course to your learning?" on a five-point Likert scale, where one is "Extremely helpful" and five was "Not helpful." Results are sorted in order of ascending mean response, from most helpful to least helpful. Although the median is more appropriate than mean for evaluating Likert scale responses, because many median scores were the same across sections, I report both median and mean in order to provide additional information on the relative rankings of student opinions.[2]

WebCT's primary value in the students' opinions was as a data repository; the four most helpful elements related to online information retrieval. Course materials, instructor in-class examples, and online syllabi were also available in the TL course offering. The most helpful component of the data repository for the students was the in-class Excel file repository. These are instructor demonstrations from each class showing how to conduct statistical analyses in Excel. Students could recapture exactly what the instructor did in class by reviewing the files saved to the website. These files were particularly useful for students in the TC section because tests and homework used the same problem solving methodology.

Multiple grading opportunities were also appreciated by the students. The automation of grading and testing outside of class were both made possible by WebCT technology and were not in the TL content dissemination website; thus, they are also considered part of the TC offering. Somewhat contradicting these opinions, prequiz capabilities (the very grading opportunities created by WebCT) were far less valued by the students. Students seem to appreciate the risk diversification of so many grading chances, but not the additional effort these chances entail. I observed that based on in-class responses to discussion questions, pre-read quizzes did not seem to improve class preparedness; rather, students seemed to use these quizzes as a grade enhancing mechanism.

Finally, the WebCT discussion board functionality was strongly encouraged in numerous open-ended discussion exercises; but without a grade incentive students were generally not motivated to participate. In sum, it appears that the technology-related components of the course were generally well-received relative to other components of the course, but not all capabilities were equally valued. Interestingly, many of the most valued components (the data repository functionalities) were also available in the TL sections (except the grade book functionality).

Standard instructor- and text-related evaluative questions were included in the survey primarily for purposes of comparison to the technology questions. Instructor-related questions were generally in the middle of the ratings. "Canned" materials such as the text and course PowerPoint slides were near the bottom of the ratings. I assign two possible interpretations to this result: These sources tied less to class coverage either because the class format varied from the text and PowerPoint slides or because students vastly preferred the active, experiential-based learning in the classroom. The latter interpretation is supported by the fact that the in-class Excel files were the most helpful learning tool in the students' view.

5.2. Course Effectiveness—Qualitative Student Assessment

In all three course sections, I conducted a Midterm Instructional Diagnostic (MID) interview session with students. In these 15–20 minute sessions held during class time, an independent third party leads a discussion with the class under the strict assurance of complete anonymity. The moderator poses three questions: In this class, (1) what is helping your learning?,

[2] Mean-based analyses of Likert scale survey responses are common in the literature; for example, Winkelmann and Winkelmann (1998) on the happiness of the unemployed; Olson (2005) and Barman et al. (2001, 1991) on journal quality; Fang (2008) on product innovation; and Hock-Eam (2008) and Dawes (2008) on the effect of scaling on the mean response in surveys, among numerous others.

Table 3 Comparison of Midsemester Discussion Group Sessions Comments and Selected Written Evaluation Comments

Technology-light course	Technology-centric course
Midsemester discussion group session comments:	
Factors helping learning:	
• Lab days are helpful • Review sessions	• In-class examples • Posting notes on WebCT • Discussion board on WebCT
Factors hindering learning:	
• Excel is helpful, but (it is) not on tests	• Book differs from his material • Terminology—what is P-hat?
Suggestions to improve learning:	
• Teach class more like review session • Better organization • Tests more like homework • Better notes • Study guide	• Keep online quiz solutions available throughout the semester • More time in class to work out problems
End-of-semester evaluation write-in comments:	
• I believe (confusion) came from trying to relate Excel and manual styles, but not sufficient of either (2) • I am not sure I have a good handle on statistics; I really didn't understand this course (2) • Exams were too hard (3)	• Book is no help—different from class (5) • In-class notes help (3) • WebCT(3)—section summaries, slides help • Tests/quizzes did not help—too complicated/no partial credit (4) • Excel was good/learning about Excel (5) • I liked having a computer in front of me/example problems in class (2)

Note. Frequency of comment specified in parentheses.

