

Urtel, M. G. (2008). Assessing academic performance between traditional and distance education course formats. *Educational Technology & Society*, 11 (1), 322-330.

Assessing academic performance between traditional and distance education course formats

Mark G. Urtel

Indiana University Purdue University Indianapolis, USA // Tel: +1 317 278-2015 // Fax: +1 317 278-2041 // murtel1@iupui.edu

ABSTRACT

The goal of this study was to explore whether differences in student academic indicators exist between taking a course face-to-face (F2F) and taking a course via distance education (DE). Three hundred and eighty five students were enrolled in a course offered, both, as F2F ($n = 116$) and as DE ($n = 269$). Course content, instructor, textbook adopted, and assessment methods were consistent between the two course delivery formats. Final grades, *DFW* rates, and end of term course and instructor evaluations were used as the outcome indicators. In addition, student demographic information was factored into data analyses. Results indicated that there was a statistically significant difference in final grade, *DFW* rates, and end of term course evaluation response rates between the course offerings. Further analysis suggested that freshman grade performance was significantly different between course offerings. Implications and policy suggestions regarding distance education will be discussed.

Keywords

Distance education, Freshmen performance, *DFW* rates, Academic indicators

Introduction

It is difficult for the *scholarship* of distance education to keep pace with the *delivery* of distance education. This is especially true given the vast increase and technological developments in (a) the definition of distance education (b) the formats of instruction and interaction (c) the entities who deliver it and (d) the demographics of who receive it. In fact, Phipps and Merisotis (1999) cautioned that the overall quality of the distance education research was suspect and the ability to predict success or explain interactions (basically generalize the findings) would be difficult, if not impossible. The focus of their argument was that most of the articles published dealt with opinion and the *how-to* of delivering distance education, absent subjects, controls, or an experimental design. As a result, it would be hard to duplicate much of the research in hopes of validating these early findings. Since Phipps and Merisotis first published those words in 1999 the amount of research on distance education has increased and improved; this is evidenced by the newly formed and peer reviewed journals of distance education which foster a decidedly international conversation regarding this topic. However, distance education practice is still outpacing the accompanying research and there still exists 'quality' of research concerns in distance education today (Moore, 2004).

Moore (2004) shares his concern using recently submitted distance education manuscripts as evidence. Specifically, Moore highlights the trend of case study research in distance education. Case study research in distance education is popular (Kozma, 2003; Poole, 2000; Zhang, 1998) but, Moore reminds the scholars of distance education that case study research too often cannot be generalized to a wider audience and, at times, the studies depend on the use of anecdotal 'evidence'. Both of which, greatly compress the ability to widely apply research findings to the policies and practices of delivering distance education in higher education.

The chasm between generalizable research and best practice in distance education can be seen explicitly when focusing on student outcomes. It was assumed, as the early research indicated, that distance education had a neutral or favorable impact on student outcomes as compared to traditional or face to face course work (Martin and Rainey, 1993; Souder, 1993; Verduin and Clark, 1991). Yet upon closer review, the research, according to Phipps and Merisotis (1999), lacked in a few areas. First, many of the studies published did not provide the proper controls needed in the experimental design. In addition, the sample size or number of subjects studied was too small to draw definitive conclusions (Ritchie and Newby, 1989). Moreover, some of the studies grouped the students who took courses as distance education as one type; without taking into account race, gender or age and their possible interactions on distance education success. Without factoring in these demographic variables when investigating course delivery format, coupled with the fact most of the previous research did not control for content, instructional format, or even instructor (Phipps and Merisotis, 1999), the ability to generalize the findings of one study to multiple applications condenses considerably.

This project sought to address some of the control issues cited as lacking in earlier research. In addition, the sample size of this project is larger than previously used which also includes multiple student demographic variables in the data collection and analysis. So, the comments of Phipps and Merisotis (1999) served as the genesis of the current project and provided solid support to the research question, “Is there a difference in student academic indicators between taking a course face to face or taking a course as distance education?”. More specifically, the indicators used were (a) course grades (b) D, F, W rates and (c) end of semester course evaluations.

Review of Literature

Definition and perspective of distance education (DE)

Geber (2000) provides a contemporary definition of distance education as any formal approach to learning in which the majority of the instruction occurs while educator and learner are at a distance from each other. This version represents earlier definitions well (Moore, 1989) and keeps distance education relative to the technology available and utilized at any one point in history.

