

Undergraduate Student Outcomes:

A Comparison of Online and Face-to-Face Courses

Report from the Comparative Outcomes Subcommittee

For the:

University of Massachusetts Amherst Faculty Senate

Ad Hoc Committee on Online Learning

Martha L. A. Stassen, Academic Planning and Assessment, Chair

Marilyn Blaustein, Institutional Research

Richard Rogers, Provost Office

Mei-Yau Shih, Center for Teaching

Introduction

One of the charges to the University of Massachusetts Amherst's Ad Hoc Committee on Online Learning (ACOL) was to pursue comparative outcomes assessment between online and face-to-face courses. Activity of the Online Comparative Outcomes Sub-Committee to date includes the following:

1. Continuing and Professional Education (CPE) collaborated with the Office of Academic Planning and Assessment (OAPA) and the Center for Educational Software Development (CESD) to develop an improved mechanism for administering the campus' course evaluation instrument (Student Response to Instruction, or SRTI) in online courses. As a result, it is now possible to compare SRTI course evaluation items across modes of instruction with greatly improved response rates.
2. Two instructors who teach an online and face-to-face version of the same course volunteered to participate in a pilot comparative learning outcomes study. This project made it possible to compare student learning outcomes while also "controlling" on instructor.

The results from these analyses follow a summary of the current literature on comparative outcomes research.

Findings from the Literature

The growing popularity of online courses and programs has been accompanied by an interest in discovering the strengths and weaknesses of this mode of instruction. Studies include explorations into what elements of online learning should be assessed, issues of student characteristics (and possible differences across instructional methods), including demographics, learning style, students' fit with the online environment, and the extent to which subject matter or activities can be adapted to online instructional environments. In addition, there have been numerous small-scale studies that have focused directly on student outcomes, comparing student satisfaction and student performance in traditionally delivered courses and online courses across numerous disciplinary fields. In this discussion, we focus on the studies that directly address comparative outcomes.

Comparative Outcomes: Student Satisfaction and Rating of Instruction

Surveys comparing the satisfaction levels of students enrolled in online courses and those enrolled in the same course taught in a face-to-face setting sometimes reveal patterns of difference between the two populations. Students in online courses have been known to give lower ratings for the amount of support they received (Lim, D. H., Morris, M. L. & Kupritz, V. W. 2006) and the availability of the instructor (Gagne, M. & Shepherd, M. 2001) than those enrolled in traditional courses. These are some of the same areas of concern that have been identified as crucial for successfully creating a sense of belonging and engagement online (Churchill 2004). Overall levels of student satisfaction with online courses have been found to be influenced by such things as the provision of prompt feedback, the amount of emphasis instructors placed on interaction, and the use of innovative ways of teaching (Finlay, W., Desmet, C. & Evans, L. 2004).

Comparative Outcomes: Student Performance and Student Learning

It is well known among educational researchers that the great majority of studies comparing traditional and online courses have reached the same basic conclusion – that there is no significant difference in the attainment of learning outcomes for students enrolled in the two types of classes (Russell 1999; Meyer 2002). These studies were conducted within a wide range of disciplines, including English composition (Finlay, Desmet & Evans 2004), accounting (Gagne & Shepherd 2001), and health education (Davies & Mendenhall 1998). The general conclusion from this research is that the two methods offer equally viable environments for student learning. As Churchill (2004) notes, proponents of online learning urge moving past this finding, arguing that innovative course design could provide a superior experience to traditional classrooms; at the same time, the validity and reliability of some of these studies may be undermined because of the inconsistent use of controlling for extraneous variables (e.g., student motivation, time on task, variability in levels of maturity, differences in preferred learning styles & pedagogical approaches, etc.).

While researchers often accede to the conclusion that there is often no significant difference in the attainment of learning outcomes, they have also presented evidence that complicates the broadly drawn message, particularly in terms of the limitations or deficiencies in the studies upon which the conclusion is based. As a notable departure from the lack of rigor of earlier studies, Collins and Pascarella (2003) randomly assigned students to a traditional course and its counterpart that was taught at a distance, and found that students in both classes demonstrated an equivalent degree of learning; however, a third group of students who *elected* to take the course remotely performed significantly better than either of the other two groups, showing that self-selection for online courses could seriously confound the results of other studies (Collins & Pascarella 2003).

