Journal of Business and Behavioral Sciences
Vol 21, No. 2; Spring 2010

LEARNING OUTCOMES AND INSTRUCTIONAL DELIVERY METHOD IN PROFESSIONAL AND BUSINESS RELATED COURSES: AN EMPIRICAL STUDY CONTROLLING FOR COURSE AND INSTRUCTOR DIFFERENCES

Gregory B. Merrill
National University
Craig S. Galbraith
University of North Carolina Wilmington

ABSTRACT

In this study we compare the student performance associated with instructional delivery methods, including pure on-line, VOIP, hybrid, and traditional classroom environments using a large sample of 2,077 students enrolled in business, criminal justice, and public administration courses. We also address the criticisms of prior empirical studies by using an explicit measure of course learning outcomes and controlling for both course and instructor differences. The results indicate that with the possible exception of hybrid classes there are generally no significant differences in student attainment of course learning outcomes associated with the different delivery methods, even when controlling for course and instructor. In addition, we examine the correlation between course grades, student perceptions of learning, and our objective measure of course learning outcome attainment.

INTRODUCTION

On-line course offerings in the professional disciplines, such as business, medicine, education, and nursing has exploded in recent years, driven, in part, by more sophisticated and available technology and, in part, by increasing demand for such courses, particularly by "non-traditional" university students. Similarly, there have been an increasing number of scholarly research efforts, both theoretical and empirical, designed to investigate the potential for learning achievement in on-line course offerings, particularly as they compare to the more traditional classroom setting.

Theoretical Considerations: Educational theory suggests that there are three major components to understanding the effectiveness of instructional activity – linking new information to old, acquiring meaningful knowledge, and stimulation the use of meta-cognitive activities (Ally, 2004). The role that instructional technologies play in achieving these three objectives has been the

subject of substantial theoretical debate over the past three decades, first with respect to computer assisted learning in the classroom, and more recently, with respect to on-line and internet-based learning (Clark, 2001, Kozma, 2001). On one hand, there is a dominant school of educational theory thought that suggests that delivery technologies do not, by themselves, significantly influence student achievement (e.g., Schramm 1977, Clark, 1983), but simply offer the mechanisms to adjust instructional strategy to fit the type of technology being used. Under this model, technologies, such as the computerized on-line learning, simply become vehicles of presentation, with content and instructional strategy remaining the primary components that influence student achievement. Even within this school of thought, however, there is substantial disagreement. While recognizing the role of both content and instructional strategy, some educational theorists appear to place greater emphasize on the role of content, while others emphasize the importance of designing an appropriate instructional strategy that corresponds with the technology (Bonk & Reynolds, 1997; Kozma, 2001).

To a large, extent this particular view of on-line learning is based upon the cognitive approach to education theory where learning is primarily an internal process that involves reflection, motivation, and thinking (Ertmer & Newby, 1993; Janicki & Liegle, 2001). Within this context, the importance of the student's processing capabilities, the amount of effort expended, the depth of processing, and the student's existing knowledge ultimately determines the amount of learning (Craik & Lockhart, 1972; Kalat, 2002). Not surprising within the cognitive theoretical perspective, on-line learning has historically been viewed suspiciously due to the lack of control regarding student motivation and effort, and the lack of immediate feedback (such as body language) regarding student processing capabilities and resulting memory. Any effective use of on-line learning must therefore focus on the content and instructional strategy that enables the student to more efficiently process the material into memory, such as "chunking" material into semi-autonomous manageable components (Ally, 2004).

On the other hand, there is a growing body of education theory that suggest that unlike the technologies associated with traditional classroom environments, the technologies associated with on-line learning create a fundamental change in the educational process. On-line instruction, according to Cole (2000) alters the normal learning process by collapsing "time and space", particularly within the context of "asynchronous" delivery. On-line learners, it is argued, simply access materials differently, thus altering both the linking process between old and new information, and the way that meta-cognitive activities are stimulated.

This view of on-line education seems to be grounded more within a constructivist perspective of learning. Constructivist theorists tend to view learners as interpreting information and knowledge within their particular reality, thus personalizing information into knowledge (Cooper, 1993; Wilson, 1997). Here technology and learning cannot be separated, but rather form a combined reality. Since the technology becomes part of the personal instructional world

reality, learners are seen as more "active" participants than "passive" recipients. Learners therefore "construct" knowledge, rather than be "given" knowledge through instruction. As such, the on-line environment fundamentally alters the construction of knowledge within the learner. In general, constructivist theory tends to view on-line, and particularly asynchronous, learning more positively, arguing that its unique characteristics can actually lead to an enhanced personal instructional reality based upon "hyper-communication" information flows (Walther, 1996, Wenger, 2001).

An underlying theme of this more constructivist view of on-line learning is that the cyberspace has its own unique culture, geographical nature and human law, Thus the key to student achievement is not so much the strategy of the instructor, but rather the particular orientation and internet "efficacy" that the student brings to the on-line environment (Eastin & LaRose, 2000).

For example, compared to the traditional classroom environment, the web provides an immense amount of supplemental material that can be instantly accessed parallel to the defined course material -- but to the inexperienced user this can be overwhelming and counterproductive (Anderson, 2004). In addition, under the lens of community-centered learning (Lipman, 1991; Wenger, 2001; Wilson, 2001) the ability to develop a "learning community" on-line is often viewed as a cultural skill that is acquired from other broader on-line activities of the student. Learners not culturally comfortable with developing on-line interactive communities will likely experience difficulties, such as lower attention rates, when asked to participate in on-line learning communities (Mason & Hart, 1997; Hine, 2000). Within this constructivist framework, it can been argued that student achievement within an on-line environment may be less dependent on the teacher's skill or instructional strategy, but more a function of the cyberspace comfort and community-building skills of the learner – and clearly, with a new generation of learners and the wide availability of the web, this new culture is rapidly developing. Under this more constructivist perspective, one would expect little or no difference in achievement between traditional and on-line environments for modern, web-experienced students, and possibly improvement given the opportunity to develop expanded learning community experiences within an on-line environment (Anderson, 2004; Harasim et al, 1995).

