Do Online Exams Facilitate Cheating? An Experiment Designed to Separate Possible Cheating from the Effect of the Online Test Taking Environment

Alan Fask · Fred Englander · Zhaobo Wang

Abstract Despite recent growth in distance education, there has been relatively little research on whether online examinations facilitate student cheating. The present paper utilizes an experimental design to assess the difference in student performance between students taking a traditional, proctored exam and those taking an online, unproctored exam. This difference in performance is examined in a manner which considers both the effect of the different physical test environments and the possible effect of a difference in the opportunity for students to cheat. This study, utilizing regression models that also account for relevant control variables, examines 44 undergraduate statistics students, finds evidence that the difference in the testing environment creates a disadvantage to students taking the online exam which somewhat offsets the advantage that the unproctored students gain from greater opportunities to cheat.

Keywords Student performance · Experimental design · Online assessment · Student cheating

Introduction

There has been an unmistakable surge in the growth of distance learning courses over the last 15 plus years. This trend has coincided with convincing evidence (McCabe et al. 2001; Brown and McInerney 2008) supporting the view that students are becoming more inclined over time to cheat on exams and other graded assignments. There have been concerns expressed by some students and faculty that the online examinations, common to distance learning courses, may be more vulnerable to student cheating than exams given in a more traditional, proctored environment. Despite this cluster of trends and their implication that the fundamental integrity of this increasing reliance on distance learning courses may be threatened by insufficient efforts to detect and deter student cheating, there has been remarkably scant research performed to measure the extent to which online exams facilitate student cheating. Among the

A. Fask · Z. Wang
Department of Information Systems and Decision Sciences, Fairleigh Dickinson University, Madison, NJ, USA

F. Englander (✉)
Department of Economics and Finance, Fairleigh Dickinson University, Madison, NJ, USA
e-mail: englandr@fdu.edu

handful of studies that have been performed to determine if online examinations are more vulnerable to student cheating, only one study (Hollister and Berenson 2009) has approached this research task with the use of a randomized experimental design.

A second goal of this paper, in addition to the empirical examination of cheating, is to study the effect of the test environment on grades. Thus the methodology utilized here is designed to isolate the separate effects on student performance of a greater possibility of cheating in an online environment from the possible differences in student comfort, familiarity and distractions that may be endemic to taking an exam online. Thus if a difference in in-class vs. online performance is observed, it is important to ascertain if such a difference is a result of cheating or the result of taking the final exam in the in-class as compared to the online environment.

Review of the Literature

Growth of Distance Learning

The extraordinary growth in postsecondary distance learning courses is reflected in the 16-fold increase in enrollment from 753,640 in academic year 1993–94 to 12,153,000 in academic year 2006–07 (Lewis et al. 1999; Parsad and Lewis 2008). The proportion of students in postsecondary institutions taking at least one distance learning course increased from 9.6 % in the fall of 2002 to 29.3 % in the fall of 2009. Over that same time period the compound annual growth rate in enrollment in online courses was 19 %, compared to the growth rate in the total student body of less than 2 %. The increase in the percentage of total enrollment represented by online enrollment is hardly slowing down. That percentage rose from 24.6 % in fall 2008 to 29.3 % in fall 2009, matching the highest year to year increase over the fall 2002 to fall 2009 interval (Allen and Seaman 2010). Evidence that this growth trend is likely to continue is provided by the percentage of college and university administrators who describe online education as being "critical to the long-term strategy of my institution." That percentage rose from 48.8 % in fall of 2002 to 63.1 % in the fall of 2010 (Allen and Seaman 2010).

Faculty Concerns About Academic Integrity

The prodigious growth of online courses and the resulting common use of online exams have raised concerns and even fears of the possible threat to the academic integrity of these courses to the extent that online courses are viewed as an assessment approach that is more vulnerable to student cheating than traditional face to face courses. Hollister and Berenson (2009: 272) reviewed a number of studies concerning the viability of distance learning and the corresponding assessment practices and conclude that concerns over the greater potential for student cheating in an online environment are paramount: "The most commonly reported challenge in online assessment is how to maintain academic integrity." Studies reporting that college faculty perceive online testing as offering a greater opportunity for cheating than traditional classroom environments have been undertaken by Kennedy et al. (2000) Stuber-McEwen et al. (2005) and Rogers (2006). Rogers (2006) reports that a majority of respondents to her survey were utilizing online testing and a majority of that group reported at least one actual or suspected occurrence of cheating during an online exam. Of some concern is that 81.8 % of the faculty respondents that report relying on online exams or quizzes administer those assessments in an unproctored environment. Although such a behavior pattern may appear paradoxical, the unwillingness of college faculty to be more aggressive in restraining student cheating has been addressed by Keith-Spiegel et al. (1998), Schneider (1999) and Coalter et al. (2007).