Another study, conducted at UMass Amherst, controlled for student interest in enrolling in the online course by randomly assigning students to either the face to face or online section of an introductory psychology course. Comparisons found that the two groups of students performed equally well in written assignments, but students in the online class performed better in the examinations (Poirier & Feldman 2004). For both instructional modes, exams were proctored. Using the same method of examination administration strengthens the researchers' findings by removing possible differences in performance related to different levels of monitoring.

There may be discipline-based variability in course effectiveness as well. One large scale, course-based (rather than student-based) study, which examined the effectiveness of online versus traditional courses for preparing for advanced courses, resulted in the finding that “the relative advantage of online is negative for Management enrollments and positive for non-Management courses” (Dominguez, P. S. & Ridley, D. R. 2001: 18).

UMass Amherst Comparative Analysis Pilot Results

UMass Amherst Comparative Outcomes: Student Ratings of Instruction (SRTI)

In Fall 2006 a new method was developed for administering the Student Response to Instruction (SRTI) course evaluation instrument online. Continuing and Professional Education (CPE) worked with the Office of Academic Planning and Assessment (OAPA) and

the Center for Educational Software Development (CESD) to develop an online SRTI administration method that greatly improved response rates within online courses.

This change in survey administration has made it possible to conduct comparative analyses of students' ratings for online course instruction and face-to-face instruction. The initial set of comparisons is provided in this report as an illustration of the kinds of analyses that will be possible down the road. However, these results must be taken as very preliminary because they reflect only one semester's worth of SRTI data, with a small number of online and face-to-face courses sponsored by CPE. In other SRTI comparisons conducted by OAPA, a full year's worth of data (or, even more typically, a rolling three year set of data) is used to guard against drawing judgments based on the anomalies of a single semester.

Tables 1A and 1B show the SRTI results for undergraduate Continuing and Professional Education (CPE) courses (1A) and for graduate CPE courses (1B), comparing the results for CPE online and CPE face-to-face courses. These comparisons are made for courses in two enrollment categories 20 or fewer and 21-50. In addition to showing the number of courses included in each analysis, these tables show the item mean and standard deviation across the two groups and two indicators of the extent of the differences between the groups.

The first measure of the strength of the differences is the statistical significance, or "p-value." This statistic indicates whether there is a statistically significant difference across the two groups (p-values of $\leq .05$ are indicated in bold). Statistical significance alone, however, does not guarantee that the difference is substantive or important. Large sample sizes (like those produced by campus-wide SRTI results) can produce statistically significant results even though the magnitude of the difference may be inconsequential, whereas differences present in small sample sizes (like the sample size we currently have for SRTI results for online courses) may not reach statistical significance, but still reflect large differences. To determine the magnitude of the effect while taking small sample sizes into consideration, we use Effect Sizes. Effect Sizes (ES) indicate the "practical significance" of the mean difference. It is calculated by dividing the mean difference by the pooled standard deviation for both groups. In practice, an effect size of .2 is often considered small, .5 moderate, and .8 large. In these tables, statistically significant differences and Effect Sizes of .5 or larger are in bold. Together the statistical significance and the Effect Size provide insight into the differences and similarities in ratings across online and face-to-face courses.

A few general observations can be made in looking over these preliminary comparisons. First, it is important to note the relatively small number of courses that make up these analyses. Again, these reflect only one semester's worth of data (Fall 2006). Second, looking at the pattern of difference, the Effect Sizes tend to be in the negative direction, indicating that the differences that do exist reflect cases where online courses are rated lower than traditional CPE courses.