Empirical Issues. Parallel to the growing theoretical debate regarding the role of technology and on-line education, the empirical examination of on-line learner achievement has also exploded. Prior empirical research appears to indicate that, in general, student performance in on-line and traditional classroom versions of the same course is similar. This conclusion appears consistent across many professional disciplines (e.g., Frydenberg, 2007; Corral et al, 2006; Buckley, 2003; Patterson & Harasym, 2001), which has now led to the "no significant difference" mantra (Russell, 2001, 2008; Coates et al, 2004).

In fact, there is some empirical evidence pointing to perhaps a slightly better learning experience coming from the on-line courses (e.g., Russell, 2001; Hersh et al, 2001; Lipman et al, 2001; Maki et al, 2000), a trend that appears to

support a more constructivist educational perspective. For example, in studies where a significant achievement difference was found between traditional and on-line classes, there was about a four to one ratio favoring on-line education versus traditional classroom. (Russell 2001, 2008). However, a recent meta-analysis of 232 studies between 1985 and 2002 indicated a high degree of variability, suggesting that in some applications distance learning, including various asynchronous forms, substantially outperform traditional classes, while other applications favored classroom instruction (Bernard et al, 2004). The authors, however, suggest that the major problem with interpreting the results of prior research is the poor experimental design of previous comparison studies, such as not controlling for course or instructor differences.

Likewise empirical research to date suggests that some subjects taught in professional schools may, indeed, do better with the traditional format. For example, the academic subjects within Business Schools where student achievement may favor traditional classroom versus on-line learning appear to be more applied mathematical in nature, such as in micro- and macroeconomics classes (Brown & Liedholm, 2002, 2004; Terry et al, 2003, Coates et al, 2004; Anstine & Skidmore, 2005), security analysis (Terry et al, 2003) and business statistics (Anstine and Skidmore, 2005) possibly due to the increased requirement for rapid instructor-student "feedback" in these types of classes. However, in their study of principles of macroeconomic classes, Harmon and Lambrinos (2006) found that on-line formats resulted in significantly higher student achievement while Navarro and Shoemaker (1999) found no difference, so it is difficult to draw any clear conclusions yet (Navarro, 2000).

While the vast majority of prior research comparing student performance between on-line and traditional face-to-face classroom courses is descriptive and exploratory in nature previously published studies that take a more experimental or larger-scale approach to the question still tend to suffer from serious design problems (see Bernard et al, 2004 and Tallent-Runnels et al, 2006 for a complete review of the validity problems associated with the vast majority of prior empirical studies). The most obvious of these problems are found in studies that pool different types of courses, without controlling for course differences – thus a marketing on-line course may end up being compared against an accounting classroom environment. Controlling for course differences may be important since student learning outcomes naturally vary by course, but should be the same for all sections of the same course (Tallent-Runnels et al, 2006).

Even studies that attempt to control for course differences, such as those examining performance among different instructional delivery methods among sections of the same course, suffer from various problems. The most important of these issues include not controlling for different instructors (e.g, Corral et al, 2006). This is important since not controlling for instructor differences (such as one instructor teaching on-line sections and another instructor teaching traditional classroom sections) introduces a potential confounding effect since we might actually be testing differences in instructors rather than differences in delivery methods. For example, in the empirical examinations of economics

classes discussed above, the often cited studies by Brown and Liedholm (2002, 2004), Terry et al (2003), and Coates et al (2004) all appear to examine courses taught by different instructors.

Another important issue is the measurement of "student performance." Most prior research has not employed objective learning outcome assessment measures, but rather employed other performance measures such as grades (e.g., Carswell et al, 2000) or more commonly, subjective student satisfaction assessments (e.g., Arbaugh, 2005; Frederickson et al, 2005; Shin & Chan, 2004; McCray, 2000). However, there is little empirical evidence to date that either overall course grades, or especially subjective student assessments, are highly correlated to objective measures of course learning outcomes (Darby, 2007; Worthington, 2002; Abrami, 1990). In fact, the usually cited criterion validation studies of student evaluations (for reviews, see Abrami et al, 1990 and Marsh, 1987) were conducted primarily during the 1980s, well before the modern notion of broader course learning outcome assessments.

The use of course learning outcomes for measuring student performance is, however, becoming increasingly relevant to instructional comparison studies. In fact, course-related learning outcome assessment has been mandated at most universities in the United States, driven in part by recent policy changes at the various professional school accrediting organizations, such as the Association to Advance Collegiate Schools of Business (AACSB), as well as the broader regional university accrediting agencies, such as the Western Association of Schools and Colleges (WASC). Recently, a few comparison studies have attempted to use standardized outcomes measures, such as the Test of Understanding College Economics (TUCE), but this has been limited to only a few more standardized "core" subjects, such as principles of economics (Coates et al, 2004).

A final problem that plagues much of the prior comparison research is sample size and research/instructor bias. Many empirical studies, particularly those that indeed do attempt to control for course and instructor differences, examine only a small number of courses, or even a single course, often with limited enrollment (e.g., Harris & Parrish, 2006; Frederickson et al, 2005; Anstine & Skidmore, 2005; Lipman et al, 2001). In addition, almost all previously published empirical studies use data where the researcher/article author is also the instructor of the courses, thus creating potential experimental bias. It should be noted that our data does not involve any courses taught by the authors of this paper.