Levels and Trends in Student Cheating

These concerns regarding the possible vulnerability of online testing to greater student cheating are concurrent with the promulgation of a substantial body of research documenting a high level of college student willingness to cheat on graded assignments and that this willingness appears to becoming more formidable over time. Representative studies by Crown and Spiller (1998) McCabe et al. (2001) and Brown and McInerney (2008) supported the view that students have been demonstrating a very high willingness to cheat over the last several decades. McCabe et al. (2001), for example, surveyed college students in 1963 and reported that 75 % of students had undertaken at least one act of 'serious cheating.' A comparable 1993 survey revealed that the percentage had risen to 82 %. As part of this research, McCabe et al. (2001) find that various types of examination-related cheating had increased from 39 to 64 % over that 30 year period. Brown and McInerney (2008) find even higher cheating prevalence rates among college students and also report an increase in these prevalence rates that was statistically significant over the 1999 to 2006 period.[1]

Earlier Studies of Vulnerability of Online Exams

An analysis of the existing literature related to efforts to determine whether online examinations facilitate student cheating indicates that several methodological strategies have been utilized. One such approach (Grijalva et al. 2006; Kidwell and Kent 2008; Stuber-McEwen et al. 2009; Kennedy et al. 2000; Lanier 2006) has been to employ a survey-based approach utilizing an anonymous survey which asks students whether and to what extent they have cheated in traditional, face to face exams and in online exams. However, there is an obvious problem. Why would the researcher expect a student who has cheated to admit it even anonymously? Furthermore, a guilt ridden student may respond positively, even to a very mild cheating infraction while a sociopathic student may respond negatively regardless of the severity of the infraction. Another relevant set of objections to the use of such self-report surveys of undergraduate students has been raised recently by Porter (2011: 45, 46) who questions whether any self-report surveys of undergraduates are valid because,

> "(a) they assume that college students can easily report information about their behaviors and attitudes, when the standard model of human cognition and survey response clearly suggests they cannot, (b) existing research using college students suggests they have problems correctly answering even simple questions about factual information, and (c) much of the evidence that higher education scholars cite as evidence of validity and reliability actually demonstrates the opposite."

Another approach is to employ an observational study of two groups of classes, one taking online examinations and the second group taking proctored, in-class exams, giving them comparable or the equivalent exams and testing for a difference in grades between the groups. This approach has been tried less often (the present authors are only aware of the research efforts of Peng (2007), Harmon and Lambrinos (2008)[2] and Yates and Beaudrie (2009)) and

---

[1] This evidence of high and increasing student cheating prevalence may be the source of even greater apprehension when one considers the evidence presented by Nonis and Swift (2001), Harding et al. (2004), Lawson (2004) and Harper (2006) to the effect that cheating behavior by college students is statistically related to a higher propensity to demonstrate dishonest behavior when such individuals move on to business or professional careers.

[2] The Harmon and Lambrinos (2008) study also made an inference regarding student cheating in online exams by examining how well a regression model (featuring student GPA as an independent variable) explaining student exam scores predicted those exam scores in classes which relied on proctored exams versus classes which utilized online exams.

would seem more desirable, especially if there is some confidence that the exams and the composition of the classes are equivalent.

A randomized experimental design method was employed by Hollister and Berenson (2009) who selected two sections taking the same course during the same semester, made an effort to verify that the students in the two classes had comparable abilities and characteristics, and then randomly assigned (utilizing a coin toss) an assessment mode involving in-class proctoring to one section and online assessment to the other section. A serious problem with this experimental technique, which has received attention to this point only from Hollister and Berenson (2009), is that the online environment is physically and psychologically different from the in-class environment. That is, the online environment is the home or the dorm; the in-class environment is the classroom. Note that Hollister and Berenson (2009) suggested that the online environment of the home or dorm is likely to handicap the exam taker because of the possible difficulties with his/her computer or network connection or because of noise factors or other distractions. It is suggested here that it can also be argued that the online exam taker has the advantage of greater comfort in less structured, less tension-inducing and more familiar surroundings. Thus the very difference in the environment could account for a difference in grades, even if no cheating is present. Furthermore the direction of the difference is indeterminate a priori, at least with the current state of knowledge. Thus from a study design point of view, the test environment is confounded with cheating.

