# Observing and Deterring Social Cheating on College Exams

Richard J. Fendler, Michael C. Yates, Johnathan M. Godbey

*College of Business, Georgia State University, Atlanta, GA 30303, USA*

(Received 19 March 2017; Accepted 23 October 2017)

This research introduces a unique multiple choice exam design to observe and measure the degree to which students copy answers from their peers. Using data collected from the exam, an empirical experiment is conducted to determine whether random seat assignment deters cheating relative to a control group of students allowed to choose their seats. Empirical results demonstrate a significant decline in measured cheating within the assigned seating sample. This study contributes to the literature by providing a measurement of actual cheating frequency among students, as opposed to relying on reported cheating in anonymous surveys, and by demonstrating that an easily implemented deterrent can significantly reduce instances of cheating.

## INTRODUCTION

Numerous academic studies indicate that a significant number of students cheat on exams in college. Research presented in Eve and Bromley (1981), Haines, Diekhoff, LaBeff and Clark (1986), Genereux and McLeod (1995), Diekhoff, LaBeff, Clark, Williams, Francis and Haines (1996) and McCabe, Butterfield and Trevino (2006) report that between 23 to 58 percent of student respondents admitted to cheating (mostly copying from, or sharing answers with, another student) on one or more major examinations during the prior academic year. Jordan (2001) found that in a particular semester, over 31 percent of students claimed to have cheated on at least one exam. And, when the definition of academic dishonesty is broadened to include activities such as plagiarism, working in groups on take home assignments without the professor's permission, accessing textbook test banks and buying research papers written by someone else, over half of the respondents in these studies self-confessed to cheating, often on multiple occasions. Indeed, Selingo (2004) and Clarke and Lancaster (2006) describe the types of cheating that they observe in their data sets as "habitual."

Cheating, especially on exams which form the basis of reporting academic success in most college courses, undermines the meaning of grades as a measure of subject matter proficiency. If, by cheating, an individual can earn the same grade in a class as a student who actually learns and masters the subject matter, course grades, and therefore GPAs, lose their meaning. Employers who hire graduates who cheat their way through college are duped into believing they are hiring, and paying a collegiate-level salary to, an individual who may in fact be unqualified for the job. Additionally, academic research shows that individuals who cheat in college are more likely to cheat at work, cheat on their spouses, and partake in other unethical activities (Novotney, 2011). With studies indicating that one-third or more of college students cheat on exams, cheating is a major problem that must be properly evaluated and appropriately addressed by colleges and universities.

However, before academics can successfully address the issue of cheating in college, we must first overcome the problematic issue that cheating is generally not observed. Most of what we know about student cheating in college comes from survey data. Obviously, there is reason to believe that survey data might be wrong (for example, see Hessing, Elffers & Weigel, 1988). Some respondents might be afraid to admit they cheat and others might brag. After all, we are asking cheaters (i.e., dishonest people) to honestly tell us whether or not they are dishonest!

In this study, we measure the amount of cheating that occurs during a proctored final exam for an undergraduate core course in corporate finance. Our study is unique because our findings are not based on survey data. Instead, using an exam design technique described in Fendler and Godbey (2015), we are able to empirically observe actual cheating behavior. And, because students have no idea that their possible cheating behavior is being "observed," our measure is highly accurate and unbiased relative to studies based on survey data. Furthermore, our study design allows us to empirically test the efficacy of a simple and easily implemented cheating deterrent, a randomly assigned seating chart. While we would expect some amount of cheating to occur on any exam, there is reason to believe that preventing students from sitting next to their companions might reduce social cheating behavior.

## LITERATURE REVIEW

Cheating can be measured either through surveys or by observation. Surveys are the most commonly used measure. Students are asked to voluntarily participate in a study where they answer a series of questions about academic dishonesty. Demographic and other activity data is also often collected. Observation is much more difficult because if students know they are being observed (e.g., cameras in the classroom), they may alter their normal behavior. Accordingly, most observed cheating data derives from carefully comparing exams taken in the same room for statistical anomalies.

The literature concerning academic dishonesty, in both the popular press and in peer-reviewed academic journals, is extensive. McCabe, Butterfield, Trevino and Klebe (2012) provides an excellent review of this literature for all levels of education. Because most cheating is self-reported (usually via anonymous surveys) and because the nature and scope of cheating varies from study to study, as well as over time, reported statistics for the amount of cheating vary tremendously. Nonetheless, most studies suggest that over fifty percent of all students admit to engaging in academically dishonest behavior in high school, college, graduate school and beyond.

The main focus of this study concerns cheating on exams. Therefore, in the following two sections we provide a review of the survey data cheating literature and the observed cheating literature, specifically focusing on cheating behaviors related to exam taking. We then list the specific research questions addressed in this study.

https://doi.org/10.20429/ijsotl.2018.120104

## Survey Data Cheating Literature Review

Eve and Bromley (1981) surveyed 681 college students about their assessment of and participation in fifteen different academic dishonesty behaviors. For "gave another student answers during exam" and "copied answers from another student during exam," 79.3 percent and 90.2 percent of students, respectively, recorded that these behaviors were very dishonest or dishonest. Nonetheless, 43.5 percent of the same students admitted to giving another student answers during an exam while in college and 41.9 percent acknowledged copying exam answers from another student. Approximately half of all students who admitted to engaging in either behavior said they did so three or more times while in school. During the most recent semester, 32.0 percent gave another student answers and 26.2 percent copied answers during an exam, with approximately one-third of these admitting to doing so three times or more.

