

A Robust Examination of Cheating on Unproctored Online Exams

Richard Fendler, David Beard and Jonathan Godbey

Department of Finance, Robinson College of Business, Georgia State University, USA

rfendler@gsu.edu (corresponding author)

dbeard@gsu.edu

jgodbey@gsu.edu

<https://doi.org/10.34190/ejel.22.5.3173>

An open access article under [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Abstract: The rapid growth of online education, especially since the pandemic, is presenting educators with numerous challenges. Chief among these is concern about academic dishonesty, especially on unproctored online exams. Students cheating on exams is not a new phenomenon. The topic has been discussed and debated within institutions of higher learning, and significant levels of cheating have been reported in the academic literature for over sixty years. Much of this literature, however, has focused on student behavior in a classroom utilizing proctored, in-class exams. Grades on exams usually determine most of a student's final grade in a course, and GPAs are used by employers and graduate schools to indicate a student's subject matter mastery. As more conventional colleges and universities expand their online course offerings it is natural to wonder if academic dishonesty is more prevalent in online classes than in face-to-face classes. In particular, are students more likely to cheat when no one is watching (i.e., on unproctored assessment assignments) than they do when someone is watching (i.e., on proctored assessment assignments)? The purpose of this study is to investigate whether students cheat more on unproctored online exams than they do on proctored in-classroom exams, and if so, is there any pattern to their cheating behavior. Our findings are derived from careful empirical analysis of 741 undergraduate students who completed three unproctored online exams, several collaboration-encouraged assignments, and a proctored in-class comprehensive final exam in the same course with the same instructor. Additionally, we collected demographic and human capital data for every student. Using bivariate and regression analysis, we find significant evidence of more cheating on unproctored online exams than on proctored in-class exams even though students were given stern honor code violation warnings. Moreover, we discover that student cheating increased with each unproctored online exam, implying that students learn how to cheat as they become more familiar with taking online assessment assignments. Finally, we find that students with certain demographic and human capital characteristics tend to cheat more than others. This research strongly supports the use of proctoring for all evaluation assignments in online classes to ensure that grades in these classes properly reflect student aptitude as opposed to merely reflecting their ability to cheat.

Keywords: Academic dishonesty, Cheating, Online education, Unproctored online exams, Proctored in-class exams

1. Introduction

Academic dishonesty, especially on exams which often determine the most significant part of a student's grade, is a major concern for higher education. When a student receives a grade that exaggerates their actual knowledge of subject matter or ability to critically think, employers who hire or graduate schools that admit the student soon realize they have been deceived. This realization not only tarnishes the reputation of the school, but it also imperils the core of the institution. Once a "cheating culture" is established, academic dishonesty flourishes (Tolman, 2017). This can lead to a perception by students that everyone is cheating and that the only way to get ahead is to cheat (Crittenden, Hanna and Peterson, 2009). Such malignant behavior must be recognized and addressed before it snowballs. Academic faculty, researchers and institutions need to study academic dishonesty not only to protect the reputation of their universities and the core of their institutions but also to protect honest students from believing that they must cheat to compete.

Much academic research (Simkin and McLeod, 2010; McCabe, Butterfield and Trevino, 2012) reports that student cheating on exams in college is rampant, and for some students cheating has become habitual (Clarke and Lancaster, 2006). Most of this research, however, has focused on traditional face-to-face (F2F) education where exams are proctored. The rapid growth in online education (Morris, et al., 2020; Guppy, et al., 2022) raises the question of whether students are more apt to cheat on unproctored online exams than on proctored in-class exams.

Interestingly, the literature concerning online exam cheating is mixed. Some of these studies use survey data (King, Guyette and Piotrowski, 2009; King and Case, 2014) where students are anonymously asked whether they would tend to cheat more on unproctored online exams or on proctored in-class exams. The most common answer to this question is "yes." Caution, however, should be exercised in strictly interpreting survey data concerning cheating. After all, surveys ask dishonest people (i.e., cheaters) to honestly report whether or not

they engaged in a dishonest activity. Since they are already dishonest, their answers may be less than forthright. Some students may fear being truthful, despite promises of anonymity; others may exaggerate their behaviors as a form of bragging, misusing the promise of anonymity as a protective shield. Indeed, information can be gleaned from student surveys about cheating, but the possibility of bias must be considered. Other studies (Harmon and Lambrinos, 2008; Beck, 2014; Dendir and Maxwell, 2020) use empirical techniques to examine unproctored versus proctored exam cheating behavior. Although many of these studies confirm that cheating on unproctored online exams is a problem, most suffer from poor “after-the-fact” design or small sample size.

The purpose of this study is to empirically investigate whether students are more likely to cheat on unproctored online exams than on proctored in-class exams and, if so, whether there is any pattern to their cheating behavior. This research is unique because the study was purposefully designed beforehand. All students in the study took the same online course taught by the same instructor. Students took multiple unproctored online exams in the class which allows observation of behavior through time. They also took a proctored in-class comprehensive final exam which allows for comparison between the two settings. Additionally, students completed several assignments (e.g., quizzes and problem sets) where they were encouraged to work in groups. This allows empirical comparison of cheating behavior on exams to assignments where collaboration was permitted. Finally, demographic and human capital data was collected for all participants to evaluate if some groups cheat more than others.

This research contributes to the literature by investigating whether the suspicions and anecdotal claims that online exams are more prone to cheating are supported by empirical evidence. It also offers a carefully designed methodology to examine online assessment assignment cheating that can be replicated by researchers in other settings. Finally, the dataset analyzed contains 741 observations, making it one of the most comprehensive studies to date devoted to examining this critically important topic.

2. Literature Review

Academic dishonesty among college students is well documented in the education literature. According to Hendershott, Drinan and Cross (2000) the practice is widespread. Clarke and Landcaster (2006) describe the cheating they find in their data as “habitual.” In a survey of 144 students, Simkin and McLeod (2010) report that 60% of business students and 64% of non-business students admitted that they cheated in college. In a study involving thousands of student surveys from 31 different schools conducted over decades, McCabe, Butterfield and Trevino (2012) report that more than 66% of college students claim that they engaged in academic dishonesty in the prior year.

Not only does the literature demonstrate that cheating is pervasive, but researchers have also found that cheating has become more prevalent over time. Vandehy, Diekhoff and LaBeff (2007) document that student academic dishonesty has linearly increased over the past 50 years. McCabe, Trevino and Butterfield (2001) find that, from 1964 to 1994, student cheating on exams increased from 39% to 64%.