In designing a study to separate the effects of environment from cheating, it is helpful to consider how an ideal research design, one which would meet Campbell and Stanley's (1966) criteria for a "true" experimental design, might be formulated. Such a design could be based on performance comparisons among the following four randomly assigned groups:

A)   In-class, proctored
B)   In-class, unproctored
C)   Online, proctored
D)   Online, unproctored

This would yield a 2×2 factorial design with interaction. The main effects would be (1) in-class vs. online and (2) proctored vs. unproctored. Thus the proctored/unproctored main effect would reflect the proctoring effect, the online/in-class main effect would reflect the environmental effect and a statistically significant positive interaction term which raised grades in group D would suggest a "cheating" effect, since such an interaction would indicate an effect on grades above and beyond the two additive main effects in the online, unproctored group.

Unfortunately, this design is virtually impossible to implement. One potential problem with the design is that group C cannot be practically implemented since it would require a proctor to be present in the home or the dorm of the various students assigned to that study group. Although students could instead be required to take online exams in separate facilities under the supervision of a third party proctor, it may be argued that a requirement that students take online exams in the presence of a third party proctor creates several types of problems. One of the principle attractions of online courses is the flexibility that they offer relative to traditional courses—the ability of the online student to undertake the necessary work at a time and place that is convenient to the online student. The research of Cluskey et al. (2011), Kitahara et al. (2011) and Kolowich (2013) suggest that various approaches to proctoring online exams offer their own set of practical limitations and threats to exam integrity.

However, the more significant problem in implementing such a pure Campbell and Stanley (1966) design would be the obvious difficulty in creating a group B. The idea of putting a whole class of students together in a single room in order to take an exam and then leaving that

group unattended during the exam for the purpose of a research study would invite scandal. For these reasons, a "true" experimental design, which would unambiguously separate out the effect of cheating from the effect of the test environment, would seem unrealistic.

Since a true experimental design cannot be constructed, it is suggested that definitive separation of environmental and cheating effects is simply not possible. The best the researcher can do is conduct a variety of different investigations, hoping that the entire body of such investigations will lead to a reasonable resolution of the problem of whether or not the test environment (in-class vs. online) plays a role in test performance. Resolution of this problem would then permit a clearer assessment of the likelihood that cheating is occurring. One such investigation is presented in the current study.

Data and Experimental Procedures

The student subjects in this study were enrollees in two separate classes of elementary statistics given in the fall 2009 semester on a campus of a private university in the northeastern region of the United States. The course represents the first semester of a two semester statistics sequence that is required of business majors. The business program is accredited by the AACSB. Both classes had the same topic coverage, lectures and homework assignments. The students in both classes were presented the course material only in a traditional, face to face environment. Students only learned of the mode (proctored, face-to face or online) in which they would take the final exam in the last week of the course. Students in both classes had the same in-class mid-term examinations. Students were not given any indication that they were participating is a study relating to student cheating. There were 22 students enrolled in each class. While the students self-enrolled for each class, there was no cause for the students to think that the classes would in any way be treated differently. Furthermore, even though there was no formal random assignment of students to classes, there was no reason to believe that there was any systematic difference in the students enrolled in the classes. That is, we believe the assignment was effectively random. Table 1 presents the mean values, by class section, of various student characteristics for the students in the two respective sections. It is noted that there were no characteristics for which a statistically significant difference between the two classes was observed.

To review, there were two classes of students who, according to the evidence presented in Table 1, were of approximate equal ability. They had the same lectures, homework and classroom/online experiences. That is, both groups were treated virtually identically. The research design endeavored to construct an approach to separate the impact of student cheating on exam performance from the impact of environment on exam performance. This was accomplished by giving a "practice test" to each class 3 days before the actual exam, with one class taking the practice exam in an online environment and the other taking it in an in-class environment.