According to Waits and Lewis (2003), 90% of two-year and four-year public institutions of higher education offer distance education. This percentage is even more striking given that just five years earlier Lucas (1998) reported that about one-half of all higher education institutes offered distance education. Both of these studies support the contention that distance education within higher education is not a passing trend; rather it is an increasingly popular delivery format that warrants as much scholarly attention as it can get.

Overall academic performance

As stated previously, distance education has been purported to have at the least a neutral effect on academic performance and at best have a positive effect on academic performance (Arbaugh, 2000; Clark 1999; Diaz, 2000; Dorbin 1999; Dutton, Dutton, and Perry, 1999; Martin and Rainey, 1993; Navarro and Shoemaker 1999; Neuhauser, 2002; Souder, 1993; Verduin and Clark, 1991; and Trinkly, 1999). When factoring in student age it has been suggested that older students tend to perform better academically in a DE format (Dille and Mezack, 1991; Souder, 1994); yet, some of the research is on graduate, not undergraduate course work. Beyond those two variables, the research on student demographics such as gender, race, and class standing with relation to academic performance within a DE format as opposed to a face-to-face (F2F) format is not apparent.

DFW rates

In review, for colleges and universities in the United States, grades earned as *D*, *F*, or *W*, have either a negative effect or no effect toward credit hours earned and subsequent semester grade point averages. Specifically, a grade of *D* equals a one out of a four point scale and grades of *F* or *W* equal a zero out of a four point scale. Regardless, these grades directly and unhelpfully impact progression toward completion of a degree.

There is general consensus that distance education (DE) course work has relatively high attrition rates as compared to face-to-face (F2F) course work (DeTure, 2004; Diaz, 2002; Morris, Wu, and Finnegan, 2005; Neuhauser, 2002). However, distance education research regarding *DFW* rates has received scant attention. Moreover, research that takes into account student demographic information (such as: age, class standing, gender, or race) in predicting who is more apt to earn a *DFW* in DE is even more scarce.

Course Evaluations

Norris and Conn (2005) report on the wide range of response rates of term-end course and instructor evaluations for the DE delivery; these return rates range from zero to 95% of student participation. In addition, the base of research on end-of-term course and instructor evaluations for DE course work is quite narrow. The evident scholarship parallels much of the academic indicator research in that it tends to be *action-research* with the intent to improve response rates in one particular setting or application (Harrington and Reasons, 2005; Johnson, 2003;).

Nonetheless, course evaluations are important tools of assessment that provide the instructor with feedback directly from the student not otherwise obtainable. In addition, information gleaned from these evaluations can help guide the planning of subsequent semesters and can inform the instructor of the most and least effective pedagogies used during any one semester. Basically, if a student feels engaged in a course, valued in their opinion, and that *someone* is listening they may be more apt to complete the end of term evaluation. Thus, response rates may signify how engaged a student was throughout the semester.

Methods

Setting

The course used in this project was historically offered as a traditional or face to face course. However, with the emergence of distance education it underwent a transformation a few years ago to be delivered as distance education and face to face concomitantly. The course to be assessed is offered at the 300 level and is required for majors in the department in which it is offered. Additionally, the course can be taken by any academic major of the university with no pre-requisite for enrollment. Course content, course instructor, textbook used, course reading schedule, lecture progression, and overall performance assessments were held constant between the two course delivery formats.

The F2F format was structured so that the instructor and students met twice weekly for 75 minutes per session for 15 weeks. The DE format delivered the content using the exact reading schedule as the F2F section and the exact lecture progression however, the lecture material was obtained using one of two mediums; (a) a weblink to the lectures (videotaped) or (b) watching the lectures on local cable TV. Again, the schedule of the lecture content and related reading was identical for both the F2F and the DE course formats.

Student interaction and discussion was emphasized in both formats. For the F2F course time was dedicated in-class for the students to accomplish the learning objectives for the reading and lecture content of each session. Similarly, for the DE student discussion forums and/or chat rooms were created and equally prompted for student interaction and discussion so that, conceptually, learning objectives from the reading and lecture content could also be discussed.