Haines et al. (1986) designed a 49-item questionnaire that was administered to 380 university students. The survey instrument included questions about student cheating behaviors on homework assignments, quizzes, and exams. Their data shows that 23.7 percent of students admitted to cheating on one or more major exams. Additionally, 54.1 percent of the students in their survey admitted to cheating on at least one item during the most recent academic year.

Genereux and McLeod (1995) surveyed 365 college students to determine their behaviors concerning cheating as well as their beliefs about the circumstances that would most likely increase or decrease their tendency to cheat. The two most common forms of cheating were reported as giving exam questions to another student (58 percent) and copying exam questions from another student (49 percent). Overall, 83 percent of the students in their survey admitted that they had cheated in college. Circumstances that promoted more cheating were inactive instructor proctoring, the perception that the exam was unfair, an instructor who did not seem to care if students cheated, and financial support tied to good grades. Items that students said highly discouraged cheating were instructor vigilance during exams, stringent penalties for getting caught, random seating or leaving an empty desk between students, fair exams and courses that a student deemed particularly valuable.

Diekhoff et al. (1996) compare a current study about student cheating to Haines et al. (1986), which was conducted ten years earlier. In the follow-up study, the authors surveyed 474 university students to determine the extent of cheating, attitudes about cheating, characteristics of cheaters, and the effectiveness of common deterrents used to dissuade cheating. They then evaluate changes in these variables over the decade. Overall cheating during the decade increased from 54.1 percent in 1984 to 61.2 percent in 1994. Cheating on major exams in both periods was constant at 23.1 percent.

McCabe et al. (2006) focus on cheating behaviors of graduate students. The authors conducted surveys at 54 colleges and universities in the U.S. and Canada, to which they received 5,331 responses. Of these, 623 were students majoring in business (mainly MBA students), and the rest were non-business majors. This comprehensive study investigates many different types of cheating and the reasons why students cheat. In terms of exam cheating, the authors report that 23 percent of business graduate students admitted to cheating on one or more exams; whereas, only 18 percent of non-business graduate students admitted to one or more instances of exam cheating.

Although survey data can provide instructive insights into thought processes and behaviors, surveys about academic dishonestly must be viewed with caution. The data for all of the studies noted above come from students anonymously completing questionnaires. Nelson and Schaeffer (1986), Scheers and Dayton (1987) and Karlins and Hargis (1988) all address the potential incongruence between actual and self-reported behavior in general. For example, Nowell and Laufer (1997) find that while 23% of students actually cheated when self-grading an assignment, their estimate of the percentage of the same students who admitted to cheating in a random response survey was only 9%.

Cheating surveys, in particular, can be significantly influenced by social desirability bias (Holbrook, Green & Krosnick, 2003). When answering a question about cheating on an exam, some students will answer no, even if in fact they did cheat, because that is what they believe is the "socially acceptable" answer, especially to academics who are administering the survey. Other students may answer yes, even if they did not cheat, because they believe most of their peers engage in the behavior and they want to be socially accepted by their peers.

Reference bias (Groot, 2000) is another major issue that must be carefully considered when interpreting the validity of survey responses about cheating. Reference bias occurs when respondents use different standards of comparison. Academic dishonesty and cheating are ambivalent concepts. Whereas one student may consider a casual glance at a neighbor's exam that confirms the answer on her/his exam to be correct as cheating, another student may think that anything less than copying the entire exam is not cheating. Mazar, Amir and Ariely (2008) demonstrate through experiments that if an individual can explain away a cheating behavior (e.g., that question was too hard or the teacher did not properly prepare me for that question) or if they can justify a cheating behavior (e.g., if I didn't cheat I would lose my scholarship and have to drop out of school or the teacher is not actively proctoring the exam), they do not consider the act as cheating.

Since social desirability bias and reference bias are highly likely for anonymous surveys concerning cheating, conclusions drawn from these studies may be misleading or even incorrect. Thus, to fully understand college cheating, it is important to study not only reported behavior, but also to measure actual behavior.

## Observed Cheating Literature Review

While the literature detailing college cheating based on surveyed student responses is extensive, studies that empirically document observed cheating are far less prevalent. Hetherington and Feldman (1964) identify three dimensions on which cheating behavior can be ordered: independent vs. social, opportunistic vs. planned, and active vs. passive. A majority of the existing studies documenting observed cheating behavior focus only on individual-opportunistic cheating. Most of these studies follow the experimental design implemented by Parr (1936), whereby completed student exams are collected and their responses are secretly recorded prior to handing the exams back in a subsequent class and requesting students to self-grade their exams given the correct answers. Parr reports that 42% of students awarded themselves higher scores than their original exams warranted. Similarly constructed studies have been performed on college students by Canning (1956), Zastrow (1970), Fakouri (1972), Tittle and Rowe (1973), Erickson and Smith (1974), Kelly and Worrell (1978), Gardner, Roper, Gonzalez and Simpson (1988), and Nowell and Laufer (1997). All of these studies find frequent cheating among sample students, ranging from a low of 15.6% of students cheating in Fakouri's experiment to a high of 50.8% in Gardner et al. Shelton and Hill (1969) performed this type of experiment with high school sophomores and juniors and found that 53% of students cheated.