The present study attempts to address the most serious of the criticisms leveled against prior empirical research that purports to compare the student performance associated with different instructional delivery environments. In particular, we control for both course and instructor differences, and employ a measure of student performance that specifically reflects the learning outcomes associated with each course examined. In addition, our study represents a large-scale sample; with a sample size over two thousand from a single university's School of Business.

DATA AND SAMPLE

The data comes from a large private university located in California. The university has traditionally used classroom instruction, but within the last decade has offered an increasing number of on-line courses. Currently the university is one of the largest providers of on-line academic courses in the United States, with on-line and traditional classroom delivery modes about equal in terms of numbers of courses offered. The university is presently offering a number of different on-line course variations, including pure on-line courses, Web-2 (VOIP) courses, and hybrid courses combining a traditional classroom setting with on-line components.

Our data comes from programs in business, criminal justice, and public administration, all administratively housed within the university's School of Business. These classes are offered at both the undergraduate and graduate (masters) level. The university makes substantial investment in on-line technologies. The School of Business has a large full-time resident faculty, but also utilizes a number of adjunct instructors under the direct supervision of a full-time faculty coordinator. Most of the students are non-traditional students, with an average age of 33 years.

The university has used course learning outcomes for every core and required course for several years, but has recently revised its learning outcome assessment process. Learning outcomes are established by a faculty committee for each core and concentration required course within the School of Business. There is an average of eight to ten learning outcomes per course.

Assessment of course learning outcome attainment by students is measured by a standardized learning outcome assessment test (SLOAT). The School's SLOAT exams are developed individually for each course in each program by a committee of content experts in the subject area, with questions designed to assess each of the stated learning outcomes. As of the 2006-2007 academic year approximately one-third of the School's core and required courses are given SLOAT exams.

SLOAT exams are administered to every student in every section of the course being assessed, regardless of instructor and delivery mode. SLOAT exams are given at the end of the course period. The exams are scored on a basis of zero (0) to four (4) points per learning outcome. For the present study, for the SLOAT score we use the mean score of all the SLOAT learning outcome questions on a particular SLOAT exam.

Depending on the course, SLOAT exams consist of multiple choice or short essay questions, or a combination of both. For the short essay questions a grading rubric is created so that there is consistency in scoring across all sections of a particular course. Approximately 10% of all SLOAT exam content is essay, with approximately 90% multiple choice. Exam design and rigor is similar to professional exams such as the C.P.A. exam. The School makes every effort that the SLOAT exams are completely objective and consistent when evaluating

the degree of learning that has been achieved by a student. SLOAT data was collected and analyzed for the entire academic year 2006-2007. A total of 2,077 SLOAT scores from 159 course sections were analyzed.

Other data collected included overall student grades and student course evaluations. As with most universities, the student course evaluations are multiple question evaluations (19 items) that gather student perceptions, with several questions directly related to perceptions of "learning experience" and "knowledge gained". For the "student perception of learning" measure we used the mean value of the seven items directly related to course learning experience (on a five-point Likert scale). For "student perception of learning", we had complete data from 127 sections (versus our full sample of 159 sections). For this analysis, we use these two measures (student grades and "student perception of learning") only to examine the degree of correlation with the more objective SLOAT measure of learning outcome assessment. Instructors were requested to submit the data to the researchers; with all student performance data being completely anonymous. Our sample consists of over 90% of all the SLOAT exams administered during the 2006-2007 academic year within the School of Business.

There are four modes of instruction used by the University, with strict policies regarding course design for all instructors. For example, all "pure on-line" instructors must follow the same delivery format and protocols. All classes meet two times per week.

Traditional Classroom: The traditional Brick and Mortar classroom approach with a live instructor in a physical classroom. There are no course websites available to supplement the classroom contact time. Students have live interaction with the professor every class session, as well as before and after class and during office hours. Exams are administered in class. The course meets a standard semester long 45 contact hours.

Pure On-line: A course instructional website (*E-college*) is used for course delivery. All graded assignments are submitted to the website. All exams are taken online. Participation in asynchronous threaded discussions and synchronous chat sessions is mandatory for every student and every course. Chat sessions are held using two techniques, traditional text based chat, and voice over IP (VOIP). In both cases a web tool for live document sharing is available to instructors and students during the chat sessions. Document sharing allows the professor to display material (Word documents, Excel spreadsheets and PowerPoint presentations) on the computer screen for all chat attendees to view and discuss just as if this was a traditional Brick and Mortar live class. The number of synchronous contact (chat) hours is approximately one-quarter to one-third that of the traditional course, the rest of the time is spent in other asynchronous online activities. The only contact the students have with the professor is through chats and email.

WEB-2: all class sessions are held over the internet using VOIP and document sharing the same as the online courses, without benefit of a web course site. The instructors and students carry on traditional classroom discussions and

board assignments using the computer as the interface instead of the restriction of the traditional classroom. Unlike online courses, Web-2 classes do not have course websites available to the students. There are no threaded discussions. All interaction between the faculty member and students is during synchronous chat sessions or via email. The course meets the standard 45 contact hours.

Hybrid: A combination of traditional Brick and Mortar classroom instruction and online instruction using a course website (*E-college*). The class meets one day per week in a traditional classroom, like a traditional course. The second weekly class meeting is an online synchronous chat session using both text chat and VOIP with document sharing like an online course. All of the traditional tools and activities of an online course are incorporated and used in the hybrid course. The number of contact (live and chat) hours is approximately 75% that of the traditional course, the rest of the time is spent in other online activities like discussion boards. The university's movement toward hybrid classes parallels the instructional literature where hybrid classes are increasingly being promoted as a way of combining the advantages of both classroom and on-line environments (Arbaugh, 2005; Easton, 2003; Coppola et al, 2002; McCray, 2000). In addition, there is some recent empirical evidence that hybrid types of courses outperform other modes of delivery in more quantitative business-related courses like information systems, finance, and economics (Terry et al, 2003; Abdullat & Terry, 2005).