Subjects

The data set, with prior approval from the university Institutional Review Board, was generated from undergraduate students at a local mid-western urban university. More specifically, the data set used in this study was comprised of 269 students enrolled in the distance education section and 116 enrolled in the face to face section for an $N = 385$. Table one illustrates the data set more fully.

Table 1. Demographics of data set

	DE	F2F
Gender		
Females	193	77
Males	76	39
Ethnicity		
White	217	97
Black	45	13
Hispanic	7	6
Class Level		
Freshmen	54	24
Sophomore	58	45
Junior	47	30
Senior	100	27
Age (M)	27	24

Data Collection

The basic assessment method involved the compilation of aggregate data found on final grade reports, electronic or hard copy grade books, and the absolute reporting rates of semester-end student course evaluations. In addition, individual student records were retrieved to collect the pertinent demographic data of the students enrolled in the courses (class standing at time of enrollment, gender, ethnicity, age at time of enrollment, major or non-major, etc). Collection methods were identical for both the face to face course and the distance education sections. Grade point averages for each group were computed by converting the letter grade earned by each student to the equivalent numeric quality point.

Statistical Analyses

Difference between groups (face to face and distance education) and with in groups (using student demographics, such as: gender, age, ethnicity, etc) was determined using factorial ANOVA and t-test analyses, respectively. Pearson-Product Moment Correlations were calculated for age and academic performance. Finally, Crosstabs was used to generate appropriate descriptive statistics. Statistical analyses were performed using SPSS (Statistical Package for the Social Sciences) with an alpha level of $p = .05$ for all tests.

Results and Conclusions

Overall academic performance

There was a statistically significant difference between the distance education and face-to-face groups regarding overall academic performance as measured by grade earned ($p = .010$). The students enrolled in the face to face section earned, overall, a 3.16/4.00 and the students who took the distance education section earned a 2.28/4.00 overall. Figure 1 depicts this relationship as well as indicating academic performance between DE and F2F formats when sorting the data by class standing.

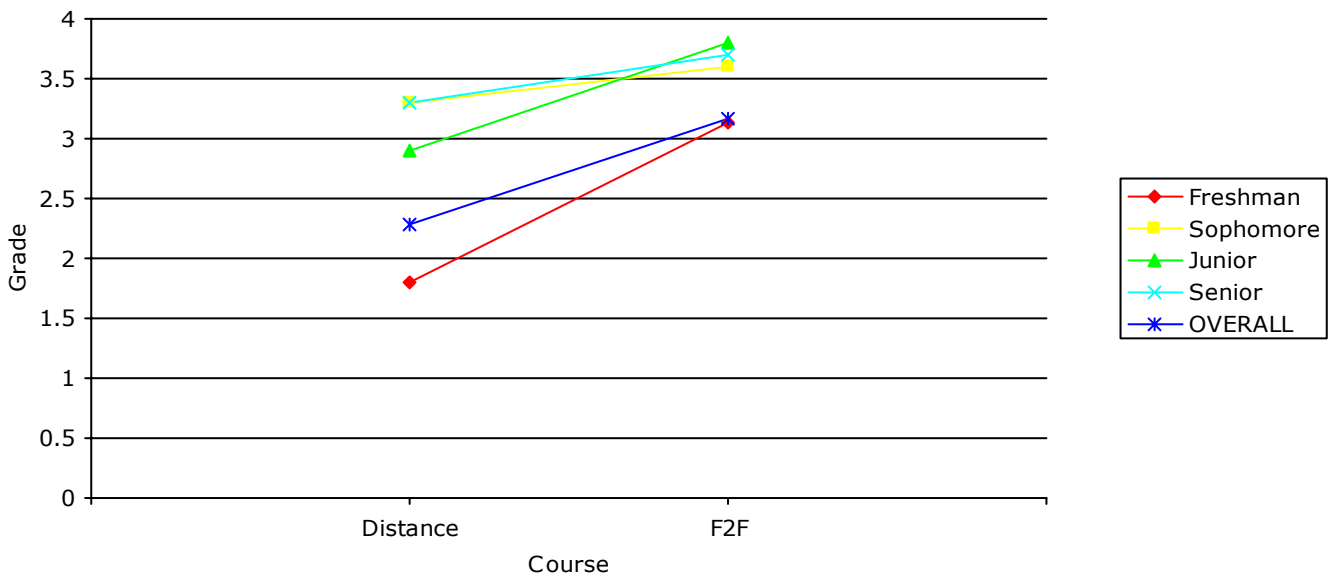


Figure 1. Final grades between DE and F2F; then disaggregated by class standing

There was a marked difference in *DFW* rates between the distance education and the face-to-face formats (see Fig. 2). This discrepancy helps to bring this grade point average down as a grade of *D* earns very little quality points (1.0) that can contribute to grade point average. More importantly, grades earned as *F* earn zero quality points toward