The determination as to which class received the practice test online versus in-class was made by random assignment (i.e., based on a coin toss), an approach similar to that used by Hollister and Berenson (2009). Note that since we believe the students were essentially randomly assigned to each class, the coin toss simply ensures random assignment of each student to each environment. The students were encouraged to work alone and do the best they could on the practice test. It was explained to the students in both sections that the nature and

Table 1 Mean values of student characteristics in both sections of elementary statistics

Variable	Mean value in-class section (N=22)	Mean value online section (N=22)	t-stat	p-stat
Attendance	22.91	21.59	1.29	0.20
Credits attempted, fall 2009	14.64	15.05	−0.71	0.48
Credits completed prior to fall 2009	36.32	42.18	−1.20	0.24
GPA	2.76	2.83	−0.39	0.70
Homework grades	72.63	83.07	−1.47	0.07
Mid-term grade	85.00	83.10	0.74	0.46
SAT math	536.70	520.50	0.72	0.48
SAT verbal	497.80	480.00	−0.84	0.41
Initial math test[a]	6.82	7.00	−0.32	0.75

[a] This 'Initial math test' is a ten question, 20-minute test of basic math skills given to students during the first class meeting of the semester

structure of the practice exam would be similar to that of the final exam. Therefore, students were told, a strong effort to take the practice exam seriously would likely pay off in providing a better foundation for the preparation process leading to the final exam. It was further emphasized that taking the practice exam was a prerequisite for taking the final exam. However, it was made very clear to all student participants that the practice test would not be counted towards their semester grade. The assumption was that the practice test would then be relatively free of student cheating. A comparison of the online and in-class practice test scores would then be an indication of the environmental effect. Nonetheless, because of the practical problems in implementing a true, Campbell-Stanley (1966) type of experimental design described above, alternative interpretations of this study's findings are possible. These problems will be addressed below in the "Limitations" section.

The actual final exam was then given to both groups in the same environment that they took the practice test. That is, the online and in-class groups took the finals in the same venue that they took the practice test. If the test environment truly had a negative (positive) effect, then it should have that negative (positive) effect for both the practice test and the final, especially since the time lapse from the practice test to the final was just 3 days.

Results

Examining the Isolated Impact of the Exam Environment on Performance

As indicated, the regression results for the two sections taking the practice exam, administered to one section in a proctored, in-class environment and administered to the other section in an online environment (i.e., in the same respective environments in which the students in each section would be taking the actual final exam), offer an indication of the influence of the respective environments on exam performance. The regression model, formally stated in Eq. (1), relates the practice score dependent variable for each student, i, $PScore_i$, to an intercept term, α, a binary variable measuring the section in which student i is enrolled,

$SDummy_i$, student i's cumulative grade point average in prior courses, GPA_i, in order to account for the student's overall academic ability and prior academic performance, and a stochastic error term, $\in_i$.

$$PScore_i = a + b_1 SDummy_i + b_2 GPA_i + \in_i \quad (1)$$

If the coefficient of the *SDummy* variable is significant and positive, such a result would suggest that the online testing environment, denoted with the dummy variable equal to one, yields an advantage to students in that environment relative to students in the proctored, in-class environment. Alternatively, if the coefficient of the *SDummy* variable is found to be negative and significant, an appropriate inference would be that students in the in-class environment had enjoyed a relative advantage over the students in the online testing environment.

The results generated from the estimation of regression Eq. (1) are presented in Table 2. It is observed that coefficient of the *SDummy* variable is −14.01. This indicates that the students taking the practice exam online performed an average of just over 14 points lower, *ceteris paribus*, than the students taking the exam in a proctored environment. The *p*-value for this finding was 0.05, indicating a statistically significant result. Therefore, it may be inferred that there is a basis for believing, as Hollister and Berenson (2009) suggest, students taking an exam in an online environment do face a disadvantage relative to other students in a traditional, proctored environment because of the absence of a proctor to provide clarifications to questions on the exam, the possibility of greater distractions in the ambient environment and possible problems with the students' computers or the connectivity of those computers. Again, inasmuch as students were clearly informed prior to the event that the practice test would not count in the determination of their semester grades, the authors believe that an appropriate interpretation of the *PScore* variable is that it reflects the influence of the testing environment on student performance, independent of the alleged connection between testing environment and student cheating.