Other articles have documented observed cheating of the social-planned variety. Karlins, Michaels and Podlogar (1988) find that 3% of upper-level business students plagiarized a particular writing assignment from students in the previous semester. Wilson (1999) reports that 117 freshman students emailed answers to one another during an exam.

While the above studies documenting observed cheating are enlightening, they overlook one of the most prevalent forms of cheating - copying another student's answer(s) during an exam. As noted in the previous section, copying on an exam or giving another student exam answers is among the most prevalent cheating behaviors reported by college students.

To our knowledge, the only work that has documented direct observation of this form of social-opportunistic observed cheating is Hetherington and Feldman (1964), which seeks to document several different types of cheating behavior. Of particular relevance to the present study, Hetherington and Feldman had five "student plants" attend class regularly in order for their presence during exams to be inconspicuous to students in the class. During an exam in which the professor was intentionally inattentive, these five observers recorded any instances of using cheat sheets, copying from another student or allowing another student to copy. Subsequently, students were also allowed to self-grade their exams in an experiment similar to Parr's. Overall, the authors report that approximately 50% of students cheated; however, as the primary focus of this study is on the individual characteristics that drive cheating, the results provided in the study do not allow us to further distinguish how many of these students copied or allowed others to copy versus those who used a cheat sheet or self-graded dishonestly.

All other measurements of observed cheating are derived from indirect observation; specifically, comparing common incorrect answers on a multiple choice exam. Fray (1996) and Haney and Clarke (2007) provide excellent reviews of this literature. For most studies cited in these works, the amount of cheating on an exam is measured through inspection of the number of identical wrong answers in the entire classroom or on exams of students sitting next to one another. Based on the number of answer choices per question and the assumed distribution of correct versus incorrect answers for the overall class or for student pairs, statistical indices can be created to indicate the probability of cheating.

These statistical indices are often computed and analyzed for national standardized examinations (e.g., SAT, MCAT, LSAT) but seldom by instructors in a class (Wesolowsky, 2000). Indeed, Cizek (2001) cautions against using such statistical ratios to measure or accuse an individual student of cheating. These measures do not indicate the direction of cheating (i.e., who copied from whom). They also effectively throw out most of the data, because they only measure the number of common incorrect answers.

For example, assume that student A copies every answer from student B on an 80 question multiple choice exam with 5 answer choices per question. If student A answered every question correctly, then student B also received a 100% grade on the exam and the statistical measures would indicate zero probability of cheating. If student A got 72 questions correct, then 8 incorrect answers on both exams would match. Random guessing would predict that 2 of the 8 should match. The fact that the other 6 match will create a high cheating probability statistic.

To accuse one or both of these students of cheating (note that there is no way of knowing whether student A copied from student B or student B copied from student A), an instructor must be willing to base an argument on the notion that the 6 common incorrect answers imply something about the 72 common correct answers. Specifically, the statistically improbable event that these incorrect answers match means that many or all of the 72 correct answers that match must have been copied by one student from the other. Such an argument is dubious at best and it would most certainly not stand up in a court of law. Perhaps the common incorrect answer chosen by both students was the next best answer choice derived by a similar error committed by both students or maybe the students studied together and made the same common error 6 times out of 72. On the other hand, if both students miss 50 questions and on all 50 have the same incorrect answer choice, then cheating almost certainly occurred, but since both students failed the exam, there is little incentive for the instructor to pursue a cheating accusation.

## Research Questions

Although observed cheating data can potentially be more accurate and useful than survey data, it is obviously very difficult to obtain. Using "student plants" may be, among other things, illegal or at least ethically questionable. Measuring cheating by only considering common incorrect answers does not indicate the direction of cheating and completely ignores the probability of cheating on common correct answers.

In this study, we use a technique proposed in Fendler and Godbey (2015) to observe cheating behavior on an in-class, proctored final exam. Our technique allows us to accurately measure the amount of cheating in the classroom by considering all of the data. It also allows us to identify the most likely cheater, if one individual covertly copies from a neighbor's exam. Additionally, by allowing one classroom of test takers to sit wherever they like and another classroom of test takers to be assigned to a random seat in the classroom when they arrive, we are able to observe whether a simple deterrent can effectively reduce cheating.

Specific research questions addressed in this paper are:

1. Can student cheating on an exam be efficiently observed without students knowing they are being watched?
2. How can the probability of cheating behavior be empirically measured for an entire class as well as for individual students?
3. Does unannounced assigned seating for an exam reduce social cheating?

# METHODOLOGY

Fendler and Godbey (2015) describe a simple way to punish cheaters and thus effectively reward honest students. The technique they describe works particularly well for multiple choice, open-ended problem solving and/or short answer exams for large section classes given in a classroom where students sit in close proximity. These are typically introductory general education or core courses where cheating behavior may be more pronounced than usual because a large percentage of the students in the class may be disinterested in the subject. Students must pass these classes to meet graduation requirements, but because the course subject matter may be unrelated to the chosen major, students merely want to pass in any way possible. If cheating allows one to pass in order to graduate, reference bias may influence what a student defines as accepted behavior (Groot, 2000).