grade point average, thus contributing to a partial explanation in the statistically significant difference between group means of taking the class face to face or taking the class as distance education. Finally, grades earned as a *W* were not factored into the computation and had no impact on earned grade point average. As a result, the data indicate that overall student academic performance declines when taking a course as distance education versus face to face. However, this data is simply reported in aggregate; with no demographic information factored in to help explain or describe this academic decrease. A more thorough analysis that includes various student demographics follows.

DFW rates

Figure 2 illustrates the finding that students enrolled in the distance education section had statistically significant higher *DFW* rates than their peers who took the face to face section (40% vs. 21%). In fact, withdrawal rates in particular were cited in Phipps and Merisotis (1999) as being knowingly higher in distance education courses and our results support this early observation. However, Diaz (2000) articulated that withdrawal rates are higher in DE courses due in part that DE enrolled students are older and tend to make withdrawal decisions based on *maturity*. While the findings from this study cannot speak to the *why* or the *maturity* issue of withdrawal rates as reasoned in Diaz, it does however cast doubt to the uniqueness of the DE student.

More specifically, it was found in this study that there was no statistically significant difference between the age of students who withdrew between the DE and the F2F course format ($p = .117$). This suggests the *maturity* of the withdrawal decision occurs whether or not one is enrolled in a DE or a F2F course. Again, given the discrepancy in withdrawal rates between the DE and F2F course formats and that the age of students who withdraw are similar, that maybe DE course delivery contributes to the withdrawal decision and is not as innocuous as originally thought.