Prior to COVID, online course offerings at colleges and universities were rapidly growing in popularity. Allen and Seaman (2008) document a 46% increase in college students taking at least one online course between 2002 to 2007. During this same period, the authors find that enrollment in online courses was growing five times faster than in traditional courses. Allen and Seaman (2013) report that 32% of higher education students in the U.S. were taking at least one online course in 2012.

Since the COVID pandemic, the number of online offerings at colleges and universities has exploded. The Digest of Education Statistics (2021) reports that in 2020, over 74% of all college students in the U.S. took at least one class online. Morris, et al. (2020) argue that online education will continue to grow in importance because university chief financial officers see e-education as a path to generate additional revenue as well as a means to widen access to traditionally underserved socio-economic groups. Guppy, et al. (2022) report that approximately 70 percent of the over 4500 college-level educators, students, administrators, and instructional designers surveyed envision fully online offerings will maintain a strong growth trajectory into the future.

This rapid growth in online education naturally raises the question of whether the amount of cheating that might occur on unproctored online assessment assignments is greater than, equal to, or less than the amount that occurs on proctored in-class assignments. On the one hand, there is the logical belief that if students are cheating when they are being watched, they will cheat even more when no one is looking. Moten, et al. (2013) argue that because exam performance in an online environment can be affected by accessing textbooks and notes, communicating with others via texting or group social networks, group test-taking, and impersonation, it is impossible to validate the results of unproctored online assessment assignments. Adzima (2020) notes that

students taking an unproctored online exam can and most likely will cheat and/or plagiarize from the Internet precisely because they are free from exam supervision. Another view is that cheating is just as likely to happen in a classroom as it is in an online environment. According to this view, exam design can be effectively used to minimize cheating on unproctored online exams. For example, password protecting exams (Rovai, 2000), randomly drawing exam questions from a pool of test questions (Shuey, 2002), and limiting the time for viewing/answering questions (Taylor, 2002) are anti-cheating techniques that can be used for online tests. Indeed, Beebe, Vonderwell and Boboc (2010) argue that educators who mechanically transfer assessment strategies they use in F2F classes to an online class are inviting students to cheat. Finally, some contend that online classes may be less prone to cheating due to the flexible schedule which mitigates “panic cheating” by students in F2F classes (Stuber-McEwen, Wiseley and Hoggatt, 2009).

Recent scholarly work reports various dimensions of academic dishonesty in unproctored online university exams. Noorbehbahani, Mohammadi and Aminazadeh (2022) systematically reviewed 58 studies on online test cheating published from 2010-2021. They categorized these articles into four themes: causal factors, cheating methods, detection techniques, and deterrence strategies. Key findings include that the leading publications are journal articles, cheating prevention and detection are the most studied themes, and the main cheating reasons involve the teacher, institution, student's internal motivations, and social/environmental factors.

Other empirical studies evaluate specific interventions aimed at reducing cheating rates. Golden and Kohlbeck (2020) tested whether paraphrasing test bank questions limits cheating compared to verbatim questions. Across undergraduate auditing classes, students performed significantly worse on paraphrased items, even with honor codes or proctoring. This implies students leveraged the internet to search answers for identical test bank content. The authors conclude that subtly modifying test banks questions can enhance academic integrity in unmonitored online exams.

Complementary experimental research by Vazquez, Chiang and Sarmiento-Barbieri (2021) examined impacts of proctoring on student achievement in university economics courses. Their findings reveal students scored over 11% higher on average when assessments were not monitored, regardless of using a live proctor versus automated software. While test anxiety or distraction at home may play a role, patterns imply collaboration amongst students represents the primary form of cheating.

Dench and Joyce (2022) conduct a field experiment to test whether informing students about the ability to detect plagiarism deters cheating in online college courses. Across four introductory business classes, students were randomly warned that software could identify copied work. Whereas warnings had little effect, notifying caught cheaters that they were now on a “watch list” dramatically reduced cheating rates by over 65% in subsequent assignments. Bilen and Matro (2021) present cases where students appeared to type out answers to complex questions in under 30 seconds. The authors interpret these results to indicate cheating will remain widespread in unproctored online assessments unless credible monitoring is implemented. Studies examining student academic dishonesty on unproctored online exams either use data derived from surveys or they attempt to extract evidence of cheating from bivariate and regression analysis of online assignment scores. Using surveys collected from 121 undergraduate business majors, King, Guyette and Piotrowski (2009) find that 73.6% of respondents believed that it is easier to cheat in an online course than to cheat in a F2F course. Watson and Sottile (2010), analyzing surveys collected from 635 undergraduate and graduate students, report that students disclosed that they engaged in academic dishonesty more in F2F classes than in online courses. King and Case (2014) collected survey data from 1817 students over a five-year period to examine possible changes in cheating behavior over time. The authors claim that, in 2013, 74% of students believed that cheating on an online exam was very easy or somewhat easy.

Though such evidence is compelling, the accuracy of anonymous survey responses has been questioned in the literature, particularly when the surveys relate to analyzing potentially sensitive behaviors. Psychologists call this phenomenon “social desirability bias” (Nederhof, 1985). Asking students if they cheated on an assessment assignment, something they are often told violates the school's honor code and could result in suspension, qualifies under this heading. Another concern about surveys is that the term “cheating” has flexible definitions (Wei, et al., 2014). For example, in a study exploring why students cheat, Perry (2010) reports that for first year undergraduates, only about one in four thought that copying word-for-word without citing the original author constituted plagiarism. If the definition of cheating is vague in students' minds, asking them if they cheated on a survey will produce vague results.

In response to these limitations, researchers are increasingly applying empirical techniques to analyze the relationship between grades on proctored versus unproctored assessment assignments to identify potential

cheating. D'Souza and Siegfeldt (2017) provide an excellent survey of this literature. The simplest approach is to compare differences between proctored and unproctored exam scores. If the difference is statistically significant for otherwise similar groups and assignments, cheating may be the cause. Boxplots and scatterplots can also be used to visualize similarities and differences between the distribution of grades in the different settings.

To more formally examine possible cheating behavior, Harmon and Lambrinos (2008) propose a regression methodology that is based on the coefficient of determination (the R^2 statistic). Beck (2014) and Dendir and Maxwell (2020) use this same model in their studies of online cheating. The procedure regresses exam grade on student characteristics that have been previously shown in the literature to be correlated with performance. Separate regressions are run for unproctored assignment grades and for proctored assignment grades. The proctored exam R^2 value shows the proportion of the variance in the dependent variable that is explained by the independent variables in the absence of cheating. Thus, it is the standard. If, for the same students, the R^2 value of an unproctored assignment regression is similar to that of the proctored exam regression, then cheating did not occur. However, if the R^2 value is significantly less for the unproctored exam, Harmon and Lambrinos (2008) argue that cheating is likely the explanation.

