Research in Higher Education, Vol. 44, No. 3, June 2003 (© 2003)

LEARNING ON CAMPUS AND LEARNING AT A DISTANCE: A RANDOMIZED INSTRUCTIONAL EXPERIMENT

Joseph Collins* and Ernest T. Pascarella‡,**

To address a major methodological problem in the body of evidence on distance learning in postsecondary education, we conducted a randomized, true-experiment paired with a quasi-experiment. Community college students randomly assigned to receive instruction at a distance via a two-way interactive telecourse demonstrated learning equivalent to that of students assigned to on-campus, face-to-face instruction. However, students choosing to take the course via telecourse at remote sites had significantly higher course learning than either randomly assigned group. Such evidence suggests that the body of evidence on distance learning could be seriously confounded by learner self-selection.

KEY WORDS: distance learning; instructional effects; community college students.

INTRODUCTION

Paralleling the recent unprecedented development of information and media technologies has been the dramatic growth of distance or remote site instructional offerings in postsecondary education (El-Khawas, 1995; Moore and Thompson, 1997). For example, a 1997 report by the National Center for Education Statistics found that approximately 60% of American public 2- and 4-year institutions offered distance education courses, usually in the form of either one-way prerecorded courses or two-way interactive video courses (Lewis, Ferris, and Alexander, 1997). Distance education has been used to deliver remote site or off-campus courses in a variety of fields, including business, library science, teacher education, general studies, medicine and nursing, social sciences, social work, and scientific/technical fields (Burgess, 1994).

*Kirkwood Community College, Cedar Rapids, IA.

‡The University of Iowa, Iowa City, IA.

**Address correspondence to: Ernest Pascarella, Educational Policy and Leadership Studies, 491 Lindquist Center North, The University of Iowa, Iowa City, IA 52242. E-mail: ernest-pascarella@uiowa.edu

0361-0365/03/0600-0315/0

Literally hundreds of studies have addressed the issue of whether instruction delivered to remote sites via various media technologies is as effective as traditional face-to-face instruction offered on campus. The clear weight of evidence from this research indicates that students who study via distance education approaches appear to learn as much course content as do their counterparts in conventional, on-campus classroom settings. This conclusion is the consensus of an impressive number of research reviews or meta-analyses of the research evidence (Barker, Frisbie, and Patrick, 1989; Jones, Simonson, Kemis, and Sorensen, 1992; Machtmes and Asher, 2000; Moore and Thompson, 1990, 1997; Olcott, 1992; Pittman, 1991; Russell, 1995, 1999; Schlosser and Anderson, 1994; Wetzel, Radtke, and Stern, 1994; Zigerell, 1991). Moreover, the weight of evidence that does exist also suggests that per-student costs of courses offered in a distance education format are not appreciably different than those offered in a conventional on-campus format (Wetzel et al., 1994).

Despite clear consensus in the evidence, there are major methodological problems in the body of research. A recent report outlines a number of these problems (Institute for Higher Education Policy, 1999; Merisotis and Phipps, 1999). Perhaps the most serious is that geographical or related constraints make it extremely difficult to conduct true-experiments that, because of random assignment of individual learners to instructional formats, maximize the ability to make causal inferences. Rather, nearly every comparative study is, understandably, characterized by students self-selecting themselves into on-campus and remote site instructional groups. Consequently, the body of evidence on distance (vs. face-to-face, on-campus) instruction and student learning is flawed by a major threat to causal inference (or internal validity)—the interaction of self-selection and course achievement (Pascarella and Terenzini, 1991). The reasons why students take courses on campus or at distant sites in the first place may represent a constellation of uncontrolled influences that bias the body of evidence in unknown ways.

This problem is particularly apparent with respect to the use of audiovisual telecourses in distance education. In an informative meta-analysis, Machtmes and Asher (2000) uncovered more than 700 studies that dealt with the use of telecourses in distance education. Out of those 700+ studies, only two were randomized, true-experiments that compared the learning of students receiving instruction via telecourses and those receiving on-campus, face-to-face instruction. The results of these two experiments are inconsistent. One experiment, conducted 40 years ago, indicated that face-to-face instruction tended to be more effective than instruction via a telecourse (Carpenter and Greenhill, 1963). The other experiment, in which the intervention was a single lecture rather than an entire course, suggested the reverse (Ritchie and Newby, 1989). Neither study reported the psychometric characteristics of the instruments used to assess achievement. Furthermore, the studies were conducted before the decade of the

1990s, during which significant advances were made in audiovisual telecourse technology.