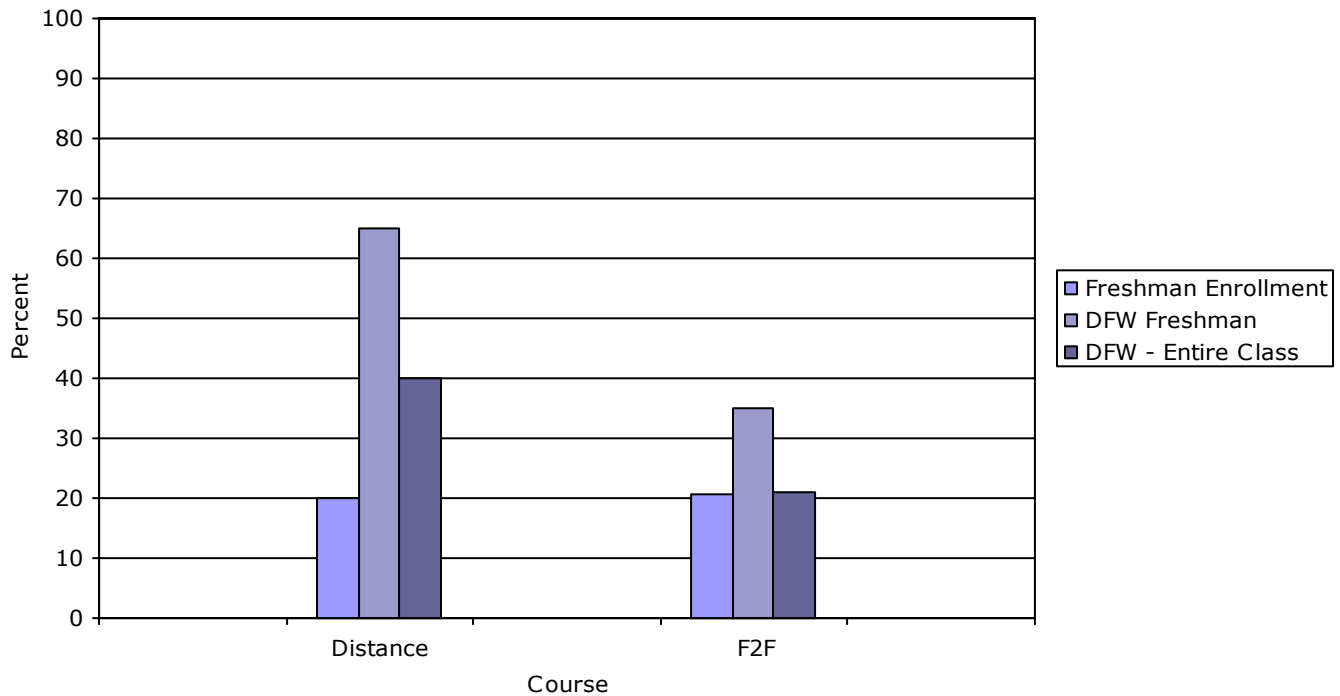


Figure 2. *DFW* rates and freshmen enrollment

Freshmen academic performance

Unsurprisingly, freshmen students underperformed compared to every other class level in grouped means for grades earned in both the face to face section and the distance education section (see Fig. 1). Moreover, the difference was

statistically significant in opposition to both sophomores ($p = .040$) and seniors ($p = .034$) and was nearing significance in opposition to juniors ($p = .060$). However, freshmen were not under- or overrepresented in either delivery format. In fact, the class composition of freshman was essentially the same with 20.7% and 20.1% being freshman in the face to face and distance education sections, respectively. Somewhat surprising, however, was the *DFW* rates between the face to face and the distance education sections for freshmen in particular (see Fig. 2). In particular, the F2F *DFW* rate was 35% and the DE *DFW* rate 65%. The significance of this finding warrants more discussion and further investigation.

Effect of ethnicity/gender/ age

There was only one between group ethnic interaction of statistical significance and that was that Black students underperformed their White counterparts (see Fig. 3) in both the face to face section and the distance education section ($p = .012$). While there was an absolute difference between the performance of Black and Hispanic students, overall, it was not statistically significant ($p = .205$).

When looking at with in group differences between DE and F2F, it was found that there was a statistically significant difference ($p = .0132$) between White students who took the course DE and F2F; with the better performance coming in the F2F format (see Fig. 3). There were no statistically significant differences between delivery format academic performance for Black ($p = .378$) or Hispanic ($p = .985$) students.

When analyzing gender interactions of the face to face and distance education sections overall, there was no statistically significant difference in the performance of females versus males ($p = .2214$). Nor were there any significant differences between the genders and delivery format; with male versus female in DE ($p = .293$) and male versus female in F2F ($p = .344$). However, there was a statistically significant finding when analyzing with in gender interactions and delivery format (see Fig. 3). It was found that females in F2F outperformed their female counterparts in DE ($p = .014$). There existed no such difference regarding males. Males in both the F2F and DE formats performed without a statistically significant difference ($p = .106$).

Finally, there was a statistically significant difference between the average ages of each delivery format ($p = .000$). The mean age of the F2F student was 24.03 years old and the mean age of the distance education student was 27.09 years old. However, there was no correlation between age and grade earned ($r = -.035$).