**Table 1A. Fall 2006 SRTI Mean Comparisons - Undergraduate Course Sections
Courses with 20 or Fewer Enrolled**

Item	Online (CE)			Traditional (CE)			p-value ^A	Effect Size ^B
	Avg. Resp. Rate = 72%			Avg. Resp. Rate = 90%				
	N	Mean	St. Dev.	N	Mean	St. Dev.		
The instructor explained course material clearly. ¹	19	4.25	0.60	13	4.47	0.42	0.47	-0.41
The instructor cleared up points of confusion. ¹	19	4.26	0.73	13	4.48	0.44	0.69	-0.35
The instructor inspired interest in the subject matter of this course.	19	4.19	0.61	13	4.40	0.66	0.44	-0.33
The instructor showed a personal interest in helping students learn	19	4.09	0.79	13	4.56	0.51	0.07	-0.68
I received useful feedback on my performance on tests, papers etc.	19	3.77	0.96	13	4.22	0.78	0.17	-0.51
The methods of evaluating my work were fair. ¹	19	4.35	0.54	13	4.54	0.50	0.25	-0.37
The instructor stimulated student participation in the class. ¹	19	3.94	0.92	13	4.43	0.55	0.16	-0.62
Overall, how much do you feel you have learned in this course? ²	19	3.70	0.63	13	3.89	0.62	0.41	-0.30
What is your overall rating of this instructor's teaching? ³	19	4.13	0.62	13	4.36	0.54	0.33	-0.38
What is your overall rating of this course? ⁴	19	3.70	0.71	13	4.03	0.67	0.23	-0.48

Courses with 21 to 50 Enrolled

Item	Online (CE)			Traditional (CE)			p-value ^A	Effect Size ^B
	Avg. Resp. Rate = 66%			Avg. Resp. Rate = 85%				
	N	Mean	St. Dev.	N	Mean	St. Dev.		
The instructor explained course material clearly. ¹	14	4.43	0.31	12	4.54	0.48	0.14	-0.27
The instructor cleared up points of confusion. ¹	14	4.49	0.40	12	4.50	0.42	0.88	-0.02
The instructor inspired interest in the subject matter of this course.	14	4.34	0.43	12	4.44	0.51	0.41	-0.22
The instructor showed a personal interest in helping students learn	14	4.41	0.52	12	4.65	0.38	0.23	-0.51
I received useful feedback on my performance on tests, papers etc.	14	4.17	0.51	12	4.51	0.49	0.07	-0.69
The methods of evaluating my work were fair. ¹	14	4.56	0.27	12	4.58	0.39	0.63	-0.06
The instructor stimulated student participation in the class. ¹	14	4.45	0.36	12	4.51	0.39	0.68	-0.15
Overall, how much do you feel you have learned in this course? ²	14	3.94	0.32	12	3.94	0.57	0.82	-0.01
What is your overall rating of this instructor's teaching? ³	14	4.41	0.39	12	4.38	0.60	1.00	0.06
What is your overall rating of this course? ⁴	14	3.97	0.40	12	4.06	0.58	0.52	-0.20

**Table 1B. Fall 2006 SRTI Mean Comparisons - Graduate Course Sections
Courses with 20 or Fewer Enrolled**

Item	Online (CE)			Traditional (CE)			p-value ^A	Effect Size ^B
	Avg. Resp. Rate = 98%			Avg. Resp. Rate = 91%				
	N	Mean	St. Dev.	N	Mean	St. Dev.		
The instructor explained course material clearly. ¹	4	4.20	0.33	25	4.53	0.45	0.08	-0.75
The instructor cleared up points of confusion. ¹	4	4.23	0.21	25	4.54	0.41	0.07	-0.79
The instructor inspired interest in the subject matter of this course.	4	4.33	0.21	25	4.55	0.32	0.11	-0.70
The instructor showed a personal interest in helping students learn	4	4.38	0.22	25	4.73	0.34	0.03	-1.08
I received useful feedback on my performance on tests, papers etc.	4	4.16	0.20	25	4.35	0.59	0.24	-0.33
The methods of evaluating my work were fair. ¹	4	4.61	0.21	25	4.64	0.40	0.41	-0.07
The instructor stimulated student participation in the class. ¹	4	4.21	0.27	25	4.62	0.33	0.03	-1.29
Overall, how much do you feel you have learned in this course? ²	3	3.73	0.06	25	3.94	0.58	0.28	-0.37
What is your overall rating of this instructor's teaching? ³	4	4.09	0.18	25	4.45	0.42	0.05	-0.92
What is your overall rating of this course? ⁴	3	3.75	0.27	25	4.13	0.59	0.13	-0.67