Given this major problem in the existing evidence, the present study had two major purposes. First, it addressed concerns about the internal validity of the existing evidence by means of a randomized, true-experiment that compared the learning of community college students receiving course instruction in a traditional face-to-face format on campus with that of students receiving course instruction at a distance via a two-way interactive (audio and visual) telecourse. The second major purpose of the study was to estimate the extent to which learner self-selection (an almost universal characteristic in the existing body of evidence) might bias the results of studies of student learning in face-to-face and distance learning modes. This was accomplished by also comparing the course learning of students in the randomly assigned, face-to-face instructional format with the learning of a third group of students who chose to receive instruction via the telecourse format at remote sites.

METHOD

Instructional Conditions

The study was conducted during the fall of 2001 in a 3-credit hour, one-semester course in the Fire Science program at a community college in Iowa. The course was titled "Firefighting Tactics and Strategy" and is one of seven technical-emphasis courses in a 2-year program leading to an Associate of Science degree in Fire Science Management. It meets once a week for 3 hours each session. The course presents methods of coordinating personnel and equipment, and deploying apparatus on the fire ground. Practical methods of controlling and extinguishing structural and other types of fires are discussed, and a number of simulation exercises are included. Instruction is mostly lecture supplemented by occasional videotapes. The course used a main text, and students were assigned readings from the text. Attendance and class participation were emphasized in the syllabus as being highly correlated with class success.

Each semester the course is offered in two instructional formats on two different evenings. In one format, the students received face-to-face instruction in a traditional classroom on campus. In the other format, students enrolled at various sites around the state received the same course instruction by two-way interactive television on the Iowa Communications Network (ICN).

To conduct this experiment, a change was made in the above arrangement. Thirty-seven students who enrolled to take the course on campus were randomly assigned to two course sections offered in different instructional formats: traditional classroom face-to-face instruction and instruction via a two-way interactive telecourse. Both instructional formats covered identical course content. The

traditionally taught section ($N = 19$) of the course met on Thursday evening from 6–9 PM in a classroom on campus. The telecourse section ($N = 18$) met on Wednesday evening from 6–9 PM in a room on campus equipped to receive interactive telecourse instruction through the ICN. However, the telecourse section instruction originated from a secondary school location about 20 miles from campus, thereby simulating a distance-learning situation. The students never saw the instructor other than over the ICN and did all homework assignments and examinations at a distance.

A third section of the course consisted of nine students who chose to take the course over the ICN on Wednesday evening at various remote sites across the state. These students in the self-selected telecourse sections received identical and simultaneous instruction over the ICN with the randomly assigned telecourse section but at different remote site locations.

The course was taught by two professional firefighters in leadership positions. The first instructor had 15 years experience as a professional firefighter and 9 years experience teaching in the community college's Fire Science program. The second instructor had 5 years experience as a professional firefighter and taught in the Fire Science program for the last three semesters before the study. The two instructors alternated on a weekly basis, so that for every week of the course the instructor of the Wednesday evening ICN sections (both randomly assigned and self-selected) was also the instructor of the Thursday evening randomly assigned face-to-face section. Thus, students in each of the three sections of the course received the same units of course content from the same instructor.

Sample

Of the original sample of 46 student participants (i.e., 19 in the randomly assigned face-to-face instruction group, 18 in the randomly assigned telecourse instruction group, and 9 in the self-selected telecourse instruction group), complete data were available for 38 students. Four students failed to complete the posttest in the randomly assigned face-to-face instruction group, and four students did not complete the posttest in the randomly assigned telecourse instruction group. Thus, the final sample consisted of 15 students in the randomly assigned face-to-face instruction group, 14 students in the randomly assigned telecourse group, and 9 in the self-selected telecourse group. The final sample was 94% male and 94% white. Their age ranged from 19 to 40, with the mean being approximately 24.

