Journal of Economic Behavior and Organization 182 (2021) 196–211

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Online cheating amid COVID-19☆

Eren Bilen[a], Alexander Matros[a,b,*]

[a] Darla Moore School of Business, University of South Carolina, Columbia, SC 29208, United States
[b] Lancaster University Management School, Lancaster LA1 4YX, United Kingdom

ARTICLE INFO

Article history:
Received 23 October 2020
Accepted 3 December 2020
Available online 28 December 2020

JEL classification:
A2
I20
K42

Keywords:
Education
Cheating
Online examinations
Online chess

ABSTRACT

We present evidence of cheating that took place in online examinations during COVID-19 lockdowns and propose two solutions with and without a camera for the cheating problem based on the experience accumulated by online chess communities over the past two decades. The best implementable solution is a uniform online exam policy where a camera capturing each students computer screen and room is a requirement. We recommend avoiding grading on a curve and giving students less time but simpler questions on tests.

1. Introduction

"No matter what the game is, when there are benefits from winning, you have cheating."

- Arkady Dvorkovich, FIDE President, 2020.**

The COVID-19 pandemic changed the lives of all people globally with most activities being forced to move online, including teaching. Most schools and universities moved from face-to-face to online delivery in March 2020. Among other difficulties related to online teaching, measuring student performance became one of the chief concerns of instructors. Many universities reported widespread cheating in online examinations that took place in Spring 2020, and the problem became so rampant that even the media addressed it. See, for example, two recent articles in Washington Post (Newton 2020) and Inside Higher Ed (Lederman 2020).[1]

☆ We would like to thank Jan Bass, Philip Brookins, Bentley Coffey, Joshua Hess, Chun-Hui Miao, Edsel Pena, Tamara Sheldon, Lindsey Woodworth, and seminar participants at the University of South Carolina for their questions and suggestions. Special thanks to the Office of Student Conduct and Academic Integrity at University of South Carolina, especially to Erin Kitchell, Director of Academic Integrity, for her feedback and comments.

* Corresponding author.
E-mail addresses: eren.bilen@grad.moore.sc.edu (E. Bilen), alexander.matros@moore.sc.edu (A. Matros).

** The International Chess Federation (FIDE) is the governing body of chess, and it regulates all international chess competitions.

[1] See the Washington Post article and the article on InsideHigherEd..