(2) what is hindering your learning?, and (3) what could be done to improve your learning? Students break into 4–5 person discussion groups to develop consensus opinions and report back to the class. The moderator then provides class discussion feedback to the instructor, as reported in Table 3. Table 3 also reports individual write-in comments from end-of-semester course evaluations and provides the frequency of each comment in parenthesis (write-ins are not literal).

The TL comments centered on requesting an increased use of technology. (These comments, in part, gave rise to my decision to move to a TC format.) The TL review sessions were problem-based and took place in a computer lab similar to the TC class sessions, which the students found valuable; however, the dichotomy of technology use between homework and tests was underscored as a disadvantage in the TL section. The classroom use of Excel in the TC class seemed to be well-received. WebCT helped address organizational and information availability issues for the students, as comments changed from "need better organization" to "like having materials available on WebCT."

Some students in the TC course voiced concerns of going too fast and not having enough time in class and on tests to work problems. This opinion could reflect novice Excel skills that slow problem solving in a time-constrained environment.

Going to an all-technology format reduced the utility of the textbook; the resulting emphasis in class was much more toward the execution of statistics in Excel reflected by the book. By taking the Excel approach to statistics for homework and tests, the algebraic expression of problems was de-emphasized, which created confusion over terminology in the text and perhaps lower appreciation and understanding of the text content. In summary, the qualitative comments were generally more positive in the TC offering than in the TL offering.

5.3. Course Effectiveness—Quantitative Student Assessment

Students completed a 30-question course evaluation survey at the end of the semester. The first 23 questions were from common student evaluations given in all undergraduate courses university-wide. The next five were questions specific to business statistics topical coverage. The last two questions were customized questions added by the instructor for the TC sections of the course. Because the last two questions apply only to the TC courses, the TC and TL are compared on only 28 questions. The scale used is a 5-point Likert scale for all questions; the full text of the questions is included as the appendix. Table 4 shows the results of the surveys. As noted before, although the median is more appropriate for Likert-scale data, the analysis is conducted on mean responses, as is common in the literature. For ease of reading, significant differences between them are indicated by bold type, with increases underlined and decreases in italics ($\alpha = 0.05$).

The first striking observation is that the TC evaluations are often lower than the TL evaluations, which is surprising given the most valued components of the course from the other survey were technology related, and qualitative comments were generally more positive in the TC course. Upon more careful inspection, however, one can see scores that are completely unrelated to technology have significant differences between the two groups, contrary to expectations that the scores would be roughly equal. For example, "instructor enthusiasm" and "students are valued in this class" were significantly lower in the TC course, when generally these should have been unaffected by the introduction of technology. (Recall that both courses were taught by the same instructor.) More strikingly, "graded promptly" and "use of technology" did not significantly increase, even though

Table 4 Student Course Evaluation Between Group Comparisons

	Technology centric		Technology light					
	Obs 73		Obs 35		TC-TL mean difference			
Question	Mean	SD	Mean	SD	Diff	SD (diff)	*P*-value	Question subject
1	3.9	0.3	3.7	0.5	0.2	0.09	0.01	**Challenging**
2	2.9	0.9	3.0	0.7	−0.1	0.16	0.19	Thinking
3	2.4	0.9	2.6	0.8	−0.2	0.17	0.13	Discussion
4	2.7	0.9	2.8	0.9	−0.1	0.18	0.22	Assignments
5	1.4	0.7	2.4	0.7	−1.0	0.14	0.00	***Writing***
6	1.4	0.9	1.9	0.5	−0.5	0.13	0.00	***Speaking***
7	3.5	0.7	3.2	0.8	0.3	0.16	0.06	Computer technology
8	2.3	1.0	2.6	1.2	−0.3	0.23	0.12	Exams
9	2.1	1.0	2.8	0.9	−0.7	0.19	0.00	***Reasonable work***
10	1.5	1.1	2.4	0.9	−0.9	0.20	0.00	***Text***
11	3.3	0.8	3.1	0.5	0.2	0.13	0.06	Effort
12	2.5	1.1	2.5	0.9	0.0	0.20	0.47	Learned
13	1.9	1.2	2.2	1.0	−0.3	0.22	0.06	Recommend course
14	2.5	0.9	2.9	1.1	−0.4	0.22	0.05	Comm. objectives
15	2.3	1.1	2.4	1.1	−0.1	0.23	0.29	Environment/learning
16	3.0	1.0	3.5	0.6	−0.5	0.15	0.00	***Enthusiasm***
17	3.2	0.8	3.5	0.7	−0.3	0.15	0.01	***Practical***
18	2.8	0.8	2.9	0.8	−0.1	0.16	0.20	Participation
19	3.0	0.8	3.3	0.5	−0.3	0.13	0.00	***Students valued***
20	2.8	1.0	3.3	0.7	−0.5	0.17	0.00	***Available***
21	3.3	0.9	3.4	0.5	−0.1	0.13	0.13	Graded promptly
22	2.4	1.1	3.2	0.7	−0.8	0.17	0.00	***Reasonable grading***
23	2.1	1.3	2.5	1.4	−0.4	0.28	0.06	Rec. instructor
24	3.0	0.7	2.9	0.8	0.1	0.16	0.35	Data-decision making
25	3.1	0.8	3.0	0.8	0.1	0.16	0.34	Graphs/tables
26	3.1	0.9	3.1	0.9	0.0	0.18	0.41	Single-variable
27	2.6	1.0	2.8	0.8	−0.2	0.18	0.10	Prob. distributions
28	2.3	1.0	2.7	0.9	−0.4	0.19	0.01	***Statistical inference***