Variables

Student learning of the course material was measured by a 61-item examination taken by all participants in the study. The test consisted of 50 multiple-choice items and 11 true-false items. It was administered on the first day of

class as a pretest and was readministered on the last day of class as a posttest. The test was comprehensive in nature and was designed to cover all the subject matter for the course. About two thirds of the questions on the examination were factual in nature; about a third of the questions required students to apply concepts. The KR-21 internal consistency reliabilities for the examination were .75 for the pretest and .78 for the posttest. Moreover, the pretest–posttest correlation was .674, suggesting that the instrument was also reasonably reliable over time. Across all three sections of the course, the pretest had a mean correct of 26.55 and a standard deviation of 6.78. The corresponding figures for the posttest were a mean correct of 40.74 and a standard deviation of 6.85. Thus, irrespective of instructional format, students in the course made a pretest–posttest improvement on the common examination of 2.09 pretest standard deviations (40.74 – 26.55/6.78). This is typically considered a large change (Bowen, 1977) and provides evidence for the content validity of the test; the test appeared to measure what was taught during the course. The posttest was the dependent variable in the study, with the pretest used as a statistical control.

In addition to the pretest, four other variables, on which all students provided complete information, were included in the study as statistical controls. These were: (a) postsecondary credit hours completed, coded 0 = none, 1 = 1–12, 2 = 13–24, 3 = 25–36, 4 = more than 36; (b) previous fire science credits earned, coded 0 = none, 1 = 1–3, 2 = 4–6, 3 = 7–9, 4 = more than 9; (c) licensed emergency medical technician (EMT), coded 1 = yes, 0 = no; and (d) certified firefighter I, coded 1 = yes, 0 = no.

The independent variable, instructional format, was represented by two dummy-coded variables (i.e., 1 and 0). The first dummy variable represented the randomly assigned telecourse group; the second dummy variable represented the self-selected telecourse group. Thus, the comparison group (always coded 0) was the randomly assigned face-to-face instructional group.

Analyses

Analysis of covariance solved by multiple regression was the primary data analytic procedure used in the study. To determine net main effects, the dependent measure, student posttest examination scores, was regressed on five covariates (pretest examination scores, postsecondary credit hours completed, previous fire science credits earned, licensed emergency medical technician, and certified firefighter I) and the two dummy variables representing the three instructional formats. Because of the small final sample size ($N = 38$), a critical α level of .10 was employed in all analyses. Since one of the assumptions of analysis of covariance, however, is the absence of covariate $\times$ instructional intervention conditional (or interaction) effects, a preliminary analysis was conducted. In this preliminary analysis, a set of 10 variables, consisting of the cross-products of the five covariates and the two dummy variables representing instructional for-

mat, was added to the main-effects equation discussed above. The addition of this set of cross-products to the main effects equation was associated with an increase in explained variance (R^2) in posttest examination scores of 9.2%, which was nonsignificant at $p < .10$ ($F = 0.81$ with 10 and 21 degrees of freedom, $p > .10$). Consequently, we could not reject the null hypothesis of no covariate × instructional format conditional effects, and a straightforward interpretation of the net, main effects equation was warranted (Pedhazur, 1982).

RESULTS

Table 1 provides, for each instructional condition, the descriptive statistics for all variables employed in the analyses. As the table indicates, the randomly assigned telecourse and face-to-face instructional groups were quite similar on all variables and particularly homogeneous on the pretest. However, the self-selected telecourse group was, on average, between one or two standard deviations higher on the pretest than the randomly assigned groups. Similarly, it also tended to average more postsecondary credits and previous fire science credits earned, and to have a higher percentage of licensed emergency medical technicians and certified firefighters. In short, self-selection led to a substantially different sample of individuals taking the telecourse than did random assignment.