RESULTS

As descriptive statistics, the mean SLOAT score among all classes was 3.06 (n=2077), with a standard deviation of 0.772. SLOAT scores ranged from a low of 0 to a high of 4. We were also interested as to the correlation between the SLOAT scores and both student course grade and "student perception of learning." This is of particular interest since that vast majority of instructional delivery comparison studies, as well as the current stream of research examining the behavior characteristics of successful on-line instructors and students, the value of student communication modes, and performance differences in student group formation, tend to use student perception measures as indicators of student or class performance (see Arbaugh, 2005).

For reference, the mean course grade (combined graduate and undergraduate) was 3.04 (4 point scale, with +/-; n=2077), with a standard deviation of 0.826. For all graduate courses, the mean course grade was 3.23, while the undergraduate mean was 2.73. As previously mentioned, "Student perception of learning" data was incomplete, with a mean class section score of 4.05 for the pooled data (5 point scale). For graduate courses the mean class section score was 4.08, while for undergraduate courses the mean class section score was 4.02. Since "student perception of learning" responses are anonymous upon collection we could not measure individual responses; thus the mean value of 4.05 represents the average for 127 sections out of the 159 sections in the full sample. The bivariate correlation between student grades and the SLOAT scores

was 0.599, statistically significant at $p<0.01$ (n=2,077). For the graduate sub-sample the bivariate correlation was 0.625, while for undergraduate courses the bivariate correlation was 0.554, both statistically significant at $p<0.01$ (two-tailed).

On the other hand, the bivariate correlation between the "student perception of learning" and the SLOAT score was only 0.312 (n=127 sections), albeit still statistically significant at $p<0.01$ (two-tailed). For the graduate sub-sample the bivariate correlation was lower, at 0.251, significant at $p<0.05$, while for undergraduate courses the bivariate correlation was 0.365, significant at $p<0.01$. As mentioned above, these correlations are for class section means, not for individual students. While not the focus of this study, the rather low correlation between "student perceptions of learning" and a specific measure of course learning outcomes (SLOAT) draws into question previous empirical comparison studies of instructional delivery that use student perceptions to measure class achievement.

The primary focus of our analysis, however, is on examining whether there were significant differences in learning outcomes associated with the different instructional delivery methods. Table 1 shows a one-way analysis of variance (ANOVA) for the pooled data, including all sections of all courses, combined graduate and undergraduate. A total of 2,077 students were examined. The most common method of delivery was the traditional classroom method, with a total of 986 students. A total of 871 students took pure on-line courses. The newer delivery methods of Web-2 and hybrid courses had substantially fewer students in the sample, with 165 and 55 respectively. As observed in Table 1, the mean SLOAT score for the hybrid courses was 3.34, statistically higher than for the other three delivery methods. For the pooled data, no significant difference was noted between Web-2, traditional, and pure on-line courses.

Table 1 also shows the ANOVA results for the graduate course sub-sample, and the undergraduate course sub-sample. For the graduate sub-sample a total of 1,287 SLOAT scores were examined, with no statistically significant differences found between the four delivery methods. However, it should be noted that only one-hybrid section, with twelve students, was included in the graduate sub-sample. For the undergraduate sub-sample (n=790) a significant ANOVA ($p<0.01$) was found, with the hybrid course showing the highest results on the SLOAT score – however, this finding should be interpreted cautiously since only two hybrid sections were included in the undergraduate sub-sample. As with both the pooled sample and the graduate sub-sample, no significant differences were noted between the other three delivery methods of Web-2, traditional, and pure one-line courses.

The above analysis, however, does not control for course differences or instructor differences. Controlling for course differences may be important since the SLOAT questions, which reflect learning outcomes, naturally vary by course, but are the same for all sections of the same course. In addition, the grading of the SLOAT questions is constant between all sections of the same course. In

other words, all sections of the same course are graded by the same rubric (for essay questions) or scored the same (for multiple choice questions).

In all, a total of eleven business courses were found that had traditional classroom sections as well as pure internet sections of the same course during the 2006-2007 academic year. Undergraduate and graduate courses were included in this analysis. The undergraduate courses with multiple delivery methods were, a) Principles of Management and Organization, b) Principles of Macroeconomics, c) Marketing Fundamentals, d) Intermediate Accounting, and e) Managerial Accounting. The first three are required core courses for all business majors, while the two undergraduate accounting courses are senior level required courses for an accounting concentration. The graduate courses with multiple delivery methods were, a) Theories, Practices, and Ethics of Leadership, b) Ethical Concerns in Business and Management, c) Marketing Management, d) Statistics for Business, e) Theory and Practice of Organizational Development and, f) Seminar in Financial Management. All of these courses are core requirements for the School's graduate degrees.

Table 1: One-Way ANOVA by Delivery Method: SLOAT Scores

	Traditional	On-line	Web-2	Hybrid	*F-statistic*
Pooled (Grad and Undergrad)	3.01 (n=986)	3.09 (n=871)	3.07 (n=165)	3.34 (n=55)	4.158***
Graduate	3.09 (n=590)	3.16 (n=585)	3.20 (n=100)	2.99 (n=12)	1.256
Undergraduate	2.89 (n=396)	2.95 (n=286)	2.86 (n=65)	3.44 (n=43)	6.217***

***=prob.<0.01, **=prob.<0.05, *=prob.<0.10

Table 2 shows the results of a paired t-test (two-tailed) comparing the SLOAT scores of traditional classroom sections with the pure internet sections for each of these eleven business related courses.