Notes. Bolded question labels significant differences; underlined values indicate TC is significantly higher; italicized labels indicate TC is significantly lower. TC1 observations = 34; TC2 observations = 39.

every assignment and test was graded nearly instantaneously and the entire course was centered on technology. Most remarkable of all, despite the constant exposure to extremely practical Excel technology and more real-world applications and data sets, the "practical" score significantly decreased.

The only significant increase of the TC course over the TL offering in the student opinion survey was in the "challenging" score of the course, which is strongly negatively correlated with the "reasonable work" score. The fact that students needed to master two new skills—Excel and statistics—in the TC course may have added to the challenge of the course relative to the TL version. I hypothesize that due to the increase in graded components of the course (prequizzes, homework quizzes), students found the workload to be unreasonable and the challenge greater. Dissatisfaction with the course workload may have negatively affected other course evaluation scores discussed above. This hypothesis is supported by discussions with a subset of students after the semester was completed; however, although plausible, this hypothesis cannot be tested with the data collected in this study.

Table 5 compares the two sections within the TC course offering. Although the student samples and the two classes offerings were essentially identical, there are more statistically significant differences in student perceptions between the TC sections (13/28 questions) than there were between TC and TL offerings (11/28). Also, all differences are of the same sign; one class seems to have had a significantly more negative experience than the other. The two questions (numbers 29 and 30) that apply only to the TC classes—"Did Excel help your learning in this class?" and "Did WebCT help your learning in this class?"—also had significantly different scores. There seem to be other forces on student opinion at play, other than course format, that have not been explained. As the instructor of both TC sections, I have no suggestions as to why the two classes might have responded so differently on the survey; there was no noticeable indication during the semester that the sections varied in their acceptance.

More data may help to reduce the variation, but there is no opportunity to collect more data in the

Table 5 Student Course Evaluation Within-Group Comparison: Technology-Centric Courses

Question	TC1-TC2	SD (TC1-TC2)	*P*-value	Question
Technology centric class student evaluation comparison TC Section 1 vs. 2				
1	0.0	0.07	0.50	Challenging
2	0.3	0.21	0.08	Thinking
3	0.0	0.22	0.50	Discussion
4	0.3	0.20	0.07	Assignments
5	0.0	0.17	0.50	Writing
6	0.3	0.20	0.07	Speaking
7	0.1	0.18	0.29	Computer technology
8	0.6	0.24	0.01	***Exams***
9	0.4	0.23	0.04	***Reasonable work***
10	1.0	0.26	0.00	***Text***
11	0.0	0.19	0.50	Effort
12	0.4	0.26	0.06	Learned
13	0.5	0.28	0.04	***Recommend course***
14	−0.1	0.22	0.33	Comm. objectives
15	0.7	0.26	0.00	***Environment/learning***
16	0.1	0.22	0.33	Enthusiasm
17	0.1	0.19	0.30	Practical
18	0.3	0.19	0.06	Participation
19	0.5	0.19	0.00	***Students valued***
20	0.7	0.23	0.00	***Available***
21	0.1	0.20	0.31	Graded promptly
22	0.6	0.25	0.01	***Reasonable grading***
23	0.7	0.31	0.01	***Rec. instructor***
24	0.3	0.17	0.04	***Data-decision making***
25	0.5	0.18	0.00	***Graphs/tables***
26	0.3	0.20	0.07	Single-variable
27	0.5	0.25	0.02	***Prob. distributions***
28	0.7	0.24	0.00	***Statistical inference***
29	0.4	0.25	0.05	***WebCT helps***
30	0.9	0.27	0.00	***Excel helps***