TABLE 1. Descriptive Statistics for All Variables

	Randomly Assigned Telecourse Group ($N = 14$)			Randomly Assigned Face-to-Face Group ($N = 15$)			Self-Selected Telecourse Group ($N = 9$)		
Variable	M	SD	%Yes	M	SD	%Yes	M	SD	%Yes
Pretest	24.50	8.17		24.60	4.81		33.00	4.18	
Posttest	38.36	5.72		38.47	4.61		48.22	6.64	
Postsecondary credit hours completed	1.57	1.55		1.93	1.71		2.22	1.39	
Previous fire science credits earned	1.07	1.59		0.67	1.45		2.44	1.74	
Licensed emergency medical technician (EMT)			28.6			26.7			77.8
Certified firefighter I			42.9			26.7			88.9

The estimated effect of instructional format on the dependent variable (posttest examination scores) is summarized in Table 2. As the table indicates, the five covariates explained 58.2% of the variance in posttest scores ($p < .01$), and instructional format was associated with an R^2 increase of 8.8% over and above the covariates ($p < .05$). The unstandardized coefficient (b) in Table 2 represents the covariate-adjusted difference between comparison group means on the posttest. Thus, net of the covariates, the randomly assigned telecourse group had, on average, less than a one-question advantage (b = .756) over the randomly assigned face-to-face instructional group in the number of questions answered correctly on the posttest. This small difference was not even close to being statistically significant. In short, when participants in the study were randomly assigned to either traditional on-campus instruction or distant-site telecourse instruction, the result was essential parity in course learning.

As Table 2 also indicates, however, this was not the case for the self-selected telecourse group. Net of the five covariates, the self-selected telecourse group had, on average, more than a six-question advantage (b = 6.092) over the randomly assigned face-to-face instructional group in the number of questions answered correctly on the posttest. This net learning advantage for the self-selected telecourse group was not only statistically significant it was also quite substan-

TABLE 2. Estimated Effects of Instructional Format on Posttest Scores

Variable	Degrees of Freedom	Unstandardized Coefficient (b)	Beta	*t*
Covariates[a]				
Pretest	1	.434	.430	2.62**
Postsecondary credit hours completed	1	.837	.191	1.58
Previous fire science credits earned	1	.728	.179	1.16
Licensed emergency medical technician (EMT)	1	5.600	.405	1.70*
Certified firefighter I	1	–3.779	–.279	1.27
$1 - R^2$ (.418)	32			
Instructional Format[b]				
Randomly assigned telecourse group[c]	1	.756	.054	0.44
Self-selected telecourse group[c]	1	6.092	.383	2.73**
$1 - R^2$ (.330)	30			

[a] R^2 increase = .582 (5, 32), $p < .01$.
[b] R^2 increase = .088 (2, 30), $p < .05$.
[c] Comparison is with the randomly assigned face-to-face instruction group.
*$p < .10$; **$p < .05$.

tial in magnitude. Dividing the regression coefficient (6.092) by the posttest standard deviation of the randomly assigned face-to-fact instruction group shown in Table 1 (4.61) converts to an average learning advantage for the self-selected telecourse group over the face-to-face instruction group of 1.32 standard deviations.

Although they had the same instructor and were receiving identical and simultaneous instruction via the ICN, it was also the case that the self-selected telecourse group had a substantial learning advantage over the randomly assigned telecourse group. The difference between the unstandardized regression coefficients for the two groups in Table 2 (6.092 – .756) provides the average covariate-adjusted learning advantage of the self-selected over the randomly assigned telecourse group (5.336 more posttest questions answered correctly). A post-hoc comparison indicated that this advantage was statistically significant at $p < .05$, and when divided by the posttest standard deviation of the randomly-assigned telecourse group (5.72), it converted to a learning advantage for the self-selected telecourse group of .93 of a standard deviation.

SUMMARY AND CONCLUSIONS

The vast body of evidence on the comparative effects on student learning of traditional face-to-face instruction delivered on campus vs. instruction delivered at a distance via various telecourse technologies is plagued by the virtual absence of randomized, true-experimental studies. This investigation sought to address this major problem by means of a true-experiment in which community college students taking a one-semester course in fire science were randomly assigned to one of two instructional conditions: traditional face-to-face instruction on campus or instruction at distance via a two-way interactive (audio and visual) telecourse. A review of the literature suggests that this is perhaps one of the few, if not the only, true-experiments assessing the comparative effects of on-campus learning vs. learning at a distance via telecourse conducted in more than a decade. The investigation also estimated the extent to which learner self-selection might bias the results of distance education studies by adding a third instructional condition in which students at remote sites chose to receive the telecourse instruction simultaneously with students randomly assigned to that instructional condition.