Courses with 21 to 50 Enrolled

Item	Online (CE)			Traditional (CE)			p-value ^A	Effect Size ^B
	Avg. Resp. Rate = 72%			Avg. Resp. Rate = 95%				
	N	Mean	St. Dev.	N	Mean	St. Dev.		
The instructor explained course material clearly. ¹	16	4.40	0.43	9	4.57	0.37	0.29	-0.43
The instructor cleared up points of confusion. ¹	16	4.36	0.45	9	4.60	0.38	0.18	-0.58
The instructor inspired interest in the subject matter of this course.	16	4.33	0.47	9	4.54	0.45	0.15	-0.46
The instructor showed a personal interest in helping students learn	16	4.39	0.40	9	4.74	0.23	0.02	-1.02
I received useful feedback on my performance on tests, papers etc.	16	3.95	0.65	9	4.32	0.41	0.11	-0.65
The methods of evaluating my work were fair. ¹	16	4.47	0.40	9	4.62	0.34	0.51	-0.37
The instructor stimulated student participation in the class. ¹	16	4.31	0.42	9	4.53	0.46	0.15	-0.51
Overall, how much do you feel you have learned in this course? ²	12	3.68	0.43	9	4.16	0.51	0.05	-1.02
What is your overall rating of this instructor's teaching? ³	16	4.28	0.48	9	4.57	0.37	0.12	-0.63
What is your overall rating of this course? ⁴	12	3.75	0.48	9	4.29	0.48	0.03	-1.13

^A p < .05 is highlighted in bold.

^B Effect size is the "practical significance of the mean difference. It is calculated by dividing the mean difference by the pooled standard deviation for both groups. In practice, an effect size of .2 is often considered small, .5 moderate, and .8 large.

¹ 1 = Almost never; 2 = Rarely; 3 = Sometimes; 4 = Frequently; 5 = Almost always.

² 1 = Much less than most courses; 2 = Less than most courses; 3 = About the same as others; 4 = More than most courses; 5 = Much more than most courses.

³ 1 = Almost never effective; 2 = Rarely effective; 3 = Sometimes effective; 4 = Usually effective; 5 = Almost always effective.

⁴ 1 = One of the worst; 2 = Worse than average; 3 = About average; 4 = Better than average; 5 = One of the best.

Undergraduate Course SRTI Results

Among undergraduate CPE courses, there are no statistically significant differences, but for both enrollment categories there are a few items that show moderately large Effect Sizes. The items reflecting these more substantial differences across the two modes of instruction seem to mirror findings from the broader literature – the differences that emerge between ratings for online and face-to-face courses tend to reflect issues related to the amount of support students receive and the availability of the instructor (Lim, D. H., Morris, M. L. & Kupritz, V. W. 2006; Gagne, M. & Shepherd, M. 2001). This appears to be the case in these UMass Amherst analyses, as well, where the items with the largest difference reflect instructor attention, feedback, and student participation. As discussed in the literature review, these findings highlight the particular challenge of creating a sense of belonging and engagement in online environments (Churchill, 2004).

Graduate Course SRTI Results

The differences across the two instructional modes in graduate CPE courses are more persistent. While only a few reach statistical significance, the majority of item comparisons show Effect Sizes of .50 or greater. With only one semester worth of data, it is impossible to know if this is an anomaly, or a pattern that will maintain when there is a larger pool of courses in the analyses. However, the results clearly suggest the importance of continuing to conduct these comparisons and monitor results.

UMass Amherst Comparative Outcomes: Student Performance

As a pilot project, we analyzed the differences and similarities in student performance in two courses that were each taught in both an online environment and in a face-to-face environment by the same instructor. The two courses are from a professional discipline, one is a primarily graduate course, the other undergraduate. By having the same instructor teaching both the online and face-to-face version of the course, we are able to “control” on differences in course content and instructor.