https://doi.org/10.1016/j.jebo.2020.12.004
0167-2681/

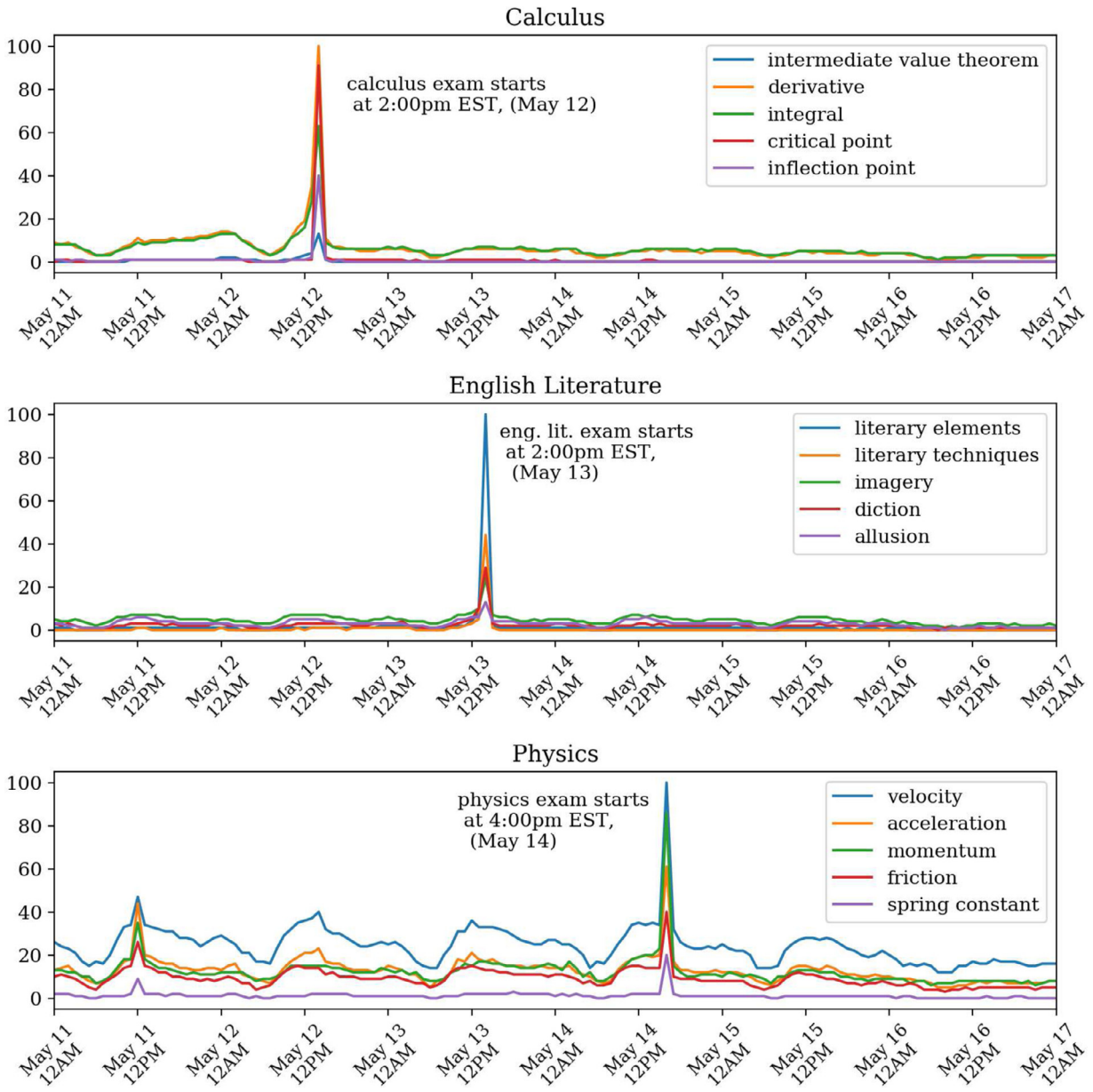


Fig. 1. Google search trends around the time of the 2020 AP Exams. *Note:* 2020 AP Exams were held online due to COVID-19 related school closures. Hourly online search data is obtained from Google Trends. The search data covers the U.S. nationwide.

The 2020 Advanced Placement (AP) examinations illustrate the difficulty of measuring true student performance on online examinations without proctoring. Fig. 1 shows surges of Google searches on keywords related to exam topics perfectly correlating with the time of the examinations. Since the online environments used for the 2020 AP exams had no proctoring, many students took advantage of having immediate access to Google search.[2]

Currently, most schools and universities are occupied with the switch to online teaching. Consequently, will the cheating problem in the fall be significant enough for schools and universities to take strict measures for the future? Will it be possible to have a fair assessment system if schools and universities decide to take no action? How can schools and universities maintain academic integrity in online examinations? We must wait several months to get clear answers. However, we can make predictions on the possible outcomes by considering theory and past evidence on cheating.

[2] College Board did not consider internet search to be cheating for the 2020 AP examinations. However, even if internet search was considered cheating, ensuring that students not use internet search during the test would be a challenging task.

The problem of cheating in online environments is not new. Online chess, in particular, has been plagued by cheating ever since chess was first introduced to the internet, with players gaining an unfair advantage by using computer assistance.[3] Online chess started when the Internet Chess Club (ICC) was established in 1995. The ICC first ran annual Dos Hermanas online blitz tournaments with monetary prizes in the early 2000s. Games were not proctored, and a whole array of cheating scandals consequently arose, with many ways to cheat in those events. In order to function, the ICC had to disqualify numerous people, including a former Junior World Champion, a top Chinese player, a top German player, hundreds of titled players, and thousands of amateurs (who enjoyed beating titled players). The main way to cheat was to use chess computer programs that found the best moves.

Now, Chess.com is the most popular online chess club with many tournaments including monetary prizes. Unsurprisingly, cheating has surfaced as a huge problem, prompting Chess.com to create its own cheating detection unit. See Chess.com Fair Play And Cheat-Detection.[4] The website states: *"Though legal and practical considerations prevent Chess.com from revealing the full set of data, metrics, and tracking used to evaluate games in our fair-play tool, we can say that at the core of Chess.com's system is a statistical model that evaluates the probability of a human player matching an engine's top choices, and surpassing the confirmed clean play of some of the greatest chess players in history."*

Cheating before the pandemic on both the ICC and Chess.com is similar to the online cheating problem that arose in Spring 2020. Thriving throughout COVID-19, cheating skyrocketed for online chess as well. For example, Chess.com announced on August 19, 2020 that it closes more than 500 accounts every day for cheating and has closed over 400,000 accounts in total, projecting to close 1,000,000 accounts by mid-2023. Of those closed accounts, nearly 400 were titled players. The only seeming positive statistic that was found indicated that female players cheat much less, only accounting for 4.57% of all title players.[5] However, recent weeks revealed an explosion of top women players cheating in both online and over the board tournaments as well. Former Women's World Champion, Anna Ushenina, was accused of cheating after her Internet 2020 Grand Prix victory. Another example is Patrycja Waszczuk, a titled young chess player, medalist of the Polish Championships, and medalist of the European Chess Championships, who was banned online and also caught cheating during her game at the Chess Festival in Ustron.[6]

If we can learn anything from online chess, then the message is very clear: online cheating will only get much worse and schools and universities will have their first hand experience already in Fall 2020. While online chess websites are private ventures and can ban any player for any reason, schools and universities will have a much more difficult task to provide clear evidence that proves students' cheating.