The Fendler/Godbey (F/G) technique is easy to replicate. Two versions of the exam are created, but every effort is made to make the two versions appear identical. On each version, corresponding questions appear in the same order and, for multiple choice exams, answer choices between the two versions are identical. The only difference is a single number (or word) in each question so that the correct answer is different for each version.

Consider the following example. Assume that two students are sitting next to one another. Unknown to the students, one of them has the Version A exam (assume this is Student A) and the other has the Version B exam (assume this is Student B). As demonstrated in Figure 1, both versions appear, at a quick glance, to be identical.

Note the unique difference between the F/G system and the regular system where all exams in a room are identical. Table 1 summarizes the possible relationships between two students' answer choices on a question for each system and the conclusions that we can draw from each outcome.

**Version A:**
Emma is saving money to buy a new pair of running shoes. On Monday, she delivered a special edition of the newspaper in her neighborhood for five hours at $7 an hour. On Wednesday, she worked as a lifeguard at the community pool for a total of nine hours. She earns $8 an hour for lifeguarding. Finally, on Thursday, she babysat for three hours at $15 an hour. Assuming she started with $0 and all money earned this week was saved, how much money does Emma currently have to buy a new pair of running shoes?
a. $152  b. $173  c. $188  d. $160  e. $201

**Version B:**
Emma is saving money to buy a new pair of running shoes. On Monday, she delivered a special edition of the newspaper in her neighborhood for five hours at $7 an hour. On Wednesday, she worked as a lifeguard at the community pool for a total of ten hours. She earns $8 an hour for lifeguarding. Finally, on Thursday, she babysat for three hours at $15 an hour. Assuming she started with $0 and all money earned this week was saved, how much money does Emma currently have to buy a new pair of running shoes?
a. $152  b. $173  c. $188  d. $160  e. $201

***Figure 1.*** *Sample Questions from Version A and Version B*

***Table 1.*** *Regular vs. F/G System for Identifying Cheaters*

| **REGULAR SYSTEM** | | |
|---|---|---|
| **Possible Outcomes** | **Cheating Indication** | **Cheater?** |
| Student A and Student B do not match. | Indicates no cheating | N.A.. |
| Student A and Student B match. Both are correct. | No conclusion possible | No conclusion possible |
| Student A and Student B match. Both are incorrect. | Suggests possible cheating | No conclusion possible |
| | | |
| **F/G SYSTEM** | | |
| **Possible Outcomes** | **Cheating Indication** | **Cheater?** |
| Student A and Student B do not match. | Indicates no cheating | N.A. |
| Student A and Student B match. Student A is correct. Student B is incorrect. | Suggests possible cheating | Student B |
| Student A and Student B match. Student B is correct. Student A is incorrect. | Suggests possible cheating | Student A |
| Student A and Student B match. Both are incorrect. | Suggests possible cheating | No conclusion possible |

Using the regular system, no conclusion can be drawn when the answers match and the correct answer is chosen. Thus, this data is effectively dropped out. Because for most exams this category will be the largest segment (i.e., exams where the average grade is greater than 50 percent), ignoring this data is significant. Also, it is impossible to determine who was the perpetrator and who was the host.

Using the F/G system, if both students answer the question correctly, then both answers must be different and this indicates that no cheating occurred. If one student's answer is correct and the other student's answer is the same (meaning it must be incorrect), possible cheating is indicated and we can specifically identify the cheater versus the host.

## Study Design

Approximately 750 students were enrolled in the fall 2015 introductory finance course, which is required of all business majors. Only about 16 percent of these students will major in finance; for the rest, this is a course that is required to graduate but it is the only finance class they will ever take. There were 16 traditional sections of no more than 45 students each and one online section with 127 students. The sections were taught by ten different instructors – three senior faculty members and seven Ph.D. students.

The final exam for all sections is given at the same time in one of seven tiered classrooms that have approximately 120 seats each and are not the same classrooms where the individual sections meet throughout the semester. During the final week of classes, each section is assigned to a particular classroom for the final exam. Two or three sections are assigned to each classroom. Given the proximity of the seating and that the final exam counts for 35% of the students' final course average, copying from a neighbor's exam is a major concern for the instructors.

Numbered exams were placed in each seat prior to the students' arrival with each successive exam being a different version (i.e., Version A or Version B). Based on the exam number we are able to determine exactly where each student sat and who, if anyone, sat in the adjacent seat or if the student was at the end of a row. Thus we are able to compare each pair of students for the possibility of cheating. Because the rows are tiered and the space between each row in the room is relatively large, it is actually very difficult to copy from someone in the next row. Nearly all copying will be from an adjacent exam.

There were 40 multiple choice questions on the exam. Seven of the questions were conceptual in nature and identical. The remaining 33 questions were computational and followed the F/G technique described earlier.