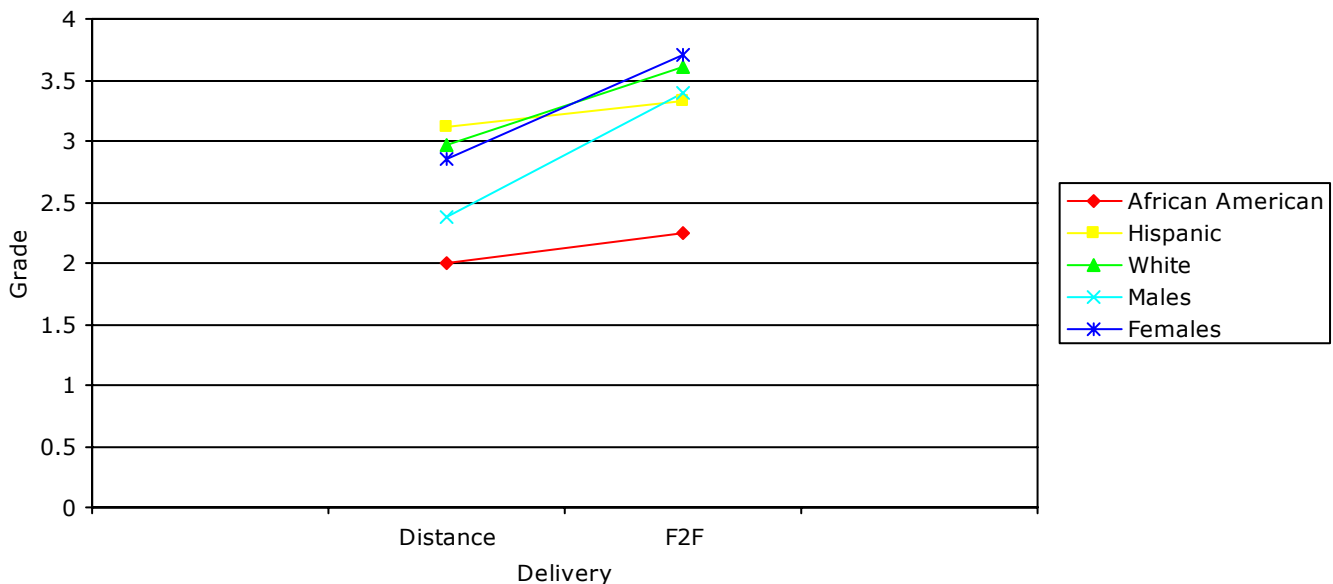


Figure 3. Race and Gender on grade earned per format

Course evaluations

The return rate for semester ending course evaluations for the face to face section was 95% and the return rate for the distance education section was less than 5% (see Fig. 4). This is in line with the findings from Norris and Conn (2005) who reported that over 200 distance education courses (of about 1000 analyzed) recently reported a zero return rate for semester ending course evaluations. Reasons for such a low DE response rate were not obtained. However, requests for responding to the end of term DE assessment were not isolative or limited.