The authors applied this technique to two different summer sections of a principal of macroeconomics course. The two sections were entirely online and identical in every respect, except for the final exam. In one section ($N = 24$), the exam was in person and proctored. In the other section ($N = 38$), it was online and unproctored. The dependent variable in their model is exam grade and the independent variables are age, grade level, major and grade point average (GPA). The R^2 value for the proctored final exam was 0.497 and the average R^2 value of the unproctored final exam was 0.081. These results suggest that cheating took place in the online class.

Beck (2014) uses the Harmon and Lambrinos model to test for cheating on monitored ($N = 80$) versus unmonitored examinations ($N = 19$). The dependent variables are grades on the mid-term exam and the final exam, and the independent variables are GPA, credit hours and major. GPA is the only significant variable in all regressions. The R^2 values for the monitored mid-term exams are higher than those for the unmonitored mid-terms exams, but this relationship is reversed for the final exam. The author concludes that these contradictory findings reject the conclusion that there is more cheating on unproctored online exams than on proctored exams.

Both of these studies suffer from small sample size and a limited number of explanatory variables. Dendir and Maxwell (2020) apply the same technique to a much larger dataset (421 unproctored exams and 227 proctored exams). Additionally, the authors recommend the Goldfeld-Quandt test and the Chow test to formally examine if a difference in the R^2 values between the different settings is significant. Their results strongly suggest that excessive cheating occurs on unproctored exams.

This research study is very similar to Dendir and Maxwell (2020), but with several important distinctions that address limitations noted by these authors. First, Dendir and Maxwell (2020) examine student data from similar sections of two different courses (principles of microeconomics and geography of North America) taught over several semesters (Fall 2014 through Spring 2019). Our study examines student behavior in a single course over multiple semesters.

Second, in Dendir and Maxwell (2020), for approximately half of the time period, all exams (three midterm exams and a final exam) in both courses were unproctored, and then starting in Spring 2019 all exams were proctored. Thus, they compare the results of proctored exams for one time period to the results of unproctored exams for a different group of students from another time period. Although the authors claim that they believe student characteristics were similar between the periods, the two populations are not the same. In our study, we examine the behavior of the same student completing unproctored and proctored assignments in the same semester.

Third, Dendir and Maxwell (2020) use only two independent variables in their regressions: age and GPA. Our regressions have ten independent variables, all of which have been validated in the academic literature as determinants of student performance in similar courses.

Finally, in Dendir and Maxwell (2020) proctoring is achieved using Respondus Lockdown Browser plus Webcam. Although studies have shown that monitoring software can effectively reduce online cheating (Hylton, Levy and Dringus, 2016), there are concerns about student online exam performance being negatively impacted because they may feel nervous about being filmed (Butler-Henderson and Crawford, 2020). In our study, in-class proctoring (i.e., the "gold standard") is used.

Using the existing empirical research literature to create a carefully predesigned research approach with 741 observations, this study answers the following research questions:

RQ1: Does empirical evidence suggest that students are more likely to cheat on unproctored online exams than they are to cheat on carefully proctored in-class exams, where in both settings students are specifically told that cheating violates the school's honor code?

RQ2: Does cheating behavior increase as the semester progresses?

RQ3: Is there a relationship between student demographic or human capital characteristics and the propensity to cheat?

3. Study Design

The 15-week course used for this study was specifically designed to answer the research questions listed above. All aspects of the course were conducted online, except for the final exam. Although the course was asynchronous, it had a specific weekly schedule. For every week, students were assigned textbook chapters to read, video lectures to view, and graded assignments to complete. Any assignments not completed by midnight on Sunday of the due week received a grade of 0. The course had 15 quizzes (one per week), five takehome problem set (THPS) assignments (due at midnight on Sunday of weeks 3, 6, 9, 12 and 15), three unproctored online midterm exams (in weeks 5, 9 and 14), and an in-class proctored comprehensive final exam.

The course used for this study was an undergraduate principles of corporate finance course. Corporate finance is a math-oriented subject in which answers to questions used in assessment assignments are either right or wrong. In such a course, it is possible to create a very large databank of questions that can either be multiple choice with random ordering of answer choices or open-ended problems. The question databank for this course was created by the course instructor over several years and it contains over 600 distinct multiple-choice questions and over 200 open-ended problems. Approximately one-third of all the questions in the databank cover the course material associated with each midterm exam. All multiple-choice questions have 5 answer choices. For open ended problems, one of more numbers in the problem changes every time the question is used such that each iteration of the problem has a unique correct answer.

The unproctored online midterm exams in the class were given over a weekend. Students were allowed to take the exam during any 150-minute period while the exam was open. Each exam had 40 total questions, 30 were multiple choice and 10 were open ended problems. To minimize cheating opportunities, the instructor employed techniques suggested by Moten et al. (2013) and Flom, Green and Wallace (2023). Questions on exams were randomly selected from the corresponding section of the question databank, questions appeared one at a time, and students had to complete a question before moving to the next. The Learning Management System used for online exams includes a countdown timer that showed students how much of the 150-minute time limit remained. Once the counter hit zero, the exam was automatically closed. Prior to each exam, students were instructed to carefully monitor the timer to be sure to answer all questions before the exam closed.

The first question on each midterm exam, that students had to certify they agreed to, stated that academic dishonesty of any form was a violation of the university honor code and that if the instructor found any evidence of cheating, the student(s) involved would be fully prosecuted. Nonetheless, because the midterm exams were unproctored, there was no way of telling whether students were using their book or notes, looking up questions or solution processes on the internet, collaborating with other students, engaging in contract cheating, or other. No students were prosecuted for cheating during the study period.

The comprehensive final exam for the class was given to all students on the same day at the same time in a single large lecture hall. The final exam had 40 questions and a strict 150-minute time limit. Students were seated such that there was an empty chair between each student and multiple versions of the exam were distributed before students entered the room so that no students seated next to each other had the same version of the exam. Fendler, Godbey and Yates (2018) show that these techniques can significantly reduce student cheating on proctored exams. Student IDs were verified before students were allowed to enter the exam room, students had to put away cellphones, and multiple proctors actively roamed the room during the exam. Finally, students had to sign that they read and understood the university honor code before they were allowed to begin their exam. Other than being in a classroom, the substance (i.e., the final exam covered the same material as the midterm exams in approximately equal amounts) and format (i.e., time limit, number of questions, etc.) of the proctored final exam was designed to be as similar as possible to the three unproctored online midterm exams.