It should also be noted that the coefficient of the *GPA* variable was, as expected, positive and statistically significant. The adjusted R-Squared result of 0.35 and the *p*-value (<0.0001) associated with the F-Statistic of 12.45, suggest that the explanatory variables of the regression equation do explain a significant part of the variation in the *PScore* variable.

Examining for the Presence of Cheating Behavior on the Actual Final Exam

Having estimated the statistical association between the testing environment and student performance in the preceding section, it is now possible to focus on the use of the experimental approach described above to separately examine whether an association between the testing mode (i.e., proctored, face to face versus unproctored, online) and student cheating can be ascertained. These experimental procedures were embodied in a regression model explaining

Table 2 Practice exam regression model

Variable	Estimate	Std error	*t*-value	*P*
Intercept	−22.61	17.14	−1.32	0.19
SDummy	−14.01	7.06	−1.99	0.05
GPA	27.89	5.94	4.69	<0.01
R-Square	0.38	F Statistic	12.45	
Adj R-Sq	0.35	Pr>F	<0.01	

the final exam scores of the combined 44 students which included the students from both the proctored, in-class section and the unproctored class which took the exam online. The final exam grade for student i ($Exam_i$), the dependent variable, was regressed against an intercept term (α), a binary variable ($SDummy_i$) reflecting the proctored/unproctored conditions under which the student completed the final exam, a variable measuring the cumulative credits ($CCompl_i$) that the student had completed at the beginning of the semester, a variable representing the score that the student received in the proctored mid-term exam ($MTerm_i$, all students took the mid-term in class) and a variable representing the score the student received in the practice test ($PScore_i$), which was taken 3 days before the actual exam and was also taken in the same proctored or unproctored environment in which the student would be taking the final exam. A stochastic error term is represented by ϵ_i. This regression model is stated more formally in Eq. (2):

$$Exam_i = a + b_1 SDummy_i + b_2 CCompl_i + b_3 MTerm_i + b_4 PScore_i + \epsilon_i \qquad (2)$$

The *CCompl* variable and the *MTerm* variable were included in the model on the basis of a stepwise regression procedure which examined a number of student characteristic variables[3] that would serve as proxies for student motivation or ability and could be expected to be substantially collinear with one another. In order to avoid the exaggeration of variable standard errors and misleading diminution of statistical significance that such collinearity fosters, the stepwise procedure was employed to cull these related student characteristic variables.

The inclusion of the *PScore* serves the purpose of adjusting for the score that a student received on his/her practice test, taken in the same environment in which the student would subsequently take the actual final exam. That is, if the dorm or home environment utilized by the students taking the online exam was more conducive to better student performance on the final exam (due to a possible student preference for a less structured, more familiar or more physically comfortable surroundings), the impact of such an advantage would accounted for by the *PScore* variable. Alternatively, if the dorm or home environment experienced by those students proved to be less conducive to a better performance on the final exam (due to possible difficulties a student might have with his/her computer connection or the absence of a proctor to clarify a possible ambiguity in the exam), that disadvantage would also be accounted for by the *PScore* variable. Therefore the sign and statistical significance of the separate *SDummy* variable coefficient would serve to measure the effect of whether the opportunity to take an exam online leads to an unfair advantage to the extent that collaboration and cheating are facilitated. Within the context of a one-tailed test examining whether online testing is conducive to student cheating, a positive and statistically significant sign for the *SDummy* variable would support the view that students taking the final exam online are more inclined to cheat. Alternatively, the absence of such a finding would be consistent with the view that there is no significant difference between proctored, in-class exams and unproctored, online exams with respect to student cheating.

---

[3] The original list of variables included in the stepwise regression procedure were: Section Dummy, score on the initial math test, hours employed, hours exercising, hours on the internet, math SAT, verbal SAT, GPA, credits completed, credits attempted, residential status, midterm grade, class attendance, homework grades, measures of procrastination in starting and submitting homework assignments and gender. The stepwise procedure uses partial correlation to determine which variable to add into the model which maximizes the coefficient of determination.