Our findings suggest two major conclusions. First, they support the prevailing view that postsecondary students can master course facts and concepts as well when they receive instruction at a distance via a two-way interactive telecourse as they can when they receive the same instruction on-campus in a traditional face-to-face format. Most important, perhaps, they underscore the robustness of this conclusion by indicating that it holds not only when students can self-select

themselves into instruction on campus vs. instruction at a distance but also under the more internally valid condition that exists when students are randomly assigned. Furthermore, since we found no covariate × instructional format interactions, it would suggest that this parity in the learning outcomes of instruction delivered on campus or at a distance holds irrespective of students' precourse level of content knowledge, their prior exposure to postsecondary education, or their related professional training and experience.

Although the course was experienced on campus by the randomly assigned face-to-face group and at a distance by the randomly assigned telecourse group, the only substantive difference in the actual instruction delivered to the two groups was the medium employed. Both randomly assigned experimental conditions received identical instruction from the same instructor. Thus, the parity in learning demonstrated by the two groups may also have implications for instructional theory in that it lends support to the argument of several scholars that the specific medium of instruction (e.g., face-to-face, television, videotapes, computer conferencing, two-way interactive video) has little impact on how much students learn (Carter, 1996; Clark, 1991, 1994; Schlosser and Anderson, 1994).

A second major conclusion is that comparative investigations of learning on campus vs. at a distance in which students self-select themselves into on-campus and remote site conditions have a high potential to yield confounded results. This conclusion stems from our finding that a group of students who self-selected themselves to receive instruction via telecourse at remote sites demonstrated significantly greater course learning than either the randomly assigned face-to-face or the randomly assigned telecourse group. It is highly unlikely that this finding is the result of instructional format. In our study, the self-selected telecourse group had the same instructor for every unit of course content as the two randomly assigned groups and received instruction simultaneously with the randomly assigned telecourse group. Rather, there is a much greater probability that the enhanced course learning demonstrated by the self-selected telecourse students is attributable to uncontrolled individual differences such as the interaction of self-selection and change. The personal characteristics and level of professional motivation that may have led such individuals to enroll in the telecourse at remote sites may also account for their learning more during the course. As shown in Table 1, the self-selected telecourse group not only had substantially higher average pre-course knowledge than the two randomly assigned groups but also more exposure to postsecondary education and higher levels of professional training and certification. Moreover, while it could not be matched to specific individuals, the randomly assigned face-to-face and telecourse groups had taken an average of 1.7 previous courses in a telecourse format. In contrast, the self-selected telecourse group had, on average, nearly twice as much experience with learning via telecourse (3.0 courses). It is quite possible that such advantages are indicative of particularly high professional

motivation and success in distance-learning modalities, which not only led to enrollment in the telecourse in the first place but also to higher levels of course achievement. Our extensive statistical adjustments may not have been particularly effective in controlling for the confounding effects of the interaction of self-selection and change. Thus, warnings that the vast body of existing evidence on distance learning is confounded by almost total reliance on quasi-experimental studies with learner self-selection to on-campus and remote-site instructional conditions (Institute for Higher Education Policy, 1999; Merisotis and Phipps, 1999) may have considerable validity.

LIMITATIONS

This study is limited in several ways. First, the measure of course learning was an instrument that measured student understanding of basic facts and concepts in an applied area—fire science. It is not clear that the findings would generalize to other disciplines or to the development of more complex higher order reasoning skills such as evaluation, synthesis, or critical reasoning. Second, the small sample size in the randomized experiment part of the study ($N = 29$) increased the probability of a Type II error (i.e., accepting the null hypothesis when it is false) by affording little statistical power to detect statistically significant differences between the face-to-face and the telecourse instructional formats. Weighed against this, however, are two factors. First, the dependent measure had adequate reliability, ranging between .75 and .78. Second, and perhaps more important, our analytic procedure included covariates that had substantial correlations with the dependent variable (posttest examination scores). Entering the covariates in the regression equation reduced the error term ($1 - R^2$) by nearly 60%. This substantially increased the likelihood of detecting real posttest performance differences between the randomly assigned face-to-face and telecourse conditions (Pedhazur, 1982).

ACKNOWLEDGMENTS

This study is based on the first author's doctoral dissertation at The University of Iowa, which was directed by the second author.