Note. Significant differences between TC sections are bolded and italicized.

same environment; since these experiments, the statistics classes have adopted a new textbook and a new version of Excel, Excel 2007, has been introduced. The next experiments would be under these new environmental changes.

5.4. Course Effectiveness—Student Learning

In order to evaluate student learning, Table 6 compares the number of correct answers out of 31 questions on identical common finals. One minor implementation difference was that TL students used pencils and scantron sheets; TC students entered their answers in WebCT. The test was designed in TL with manual test taking in mind; thus, there was no technology bent to the questions. Table 6 shows that the TC classes scored almost two full correct answers higher than the TL classes on the final, a significant difference for $\alpha = 0.05$ (p-value $= 0.0307$). There was no significant statistical difference in scores between the TC sections, and both TC sections scored significantly higher than the TL section. This finding supports the hypothesis of increased learning in the technology environment. This favorable result is particularly noteworthy because TC students faced a slightly different testing format on the final than on their prior tests.

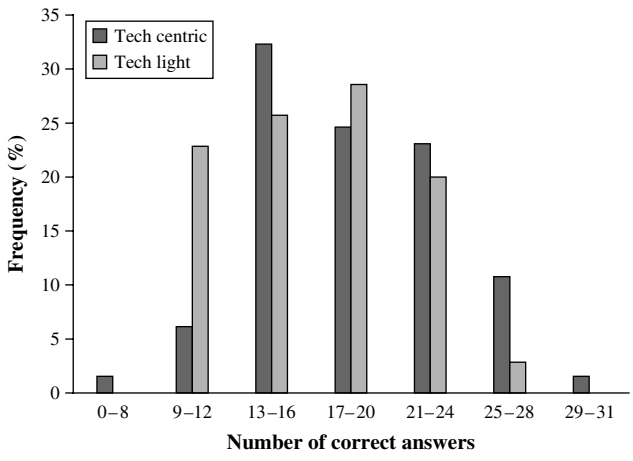
Table 6 Comparison of Final Test Scores: Technology-Centric Versus Technology-Light Offerings

	Tech centric	Tech light
Summary descriptive statistics final exam correct answers		
Min	8	10
Max	30	25
Mode	16	19
Mean	18.57	16.80
Std. dev.	4.97	4.48
Count	65	35
Diff. of means	1.77	
T statistic	1.81	
P-value	0.037	

The increase in test scores in the TC sections may not be due entirely to the introduction of technology. Most notably, the increased outside-of-class workload for students (that was, to be sure, to some degree caused or enabled by technology) may have likely in itself caused an increase in scores. Other potentially confounding influences include a more experienced and organized instructor (from additional experience of offering this course) and more active learning in a lab setting in the TC offering.

Figure 1 shows the distribution of the number of final examination questions correct in the two offerings. There was a slightly higher variance in final scores in the TC sections, with both a higher maximum and a lower minimum score. The spread may imply less consistency of learning in TC sections, perhaps because of the challenge of acquiring both Excel and statistical skills during the course. This observation is somewhat troublesome, given the larger spread in mathematical ability in the TL sections coming into

Figure 1 Distribution of Correct Answers on the Common Final Exam



the class when measured by SAT and ACT math scores (Table 1(a)). On the other hand, a point that bolsters the improved learning claim and reduces concerns of increased variance in TC sections is that two students who were underperforming in TL dropped the class, ostensibly increasing the TL average score and reducing its variance. (No students dropped the TC course.) This "self censoring" may have also contributed to the greater variance noted in the TC final scores and makes the mean score improvement in the TC format more impressive. Also, the maximum score was five points higher in the TC section, which increases the spread but adds to the claims of increased learning in those sections. Finally, the higher variance in the TC course lends evidence that cheating was not widespread in the TC course; if cheating were prevalent, then the variance of scores would be lower.