Results Relating to Cheating on the Actual Final Exam

The results from the regression model of Eq. 2 are presented in Table 3. Note that the *SDummy* variable is statistically significant, with the online class scoring 10.13 percentage points higher than the in-class group (after controlling for the independent variables). As indicated from the earlier *t*-test analysis (Table 2), the data significantly suggest that the online environment is detrimental to student performance. On the other hand, this current analysis of the final exam scores suggests that the online environment scores are significantly higher than the in-class test scores. It is therefore concluded that this extraordinary change in the online test scores from adversely affecting performance on the practice final exam to bolstering performance on the actual final exam is likely the result of cheating in the online class.

It is also observed that the coefficient of the *MTerm* variable is positive and significant, indicating that students who performed better on the mid-term exam also performed better on the final exam. The positive and statistically significant impact of the *PScore* variable indicates that students who achieved a higher score on the practice text, taken 3 days in advance of the actual final, also performed better on the actual final. The significant and negative result for the coefficient of the *CCompl* variable indicates that students who are farther along in their academic program performed systematically worse on the final than students with fewer cumulative credits. This may seem counter-intuitive, especially in light of prior research (Okpala et al. 2000; Durden and Ellis 2003) that have shown that students farther along in their academic program tend to outperform students closer to the outset of their college careers. However, it may be the case that many students find quantitative courses such as elementary statistics to be more formidable than most other courses in the business curriculum. That is, such students may be affected by *statistics anxiety*. Rodarte-Luna et al. (2006) has suggested that one common response to statistics anxiety for undergraduates is to delay taking even the first statistics course until later in one's academic program. To the extent that students with greater statistics anxiety are justified in their lack of confidence in their abilities to perform well in the first statistics course, the inverse relationship between the final exam score and the *CCompl* variable may be better understood.

Limitations

Although the sample size of 44 students is relatively small, a smaller sample size generally decreases the investigator's ability to find statistical significance. However,

Table 3 Final exam regression model

Variable	Estimate	Std error	*t*-value	*P*	VIF score
Intercept	12.63	24.01	0.53	0.60	
SDummy	10.13	4.80	2.11	0.02*	1.09
MTerm	0.83	0.29	2.90	< 0.01	1.06
PScore	0.19	0.09	2.16	0.04	1.10
CCompl	−0.48	0.14	−3.30	< 0.01	1.04
R-Square	0.41	F Value	6.76		
Adj R-Sq	0.35	Pr>F	0.00		

*This *p*-value is for a one-tailed test

the sample size was adequate to yield statistical significance regarding the central issue of the relationship between online testing and student cheating. The approach that is offered in this study to diagnose the presence of cheating among students in two (or more) sections of a class taught by the same instructor would appear to be useful to detect cheating in either small or large classes (i.e., in small sample size or large sample size environments). It may be hypothesized that the nature and diagnosis of cheating would be different in these two environments. Therefore, this research effort potentially makes a meaningful contribution by applying this experimental methodology and demonstrating its potential effectiveness in a small class environment. It is believed this methodology also can be used to detect cheating in large class situations, but this remains to be investigated.

A potential limitation of the present study flows from the practical difficulties in implementing a true, experimental (Campbell-Stanley 1966) approach in separating the impacts on the final exam attributable to a proctor/non-proctor effect and an environmental effect. That is, in the absence of a true experimental design, alternative explanations of results are almost always to be expected. Here, the use of a practice exam given separately to the proctored section and the online section as a technique for gauging the effect of the physical testing environment potentially raises questions. It is possible that the difference in exam scores between the students in those two sections could alternatively result from differences in the earnestness applied to the practice exam by the students in the respective sections. In the same vein, it was explained to students in both sections that the practice exam would not count in the computation of the students' semester grades. That is, it would be pointless to cheat on the practice exam. Additionally, the students were told that the actual final exam would be quite similar to the practice exam. The authors made sure that these facts were emphasized to the student subjects in both sections. The present authors also believe that because there was only a 3 day interval between the time that students were presented with the practice exam and when students would be taking the actual final exam, then students would be already adopting a more serious disposition or mindset for the practice exam—i.e., would already be in 'final exam mode.'

While our results cannot be viewed definitively (this was not a true experimental design), we believe that the negative affect of the online environment on the online practice scores, coupled with the strongly positive affect of the online test-taking mode on the final exam score is strongly suggestive of student cheating. Said another way, in light of the practice test scores, it would be expected that the online group should score significantly below the in-class group on the final exam. However, the online group scored significantly above the in-class group on the final exam, thus suggesting a cheating effect quite separate from the environmental effect.