Of the eleven courses examined, eight courses were found to have significant differences between delivery methods. Of these eight courses, five resulted in the statistically significantly higher outcome SLOAT scores for the traditional classroom delivery versus the pure internet delivery. The undergraduate course "Marketing Fundamentals" had a total of 103 students, with 29 students taking pure internet sections while 74 students took the traditional classroom delivery. Here the SLOAT scores for the traditional class room delivery were significantly higher than the pure-internet delivery (3.18 v. 2.47, $p<0.01$). The graduate course, "Seminar in Financial Management" had a total of 83 students, with the traditional class room scoring a mean of 3.23 (n=60)

while the pure-internet delivery scored 3.08 (n=23), a statistically significant difference ($p<0.10$). Similarly the graduate course, "Ethical Concerns in Business and Management" had significantly higher ($p<0.01$) classroom delivery SLOATS (3.84, n=12) compared to pure internet courses (3.21, n=26). With a total of 34 students, the graduate course "Theory and Practice of Organizational Development" also showed mean SLOAT scores significantly ($p<0.05$) higher for the traditional classroom section (3.26, n=14) versus the pure internet section (2.78, n=20). For the upper division undergraduate accounting course, "Intermediate Accounting" traditional classroom sections had significantly higher ($p<0.05$) outcome scores (2.91, n=41) versus the pure internet sections (2.52, n=50).

In contrast three courses showed significantly higher learning outcomes for the pure internet sections. The graduate "Marketing Management" course had a total of 201 students, with 52 students taking on-line sections and 149 students taking traditional classroom sections. The mean SLOAT score for the on-line sections was 3.52, while the mean score for the traditional classroom sections was 3.20 ($p<0.01$). Similarly the graduate "Statistics for Business" had a total of 208 students, with 99 taking internet sections and 109 in traditional classroom sections. The mean SLOAT score was 3.10 for on-line sections versus 2.61 for traditional classroom delivery ($p<0.01$). For the core undergraduate management course, "Principles of Management and Organization" the traditional classroom section had significantly ($p<0.01$) lower outcome scores (2.61, n=25) than the pure internet section (3.68, n=35).

No significant difference was noted for the classes, "Theories, Practices, and Ethics of Leadership," "Managerial Accounting," and "Principles of Macroeconomics."

It should be noted, however, that an important concern of the analysis presented above (and in Table 2) is that even though we have controlled for course differences, we still might be testing differences in instructors rather than differences in delivery methods. This is a major criticism of all of the previous studies that examine outcome differences between delivery methods, where the different delivery methods for the same course are taught by different instructors. In fact, of the eleven course sub-samples analyzed above, all of the courses involved sections taught by different instructors. Thus as both Bernard et al (2004) and Tallent-Runnels et al (2006) note, the high variability in the achievement results may be more a function of the confounding effects of instructor differences rather than delivery methods. Thus to control for both course differences and instructor differences our final analysis examined only those courses that were taught by the same instructor, with one or more sections using one delivery method and one or more sections using another delivery method. A review of all instructors in all sections revealed a total of ten courses where one instructor taught different sections of the same course with different delivery methods during the same academic year. The courses where this condition was satisfied included five of the courses discussed above (one course had two different instructors, each teaching different sections with different

delivery methods). In order to not identify individual instructors, the courses are referred to as SB1 to SB5

In addition, three other courses taught in the School of Business, but not in traditional business disciplines, were identified that also met the common course and instructor criteria. These were a public administration course, referred to as "PA1" and two criminal justice courses, "CJ1" and "CJ2".

Table 2: Mean SLOAT Scores by Course

	Instructional Method		
Course Title (Grad, Undergrad)	Traditional (mean)	On-line (mean)	*t-stat* (2-tailed)
Principles in Management and Organization (u) [m]	2.61 (n=25)	3.68 (n=35)	5.497***
Principles of Macro-economics (u) [m]	2.65 (n=103)	2.74 (n=74)	0.683
Marketing Fundamentals (u) [m]	3.18 (n=74)	2.47 (n=29)	3.852***
Intermediate Accounting (u) [m]	2.91 (n=41)	2.52 (n=50)	2.489**
Managerial Accounting (u) [m]	3.29 (n=20)	3.35 (n=27)	0.467
Theories, Practices, and Ethics of Leadership (g) [m]	3.20 (n=42)	3.26 (n=120)	0.551
Ethical Concerns in Business and Management (g) [m]	3.86 (n=12)	3.21 (n=26)	4.258***
Marketing Management (g) [m]	3.20 (n=149)	3.52 (n=52)	3.010***
Statistics for Business (g) [m]	2.61 (n=109)	3.10 (n=99)	4.146***
Theory and Practice of Organ-izational Development (g) [m]	3.26 (n=14)	2.78 (n=20)	2.603**
Seminar in Financial Management(g) [m]	3.23 (n=60)	3.08 (n=23)	1.903*

[m]=multiple instructors; ***=prob.<0.01, **=prob.<0.05, *=prob.<0.10

Table 3 shows the results of a paired t-test (two-tailed) comparing the mean SLOAT scores for each of these courses where the same instructor taught with different delivery methods in different sections. This analysis therefore controls for both course differences and instructor differences. Our analysis, however, revealed only one course where there was a significant difference between delivery methods – the graduate level criminal justice course, "CJ2"

Here the mean SLOAT score of 3.22 (n=25) for the pure-internet section was significantly higher (p<0.05) than the 2.64 mean SLOAT score for the traditional classroom. No significant differences were noted for the other nine cases where one instructor taught with different delivery methods in different sections of the same course.