Table 7 Comparison of Final Test Scores: Technology-Centric Versus Technology-Light Offerings

Question type	Question number	Concept question (%)	Problem question (%)	All questions (%)
Descriptive stats	1		13	
	2		23	
	3		1	
	4		20	
	5		2	
	6		−10	
	7	−11		
Descriptive stats total		−11	8	5
Inferential stats	1	7		
	2		3	
	3		5	
	4		−7	
	5		2	
	6	9		
	7		31	
	8	21		
	9	10		
	10		17	
	11		−8	
	12		21	
	13		21	
	14		−1	
Inferential stats total		12	9	10
Probability	1		23	
	2		−6	
	3		−17	
	4		8	
	5		−2	
	6		22	
	7		8	
	8		−13	
	9		−12	
	10	−13		
Probability total		−13	1	0
Grand total		4	6	6

Notes. Observations: TC format = 65; TL format = 35. All percentages shown are percent correct TC-TL.

Table 7 furthers the analysis of test score differential on a question-by-question basis, evaluating both the type of question (statistical concept or problem solving) as well as the subject area of the question (descriptive statistics, probability, inferential statistics). Generally, the questions are weighted more towards inferential statistics because inferential statistics draws from skills learned in both descriptive statistics and probability sections. The TC classes generally did better in all question types and subject areas. Generally, the improvement in TC performance was greater where Excel tools apply more directly; TC students' performance was higher relative to TL students in calculation problems than in concept questions and higher in descriptive statistics and inferential statistics questions than in probability questions.

5.5. Course Effectiveness—Instructor Assessment

As pointed out in research on WebCT adoption by White and Myers (2001), instructor buy-in is critical to technology introduction because the instructor has such a large stake in a course redesign. In this case, extensive instructor time was invested in course redesign (new technology-centric notes, prequizzes, exercises, data sets, website development, etc.); preparing for the TC course was akin to the time requirements for a new course preparation. The time costs could be mitigated through the purchase of Web-enabled test bank questions or a more gradual technology adoption. Ongoing course delivery will require approximately the same time requirements of the more traditional offering, although the mix of how the time spent will shift from question preparation and grading to more student interaction. Overall, I was highly satisfied with the new structure. Class time was spent in hands-on application of concepts rather than "board work"; PowerPoint notes became reference material for the students rather than the subject of a lecture. The lab setting and activity-based learning improved the liveliness of the classroom, including students asking problem-based questions and students helping each other. In-class time was better utilized because electronic assignments obviate the need for returning papers. In fact, no paper was distributed or collected at any time during the semester. This allowed for an estimated savings of approximately 60 minutes of class time per semester for returning papers (5 minutes per graded assignment). I also saved 45 minutes of in-class quiz time (15 minutes per quiz), and some time on quiz and homework review (that the students conducted individually online outside of class on an as-needed basis). The WebCT grade book functionality reduced the number of grade queries by students to zero. More of my out-of-class time was spent developing quality in-class exercises and test questions and participating in discussion boards rather than grading

responses. Generally, there was a shift from lower value-added administrative activities to those that were more geared toward learning, a switch that benefits both student and instructor.

6. Conclusions and Limitations

In this research, I evaluate the use of technology (Excel spreadsheets and WebCT content delivery and student evaluation mechanism) as an enabler of higher quality pedagogy, more efficient classroom time allocation, and increased student contact time.

Somewhat paradoxically, students' assessment of the functionalities of the technology is positive, but course evaluations are generally lower in the technology-centric class. I hypothesize the lower evaluations may be explained in part by increased student workload that was enabled by technology (through automated grading of outside-of-class-time assignments).

Of course, student evaluations are not the ultimate measure of success. Students might be viewed as patients rather than customers; what hurts now may benefit them in the long run. Student learning, when measured by scores on a common final, improves significantly when aided by these technologies, perhaps in large part because the quantity and quality of time spent with the material increases with Web-enabled outside-of-class student evaluation and feedback mechanisms. In particular, I observe no loss of conceptual understanding when spreadsheets are used for statistical calculations.