Comparisons on Background and Academic Status Differences

One of the challenges in conducting comparisons in student performance across online and face-to-face courses is that, in many cases, the two student populations vary in important ways. This is the case in the comparison populations for this study. Table 2A shows selected background and demographic information for the two undergraduate course populations as well as an indication of which differences reach statistical significance. For this group, there are no significant differences on number of credits accrued, cumulative college GPA or in SAT scores. However, the online student population has a lower high school GPA, is older, and has more non-degree and transfer students.

Table 2B shows a similar analysis of relevant demographic and academic preparation variables for the graduate course population. Here we also see no significant difference in cumulative college GPA and the standardized test (GMAT); there is also no significant difference in age. However, with this population, the significant differences are in the cumulative credits (online students have collected fewer credits) and, again, in academic status (more online students are non-degree students).

These tables illustrate another complexity in conducting statistical comparisons across these populations. Note the differences in the “n” for various demographic and academic preparation variables. Many fewer online students have indicators of previous

academic performance (standardized test results, high school GPA, and in the case of the Undergraduate course, cumulative College GPA). This is at least in part a function of the larger number of non-degree students who make up the online course population. The missing data complicates the task of developing a multivariate statistical model that controls on characteristics particularly relevant to understanding the effect online course enrollment has on student performance because some variables cannot be used without significantly decreasing the number of students included in the analysis.

Table 2. Demographic/Academic Preparation Variables

2A. Undergraduate Course: Online vs. Traditional

CONTINUOUS VARIABLES		Online (n=21)		Traditional (n=46)		p-value
		n	Means	n	Means	
Cumulative Credits		20	31.4	46	34.8	NS
Cumulative GPA		11	3.3	46	3.4	NS
Math SAT		7	625.7	46	598.3	NS
Verbal SAT		7	567.1	46	568.9	NS
High School GPA		13	3.3	46	3.7	0.02
Age		21	27.6	46	19.8	<.001

CATEGORICAL VARIABLES		Online (n=21)		Traditional (n=46)		p-value
		n	Percent	n	Percent	
Academic Career	Undergraduate	13	61.9	46	100	<.001 ^a
	Non-degree	8	38.1	0	0	
Admit Type	First-year	4	19.1	46	100	<.001 ^a
	Transfer	7	33.3	0	0	
	Other	10	47.6	0	0	

2B. Graduate Course Online vs. Traditional

CONTINUOUS VARIABLES		Online (n=33)		Traditional (n=12)		p-value
		n	Means	n	Means	
Cumulative Credits		33	19.1	12	28.8	0.002
Cumulative GPA		28	3.7	11	3.5	NS
GMAT		15	558.7	10	600.0	NS
AGE		33	32.83	12	29.8	NS

CATEGORICAL VARIABLES		Online (n=33)		Traditional (n=12)		p-value
		n	Percent	n	Percent	
Academic Career	Graduate	19	57.6	12	100	0.02 ^a
	Undergraduate	1	3.0	0	0	
	Non-degree	13	39.4	0	0	

^a Due to multiple cells with small N's, Chi-Square results may be unstable.

Comparing Student Performance: Course Outcomes

We pursue the question of whether there are differences in students' performance across the instructional methods in two ways: (1) simple bivariate comparisons in student performance on a number of indicators of student learning and (2) multivariate analyses that explore the effect of online instruction on student performance after controlling on potentially relevant student characteristics.

Outcomes Comparisons: Bivariate Analysis

Table 3 shows the bivariate comparisons of student performance on a number of different outcomes. The analysis for the undergraduate course (Table 3A) uses five measures of student performance: total exam score, score on online homework, grades on two project-based assignments, and the final course grade. As the p-values column indicates, there are no statistically significant differences in student performance between online and face-to-face instruction on any of these outcomes.