For weekly quizzes, students were allowed to retake these as many times as they wanted and only the highest grade counted. Additionally, students were encouraged to collaborate with others in the class if they encountered any questions or problems that they could not correctly answer on their own. Students were given two weeks to complete THPS assignments. They were also encouraged to work in groups on these assignments. At the end of the semester, an average quiz grade and an average THPS grade were computed for each student.

During the first week of class, all students took an online algebra math quiz. Students also took an online risk tolerance quiz (pfp.missouri.edu, n.d.) that scored their attitude about investment risk. To collect the necessary demographic data, students were asked to complete a pre-course and a post-course survey quiz. Those who properly completed both survey quizzes were given extra credit points towards their final course grade. Students who did not wish to participate were offered an alternative assignment to receive the same number of bonus points. The university provided student GPA and gender data.

The instructor in charge of the class received IRB approval to collect and use the data analyzed in this study. Once collected and merged, all student identifier data was removed to protect the identity of all participants.

Data was collected for students taking the course over six consecutive semesters (spring 2017, fall 2017, spring 2018, fall 2018, spring 2019 and fall 2019). For each of these semesters, all aspects of the course were conducted as described above. A total of 741 students completed all assignments in the classes, the two survey quizzes, the math quiz, and the risk tolerance quiz. T-tests were conducted on all variables for each semester to confirm that all semester samples were derived from the same population. Thus, compiling all data into one large dataset is appropriate.

4. Data

Descriptive statistics for the entire sample are provided in Table 1.

Table 1: Descriptive Statistics for All Sample Variables

Variable	Mean	S.D.	Min.	Max.	# Obs.
Proctored Assessment Assignment					
Comprehensive Final Exam	63.6%	19.6	20.0	100	741
Unproctored Assessment Assignments					
Exam 1	71.7%	16.7	23.0	100	741
Exam 2	75.9%	13.9	36.3	100	741
Exam 3	80.2%	10.8	46.5	100	741
Unproctored Collaboration Encouraged Assessment Assignments					
Quiz-Avg	95.1%	6.3	52.3	100	741
THPS-Avg	84.9%	11.5	19.6	100	741
Demographic and Personal Characteristic Variables					
Math Quiz	66.4%	21.7	6.0	100	741
Age	24.0	4.9	18.0	61.0	741
GPA	3.18	0.43	2.0	4.19	741
TrmHrs	12.8	3.5	3.0	24.0	741
Gender (F = 1)	56.3%	-	0	1	741
Major (Fin/Acct = 1)	33.9%	4.96	0	1	741
Risk Score	27.2	5.3	13.0	50.0	741
MathAnx(HiAnx = 5)	3.1	1.4	1	5	741
#Online	3.5	2.2	0	14	741
Job-SptsHrs	23.0	14.1	0	40.0	741

In addition to the variables described in the prior section:

- TrmHrs is the total number of course hours that the student was taking in the same semester that they took this course.

- Major indicates whether a student is a finance or accounting major (basically, “sister” subjects) or some other major.
- MathAnx is a Likert scale variable for the student answer to the question: “Which of the following BEST describes how much you agree (or disagree) with the following statement: I get nervous and/or anxious when taking a class that covers or uses math.” For this variable a value of 1 indicates the student strongly disagrees with the statement and a value of 5 indicates strong agreement with the statement.
- #Online is the total number of online classes that the student took prior to this class.
- Job-SptsHrs are the self-reported number of hours per week that the student spends working in a job and/or participating in a university sports activity.

5. Results

5.1 Bivariate Analysis

The first step in examining whether the data indicates cheating may have occurred on the unproctored exams in the class is to compare the descriptive statistics of the grade on each unproctored exam with the proctored exam grade. The proctored final exam in the class was comprehensive and similarly structured to the midterm exams in the number of questions and degree of difficulty. If the grades on each successive exam, including the final exam, improved, then the data may be reflecting student learning, a highly desired outcome. If, however, the score on the proctored final exam is significantly lower than the scores on the unproctored midterm exams, cheating is a possible explanation.

Table 2 shows the mean and standard deviation of each of the unproctored assessment assignments compared to the same statistics for the proctored exam. The last two columns in Table 2 list the results of a two-sample t-test for the equality of means between each of the unproctored vs. proctored exam pairings.

Table 2: Mean Comparisons Between Unproctored and Proctored Exams

Assessment Assignment	N	Mean	St. Dev.	T	P > t
Unproctored Exam 1	741	71.7	16.7	11.21	0.00
Proctored Final Exam	741	63.6	19.6		
Unproctored Exam 2	741	75.9	13.9	16.01	0.00
Proctored Final Exam	741	63.6	19.6		
Unproctored Exam 3	741	80.2	10.8	23.27	0.00
Proctored Final Exam	741	63.6	19.6		

The average grades on all of the unproctored exams are significantly higher than the average grade on the proctored exam. A possible non-cheating explanation for a lower final exam average is that students fear final exams, especially in a classroom with proctors, more than they do online exams. This fear could cause students to perform poorer on the final exam. However, although course subject matter tends to increase in difficulty as the course progresses, the average grade on the unproctored exams increases with each successive exam. Additionally, the standard deviation of grades is less for the unproctored exams versus the proctored exams, and the standard deviation of grades decreases with each successive unproctored exam. These relationships are consistent with the conclusion that cheating did occur on the unproctored exams, and that the degree of cheating may have increased with each successive unproctored event.

Figure 1 shows boxplot comparisons between each unproctored exam and the proctored final exam. The median of each unproctored exam is greater than the median of the proctored exam, and both the range and dispersion of the unproctored exams are less than those of the proctored exam. Additionally, the range decreases, the median grade increases and the dispersion of grades decreases with each successive unproctored exam taken. These relationships correspond with the trends noted above.