Table 3: Mean SLOAT Scores for Single Instructor

	Instructional Method			
Course Title (Grad, Undergrad)	Traditional (mean)	On-Line (mean)	Web-2 (mean)	*t-stat* (2-tailed)
SB1 (u)	2.94 (n=31)	2.66 (n=29)		1.231
SB2 (u)	2.95 (n=16)	2.59 (n=12)		0.360
SB3 (u)		3.01 (n=22)	3.33 (n=22)	1.449
SB4 (g)	3.13 (n=27)	3.33 (n=120)		1.406
CJ1 (u)	2.60 (n=17)		2.80 (n=12)	0.996
SB5 (g) – Instructor 1	2.73 (n=56)	2.94 (n=12)		0.824
SB5 (g) – Instructor 2	2.42 (n=39)	2.47 (n=38)		0.254
CJ2 (g)	2.64 (n=10)	3.22 (n=25)		4.102***
PA1 (g) – Instructor 1	3.45 (n=20)	3.32 (n=16)		1.048
PA1 (g) – Instructor 2	3.42 (n=8)		3.40 (n=2)	0040
***=prob.<0.01, **=prob.<0.05, *=prob.<0.10				

DISCUSSION

In this study we examined the effectiveness of different instructional delivery methods for a large sample of undergraduate and graduate business, criminal justice, and public administration classes. In developing this analysis we attempted to address many of the criticisms leveled against prior empirical

research. First, we employed an explicit measure of the multiple learning outcomes developed for each course as part of the university's accreditation process rather than course grades, a single final exam, satisfaction surveys or student course evaluations. Second, we controlled for course differences by examining different instructional delivery methods for the same class. And third, we controlled for instructor differences by comparing courses where the same instructor taught the same course, but with different delivery methods in different sections. In addition, unlike most prior empirical studies, the researchers were not instructors of the courses within our data sample, thus we control for this common form of experimental bias.

Four different instructional delivery methods were examined in our sample of 2,077 students from 159 class sections. The two most common delivery methods were the traditional classroom and pure on-line internet delivery methods, with both being somewhat equal in numbers (986 and 871 respectively). A third, newer method, called Web-2 was next in offerings with 165 students. Another, more experimental approach was the "hybrid" method with only 55 students.

Our initial examination of the pooled data indicated little or no differences between instructional delivery methods with respect to learning outcome measures. A possible exception may be the "hybrid" delivery method; however, since only three "hybrid" sections were examined, it is too speculative to draw any definitive conclusions about this instructional method from our data. But it is intriguing to note that a few recent, albeit much smaller-scale studies of business-related courses have also found the hybrid form of delivery to be the most effective (e.g., Terry et al, 2003; Abdullat & Terry, 2005).

When controlling for course differences, eleven different business-related courses were examined that were offered as both pure on-line classes and traditional classroom experiences during the same academic year. The Web-2 method was not examined in this analysis since this method was used in only one of the eleven business-related courses. Controlling for course differences is important since outcome measures tend to be course-specific. In our study, for example, the SLOAT measures were developed specifically for each course, and were ultimately scored across all sections of the same course by the same grading process.

Examining these eleven courses revealed eight had significant differences between delivery methods with respect to the SLOAT outcome measures. However, the results were mixed with five courses indicating better outcome performance from the traditional classroom setting, while three courses appeared to favor the pure-internet delivery method. A review of the results distribution did not reveal any consistent reason (i.e., graduate versus undergraduate, core versus elective, qualitative versus quantitative) why one delivery method appeared to outperform the other delivery method. Therefore it is likely that the high variation in achievement results may be due to the confounding effects of instructor differences, a common problem now recognized in previous empirical research (Bernard et al, 2004; Tallent-Runnels et al, 2006).

Thus our final analysis controlled for both course and instructor differences. Controlling for just course differences is still open to some criticism since significant differences between instructional delivery modes may actually represent instructor-specific characteristics if different instructors teach the different delivery modes – particularly since the impact of instructor delivery style and technique on student satisfaction in both traditional and on-line courses has been well established in the literature (Arbaugh, 2005).

A total of ten different instructors were identified that offered different sections of the same course using different instructional delivery methods. Our analysis indicated that in nine out of ten cases there were no significant differences between instructional delivery methods. The one exception was an undergraduate criminal justice class, were the pure on-line delivery method significantly out-performed the traditional classroom instruction.

Overall, the results of our study suggest that traditional, on-line, and other delivery methods all have similar teaching and learning outcome capabilities. As discussed above, most prior studies indicate that student performance in on-line and traditional classroom versions of the same courses is similar. These findings appear consistent across many professional disciplines including business related courses (McCray, 2000), education courses (Knight, 2007), medical courses (Corral et al, 2006; Anderson & Mercer, 2004; Patterson & Harasym, 2001), and nutritional science courses (Buckley, 2003), with evidence pointing to perhaps a slightly better learning experience coming from the on-line courses (Hersh et al, 2001; Lipman et al, 2001; and Maki et al, 2000). But as previously mentioned, almost all of these prior studies that compare student performance in on-line and traditional classroom versions of the same courses suffer from important experimental design problems, including not controlling for different instructors, small sample size, or not using a specific learning outcome assessment scale. However, in spite of the limitations associated with prior empirical studies, our findings still remain consistent with the general conclusions of this prior research and thus add to the argument of "no significant difference" as suggested by Russell (2001, 2008) and others.