Two Ph.D. student instructors with similar backgrounds and teaching two sections each assigned one of their sections to a room we will call Unassigned and the other to a room we will call Assigned, such that both exam classrooms contained one section of students from each of the instructors. A total of 86 students were assigned to and arrived in Unassigned and were allowed to sit anywhere in the first eight rows. A total of 74 students were assigned to and arrived in Assigned and were given assigned seats in the first eight rows. To create the seat assignments, students from one instructor's section were assigned random odd seat numbers, while students from the other instructor's section were given random even seat numbers. Thus, each student was necessarily seated next to students from a different class section, further reducing the likelihood that neighboring students would be familiar with one another. The students in Assigned did not know that they would be given a seat assignment until they arrived. In past semesters, students were assigned to a classroom and allowed to choose their seat.

There are many ways that familiarity with neighboring students can lead to increased cheating frequency. A poor student may purposely seek out and sit next to a top-performing student in the class and covertly copy as many questions as possible. Even in the absence of premeditation, students who feel the need to cheat during the exam may be more likely to copy from a student whom they know, reasoning that their behavior is less likely to be reported if discovered by the other student. Similarly, stronger students may be more likely to help a neighboring student by positioning their answer sheets in plain sight if the neighboring student is a friend. Even more alarmingly, students allowed to select their own seats may conspire to cheat, with different students studying specific material in the class and allowing their co-conspirator to copy answers to questions for which each is most prepared. Comparing the amount of observed cheating in both rooms allows us to determine whether assigning students to seats in an exam room can reduce such instances of social cheating.

## DATA

Prior approval from the university's IRB was received for this research project. Individual instructors provided copies of final exam answer sheets to the authors. The university provided student-level data on semester GPA, overall GPA, gender, age, and major. This data was linked with individual exams and then all student-specific identifiers were removed to strictly protect student anonymity.

Table 2 provides summary statistics for the students populating the two classrooms. At the 95% confidence level, there are no statistically significant differences in means between the two samples in terms of overall GPA (2.99 in Assigned; 3.11 in Assigned), semester GPA (3.11 in Assigned; 3.20 in Unassigned), age (24.5 in Assigned; 23.8 in Unassigned), or the percentage of finance majors (30.3% in Assigned; 34.1% in Unassigned).

***Table 2:*** *Group Demographic Characteristics*

| | Unassigned (N=86) | Assigned (N=74) | Unassigned minus Assigned |
|---|---|---|---|
| Characteristic | Mean | Mean | Difference |
| GPA – term | 2.99 | 3.11 | -0.12 |
| GPA – school | 3.15 | 3.20 | -0.05 |
| Gender (F=1) | 44.94% | 63.53% | -18.59%* |
| Age (years) | 24.47 | 23.77 | 0.7 |
| %Fin/Acct Major | 30.34% | 34.12% | -3.78% |

**indicates significance at 95% confidence level*

There is, however, a statistically significant difference in the means between the two samples with regards to gender (44.9% in Unassigned; 63.5% in Assigned). That is, female students comprise a much greater percentage of the population in Assigned. Prior research has provided mixed results regarding the role of gender on cheating. For example, McCabe and Trevino (1997) find that men self-reported more cheating than women, while Baird (1980) and Haines et al. (1986) report no statistically significant differences in cheating between men and women. Thus, we do not believe this difference influences our results, but the role that gender plays in observed cheating is an interesting area for future research.

## RESULTS AND ANALYSIS: CHEATING AT THE CLASSROOM LEVEL

As explained earlier, given the Fendler/Godbey exam design, any instance of a student answering a question incorrectly and selecting the same answer as his/her neighbor suggests the possibility of copying. To examine the differences in behavior between the assigned and unassigned seating classrooms, we calculate the total percent of wrong and matching answers in each classroom relative to all wrong answers for every possible pairing of students. We call this value Actual Match Percentage. Actual Match Percentage is then compared to the percentage of similar answers that we would expect if no cheating occurred (Expected Match Percentage).

To illustrate the calculation of Actual Match Percentage, consider the simplified example of a classroom with one row and four students presented in Figure 2. The position of each student in the figure corresponds to his or her position on the row, with Marge sitting on the far left end of the row and Bart sitting at the far right.

| | Marge | Homer | Lisa | Bart |
|---|---|---|---|---|
| Wrong Answers | 5 | 20 | 1 | 10 |
| Wrong & Match with Left | N.A. | 4 | 0 | 6 |
| Wrong & Match with Right | 2 | 6 | 1 | N.A. |

***Figure 2:*** *Simple Classroom Example to Measure Potential Cheating Behavior*

The situation in Figure 2 provides six opportunities for potential copying, as Homer, Lisa and Bart can copy from the student to their left and Marge, Homer and Lisa can copy from the student to their right. The possible copying pairs, number of questions that are wrong and match for each pair, and total number of questions each potential "copier" got wrong, are displayed in Figure 3.

| Possible Copying Pairs | Wrong & Match Neighbor | Total Number of Wrong Answers |
|---|---|---|
| Marge copies from Homer | 2 | 5 |
| Homer copies from Marge | 4 | 20 |
| Homer copies from Lisa | 6 | 20 |
| Lisa copies from Homer | 0 | 1 |
| Lisa copies from Bart | 1 | 1 |
| Bart copies from Lisa | 6 | 10 |
| **Total** | **19** | **57** |

***Figure 3:*** *Possible copying pairs, and the resulting incorrect answers that match with a neighbor*

Note that Homer and Lisa can possibly copy from two different individuals, so the number of wrong answers by Homer and Lisa that match their right-hand neighbor and their left-hand neighbor are both counted in the numerator, and their total number of wrong answers is counted twice in the denominator. Marge and Bart's corresponding values appear only one time in the numerator and one time in the denominator, because they had only one student sitting next to them from whom they could possibly copy.