Discussion

The central question of this study is whether the use of online examinations, a common assessment tool in distance learning and other courses, is more susceptible to student cheating than traditional proctored exams. The argument has been made that this has become a critical question over the last 15 plus years as the popularity of distance learning has grown at an exponential pace at the same time that evidence suggests that the propensity of undergraduate students to exhibit dishonest behavior on graded assignments has reached very high levels and continues to grow. It has been further argued that despite the urgency of this question, there has been remarkably little research undertaken to shed light on this issue. Only a handful of studies have relied on statistical approaches other than surveys of students that ask them to

anonymously report on their own potentially illicit behaviors. Among that handful of studies which have attempted to apply statistical models, the attempt to detect student cheating on online versus proctored exams has been handicapped by the confounding relationships (1) between student performance and the possible differences in the opportunity to cheat in the two testing environments and (2) between student performance and possible differences in the two environments with respect to the level of distractions, student comfort, technical problems related to the use of computer technology and the opportunity to have the content of exam questions clarified. Leading up to the current study, only the Hollister and Berenson (2009) study even recognized this possible confounding of effects. However, the current study is the first effort to explicitly account for these potentially confounding effects in the design methodology.

The current study offers an experimental method designed to separate the effect of the testing environment (in-class versus online) on exam performance from the distinct effect of a possible greater propensity to cheat on unproctored, online exams on exam performance. This method is applied to 44 students in two sections of an introductory statistics class taking that course as part of the business curriculum at a university in the northeast. The evidence suggests that the negative effect of the online testing environment on performance, reflecting the possible disadvantages associated with greater ambient distractions, differences in student comfort, differences in technical problems and differences in the opportunity to solicit clarifications for possible ambiguous exam questions, is statistically significant. This negative effect of the online testing environment may be partially offsetting the measured positive and statistically significant effect on exam performance that may reasonably be attributed to student cheating. This may explain the lack of a significant cheating effect reported in studies that only compare online to in-class exam scores, in the absence of a measured environmental effect.

Of course, to the extent that these results showing that online testing does facilitate student cheating are representative, then the growing reliance of distance learning on online testing does suggest that professors and deans must take affirmative steps to suppress student cheating in those courses relying on online testing. Although previous research does not offer a clear set of findings suggesting that online testing facilitates cheating, the research presented here is consistent with such a view. It is hoped that additional research will help to clarify this question. The central issue related to our findings is that the existence of student cheating on online exams should not only be viewed in the context of the moral failings of students. Given the evidence that more and more professors in more colleges and universities are relying on and promoting distance learning, it may be argued that additional research findings similar to those presented here would implicitly impose a moral burden that would fall on professors and institutions of higher education to assure students, potential employers, graduate admissions departments and other consumers of grade information that the grades that students are receiving are truly reflective of the learning that has been accomplished.

References

Allen, I., & Seaman, J. (2010). *Class differences: online education in the United States, 2010*. Babson Park, MA: The Sloan Consortium, Babson College.

Brown, B., & McInerney, M. (2008). Changes in academic dishonesty among business students in the United States, 1999–2006. *International Journal of Management, 25*(4), 621–632.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Skokie, IL: Rand McNally.

Cluskey, G., Ehlen, C., & Raiborn, M. (2011). Thwarting online exam cheating without proctor supervision. *Journal of Academic and Business Ethics, 4*, 1–7.

Coalter, T., Lim, C. L., & Wanorie, T. (2007). Factors that influence faculty actions: a study of faculty responses to academic dishonesty. *International Journal for the Scholarship of Teaching and Learning, 1*(1), 1–19.

Crown, D., & Spiller, M. (1998). Learning from the literature on collegiate cheating: a review of empirical research. *Journal of Business Ethics, 17*(6), 683–700.

Durden, G., & Ellis, L. (2003). Is class attendance a proxy variable for student motivation in economics classes? An empirical analysis. *International Social Science Review, 78*(1 & 2), 42–46.

Grijalva, T., Nowell, C., & Kerkvliet, J. (2006). Academic honesty and online courses. *College Student Journal, 40*(1), 180–185.

Harding, T., Carpenter, D., Finelli, C., & Passow, H. (2004). Does academic dishonesty relate to unethical behavior in professional practice? An exploratory study. *Science and Engineering Ethics, 10*(2), 311–324.