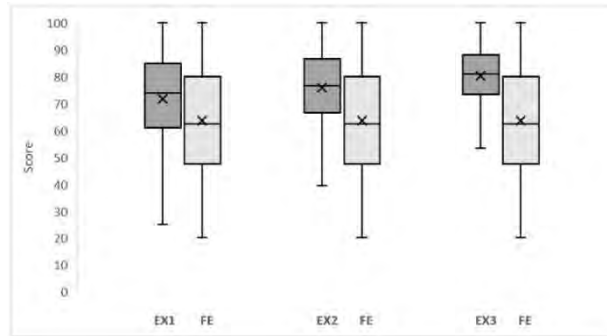


Figure 1: Boxplot Comparisons of Unproctored and Proctored Exams

Scatterplots showing the relationship for all observations between each unproctored exam and the proctored exam are shown in Figure 2. As one would expect there is a large number of plots in the upper, right hand corner (students who did well on the midterm exams and had equivalent performance on the final exam). Though a trend exists, the correlation coefficients for Exam 2 and Exam 3 relative to Final Exam are very low (0.2488 and 0.2940, respectively). However, the curious take away from the scatterplots is that the data points become more compact with each successive unproctored exam, with a large number of plots in the upper left quadrant of each graph (students who did very well on the unproctored midterm exams and performed very poorly on the proctored final exam). This pattern is consistent with students cheating on the unproctored exams. Although cheaters achieved high grades on the midterm exams, because they did not actually learn the course material, these students performed very poorly on the carefully proctored comprehensive final exam.

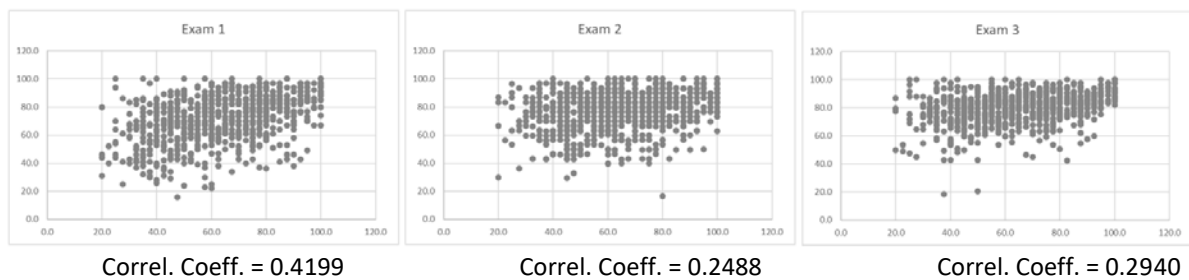


Figure 2: Scatterplots for Unproctored and Proctored Exams

5.2 Regression Analysis: Comparing Unproctored to Proctored Assessment Assignments

A more formal analysis of the data involves regressing exams scores on appropriate independent variables. All the independent variables used in this study (i.e., the Demographic and Personal Characteristic Variables listed in Table 1) have been validated in the academic literature as factors that impact exam performance in a finance course. Borde, Byrd and Modani (1998) report that age, GPA, gender, declared major, number of course hours, and outside distractions (i.e., work or sports hours) influence performance in finance courses. Ross and Wright (2020) find that quantitative skill is a significant determinant of success in finance. We use two similar, yet unique, variables to capture quantitative ability: score on the pre-course algebra quiz and the answer to a survey question that quantifies a student’s level of anxiety towards math. The algebra math quiz that students took at the beginning of each semester did not count towards the final grade in the course, so it was a “no pressure” assignment. The MathAnx variable is the same measure used by Pekrun, et al. (2017) who find that negative emotions about math (i.e., a high MathAnx value) significantly negatively impact student performance on test scores in math-based subjects. Fendler and Bredthauer (2016) show that the number of previously taken online courses is an additional important determinant of grades for students taking an online finance course. Sarmiento and Manaloto (2018) find that students with a high level of risk tolerance are more likely to cheat. We surmise that if a student is more likely to cheat, their grade on unproctored exams will be higher, but on a comprehensive proctored final exam it will be lower (because they did not actually learn the course material).

The specific regression model that we estimate is:

$$\text{Grade}_i = a_{0,i} + a_{1,i}(\text{Math Quiz}) + a_{2,i}(\text{Age}) + a_{3,i}(\text{GPA}) + a_{4,i}(\text{TrmHrs}) + a_{5,i}(\text{Gender}) + a_{6,i}(\text{Major}) + a_{7,i}(\text{Risk Score}) + a_{8,i}(\text{MathAnx}) + a_{9,i}(\text{\#Online}) + a_{10,i}(\text{Job-SptsHrs}) + \epsilon_i$$

where $Grade_i$ is score on an assessment assignment and ϵ_i is the regression error.

Regression coefficients for the Proctored Final Exam, Unproctored Exam 1, Unproctored Exam 2, Unproctored Exam 3, collaboration allowed Average Quiz and collaboration allowed Average THPS models are presented in Table 3.

Table 3: Regressions for Proctored ($Grade_p$) and Unproctored ($Grade_{up}$) Exams

Variable	Proctored	Unproctored			Collaboration Allowed	
	Final Exam	Exam 1	Exam 2	Exam 3	Quiz-Avg	THPS-Avg
Constant	35.12*** (4.666)	28.18*** (3.892)	32.12*** (5.133)	57.27*** (11.561)	79.86*** (27.859)	71.33*** (13.048)
Math Quiz	0.14*** (5.027)	0.08*** (2.823)	0.06*** (2.575)	0.03 (1.593)	0.00 (0.411)	0.00 (0.227)
Age	0.01 (0.100)	-0.01 (0.034)	-0.09 (0.881)	-0.08 (0.984)	0.10*** (2.096)	-0.11 (1.254)
GPA	16.96*** (11.852)	12.32*** (8.943)	10.53*** (8.851)	5.30*** (5.628)	4.24*** (7.771)	5.67*** (5.451)
TrmHrs	-0.83*** (4.803)	-0.25 (1.512)	-0.17 (1.178)	-0.21* (1.881)	-0.05 (0.696)	-0.12 (0.933)
Gender (F=1)	-5.20*** (4.492)	-0.18 (0.162)	-1.66* (1.725)	1.23 (1.619)	0.14 (0.326)	0.20 (0.233)
Fin-Acc=1	4.07*** (3.289)	2.92*** (2.454)	-1.02 (0.991)	1.80*** (2.213)	1.20*** (2.555)	1.26 (1.397)
RiskScore	-0.64*** (5.858)	0.30*** (2.846)	0.38*** (4.209)	0.32*** (4.397)	0.03 (0.790)	0.11 (1.373)
MathAnx (HiAnx=5)	-1.31*** (3.015)	-0.71* (1.686)	0.42 (1.153)	-0.35 (1.223)	-0.28* (1.706)	-0.40 (1.266)
#ONLINE	0.37 (1.416)	0.19 (0.736)	0.56*** (2.572)	0.63*** (3.632)	0.13 (1.229)	0.02 (0.102)
JOB-SPTS- HRS	-0.13*** (2.965)	-0.21*** (5.098)	-0.07*** (2.031)	-0.10*** (3.371)	-0.05*** (3.296)	-0.10*** (3.077)
R square	0.4001	0.2311	0.1707	0.1407	0.1481	0.0889
F ratio	48.70	21.94	15.03	11.95	12.69	7.122
N	741	741	741	741	741	741

Note. |t| ratios are shown in parentheses below each parameter estimate; *** shows significance at $p < 0.01$; ** shows significance at $p < 0.05$; * shows significance at $p < 0.10$.