Thus, for this example, Actual Match Percentage (AMP) is calculated as follows:

$$\text{AMP} = \frac{2+4+6+0+1+6}{5+20+20+1+1+10} = \frac{19}{57} = .3333 = 33.33\%$$

Each question on the multiple choice exam has five possible answers – the correct answer for both versions and three distractors that are close to the correct answer numerically but are not the results of common mistakes. As described in the methodology section, an honest student who answers incorrectly has a one in five chance of matching the student in the adjacent seat. Therefore, in the absence of any cheating in the classroom, the Expected Match Percentage (EMP) is 20%.

Table 3 presents an analysis of the Actual Match Percentage and Expected Match Percentage for both classrooms – Unassigned and Assigned. The AMP in the Unassigned room is 35.6% and the AMP in the Assigned room is 31.7%. The difference between AMP and EMP is highly significant for both rooms. Thus, we can reject the null hypothesis that the match percentage in either classroom was 20% with approximately 100% certainty. In other words, we can state with almost total certainty that some cheating took place in each classroom.

***Table 3:*** *Estimate of Cheating in each Classroom*

| | Unassigned | Assigned |
|---|---|---|
| Actual Match Percent (AMP) | 35.6% | 31.7% |
| Expected Match Percent (EMP) | 20.0% | 20.0% |
| Difference (AMP – EMP) | 15.6% | 11.7% |
| Z-statistic for Signif. of Diff. | 15.69 | 9.20 |
| Significance level | 100.0% | 100.0% |

Furthermore, we observe that the students in Unassigned had a significantly higher Actual Match Percentage (35.57%) than their peers in Assigned (31.69%). As shown in Table 4, the difference between these values is statistically significant at the 99% level. That is, statistical testing reveals that the chances of observing a difference this large between the match percentages in the classrooms due to chance is less than 1%, strongly supporting the hypothesis that the implementation of assigned seating significantly reduced instances of cheating on the exam.

***Table 4:*** *Estimate of Cheating in Unassigned vs. Assigned Classroom*

| | Difference |
|---|---|
| AMP in Unassigned classroom minus AMP in Assigned classroom | 3.88% |
| Z-statistic | 2.41 |
| Significance level | 99.2% |

## Results and Analysis: Cheating at the Student Level

Having documented our findings at the classroom level, we next turn our attention to identifying cheating at the student level. With each question having five answer choices, we know that an honest student's wrong answer should match with a particular neighbor only 20% of the time. We can thus use the binomial probability formula (DeGroot and Schervish, 2012) to estimate the probability that a student has engaged in copying from a neighbor's exam.

Specifically, given that a student who answers a question incorrectly has a 20% chance of randomly matching his neighbor's answer, the probability that a student randomly selected exactly *m* answers that matched the answers of his/her neighbor given that *w* answers were wrong is:

$$P(x = m) = \frac{w!}{m!\,(w-m)!} 0.2^m (1-0.2)^{(w-m)} \qquad (1)$$

The probability that this student randomly matched *M* or more of his neighbor's answers is given by:

$$P(x \geq M) = 1 - \sum_{m=0}^{M-1} P(x = m) \qquad (2)$$

Also, the probability that this student cheated is one minus the probability of observing this much matching by chance:

$$P(copied) = 1 - \left(1 - \sum_{m=0}^{M-1} P(x=m)\right) = \sum_{m=0}^{M-1} P(x=m) \qquad (3)$$

For example, returning to Figure 2, both Bart and Homer had six instances of a wrong answer choice that also matched the answer given on Lisa's exam for the same question. However, since Homer missed 20 questions on his exam, we would expect a greater number of matching answers simply by chance. Indeed, for Homer:

$$P(x=m) = \frac{20!}{m!\,(20-m)!} 0.2^m (1-0.2)^{(20-m)}$$

And the calculated probability that Homer copied from Lisa is:

$$P(copied) = \sum_{m=0}^{5} P(x=m) = 0.8042 = 80.4\%$$

Thus, the probability that Homer copied from Lisa is approximately 80%.

On the other hand, Bart matched with Lisa on 6 of his 10 missed questions. Setting *w* equal to 10 in the expression for Bart, the calculated probability that Bart copied from Lisa is:

$$P(copied) = \sum_{m=0}^{5} P(x=m) = 0.9936 = 99.4\%$$

Thus, the probability that Bart copied from Lisa is over 99%.

Table 5 lists the percentage of students in each classroom that engaged in probable copying from their left or right neighbors at various levels of confidence. Specifically, formula (3) was computed for every pair of students in each classroom. Then the number of students for whom P(copied) ≥ confidence level was divided by the total number of pairs in the room.