REFERENCES

Barker, B., Frisbie, A., and Patrick, K. (1989). Broadening the definition of education in light of the new telecommunications technologies. *Am. J. Distance Educ.* **3:** 20–29.
Bowen, H. (1977). *Investment in Learning*, Jossey-Bass, San Francisco.

Burgess, W. (1994). *The Oryx Guide to Distance Learning*, Oryx Press, Phoenix.
Carpenter, C., and Greenhill, L. (1963). *Comparative Research on Methods and Media for Presenting Programmed Courses in Mathematics and English*, The Pennsylvania State University, The University Division of Instructional Services and the Departments of Mathematics and English, University Park.
Carter, V. (1996). Do media influence learning? Revisiting the debate in the context of distance education. *Open Learn.* **11:** 31–40.
Clark, R. (1991). When researchers swim upstream: Reflections on an unpopular argument about learning from media. *Educ. Technol.* **31:** 34–38.
Clark, R. (1994). Media will never influence learning. *Educ. Technol. Res. Dev.* **42:** 21–29.
El-Khawas, E. (1995). *Campus Trends, 1995*, American Council on Education, Washington, DC.
Institute for Higher Education Policy (1999). *What's the Difference: A Review of Contemporary Research on the Effectiveness of Distance Learning in Higher Education*, Author, Washington, DC.
Jones, J., Simonson, M., Kemis, M., and Sorensen, C. (1992). *Distance education: A cost analysis*, unpublished manuscript, Iowa State University, Ames, IA.
Lewis, L., Ferris, E., and Alexander, D. (1997). *Distance education in higher education institutions* (NCES Report, No. 98-062). U.S. Department of Education, Office of Educational Research and Improvement.
Machtmes, K., and Asher, J. (2000). A meta-analysis of the effectiveness of telecourses in distance education. *Am. J. Distance Educ.* **14:** 27–46.
Merisotis, J., and Phipps, R. ((1999). What's the difference? Outcomes of distance vs. traditional classroom-based learning. *Change* **31:** 12–17.
Moore, M., and Thompson, M. (1990). *Effects of Distance Learning: A Summary of Literature*, The American Center for the Study of Distance Education, University Park, PA.
Moore, M., and Thompson, M. (1997). *The Effects of Distance Learning: Revised Edition*, The American Center for the Study of Distance Education, University Park, PA.
Olcott, D. (1992). *Instructional television: A review of selected evaluation research*, unpublished manuscript, Oregon State University, Corvallis, OR.
Pascarella, E., and Terenzini, P. (1991). *How College Affects Students: Findings and Insights from Twenty Years of Research*, Jossey-Bass, San Francisco.
Pedhazur, E. (1982). *Multiple Regression in Behavioral Research: Explanation and Prediction* (2nd Ed.), Holt, Rinehart & Winston, New York.
Pittman, V. (1991). Academic credibility on the "image problem": The quality issue in collegiate independent study. In: Watkins, B., and Wright, S. (eds.), *The Foundations of American Distance Education: A Century of Collegiate Correspondence Study*, Kendall/Hunt Publishing Company, Dubuque, IA, pp. 109–133.
Ritchie, H., and Newby, T. (1989). Classroom lecture/discussion vs. live televised instruction: A comparison of effects on student performance, attitude, and interaction. *Am. J. Distance Educ.* **3(3):** 36–45.
Russell, T. (September 1995). *The "No Significant Difference" Phenomenon as Reported in 214 Research Reports, Summaries, and Papers*, Office of Instructional Telecommunications, North Carolina State University, Raleigh.
Russell, T. (January 1999). *The "No Significant Difference" Phenomenon as Reported in 248 Research Reports, Summaries, and Papers* (4th Ed.), Office of Instructional Telecommunications, North Carolina State University, Raleigh.

Schlosser, C., and Anderson, M. (1994). *Distance Education: Review of the Literature*, Research Institute for Studies in Education, Iowa State University, Ames.
Wetzel, C., Radtke, P., and Stern, H. (1994). *Instructional Effectiveness of Video Media*, Lawrence Erlbaum Associates, Hillsdale, NJ.
Zigerell, J. (1991). *The Uses of Television in American Higher Education*, Praeger, New York.

Received February 25, 2002.