As shown in Table 3, the R^2 value for the Proctored Final Exam is 0.4001, for Unproctored Exam 1 is 0.2311, for Unproctored Exam 2 is 0.1707, and for Unproctored Exam 3 is 0.1407. And the R^2 value for the collaboration allowed quiz and THPS assignments are 0.1481 and 0.0889, respectively.

To examine whether the R^2 values for these regression equation differences are significant, Dendir and Maxwell (2020) propose using the Goldfeld-Quandt (GQ) test. The GQ test checks for heteroscedasticity in regression analysis by comparing the variance of the error term for two regressions to check whether they both derive from the same population. For this study, if the GQ test suggests that heteroscedasticity is present between the proctored exam regression and an unproctored exam regression, then cheating most likely occurred on the unproctored exam.

The Goldfeld-Quandt F-ratio statistic for each exam pairing is shown in Table 4. The GQ test indicates that the error variances are significantly different between the proctored Final Exam and the unproctored Exam 2

regressions and between the proctored Final Exam and the unproctored Exam 3 regressions, suggesting that cheating occurred on both of these unproctored exams. Additionally, the larger F-ratio statistic for Exam 3 versus Exam 2 suggests that the amount of cheating increased as students took additional exams in the class.

Table 4: Formal Tests of Unproctored versus Proctored Exam Regressions

	Unproctored Exam 1 versus Proctored Final Exam	Unproctored Exam 2 versus Proctored Final Exam	Unproctored Exam 3 versus Proctored Final Exam
Goldfeld-Quandt Test F(730, 730)	1.08	1.45***	2.31***
Chow Test F(10,1472)	19.33***	84.25***	392.13***
Note. *** indicates significance at $p < 0.01$.			

Dendir and Maxwell (2020) also suggest using a Chow test. The Chow test can be used to determine whether one regression line or two separate regression lines best fit a split set of data. For this study, if the Chow F-statistic is significant, then the estimated coefficients in the unproctored exam regression, as a group, are statistically different from those in the proctored exam regression. Cheating is a plausible explanation for this difference. The Chow test statistics shown in Table 4 are highly statistically significant ($p < .01$) for each unproctored exam, and the statistic increases with each successive exam. These results support the conclusion that cheating occurred on the unproctored exams and that the degree of cheating increased as students took additional exams in the class.

5.3 Regression Analysis: Comparison of Significant Parameters

Trends in the size and significance levels of the coefficients in the Table 3 regressions lend further support to a conclusion that the amount of cheating most likely increased as the semester progressed. For nearly every significant variable, the size of the regression coefficients decreased with each successive exam. And for most of these, the impact of the Exam 3 regression on grade is more closely related to the “collaboration allowed” assessments assignments than to the proctored Final Exam.

Similarly, for GPA the $|t|$ ratio follows the same pattern. The $|t|$ ratio for GPA decreases from 11.852 for the Final Exam to 8.943 for Exam 1, 8.851 for Exam 2, and 5.628 for Exam 3. In fact, the impact of GPA on grade for the “collaboration allowed” THPS assignments is essentially the same as the impact of GPA on Exam 3.

The most interesting coefficient relationship trends, however, are for #Online, Math Quiz, MathAnx and RiskScore. #Online is insignificant for the proctored Final Exam, but highly positively significant for Exam 2 and Exam 3. This trend implies that students with more experience taking online classes may be more likely to cheat on exams. For this class, the warning against cheating may have prevented these students from cheating on Exam 1, but when they discovered that the warning was mostly hollow, they were more emboldened to cheat on Exam 2 and Exam 3.

Whereas both Math Quiz and MathAnx are, as expected, highly significant for the proctored exam equation, they are insignificant for unproctored Exam 3. One possible reason for this change might be that students with poorer math skills and those with math phobia felt more compelled, or perhaps justified, to cheat by the third unproctored exam in the class. As these students experienced the importance of math in finance, they may have chosen to cheat to overcome their weakness(es) in this area.

The trend in RiskScore is perhaps the most compelling evidence that cheating occurred on the unproctored exams in the class. As previously noted, Sarmiento and Manaloto (2018) report that students with a high degree of risk tolerance are more apt to cheat. Thus, on unproctored exams, these students would be expected to get a higher grade. But because this grade derives from cheating instead of from learning, when they are tested on this same material on a comprehensive final exam where they cannot cheat, they perform significantly worse. In fact, that is exactly what the regressions in Table 3 show. Whereas RiskScore is significantly positively related to grade on all three unproctored exams, it is significantly negatively related to the grade on the proctored comprehensive Final Exam.

5.4 Answers to Research Questions

The bivariate and regression analysis presented above provides answers to the three research questions. Specifically, students are more likely to cheat on unproctored online exams than they are to cheat on carefully proctored in-class exams, where in both settings students are specifically told that cheating violates the school's

honor code. As experience taking unproctored online exams increases, the likelihood of student cheating also increases. In fact, the degree of cheating on the third unproctored exam in the class was so extensive that the relationship between grade and mastery of subject matter was essentially the same as what would be expected for a collaboration encouraged assignment. Finally, students with more experience taking online classes seem to cheat more, students who are weak or believe they are weak in a topic necessary to succeed in the class tend to cheat more, and students who considered themselves to be risktakers tend to cheat more.

6. Discussion and Conclusion

Academic dishonesty, especially on exams which are used in most courses for assessment, is a major concern for higher education. Students who cheat on exams often receive a grade for a class that does not reflect their knowledge or understanding. Improper evaluation of student ability can tarnish a school's reputation. Even more concerning, left unchecked, cheating can become established as part of an institution's culture, causing honest students to consider cheating as necessary to maintain their position. The rapid growth in online education, especially since COVID, raises the question of whether students are more apt to cheat on unproctored online exams than they do on proctored in-class exams.