For the graduate course (Table 3B) there are three outcomes used: total exam score, performance on a class presentation, and final grade. There are no significant differences in student performance on the exams or on the class presentation. However, differences do emerge in the final grade, with students in the online course achieving a higher grade than those in the face to face course. This difference is most likely related to the students' grades in a project-based assignment that is not included among the outcomes studied here. For this assignment, the instructor used different grading formulas across the two modes of instruction (related to differences in the assignment itself across the two instructional modes). This difference in how grades are calculated probably accounts for the differences in final grade evident in Table 3B.

Table 3. Bivariate Comparisons of Undergraduate Student Performance Outcomes

3A. Undergraduate Course: Online vs. Traditional

	Online (n=21)	Traditional (n=46)	
	Means	Means	p-value
Exam Total	77.83	78.94	NS
Online Homework	95.52	98.22	NS
Project 1	0.86	0.90	NS
Project 2	0.96	0.95	NS
Final Grade	83.92	86.55	NS

3B. Graduate Course: Online vs. Traditional

	Online (n=33)	Traditional (n=12)	
	Means	Means	p-value
Exam Total	79.64	79.42	NS
Class Presentation	93.06	92.25	NS
Final Grade¹	85.57	74.19	0.001

¹Instructor indicates differences in grading formula for one assignment probably explains this difference in final grade across the two courses.

Outcomes Comparisons: Multivariate Analyses

Of course, given the differences in the online and traditional student populations, the simple bivariate analyses may mask differences in performance across the two instructional formats. To test for this, we ran multiple regression analyses, looking at the effect of online course enrollment on student performance after controlling on student characteristics. It is important to note that our capacity to conduct the most complete multivariate analysis is hampered by the reality and limits of this pilot project. The problem of missing data on key predictor variables has already been discussed. In addition, the

relatively small number of subjects (students) in the study limits the number of variables we can put in our model.

With this in mind, we developed fairly parsimonious models, selecting predictor variables using the following criteria: variables that are moderately correlated with at least two of the outcome variables and variables with fewer missing data within the study population. One variable, student academic status (undergraduate, graduate, or non-degree), did not make the multivariate analysis, even though it was correlated with two outcomes, because of problems with collinearity. Academic status is highly correlated with online enrollment status (non degree students are much more likely to be in the online section than the face-to-face section, see Tables 2A and 2B). As a result, this variable was removed from the model.

The undergraduate course regression results are presented in Table 4A. The Online Course effect (after controlling on students' ALANA status, Cumulative College GPA, and Cumulative College Credits) is shown, with statistically significant effects highlighted in bold. As Table 4A indicates, online instruction has no significant effect on three of the four outcomes (Total Exam score, Project 2, or Final Grade). However, it does show a significant positive effect on the Project 1 (meaning, after controls, students in online courses perform somewhat better on Project 1 than students in the non-online course).

Table 4. Multivariate Linear Regression Analyses of Learning Outcome Variables 4A. Undergraduate Courses: ^b

	Multivariate Model: Online Course Estimates for Each Outcome (Controlling on ALANA status, Cumulative GPA, and Cumulative Credits)			
	Exam	Project 1	Project 2	Final Grade
Online Course Effect	3.13	.04 **	.02	.69
Summary Statistics				
R-Square	.762	.441	.553	.749
Number of Observations	(57)	(57)	(57)	(57)

^bDue to missing data on Cumulative GPA for Online Students, these models includes only 11 of the 21 online students and all 46 of the traditional students. Significant at *p≤ .05, **p≤ .01; ***p≤.001.

Table 4B shows the graduate course regression results. The Online Course effect (after controlling student Age and Cumulative College Credits) is shown with statistically significant effects highlighted in bold. In this case, online course status has no relationship with Total Exam score or with students' Class Presentation. However, it does have a significant relationship with students' Final Grade. As indicated earlier, the significant difference in final grade is most likely the function of the differential grading procedure (and assignment structure) of the project-based assignments in this class.

4B. Graduate Course:

	Multivariate Model: Online Course Estimates for Each Outcome (Controlling on Age and Cumulative College GPA)		
	Exam	Class Presentation	Final Grade
Online Course Effect	-4.37	-.57	11.08 ***
Summary Statistics			
R-Square	.389	.395	.570
Number of Observations	(39)	(39)	(39)

Significant at *p≤.05, **p≤.01; ***p≤.001.