There are several other important implications in our study. First, there is a stream of literature that suggests that experienced on-line instructors, "may have an advantage in the on-line environment" (Arbaugh, 2005: 136, see also Easton, 2003; Coppola et al, 2002). While this is no doubt true, all of the on-line instructors in this sample were experienced, having taught multiple previous courses in the on-line environment. Yet, when controlling for course and instructors, it is apparent that there still remains little or no differences in achievement of learning outcomes between the traditional classroom delivery and the on-line environment, even when the course is being taught by an experienced on-line instructor. Thus, while an experienced on-line instructor will most probably outperform a novice (as is true in any instructional setting), on-line experience does not make the on-line course outperform its traditional counterpart even when taught by the same instructor. Second, there is increasing interest in hybrid courses that combine elements of the traditional classroom and

the on-line experience. While our sample did include some hybrid courses, with generally good outcome performance, there was not a sufficient sample size to draw a definitive conclusion.

There are clearly limitations to our study. First, there is naturally a self-selection process by the students in choosing which format they enroll in. There are some indications that more engaged and non-traditional students will tend to enroll in on-line courses (Skorga, 2002). However, most of the students in our sample, in all forms of delivery methods, were non-traditional students due to the university's education strategy. In addition, Corral et al (2006) found that students that enrolled in on-line courses initially had "more knowledge" than those that enrolled in the traditional classroom counterpart. This leads to the second possible limitation of our study. While unlike most previous research, our study does employ a specific scale designed to measure learning outcomes, we did not measure the student's knowledge level prior to taking the course. Thus, we can not determine "increase" in knowledge, but only the final level of attainment of the learning outcomes associated with the course.

While there is a plethora of comparison on-line and traditional classroom research regarding differences in perceptions and satisfaction, instructor and student characteristics, student engagement, communication processes and other behavioral themes, clearly much more research needs to focus on the basic question of learning outcomes, that is, "are there differences in the performance and learning associated with different delivery methods?" This question can only be answered by well-designed, large scale studies that utilize measures directly associated with the learning outcomes associated with the respective courses. It is our hope that the present study has provided useful results in this direction.

REFERENCES

Abdullat, A., & Terry, N. (2005). Assessing the effectiveness of virtual learning in a graduate course in computer information systems. *Information Systems Education Journal*, 3(34), 1-8.

Abrami, P., d'Apollonia, S., & Cohen, P. (1990). Validity of student ratings of instruction: What we know and what we do not know. *Journal of Educational Psychology*, 82, 285-296.

Ally, M. (2004). Foundations of educational theory for online learning. In T. Anderson and F. Elloumi (Ed.), *Theory and practice of online learning*. Athabasca, AB, Canada: Athabasca University, 3-25.

Anderson, T. & Elloumi, F. (2004). *Theory and practice of online learning*. Athabasca, AB, Canada: Athabasca Theory and Practice of Online Learning.

Anderson, E., & Mercer, Z. (2004). Impact of community health content on nurse practitioner practice: A comparison of classroom and web-based teaching. *Nursing Education Perspectives*, 25(4), 171-175.

Anstine, J., & Skidmore, M. (2005). A small sample study of traditional and online course with sample selection adjustment. *Journal of Economic Education*, 36(2), 107.

Arbaugh, J., (2005). Is there an optimal design for on-line MBA courses? *Academy of Management Learning & Education*, 4(2), 135-149.

Bernard, R., Abrami, P., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., Wallet, P., Fiset., M., & Huang, B. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. *Review of Education Research,* 74(3), 379-439.

Bonk, C., & Reynolds, T. (1997). Learner-centered Web instruction for higher-order thinking, teamwork, and apprenticeship. In B. H. Khan (Ed.), *Web-based Instruction*. Englewood Cliffs, NJ: Educational Technology Publications, 167-178.

Brown, B., & Liedholm, C. (2002). Can web courses replace the classroom in principles of microeconomics? *The American Economic Review*, 92(2), 444-448.

Brown, B., & Liedholm, C. (2004). Student preferences in using online learning resources. *Social Science Computer Review*, 22(4), 479-492.

Buckley, K. (2003). Evaluation of classroom-based, Web-enhanced, and Web-based distance learning nutrition courses for undergraduate nursing. *Journal of Nursing Education*, 42, 367-370.

Carswell, L., Thomas, P., Petre, M., Price, B., & Richards, M. (2000). Distance education via the Internet: The student experience. *British Journal of Educational Technology,* 31(1), 29-46.

Clark, R. (1983). Reconsidering research on learning from media. *Review of Educational Research*, 53(4), 445-459.

Clark, R. (2001). A summary of disagreements with the "mere vehicles" argument. In R. Clark (Ed.), *Learning from media: Arguments, analysis and evidence*. Greenwich, CT: Information Age Publishing Inc, 125-136.

Coates, D., Humphreys, B., Kane, J., & Vachris, M. (2004). "No Significant Distance" Between face-to-face and online instruction: evidence from principles of economics. *Economics of Education Review*, 23, 533-46.

Cole, R. (2000). *Issues in Web-based pedagogy: A critical primer*. Westport, CT: Greenwood Press.

Cooper, P. (1993). Paradigm shifts in designing instruction: From behaviorism to cognitivism to constructivism. Educational Technology, 33(5), 12-19.

Corral, M., Guevara, J., Luquin, P., Pena, H., & Otero, J. (2006). Usefullness of an Internet-based thematic learning network: Comparison of effectiveness with traditional teaching. *Medical Informatics and the Internet in Medicine*, 31(1), 59-66.

Craik, F., & Lockhart, R. (1972). *Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior,* 11, 671-684.

Darby, J. (2007). Evaluating course evaluations: The need to establish what is being measured. *Assessment and Evaluation in Higher Education*, 27, 441-455.

Eastin, M., & LaRose, R. (2000). Internet self-efficacy and the psychology of the digital divide. *Journal of Computer Mediated Communications*, 6(1). Retrieved March 5, 2008, from http://jcmc.indiana.edu/vol6/issue1/eastin.html.