Finding the answer to this question is important because as experience with e-education grows, both faculty and students are discovering significant benefits with online testing. Students prefer taking online exams because they like the immediate feedback, they are confident that computer grading is more equitable and transparent than instructor hand grading, and they prefer the flexibility of being able to take an exam when it is best for them as opposed to when the course is scheduled by the university (Baleni, 2015). Indeed, in a survey of 220 students, Baleni (2015) reports that an overwhelming majority (83%) of respondents indicated a strong preference for taking online exams, with "only a few students declaring a preference for traditional assessments" (p 232). Dreher, Reiners and Dreher (2011) report that faculty also like giving online exams because they eliminate time spent printing, distributing and collecting paper exams, and they eliminate the tedium of grading the same question numerous times which frees faculty to provide more meaningful feedback to students.

The purpose of this study is to investigate whether students cheat more on unproctored online exams than on proctored in-class exams in order to help educators protect their institutions as well as to defend honest students. Data was collected from a 15-week long, online undergraduate course that was specifically designed for this project. The dataset consists of student grades on three unproctored mid-term exams (given in weeks 5, 9 and 14), several collaboration-encouraged assessments assignments, a proctored in-class comprehensive final exam, human capital data, and demographic information. The study ran for 6 consecutive semesters from spring 2017 to fall 2019. Course design, delivery, assignments, directions given, and instructor were consistent across all semesters. The final dataset includes 741 observations.

To determine whether student cheating behavior differs between exam settings (i.e., unproctored versus proctored), we estimate regressions using the grade on each of the three unproctored exams (where cheating is possible), the grade on the class collaboration-encouraged assignments (where "cheating" is expected to be observed), and the grade on the proctored final exam (where cheating is not highly unlikely due to exam design) as the dependent variable. We use independent variables of student characteristics that are identified in the literature as affecting performance. As discussed in Harmon and Lambrinos (2008), comparing the R^2 values of these regressions can provide insights into academic dishonesty. The R^2 value of the proctored exam is considered to be the standard for acceptable honest behavior. The R^2 values of the collaboration-encouraged assignments should reflect open cooperation (i.e., the expected value when cheating is rampant).

We find a significant amount of cheating on each of the unproctored online exams. The R^2 values for each of these regression equations is significantly lower than the model for the proctored final exam. We also find that as the semester progresses, online students learn how to cheat more effectively. The R^2 value for each successive exam decreases. In fact, by the third unproctored online exam in the class, the R^2 value is essentially the same as the R^2 value for the collaboration-encouraged assessment assignments.

This study offers several contributions for academics in higher education. First, we provide compelling empirical evidence from a purposefully designed study showing that students will cheat more on unproctored online exams compared to proctored in-class exams. This confirms suspicions and anecdotal claims that online exams are more prone to cheating. Second, this study demonstrates through quantitative analysis that student cheating actually increases on successive unproctored online exams, implying that students learn how to cheat more effectively as they take more online classes. Thus, cheating is a progressive threat, not a static one. Third,

we identify specific demographics and human capital factors, such as more online experience, poorer math skills, and higher risk tolerance, to be associated with more cheating. These are concrete student characteristics possibly associated with cheating of which academics should be aware. Fourth, our study makes a forceful case, through a dataset of over 700 students, that all online assessments should be proctored to safeguard integrity. Finally, we offer a model methodology for other researchers to copy and hopefully improve. In summary, this study makes both an empirical and methodological contribution to the literature, providing compelling evidence to justify proctored online exams and a framework for further research on this important topic.

Limitations of this study that represent areas for future research include the following. The proctoring used for this study was in-class, human proctors. Requiring online students to come to a classroom to take an exam may be impractical. The study should be replicated with software-proctored exams. This study examines student behavior in a single course (principles of corporate finance). Analyzing multiple courses could reveal whether our findings generalize across disciplines. Some of the variables used in the study are self-reported (e.g., job/sports hours). More objective measures may improve reliability. Additional student characteristics, such as family income level, citizenship status, or instructor rating, could be incorporated into the study. Finally, the emergence of other technical tools, such as ChatGPT, and their impact on student cheating need to be carefully studied.

Overall, the accumulating evidence makes clear that without credible oversight, cheating is likely to remain extensive on unproctored online university exams. Multiple detection approaches and deterrence strategies show promise in helping to ensure integrity. Further research is warranted to better understand student behavior in digital environments as well as to determine how to best evaluate student learning in this evolving academic arena.

Declarations of interest: none

References

- Adzima, K., 2020. Examining online cheating in higher education using traditional classroom cheating as a guide. *Electronic Journal of E-Learning*, 18(6), pp.476-493.
- Allen, I.E. and Seaman, J., 2008. Staying the course: Online education in the United States, 2008. Sloan Consortium. PO Box 1238, Newburyport, MA 01950. Retrieved from <https://www.bayviewanalytics.com/reports/staying-the-course.pdf>
- Allen, I.E. and Seaman, J., 2013. Changing course: Ten years of tracking online education in the United States. Sloan Consortium. PO Box 1238, Newburyport, MA 01950. Retrieved from <http://www.onlinelearningSurvey.com/reports/changingcourse.pdf>
- Baleni, Z.G., 2015. Online formative assessment in higher education: Its pros and cons. *Electronic Journal of e-Learning*, 13(4), pp.228-236.
- Beck, V., 2014. Testing a model to predict online cheating—Much ado about nothing. *Active learning in higher education*, 15(1), pp.65-75.
- Beebe, R., Vonderwell, S. and Boboc, M., 2010. Emerging patterns in transferring assessment practices from f2f to online environments. *Electronic Journal of e-learning*, 8(1), pp.1-12.
- Bilen, E. and Matros, A., 2021. Online cheating amid COVID-19. *Journal of Economic Behavior & Organization*, 182, pp.196-211.
- Borde, S.F., Byrd, A.K. and Modani, N.K., 1998. Determinants of student performance in introductory corporate finance courses. *Journal of Financial Education*, pp.23-30.
- Butler-Henderson, K. and Crawford, J., 2020. A systematic review of online examinations: A pedagogical innovation for scalable authentication and integrity. *Computers & Education*, 159, p.104024.
- Clarke, R. and Lancaster, T., 2006, June. Eliminating the successor to plagiarism? Identifying the usage of contract cheating sites. In *Proceedings of 2nd international plagiarism conference* (pp. 19-21). Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.120.5440&rep=rep1&type=pdf>
- Crittenden, V.L., Hanna, R.C. and Peterson, R.A., 2009. The cheating culture: A global societal phenomenon. *Business Horizons*, 52(4), pp.337-346.
- Dench, D. and Joyce, T., 2022. Information and credible sanctions in curbing online cheating among undergraduates: a field experiment. *Journal of Economic Behavior & Organization*, 195, pp.408-427.
- Dendir, S. and Maxwell, R.S., 2020. Cheating in online courses: Evidence from online proctoring. *Computers in Human Behavior Reports*, 2, p.100033.
- Digest of Education Statistics, 2021. National Center for Education Statistics (NCES) Home Page, a part of the U.S. Department of Education. (n.d.). Retrieved from <https://nces.ed.gov/programs/digest/d21/>
- Dreher, C., Reiners, T. and Dreher, H., 2011. Investigating factors affecting the uptake of automated assessment technology. *Journal of Information Technology Education: Research*, 10(1), pp.161-181.
- D'Souza, K.A. and Siegfeldt, D.V., 2017. A conceptual framework for detecting cheating in online and take-home exams. *Decision Sciences Journal of Innovative Education*, 15(4), pp.370-391.