***Table 5:*** *Estimate of Individual Student Cheating in Unassigned vs. Assigned Classroom*

| | **Unassigned** | | **Assigned** | |
|---|---|---|---|---|
| **Confidence level** | *Left* | *Right* | *Left* | *Right* |
| 80.0% | 41.33% | 44.00% | 32.26% | 32.26% |
| 90.0% | 33.33% | 30.67% | 22.58% | 20.97% |
| 95.0% | 17.33% | 21.33% | 14.52% | 14.52% |
| 99.0% | 13.33% | 13.33% | 6.45% | 6.45% |
| 99.9% | 5.33% | 9.33% | 1.61% | 1.61% |

For example, at the 80% confidence level, we find that 41.33% of the students in the Unassigned room copied from a student to their left and 44.00% copied from a student to their right. If there was no copying, these values should be only 20%. Thus, we can say with 80% confidence that the probability that cheating occurred in the Unassigned room was greater than random chance by a factor of more than 2.

Comparing the two classrooms, we find that for any confidence level above 80%, fewer of the students in the Assigned room appear to have copied their neighbors' exams than in the Unassigned room. For example, the data allows us to be nearly certain, with a confidence level of 99.9%, that 9.33% (5.33%) of students in the Unassigned classroom copied answers from the student to their right (left). On the other hand, in the Assigned room, only 1.61% of students can be identified as having copied at this certainly level. Thus, randomly assigning students, without pre-knowledge, to specific seats in the exam room significantly lowered the amount of cheating.

## CONCLUSION

This research study presents a novel methodology for measuring the observed frequency of possible cheating on a multiple choice exam. Using this method to compare an exam classroom where students are allowed to choose their own seats to one where students are unexpectedly given an assigned seat next to a student from another section, we are able to test whether the implementation of a seating chart leads to reduced cheating. Our results show that the unassigned seating classroom as a whole indeed had a significantly higher percentage of probable copying.

Next, we examine the data at the student level to see how many students engaged in copying behavior egregious enough to allow us to identify them as probable cheaters for varying levels of statistical confidence. Regardless of our chosen confidence level, we are able to identify a higher percentage of cheaters in the unassigned classroom as compared to the assigned seating room. These findings strongly support the hypothesis that separating test-takers from their acquaintances can effectively reduce instances of copying on a multiple choice exam.

To be sure, the seating chart is not a panacea for all cheating, as our findings show that even in the assigned seating classroom, students' incorrect answers matched with their neighbors more often than can be explained by chance. It is hardly surprising to observe that in a moment of weakness or desperation, and when the opportunity is available, students will sporadically copy from an adjacent student's exam, even when the cheater does not necessarily know the other student or even believe that the other student better knows the correct answer. This type of opportunistic cheating, as discussed in Mazar, Amir and Ariely (2008) and Ariely (2012), is a common behavior throughout society.

While all cheating is unacceptable, intermittent instances of opportunistic copying are unlikely to have large effects on grading outcomes. Of greater importance to the integrity of the grading system is systematic cheating arising from a student's social familiarity with other students in the class. This form of cheating can arise through collusion between students, whereby one student explicitly or implicitly agrees to allow another student to easily see his or her answers. Even in the absence of collusion, social cheating can arise when an unprepared student intentionally positions himself next to a student whom he believes to be better prepared, with the intention of copying. Finally, even in the absence of premeditation, familiarity with neighboring students can increase the frequency of opportunistic cheating when the copying student believes that an acquaintance who notices his attempt to copy is less likely to report this behavior to the instructor.

Our findings that the implementation of a seating chart reduces the overall level of cheating in the classroom, and particularly seems to reduce the percentage of students who consistently and repeatedly copy from the same neighbor, is consistent with the notion that by eliminating students' ability to choose where they position themselves in the classroom, these instances of social cheating can be significantly reduced.

Given the distortive nature of cheating on the usefulness of grades as a signal of student quality, these results are particularly encouraging for instructors. Our findings show that a relatively costless and easily-implemented system, such as an unexpected random seating chart, can help to reduce instances of cheating in the classroom. Moreover, this simple deterrent can be implemented even in courses that use exam formats other than multiple choice.

Because observed cheating behavior is most likely a more accurate measure of academic dishonesty than conclusions drawn from survey data, this study suggests many areas for future research. For example, the effectiveness of other often used cheating deterrents, such as signing the university's honor code or the instructor speaking about cheating before students take an exam, can be examined in a manner similar to the random seating procedure discussed in this paper. Also, observed cheating behavior in a class can be compared to student surveys about cheating to determine potential bias areas in survey data on cheating. Finally, the characteristics of observed cheaters (such as gender, age, GPA, major, outside distractions, and others) can be examined to better understand who cheats and why.

## REFERENCES


Ariely, D. (2012). *The (honest) truth about dishonesty.* New York: HarperCollins Publishing Company.

Baird, J.S. (1980). Current trends in college cheating. *Psychology in the Schools*, 17, 515-522.

Canning, R.R. (1956). Does an honor system reduce class-room cheating? An experimental answer. *The Journal of Experimental Education,* 24(4), 291-296.