Ertmer, P., & Newby, T. (1993). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly,* 6(4), 50-70.

Frederickson, N., Reed. P., & Clifford, V. (2005). Evaluating web-supported learning versus lecture-based teaching: Quantitative and qualitative perspectives. *Higher Education*, 50, 645-664.

Frydenberg, J. (2007). Persistence in university continuing education online classes. *The International Review of Research in Open and Distance Learning* 8(3), 1-15 Retrieved March 5, 2008, from http://www.irrodl.org/index.php/irrodl/article/view/375/957

Harasim, L., Hiltz, S., Teles, L., & Turoff, M. (1995). *Learning networks: A field guide to teaching and learning online*. London: MIT Press.

Harmon, O., & Lambrinos, J. (2006). Online format vs. live mode of instruction: Do human capital differences or differences in returns to human capital explain the differences in outcomes? Working Paper 2006-07. Department of Economics, University of Connecticut.

Harris, D., & Parrish, D. (2006). The art of on-line teaching: Online instruction versus in-class instruction. *Journal of Technology in Human Services*, 24, 105-117.

Hersh, W., Junium, K., Mailhot, M., & Tidmarsh, P. (2001). Implementation and evaluation of a medical informatics distance education program. *Journal of the American Medical Informatics Association*, 8, 570-584.

Hine, C. (2000). *Virtual ethnography*. London: Sage.

Janicki, T. & Liegle, J. (2001). Development and evaluation of a framework for creating Web-based learning modules: A pedagogical and systems approach. *Journal of Asynchronous Learning Networks*, 5(1), 58-84. Retrieved March 5, 2008, from http://www.sloan-c.org/publications/jaln/v5n1/pdf/v5n1_janicki.pdf

Kalat, J. (2002). *Introduction to psychology*. Pacific Grove, CA: Wadsworth-Thompson Learning.

Knight, P. (2007). Promoting retention and successful completion on Masters courses in education: A study comparing e-tuition using asynchronous conference software with face-to-face tuition. *Open Learning*, 22, 87-96.

Kozma, R. (2001). Counterpoint theory of "learning with media." In R. Clark (Ed.), *Learning from media: Arguments, analysis, and evidence*. Greenwich, CT: Information Age Publishing Inc, 137-178.

Lipman, A., Sade. R., Glotzbach, A., Lancaster, C. & Marshall, M. (2001). The incremental value of Internet-based instruction as an adjunct to classroom instruction: a prospective randomized study. *Academy of Medicine*, 76, 1060-1064.

Lipman, M. (1991). *Thinking in education*. Cambridge: Cambridge University Press.

Maki, R., Maki, W., Patterson., M., & Whittaker, P. (2000). Evaluation of a Web-based introductory psychology course: Learning and satisfaction in one-line versus lecture courses. *Behavior Research Methods, Instruments, and Computers*, 32, 230-239.

Marsh, H., (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11, 253-388.

Mason, J. & Hart, G., (1997). *Effective use of asynchronous virtual learning communities*. Retrieved March 5, 2008 from: http://www.arch.usyd.edu.au/kcdc/conferences/VC97/papers/mason.html

McCray, G. (2000). The hybrid course: Merging on-line instruction and the traditional classroom. *Information Technology and Management*, 1, 307-327.

Navarro, P. (2000). Economics in the cyberclassroom. The *Journal of Economic Perspectives*, 14(2), 119-132.

Navarro, P. & Shoemaker, J. (1999). The power of cyberlearning: An empirical t est. *Journal of Computing in Higher Education*, 1(1), 29-54.

Patterson, R., & Harasym, P. (2001). Educational instruction on a hospital information system for medical students during their surgical rotations. *Journal of the American Medical Informatics Association*, 8, 111-116.

Russell, T. (2001). The No Significant Difference Phenomenon. IDECC, fifth edition.

Russell, (2008) http://nosignificantdifference.wcet.info/

Schramm, W. (1997). *Big media, little media*. Beverly Hills, CA: Sage.

Shin, N., & Chan, J. (2004). Direct and indirect effects of online learning on distance education. *British Journal of Educational Technology*, 35, 275-288.

Skorga, P. (2002). Interdisciplinary and distance education in the Delta: The Delta Health Care Partnership. *Journal of Interprofessional Care*, 16, 149-157.

Tallent-Runnels, M., Thomas, J., William, L., Cooper, S., Ahem, T., Shaw, S., & Xiaoming, L. (2006). Teaching courses on-line: A review of the research. *Review of Educational Research,* 76, 93-135.

Terry, N., Lewer, J. & Macy J. (2003). The efficacy of alternative instruction modes in economics. *Journal of Economics and Economic Education Review*, 4(1), 23-34.

Thatcher, A. (2007). Using the World Wide Web to support classroom lectures in a psychology course. *South African Journal of Psychology*, 37, 348-353.

Walther, J. (1996). Computer mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, 23 (1), 3-43.

Wenger, E. (2001). *Supporting communities of practice: A survey of community-oriented technologies* (1.3 ed.) (Shareware). Retrieved March 5, 2008, from: http://www.ewenger.com/tech

Wilson, B. (1997). Refelctions on constructivism and instructional design. In C. Dills and A. Romiszowski (Eds.), *Instructional development paradigms*. Englewood Cliffs, NJ: Educational Technology Publications, 63-80

Wilson, B. (2001). Sense of community as a valued outcome for electronic courses, cohorts, and programs. Retrieved March 5, 2008, from: http://carbon.cudenver.edu/~bwilson/SenseOfCommunity.html

Worthington, A. (2002). The impact of student perceptions and characteristics on teaching evaluations: A case study in finance education. *Assessment and Evaluation in Higher Education*, 27, 49-64.