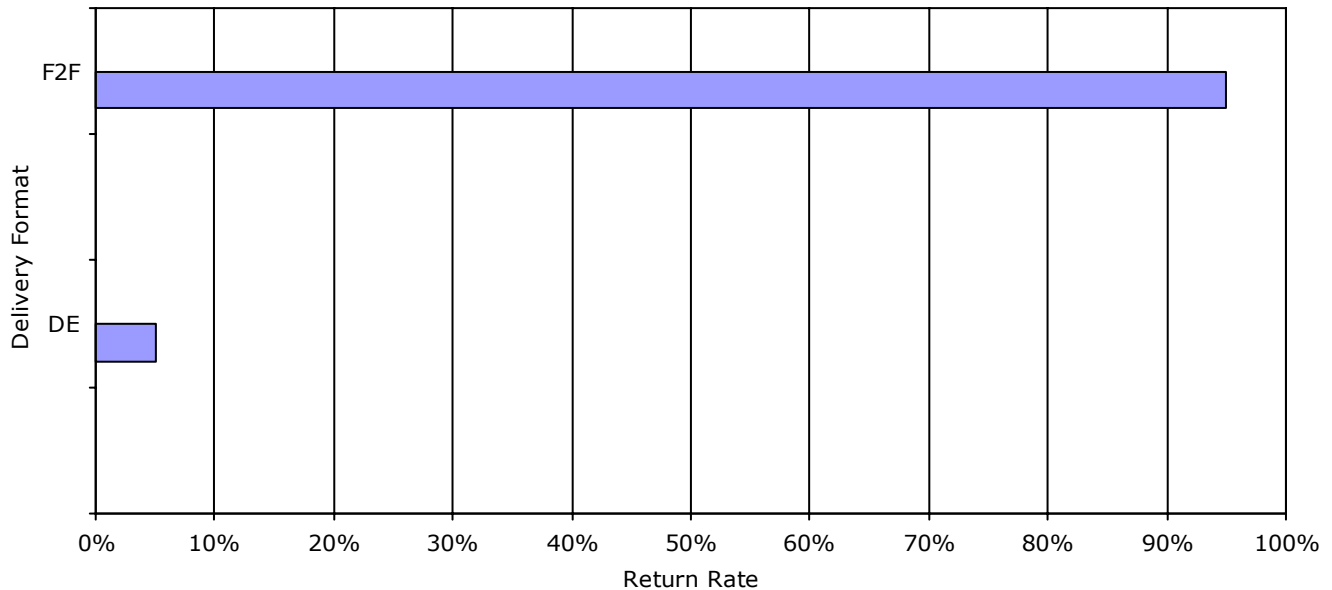


Figure 4. End-of-term course and instructor evaluation return rates

Conclusions

The results of this study are far from comprehensive or exhaustive. However, there are a few noteworthy findings that may serve to inform future research and address previously overlooked implications of converting classes to a distance education format. This is especially true regarding the knowledge base of distance education and how it impacts freshmen academic performance.

Additionally, these results provide an alternate and antithetical view to the unequivocal findings reported previously regarding the positive impact of distance education on student outcomes. This study suggests that converting a class to distance education has statistically significant impact on academic indicators, such as final grade, for students.

More specifically, students in a distance education course do not automatically perform equally as well, or even better, than as if they were in a face to face course. This is contrary to earlier studies. In addition, this study suggests that while age is a predictor in enrolling in a DE versus a F2F course, age is not a predictor in the subsequent academic success relative to format taken. Specifically, contrary to earlier reports older students do not automatically outperform their younger counterparts in DE. And in fact, the students who withdrew from a DE class are about the same age as students who withdrew from a F2F class; this is contrary to previously published assertions.

Next, it can be argued that the results of this study, while admittedly preliminary, can certainly be included in future discussions regarding what the risks are for freshmen who take a course as distance education. Notably, the findings of this study indicate that students classified as a freshmen earn a disproportionate rate of *DFW*'s when taking the class as distance education than their freshmen peers who take the class as a face to face course (i.e. 65% vs. 35%). The apparent risk of earning a *DFW* in DE, as a freshman, is not evident in the current literature. It may not be safe

to assume that the most *tech-savvy* student (freshmen) is comfortable in a course that depends on the technology to complete it.

While Black or Hispanic students did not display significant academic performance differences between delivery formats, White students did. The reasons for the performance difference, while worthy, were not pursued in this project. Similarly, while it is interesting to note the significant gender findings in this study, the reasons why females in F2F outperformed their female DE counterparts were not investigated. Course delivery format and the subsequent interactions on academic performance with regard to student ethnicity and gender do warrant further investigation.

This study supports earlier findings that DE *end of term* course and instructor evaluations, by default, have lower return rates. This is easily explained as the F2F course students are a ‘captive’ audience when these evaluations are administered where as DE students need to independently complete then submit similar instruments. In addition, F2F students do not need to be initially informed or reminded that the assessment will be given. Nor do they need to be independently driven to complete the assessments; all they need to do is be present the day they are administered. In contrast, DE students need prior notice that the instrument will be made available to them. They may even need a reminder as the semester progresses and even then, it may be forgotten or not completed for a variety and often unknown number of reasons. Regardless, the low response rates may indicate a lack of engagement by the student and, indirectly, negatively impacts the instructor as any suggestions, confirmations, or general comments are not available and hence cannot be factored into future planning of the course.

Finally, it is hoped that by addressing some concerns of the earlier published research on distance education that the findings from this current study will help to close the gap between distance education practices and research. This is acutely important for freshmen students. By establishing controls between each delivery format (via instructor, textbook, course content, reading schedule, lecture progression, and assessment techniques) combined with using a fairly robust *N*, and subsequently, factoring the demographic information into the data analysis, the findings of this study will help update and shape future research on distance education.