Both the ICC and Chess.com have been successful to some degree in dealing with the cheating problem although it is nowhere near to being solved. Interestingly, there are similarities in addressing the problem by both chess websites. First, they do not reveal their detection systems. Second, their disqualification decision is final. This approach admits that the detection system is vulnerable to knowledgeable cheaters. Since the websites do not have the resources to check millions of games, they implement a simple way to address the problem: a chess website monitors particular characteristics of play, and its methods of analyzing these characteristics are not revealed to the players. Players do not know what the website is looking for, making cheating more difficult.[7]

This cheating behavior supports the mounting evidence that the lack of "perfect honesty" exists in situations where the returns to dishonesty are high. Numerous studies using different settings and samples investigated in Gneezy (2005), Fischbacher and Föllmi-Heusi (2013), Gächter and Schulz (2016), and Vanberg (2017) show that people are more likely to deceive if the marginal benefit from deception is significantly large. Therefore, professional competitions and examinations have to use extensive monitoring to prevent cheating. However, lockdowns due to the COVID-19 pandemic in early 2020 made monitoring very difficult (or impossible) in many situations. Online tests are done without face-to-face proctoring, which implies that students can potentially use their notes, internet search, and any other assistance to help them solve questions. Furthermore, they can communicate via teleconference (or some other method) and collaborate during their exams. This cheating behavior on online examinations imposes a negative externality on students who do not cheat, especially if the instructor curves the exam scores.

In this paper, we first consider two simple models of face-to-face and online examinations. These models suggest that unlike the face-to-face examination, cheating should be expected in the online examination, with the reason being very intuitive: the instructor can observe cheating evidence in the face-to-face examination, but there is only indirect cheating evidence in the online examination. Therefore, cheating is a part of the student equilibrium strategy in the online examination.

We then present evidence of cheating that took place in an online examination held as part of a course taught at a large public university in Spring 2020 during COVID-19 lockdowns. Using a simple way to detect cheating - timestamps from the students' Access Logs - we identify cases where students were able to type in their answers under thirty seconds per

[3] Cheating in chess is a big issue in both online and over-the-board settings. This problem is relevant even in scholastic chess events. See https://en.chessbase.com/post/promoting-fair-play-among-child-chess-players.

[4] https://www.chess.com/article/view/chess-com-fair-play-and-cheat-detection.

[5] See https://www.chess.com/article/view/online-chess-cheating.

[6] See https://www.spraggettonchess.com/the-game-of-cheating-part-i/.

[7] For example, there are numerous cases of titled players admitting they had cheated and were correctly identified and caught by Chess.com.

question. We found that the solution keys for the exam were distributed online, and these students typed in the correct as well as incorrect answers using the solution keys they had at hand.

Now we suggest how to mitigate cheating based on the experience accumulated by online chess communities in the last two decades. Currently, many universities are requiring students to purchase and use a camera to record themselves while taking an exam in order to crack down on cheating, but these rules conflict with privacy rights from some students perspectives. In order to address this issue based on our theoretical models, we suggest that instructors present their students with two options: (1) If a student voluntarily agrees to use a camera to record themselves while taking an exam, this record can be used as evidence of innocence if the student is accused of cheating; (2) If the student refuses to use a camera due to privacy concerns, the instructor should be allowed to make the final decision on whether or not the student is guilty of cheating, with evidence of cheating remaining private to the instructor. Both options are designed to "implement" the outcome of the face-to-face examination when cheating is not expected in the equilibrium. Of course, there are other ways to achieve the same outcome. For example, students can take exams at proctoring centers.

The implications of this paper are simple: if no action is taken for online exams in the upcoming semester, there will be widespread cheating. Students have much to gain while the probability of being caught with definitive evidence is close to zero. Using online proctoring services involving the use of a camera is one solution - albeit imperfect - to the problem of cheating. We believe that cheating can never be fully detected online and therefore recommend that instructors stay away from curving their grades in order to not punish honest students.

The rest of the paper is organized as follows: Section 2 presents mixed evidence about online and face-to-face cheating, Section 3 provides two theoretical models for face-to-face and online exams, Section 4 presents examples of online cheating, and Section 5 concludes.