- Fendler, R. J. and Bredthauer, J. (2016). Predictors of success in an online undergraduate core course in finance. *Journal of Economics and Finance Education*, 2016, Vol. 15, No. 1, pp.101-111.
- Fendler, R.J., Godbey, J.M. and Yates, M.C., 2018. Observing and Deterring Social Cheating on College Exams. *International Journal for the Scholarship of Teaching and Learning*, 12(1), p.4.
- Flom, J., Green, K. and Wallace, S., 2023. To cheat or not to cheat? An investigation into the ethical behaviors of generation Z. *Active Learning in Higher Education*, 24(2), pp.155-168.
- Golden, J. and Kohlbeck, M., 2020. Addressing cheating when using test bank questions in online classes. *Journal of Accounting Education*, 52, p.100671.
- Guppy, N., Verpoorten, D., Boud, D., Lin, L., Tai, J. and Bartolic, S., 2022. The post-COVID-19 future of digital learning in higher education: Views from educators, students, and other professionals in six countries. *British Journal of Educational Technology*, 53(6), pp.1750-1765.
- Harmon, O.R. and Lambrinos, J., 2008. Are online exams an invitation to cheat? *The Journal of Economic Education*, 39(2), pp.116-125.
- Hendershott, A., Drinan, P. and Cross, M., 2000. Toward enhancing a culture of academic integrity. *NASPA journal*, 37(4), pp.587-598.
- Hylton, K., Levy, Y. and Dringus, L.P., 2016. Utilizing webcam-based proctoring to deter misconduct in online exams. *Computers & Education*, 92, pp.53-63.
- King, C.G., Guyette Jr, R.W. and Piotrowski, C., 2009. Online exams and cheating: An empirical analysis of business students' views. *Journal of Educators Online*, 6(1), p.n1.
- King, D.L. and Case, C.J., 2014. E-cheating: Incidence and trends among college students. *Issues in Information Systems*, 15(1).
- McCabe, D.L., Butterfield, K.D. and Trevino, L.K., 2012. *Cheating in college: Why students do it and what educators can do about it*. JHU Press.
- McCabe, D.L., Treviño, L.K. and Butterfield, K.D., 2001. Cheating in academic institutions: A decade of research. *Ethics & Behavior*, 11(3), pp.219-232.
- Morris, N.P., Ivancheva, M., Coop, T., Mogliacci, R. and Swinnerton, B., 2020. Negotiating growth of online education in higher education. *International Journal of Educational Technology in Higher Education*, 17, pp.1-16.
- Moten Jr, J., Fitterer, A., Brazier, E., Leonard, J. and Brown, A., 2013. Examining online college cyber cheating methods and prevention measures. *Electronic Journal of E-learning*, 11(2), pp.139-146.
- Nederhof, A.J., 1985. Methods of coping with social desirability bias: A review. *European journal of social psychology*, 15(3), pp.263-280.
- Noorbebahani, F., Mohammadi, A. and Aminazadeh, M., 2022. A systematic review of research on cheating in online exams from 2010 to 2021. *Education and Information Technologies*, 27(6), pp.8413-8460.
- Pekrun, R., Lichtenfeld, S., Marsh, H.W., Murayama, K. and Goetz, T., 2017. Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child development*, 88(5), pp.1653-1670.
- Perry, B., 2010. Exploring academic misconduct: Some insights into student behaviour. *Active Learning in Higher Education*, 11(2), pp.97-108.
- pfp.missouri.edu. (n.d.). Investment Risk Tolerance Assessment // Personal Financial Planning. [online] Available at: <https://pfp.missouri.edu/research/investment-risk-tolerance-assessment/>.
- Ross, M.M. and Wright, A.M., 2020. Quantitative Skill and Introductory Finance. *Journal of Financial Education*, 46(2), pp.193-220.
- Rovai, A.P., 2000. Online and traditional assessments: what is the difference? *The Internet and higher education*, 3(3), pp.141-151.
- Sarmiento, P.J.D. and Manaloto, A.C., 2018. The Perception of Faculty and College Students on Cheating: Envisioning a Cheat-Free Academic Community. *International Journal of Learning in Higher Education*, 25(3).
- Shuey, S., 2002. Assessing online learning in higher education. *Journal of Instruction Delivery Systems*, 16(2), pp.13-18.
- Simkin, M.G. and McLeod, A., 2010. Why do college students cheat? *Journal of business ethics*, 94, pp.441-453.
- Stuber-McEwen, D., Wiseley, P. and Hoggatt, S., 2009. Point, click, and cheat: Frequency and type of academic dishonesty in the virtual classroom. *Online Journal of Distance Learning Administration*, 12(3), pp.1-10.
- Taylor, S., 2002. Education Online: Off Course or On Track? *Community College Week* 14(20), pp.10-12.
- Tolman, S., 2017. Academic dishonesty in online courses: Considerations for graduate preparatory programs in higher education. *College Student Journal*, 51(4), pp.579-584.
- Vandehey, M., Diekhoff, G. and LaBeff, E., 2007. College cheating: A twenty-year follow-up and the addition of an honor code. *Journal of College Student Development*, 48(4), pp.468-480.
- Vazquez, J.J., Chiang, E.P. and Sarmiento-Barbieri, I., 2021. Can we stay one step ahead of cheaters? A field experiment in proctoring online open book exams. *Journal of Behavioral and Experimental Economics*, 90, p.101653.
- Watson, G. and Sottile, J. (2010). Cheating in the digital age: Do students cheat more in online courses? *Online Journal of Distance Learning Administration*, 13(1). Retrieved from https://mds.marshall.edu/cgi/viewcontent.cgi?article=1000&context=eft_faculty
- Wei, T., Chesnut, S.R., Barnard-Brak, L. and Schmidt, M., 2014. University students' perceptions of academic cheating: Triangulating quantitative and qualitative findings. *Journal of Academic Ethics*, 12, pp.287-298.