Summary: Initial Findings

In this report, we have analyzed the similarities and differences between online and face-to-face courses in two ways: (2) comparing students’ ratings of the course and course instruction and (2) comparing student performance on common course outcomes. These analyses represent an initial step in the University’s capacity to conduct comparative analyses and, as a result, should be considered preliminary at best. As the availability of comparative data expands over the next year, we will be able to make more conclusive judgments about the similarities and differences across these instructional modes.

The preliminary results from one semester’s worth of Student Response to Instruction (SRTI) course evaluation data do suggest differences in students’ ratings of online and face-to-face courses, with online courses generally receiving lower ratings on a number of SRTI items. Some of these differences (on items related to support and instructor availability) are not particularly surprising since they echo differences found in larger comparative studies found in the literature. These differences, which are the only substantial differences found among undergraduate courses, reflect some of the particular pedagogical challenges that instructors of online courses face.

The graduate course SRTI comparisons show more substantial Effect Size differences (although most do not reach statistical significance) and cover a wider range of SRTI items than is true for undergraduate courses. While the small number of courses makes it difficult to draw conclusions, these results do suggest the importance of continuing to pursue analyses that explore the possibilities of differences in student ratings of courses taught in these two instructional modes, particularly with respect to graduate course offerings.

The comparative analysis of student learning outcomes focused on the differences and similarities in student characteristics and performance in two courses in a professional discipline, one an undergraduate course and one graduate course, that were each taught in both an online and face-to-face

environment. The research design made it possible to control on one important factor in conducting course-based comparisons of student performance – differences in instruction and course content. In this case, the online and face-to-face courses used in the comparisons were taught by the same instructor and used the same exams as well as other course materials.

However, the pilot study results also illustrate some of the other complexities in conducting these types of comparative analyses. Specifically, the student populations across the two instructional modes vary in some important ways. These differences affected the extent to which some academic preparation variables (like standardized tests, previous GPA, and academic standing) could be used as controls in testing for differences in academic performance.

In simple bivariate comparisons of student performance on various outcomes across the two instructional modes, only one statistically significant result emerged (final grades in the graduate course). And this difference can be explained by a difference in grading procedures on one set of assignments that varied across the two instructional modes. In the multivariate analyses that looked at the effect of online instruction after controlling on relevant predictors of performance, online instruction showed no significant effect on most of the outcomes. The two exceptions were, again, the final grade in the graduate course and, in the undergraduate course, a difference in performance on one of two project-based assignments. With the limits inherent in this data set, it is difficult to determine what this second results might mean – particularly since no statistically significant differences emerged in the simple bivariate analyses (see Table 3A).

While preliminary, these results generally support findings from the broader research on online instruction – that there are few differences in student learning outcomes across the two instructional modes. Of course, it would be beneficial to continue this line of investigation and explore whether differences exist in certain disciplines or certain types of courses. With the limited sample of courses available for this study, this type of more detailed analysis is not possible. However, with more instructors willing to volunteer their courses for this type of analysis, we could expand this study.

Bibliography

Arbaugh, J. B. (2000). Virtual Classroom Versus Physical Classroom: An Exploratory Study of Class Discussion Patterns and Student Learning in an Asynchronous Internet-based MBA course. *Journal of Management Education* 24(2), 213-233.

Arbaugh, J. B. and Rau, B. L. (2007). A Study of Disciplinary, Structural, and Behavioral Effects on Course Outcomes in Online MBA Courses. *Decision Sciences Journal of Innovative Education*, 5(1), 65-95.

Brown, B. W. & Liedholm, C. E. (2002). Can Web Courses Replace the Classroom in Principles of Microeconomics? *The American Economic Review* 92(2) 444-448.

Churchill, A. (2004). *Ensuring Quality in Online Higher Education Courses*. Amherst, MA: University of Massachusetts Center for Education Policy.

Collins, J. & Pascarella, E. T. (2003). Learning on Campus and Learning at a Distance: A Randomized Instructional Experiment. *Research in Higher Education* 44(3), 315-326.