References

- Arbaugh, J.B. (2002). Virtual classroom versus physical classroom: An exploratory comparison of class discussion patterns and student learning in an asynchronous internet-based MBA course. *Journal of Management Education*, 24(2), 207-227.
- Clark, D. (1999). Getting results with distance education. *The American Journal of Distance Education*, 12(1), 38-51.
- DeTure, M (2004). Cognitive style and self-efficacy: Predicting student success in online distance education. *The American Journal of Distance Education*, 18(1), 21-38.
- Dille, B., & Mezack, M. (1991). Identifying predictors of high risk among community college telecourse students. *The American Journal of Distance Education*, 5(1), 24-35.
- Diaz, D. P. (2002). Online drop rates revisited. *The Technology Source Archives*, retrieved December 10, 2007 from http://technologysource.org/article/online_drop_rates_revisited/.
- Dorbin, J. (1999). Who’s teaching online? *ITPE News*, 2(12), 6-7.
- Dutton, J., Dutton, J., & Perry, J. (1999). Do online students perform as well as traditional students? In T. L: Russell (Ed.), *The No Significant Difference Phenomenon*, IDECC, Montgomery.
- Geber, B., (2000). Distance Education. *Technology and Learning*, 20(6), 20.
- Harrington, C.F., & Reasons, S.G. (2005). Online student evaluation of teaching for distance education: A perfect match? *The Journal of Educators Online*, 2(1), 1-12.
- Johnson, T.D. (2003). Online student ratings: Will students respond? *New Directions for Teaching and Learning*, 96, 49-59.

- Kozma, R.B. (2003). Technology and classroom practices: An international study. *Journal of Research on Technology in Education*, 36(1), 1-14.
- Lucas, R. (1998). An ecology of distance learning. *Syllabus*, 11(10), 14-16, 22.
- Martin, E.E., & Rainey, L. (1993). Student achievement and attitude in a satellite delivered high school science course. *The American Journal of Distance Education*, 7(1), 54-61.
- Moore, M.G. (2004). Editorial: Research worth publishing. *The American Journal of Distance Education*, 18(3), 127-130.
- Moore, M.G. (1989). Three types of interaction. *The American Journal of Distance Education*, 3(2), 1-6.
- Morris, L.V., Wu, S.S., & Finnegan, C.L. (2005). Predicting retention in online general education courses. *The American Journal of Distance Education*, 19(1), 23-26.
- Navarro, P., & Shoemaker, J. (1999). *Economics in cyberspace: A comparison study*, Discussion paper, University of California: Irvine Graduate School of Management, 33.
- Neuhauser, C. (2002). Learning style and effectiveness of online and face-to-face instruction. *The American Journal of Distance Education*, 16(2), 99-113.
- Norris, J., & Conn, C. (2005). Investigating strategies for increasing student response rates to online-delivered course evaluations. *The Quarterly Review of Distance Education*, 61(1), 13-29.
- Pascarella, E.T., & Terenzini, P.T. (1980). Predicting Freshman persistence and voluntary dropout decisions from a theoretical model. *Journal of Higher Education*, 51, 60-75.
- Pascarella, E.T., & Terenzini, P.T. (1991). *How college affects students: Findings and insights from twenty years of research*, San Francisco: Jossey-Bass.
- Phipps, R.A., & Phipps, J.P. (1999). What's the difference? Outcomes of distance education versus traditional classroom based learning. *Change*, 31(3), 12-17.
- Poole, D.M. (2000). Student participation in a discussion oriented online course: A case study. *The Journal of Research on Computing in Education*, 33(2), 162-177.
- Ritchie, H., & Newby, T.J. (1989). Classroom lecture/discussion vs. live televised instruction: A comparison of effects on student performance, attitude, and interaction. *The American Journal of Distance Education*, 3(3), 36-41.
- Souder, W.E. (1994). The effectiveness of traditional versus satellite delivery in three management of technology master's degree programs. *The American Journal of Distance Education*, 7(1), 37-53.
- Tinto, V. (1999). Taking retention seriously: Rethinking the first year of college. *NACADA Journal*, 19(2), 5-9.
- Trinkle, D.A. (1999). Distance Education: A means to an end, no more, no less. *The Chronicle of Higher Education*, A60, 6 August.
- Verduin, J. R., & Clark, T. (1991). *The foundations of effective practice*, San Francisco: Jossey-Bass.
- Waits, T., & Lewis, L. (2003). *NCES Report, Distance Education at Degree-Granting Post-Secondary Institutions; 2000-2001*, retrieved December 10, 2007 from <http://nces.ed.gov/pubs2003/2003017.pdf>.
- Williams, P.E. (2003). Roles and competencies for distance education programs in higher education institutions. *The American Journal of Distance Education*, 17(1), 45-57.
- Zhang, P. (1998). A Case study on technology use in distance learning. *The Journal of Research on Computing in Education*, 30(4), 398-420.