Harmon, O., & Lambrinos, J. (2008). Are online exams an invitation to cheat? *Journal of Economic Education, 39*(2), 116–125.

Harper, M. (2006). High tech cheating. *Nurse Education Today, 26*, 672–679.

Hollister, K., & Berenson, M. (2009). Proctored versus unproctored online exams: studying the impact of exam environment on student performance. *Decision Sciences Journal of Innovative Education, 7*(1), 271–294.

Keith-Spiegel, P., Tabachnick, B., Whitley, B., & Washburn, J. (1998). Why professors ignore cheating: opinions of a national sample of psychology instructors. *Ethics & Behavior, 8*(3), 215–227.

Kennedy, K., Nowak, S., Raghuraman, R., Thomas, J., & Davis, S. (2000). Academic dishonesty and distance learning: student and faculty views. *College Student Journal, 34*(2), 309–314.

Kidwell, L., & Kent, J. (2008). Integrity at a distance: a study of academic misconduct among university students on and off campus. *Accounting Education: An International Journal, 17*(Supplement), S3–S16.

Kitahara, R., Westfall, F., & Mankelwicz, J. (2011). New, multifaced hybrid approaches to ensuring academic integrity. *Journal of Academic and Business Ethics, 3*, 1–12.

Kolowich, S. (2013). Behind the webcam's watchful eye, online proctoring takes hold. *Chronicle of Higher Education*. https://chronicle.com/article/Behind-the-Webcams-Watchful/138505/.

Lanier, M. (2006). Academic integrity and distance learning. *Journal of Criminal Justice Education, 17*(2), 244–261.

Lawson, R. (2004). Is classroom cheating related to business students' propensity to cheat in the "real world"? *Journal of Business Ethics, 49*(2), 189–199.

Lewis, L., Snow, K., Farris, E., & Levin, D. (1999). *Distance education at degree-granting education institutions: 1997–98*. Washington, D.C.: National Center for Education Statistics, U. S. Department of Education.

McCabe, D., Treviño, L., & Butterfield, K. (2001). Cheating in academic institutions: a decade of research. *Ethics & Behavior, 11*(3), 219–232.

Nonis, S., & Swift, C. (2001). An examination of the relationship between academic dishonesty and workplace dishonesty: a multicampus investigation. *The Journal of Education for Business, 77*(2), 69–77.

Okpala, A., Okpala, C., & Ellis, R. (2000). Academic efforts and study habits among students in a principles of macroeconomics course. *Journal of Education for Business, 75*(4), 219–224.

Parsad, B., & Lewis, L. (2008). *Distance education at degree-granting postsecondary institutions, 2006–07*. Washington, D.C.: National Center for Education Statistics, U.S. Department of Education.

Peng, Z. (2007). Giving online quizzes in corporate finance and investments for a better use of seat time. *The Journal of Educators Online, 4*(2), 1–18.

Porter, S. R. (2011). Do college student surveys have any validity? *The Review of Higher Education, 35*(1), 45–76.

Rodarte-Luna, B., Sherry, A., & Kim, S. H. (2006). Statistics anxiety, procrastination and the use of self-regulation. Presentation to the American Psychological Association.

Rogers, C. (2006). Faculty perceptions about e-cheating during online testing. *Journal of Computing Sciences in Colleges, 22*(2), 206–213.

Schneider, A. (1999). Why professors don't do more to stop students who cheat. *Chronicle of Higher Education, 45*(20), A8–A10.

Stuber-McEwen, D., Wiseley, P., Masters, C., Smith, A., & Mecum, M. (2005). *Faculty perceptions versus students' self-reported frequency of academic dishonesty*. Paper presented at the 25th Annual Meeting of the Association for Psychological & Educational Research in Kansas, Emporia, KS.

Stuber-McEwen, D., Wiseley, P., & Hoggat, S. (2009). Point, click, and cheat: frequency and type of academic dishonesty in the virtual classroom. *Online Journal of Distance Learning Administration, 12*(3). http://www.westga.edu/~distance/ojdla/fall123/stuber123.html.

Yates, R., & Beaudrie, B. (2009). The impact of online assessment on grades in community college distance education mathematics courses. *The American Journal of Distance Education, 23*(2), 62–70.