Cizek, G. J. (2001). An overview of issues concerning cheating on large-scale tests. In *National Association of Test Directors 2001 Symposia*, 1-30. Retrieved from: http://nationalassociationoftestdirectors.org/wp-content/uploads/2012/12/natd2001.pdf#page=4

Clarke, R., & Lancaster, T. (2006). Eliminating the successor to plagiarism? Identifying the usage of contract cheating sites. In *Proceedings of 2nd International Plagiarism Conference*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.120.5440&rep=rep1&type=pdf

DeGroot, M. H., & Schervish, M. J. (2012). *Probability and statistics*. Addison-Wesley,.

Diekhoff, G. M., LaBeff, E. E., Clark, R. E., Williams, L. E., Francis, B., & Haines, V. J. (1996). College cheating: Ten years later. *Research in Higher Education*, 37(4), 487-502.

Erickson, M.L., & Smith, W.B. (1974). On the relationship between self-reported and actual deviance: An empirical test. *Humbolt Journal of Social Relations,* 1, 106-113.

Eve, R. A., & Bromley, D. G. (1981). Scholastic dishonesty among college undergraduates: Parallel tests of two sociological explanations. *Youth & Society*, 13(1), 3-22.

Fakouri, M.E. (1972). Achievement motivation and cheating. *Psychological Reports*, 31, 629-630.

Fendler, R. J., & Godbey, J. M. (2015). Cheaters should never win: Eliminating the benefits of cheating. *Journal of Academic Ethics*, 1-15.

Frary, R. B. (1993). Statistical detection of multiple-choice answer copying: Review and commentary. *Applied Measurement in Education*, 6(2), 153-165.

Gardner, W.M., Roper, J. T., Gonzalez, C., & Simpson, R. G. (1988). Analysis of cheating on academic assignments. *The Psychological Record*, 38, 543-555.

Genereux, R. L., & McLeod, B.A. (1995). Circumstances surrounding cheating: A questionnaire study of college students. *Research in Higher Education,* 36(6), 687-704.

Groot, W. (2000). Adaptation and scale of reference bias in self-assessments of quality of life. *Journal of Health Cconomics,* 19(3), 403-420.

Haines, V. J., Diekhoff, G. M., LaBeff, E. E., & Clark, R. E. (1986). College cheating: Immaturity, lack of commitment, and the neutralizing attitude. *Research in Higher Education*, 25(4), 342-354.

Haney, W. M., & Clarke, M. J. (2007). Cheating on Tests: Prevalence, detection and implications for online testing. *Psychology of academic cheating,* 255-287.

Hessing, D. J., Elffers, H., & Weigel, R. H. (1988). Exploring the limits of self-reports and reasoned action: An investigation of the psychology of tax evasion behavior. *Journal of Personality and Social Psychology,* 54(3), 405.

Hetherington, E. M., & Feldman, S.E. (1964). College cheating as a function of subject and situational variables. *Journal of Educational Psychology,* 55(4), 212-218.

Holbrook, A. L., Green, M. C., & Krosnick, J. A. (2003). Telephone versus face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability response bias. *Public Opinion Quarterly*, 67(1), 79-125.

Karlins, M., & Hargis, E. (1988). Inaccurate self-perception as a limiting factor in managerial effectiveness. *Perceptual and Motor Skills*, 66, 665-666.

Karlins, M., Michaels, C., & Podlogar, S. (1988). An Empirical Investigation of Actual Cheating in a Large Sample of Undergraduates. *Research in Higher Education*, 29(4), 359-364.

Kelly, J.A., & Worrell, L. (1978). Personality characteristics, parent behaviors, and sex of subject in relation to cheating. *Journal of Research in Personality*, 12, 179-188.

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, 45(6), 633-644.

McCabe, D.L., & Trevino, L.K. (1997). Individual and contextual influences on academic dishonesty: A multicampus investigation. *Research in Higher Education*, 38, 379-396.

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

Nelson, T., & Schaeffer N. (1986). Cheating among college students estimated with the randomized-response technique. *College Student Journal*, 20, 321-325.

Novotney, A. (2011). Beat the Cheat. *American Psychological Association*, 42(6), 54.

Nowell, C., & Laufer, D. (1997). Undergraduate student cheating in the fields of business and economics. *The Journal of Economic Education*, 28(1), 3-12.

Scheers, N., & Dayton, C.M. (1987). Improved estimation of academic cheating behavior using the randomized response technique. *Research in Higher Education*, 26, 61-69.

Selingo, J. (2004). The cheating culture. ASEE Prism, 14(1), 24-30.

Shelton, L., & Hill, J. P. (1969). Effect on cheating of achievement anxiety and knowledge of peer performance. *Developmental Psychology*, 1, 449-455.

Tittle, C.R., & Rowe, A.R. (1973). Moral appeal, sanction threat, and deviance: an experimental test. *Social Problems*, 20(4), 488-498.

Wesolowsky, G. O. (2000). Detecting excessive similarity in answers on multiple choice exams. *Journal of Applied Statistics*, 27(7), 909–921.

Wilson, R. (1999). Colleges urged to better define academic integrity and to stress its importance. *Chronicle of Higher Education*, 46(8), A18.

Zastrow, C.H. (1970). Cheating among college graduate students. *The Journal of Educational Research*, 64, 157-160.