This research has limitations that should be recognized. Clearly, the students in each course format could not compare their experience to what the other format "would have been." In this sense, student evaluation of each format is unavoidably blind of the alternative. All course comparisons are univariate in nature; because of student survey anonymity, it is impossible to tie survey responses with individual student attributes, hindering multivariate analyses. My evaluation combines the effects of WebCT and Excel; their individual effects on learning cannot be discerned in this study. I also observe student evaluations that differ statistically between sections of the course with the same technology-centric format, indicating (not surprisingly) that students' opinions were affected by factors other than the course format. Additional replications of this experiment could help produce more clear and conclusive results, but due to changes in course textbook and teaching schedules, clean replications are unfortunately not possible. Finally, it should be noted that various confounding elements limit the interpretation of the results. For example, in the TC course the work content was higher and the instructor more experienced than in the TL course, both of which could affect results.

Having new avenues for cheating is a potential problem for the TC course, both for information retrieval (e.g., notes, worked problems) and for information sharing (file passing). In class, products such as SynchronEyes help increase detection and reduce the incidence of such behavior. The benefits of cheating were reduced through random test question and option ordering and randomized data sets. I see room to improve controls for plagiarism risk when utilizing an electronic medium for testing. There is no adequate control for assuring that work files were not passed between sections. For example, automated random data sets for each test taker would eliminate answer passing (Tsai and Wardell 2006b). A product that could act as a Turnitin.com equivalent for automation of plagiarism detection for Excel spreadsheets would be a tremendous help for deterring shared work files. Plagiarism was minimized for the final exam by having multiple test layouts and not providing electronic versions of the questions.

Overall, I find Excel to be a convenient, practical, and effective medium for teaching statistics. WebCT technology extends the classroom learning environment, allowing for out-of-class evaluation and testing with immediate student feedback. Due to the expanded time capacity and efficiency of testing and grading made possible by WebCT, care must be taken by instructors not to overuse this classroom extension and cause adverse reaction from students who may feel overworked. More broadly, careful expansion of the student learning experience using new technologies should continue to be explored and evaluated, maximizing their opportunities while minimizing their risks. Introducing these technologies has returns worth the risks. Students gain two valuable skills: understanding statistics concepts and learning how to implement them with a ubiquitous technology they will use in subsequent classes and in their work lives.

Appendix. Student Course Evaluation Form University of Dayton Student Assessment of Instructors

Mark whether or not you agree or disagree with each of these statements. If you neither agree nor disagree with the statement, mark "Neutral." If you believe that the statement is not applicable to you, the instructor, or the course, fill in the circle under "Not applicable."

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Not applicable
5	4	3	2	1	

Course-Related Statements

1. This course was academically challenging.
2. This course improved my critical thinking skills.

3. Students' classroom discussion contributed to my understanding of the course material.
4. Assignments contributed to my understanding of the course material.
5. Assignments in this course improved my writing.
6. Assignments in this course improved my speaking.
7. My use of computer technology contributed to my understanding of the course material.
8. Examinations/evaluations related to the material emphasized in the course.
9. Given the course level, the quantity of work required was reasonable.
10. The textbook and/or required readings were an asset to this course.
11. My effort in this course was substantial.
12. I learned a great deal from this course.
13. I would recommend this course to other students.

Instructor-Related Statements

14. The instructor clearly communicated course objectives and requirements.
15. The instructor presented the material in a manner conducive to learning.
16. The instructor's enthusiasm inspired interest in the course.
17. The instructor encouraged students to relate course material to practical situations.
18. The instructor encouraged classroom participation.
19. The instructor created an environment in which all students were valued.
20. The instructor was available to help students outside the classroom.
21. The instructor graded and returned assignments and examinations in a reasonable time.
22. The instructor's grading was reasonable, given the quality of my work.
23. I would recommend this instructor to other students.