Davies, R. S. & Mendenhall, R. (1998). Evaluation Comparison of Online and Classroom Instruction for HEPE 129 – Fitness and Lifestyle Management Course. Evaluation project for Department of Instructional Psychology and Technology, Brigham Young University. Accessed 4/1/07, ERIC document #427752.

Diaz, D. P. & Cartnal, R. B. (1999). Students' Learning Styles in Two Classes. *College Teaching* 47(4), 130-141.

Dominguez, P. S. & Ridley, D. R. (2001). Assessing Distance Education Courses and Discipline Differences in their Effectiveness. *Journal of Instructional Psychology* 28(1): 15-19.

Drago, W. A. & Wagner, R. J. (2004). Vark Preferred Learnign Styles and Online Education. *Management Research News* 27(7), 1-13.

Edmonds, C. L. (2006). The Inequivalence of an Online and Classroom Based General Psychology Course. *Journal of Instructional Psychology* 33(1), 15-19.

Finlay, W., Desmet, C., & Evans, L. (2004). Is it the Technology or the Teacher? A Comparison of Online and Traditional English Composition Classes. *Journal of Educational Computing Research* 31(2), 163-180.

French, D., Olrech, N., Hale, C., & Johnson, C. (2006). *Blended Learning: An Ongoing Process for Internet Integration*. Trafford and E-linkages Inc.

Gagne, M. & Shepherd, M. (2001). Distance Learning in Accounting: A Comparison between a Distance and a Traditional Graduate Accounting Class. *T.H.E. Journal*, 28(9), 58-60.

Greenlaw, S. A. & DeLoach, S. B. (2003). Teaching Critical Thinking with Electronic Discussion. *Journal of Economic Education*, 34(1), 36-52.

Holian, R. (2004). The Practice of Management Education in Australian Universities. *Management Decision*, 42(3/4), 396. Accessed 5/1/07 at <http://proquest.umi.com.silk.library.umass.edu:2048/pqdweb?did=640554871&Fmt=4&clientId=2724&RQT=309&VName=PQD>

Johnson, D., Burnett, M., & Rolling, P. (2002). Comparison of Internet and Traditional Classroom Instruction in a Consumer Economics Course. *Journal of Family and Consumer Sciences Education*, 20(2), 20-28.

Meyer, K. A. (2002). Quality in Distance Education. ERIC Digest #ED470542.

Mupinga, D. M., Nora, R. T., & Yaw, D. C. (2006). The Learning Styles, Expectations, and Needs of Online Students. *College Teaching* 54(1), 185-189.

Neuhauser, C. (2002). Learning Style and Effectiveness of Online and Face-to-Face Instruction. *The American Journal of Distance Education* 16(2), 99-113.

Poirier, C. R. & Feldman, R. S. (2004). Teaching in Cyberspace: Online Versus Traditional Instruction Using a Waiting-List Experimental Design. *Teaching of Psychology* 31(1), 59-62.

Russell, T. L. (1999). *The No Significant Difference Phenomenon*. Raleigh: North Carolina State University Press.

Scheines, R., Leinhardt, G., Smith, J. & Cho, K. (2005). Replacing Lecture with Web-based Course Materials. *Journal of Educational Computing Research* 32(1): 1-26.

Schoenfeld-Tacher, R., McConnell, S., & Graham, M. (2001). Do No Harm – A Comparison of the Effects of Online vs. Traditional Delivery Media on a Science Course. *Journal of Science Education and Technology* 10(3), 257-265.

Wang, M. (2007). Designing Online Courses that Effectively Engage Learners From Diverse Cultural Backgrounds. *British Journal of Educational Technology*, 38(2), 294-311.

White, S. & Liccardi, I. (2006). Harnessing Insight into Disciplinary Differences to Refine e-learning Design. 36th ASEE/IEEE *Frontiers in Education Conference*, Oct. 28-31, Session S3J.

Yuretich, R. & Schneider, S. (Under Review). Planet Earth: Hands-On Earth Science Online. 2006 Association of Science Teacher Education Annual Conference Proceedings.