Instructor/Department/School Statements

24. In this course, I developed an understanding of how data analysis and statistical techniques can help in business decision-making.
25. From this course, I developed skills in describing data sets with tables and graphical displays.
26. In this course, I developed skills in calculating and interpreting descriptive statistics (e.g., means, standard deviations, and correlation coefficients).
27. From this course, I can use probability distributions (like Normal or Binomial) to make risk or likelihood statements about the value that a certain variable may have.
28. From this course, I understand how to estimate an important property of a target population (like an average or a proportion) using sample results—to include specification of possible estimation error.
29. The use of WebCT aided my learning in this course.
30. The use of Excel aided my learning in this course.

References

Albritton, M., P. McMullen. 2006. Classroom integration of statistics and management science via forecasting. *Decision Sci. J. Innovative Ed.* **4**(2) 331–336.

Arbaugh, J. B. 2005. Is there an optimal design for on-line MBA courses? *Acad. Management Learn. Ed.* **4**(2) 135–149.

Barman, S., M. D. Hanna, R. L. LaForge. 2001. Perceived relevance and quality of POM journals: A decade later. *J. Oper. Management* **19**(3) 367–385.

Barman, S., R. J. Tersine, M. R. Buckley. 1991. An empirical assessment of the perceived relevance and quality of POM-related journals by academicians. *J. Oper. Management* **10**(2) 194–212.

Basile, A., J. D'Aquila. 2002. An experimental analysis of computer-mediated instruction and student attitudes in a principles of financial accounting course. *J. Ed. Bus.* **77**(3) 137–143.

Dawes, J. 2008. Do data characteristics change according to the number of scale points used? *Internat. J. Market Res.* **50**(1) 61–77.

Fang, E. 2008. Customer participation and the trade-off between new product innovativeness and speed to market. *J. Marketing* **72**(4) 90–104.

Haywood, T. 2000. Defining moments: The tension between richness and reach. *Inform. Comm. Soc.* **3**(4) 648–654.

Hock-Eam, L. 2008. Use of different happiness rating scales: Bias and comparison problem? *Soc. Indicators Res.* **87**(2) 259–267.

Johnson, A., A. Drougas. 2004. Illustrating type I and type II errors via spreadsheet simulation in the business statistics course. *Decision Sci. J. Innovative Ed.* **2**(1) 89–95.

Keller, G., B. Warrack. 2003. *Statistics for Management and Economics*, 6th ed. Duxbury Press, Boston.

Martins, L., F. Kellermanns. 2004. A model of business school students' acceptance of a web-based course management system. *Acad. Management Learn. Ed.* **3**(1) 7–26.

Olson, J. E. 2005. Top-25-business-school professors rate journals in operations management and related fields. *Interfaces* **35**(4) 323–338.

Price, B., X. Zhang. 2007. The power of doing: A learning exercise that brings the central limit theorem to life. *Decision Sci. J. Innovative Ed.* **5**(2) 405–411.

Saadé, R., D. Kira. 2004. Effectiveness of an interactive application to assist learning: A test case. *J. Inform. Systems Ed.* **15**(4) 357–363.

Troxell, D. S. 2006. Introduction to the ITE special issue: Effective use of software in the MS/OR classroom. *INFORMS Trans. Ed.* **7**(1). http://archive.ite.journal.informs.org/Vol7No1/Troxell/.

Tsai, W., D. G. Wardell. 2006a. An interactive Excel VBA example for teaching statistics concepts. *INFORMS Trans. Ed.* **7**(1). http://archive.ite.journal.informs.org/Vol7No1/TsaiWardell/.

Tsai, W., D. G. Wardell. 2006b. Creating individualized data sets for student exercises using Microsoft Excel and Visual Basic. *INFORMS Trans. Ed.* **7**(1). http://archive.ite.journal.informs.org/Vol7No1/TsaiWardell2/.

White, J. T., S. D. Myers. 2001. You can teach an old dog new tricks: The faculty's role in technology implementation. *Bus. Comm. Quart.* **64**(3) 95–101.

Winkelmann, L., R. Winkelmann. 1998. Why are the unemployed so unhappy? Evidence from panel data. *Economica* **65** 1–15.

Yin, L. R., L. Urven, R. Schramm, S. Friedman. 2002. Assessing the consequences of online learning: Issues, problems, and opportunities at the University of Wisconsin—Whitewater. *Assessment Update* **14**(2) 4–6.