2. Related literature

Educational institutions have traditionally been using proctoring in order to ensure academic integrity on examinations. Online education, however, typically relies on unproctored examinations that are held online. Previous surveys exploring cheating in online examinations generally claim that there is little to no difference between face-to-face and online examinations in terms of cheating. Watson and Sottile (2010) surveyed 635 students from a medium-sized university and asked whether or not they had previously cheated on an examination. They found that 32.1% of students from face-to-face courses admitted to cheating. For students from online courses, the admitted cheating rate was 32.7%. Observing that these rates are very similar, they claim that online courses do not involve more cheating. However, the main concern regarding their methodology is that they rely on self-reporting which requires students admitting they have cheated rather than actually using a mechanism to detect cheating.

The next set of research addressing cheating concerns in online education includes Fask et al. (2014) which used random assignment of students to face-to-face and online examinations. They first assessed student performance using practice tests and found that the online test-takers scored 14% lower than those who took the practice test held proctored in class. However on the actual test, online test-takers scored 10% higher than the face-to-face test-takers. While their methodology had limitations in terms of detecting cheating, it provides suggestive evidence on cheating for students who take unproctored online examinations.

More concrete evidence on cheating in online environments is presented in Dee and Jacob (2012), Karim et al. (2014), and Diedenhofen and Musch (2017). Using a text-based algorithm that detects plagiarism, Dee and Jacob (2012) show that 112 out of 1200 papers submitted on the Blackboard from a sample of 28 universities were plagiarized. They suggest that giving a quick tutorial explaining how plagiarism jeopardizes academic integrity at the beginning of the semester is an effective tool in preventing plagiarism. However, this may not be as effective for more direct cases of cheating. Using evidence from laboratory and online experiments, Diedenhofen and Musch (2017) show that participants cheat more (via Google search) when monetary incentives are higher. They use a computer program that triggers a pop-up message if a participant frequently changes browser tabs in a short period of time with the message stating that the researchers are aware of the participant's cheating activity, which reduces cheating sharply. In another experimental setting using participants from Amazon's Mechanical Turk, Karim et al. (2014) show that low-cost technology, such as web-cameras, are effective at decreasing cheating without necessarily impacting test performance. However, they observe negative reactions from a portion of the participants pointing out that these web-cameras may be viewed as invasive and thus raise feelings of pressure and tension.

There is vast literature exploring cheating and deception. Becker (1968) was the first to provide the rationale for individuals who take part in illegal activities. An empirical investigation on cheaters and their incentives in Duggan and Levitt (2002) show that individuals are indeed more likely to cheat if they view returns to cheating are high and that these returns come with little cost.[8] Field experiments using different settings reveal that individuals deceive more if the cost of deception is small; see Gneezy (2005), Erat and Gneezy (2012), Gächter and Schulz (2016), Vanberg (2017), Martinelli et al. (2018), Charness et al. (2019), Alan et al. (2020), and Maggioni and Rossignoli (2020).

[8] Along with Jacob and Levitt (2003) these papers were later included in Levitt and Dubner (2005): the Freakonomics book, documentary, and podcast series.

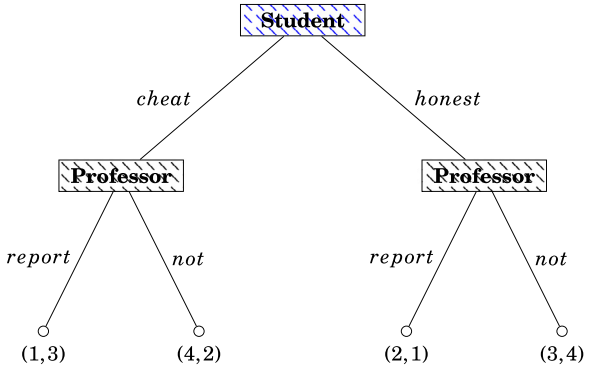


Fig. 2. Game Tree for the sequential move game.

Table 1
Payoff Matrix for the simultaneous move game.

		Professor	
		Report	*Not*
Student	*Cheat*	1,3	4,2
	Honest	2,1	3,4

2.1. Sequential-move game

In the sequential-move game, the student chooses to either cheat or be honest. The professor observes the student choice and decides either to report the student for cheating or not.

There are four outcomes in this game, but the professor and the student rank these outcomes differently. The professors outcomes from the best to the worst are (honest, do not report), (cheat, report), (cheat, do not report), (honest, report), where we record paths of play in brackets. The students outcomes from the best to the worst are (cheat, do not report), (honest, do not report), (honest, report), (cheat, report). See Fig. 2.

It is easy to find a unique subgame perfect equilibrium outcome, where the student is honest and the professor does not report the student. Note that this is the best outcome for the professor and the second best outcome for the student.

This sequential-move game is supposed to be played between the student and the professor during in-class exams. In the current situation, one way to implement this game is to use a camera to record the student while taking an exam. However, many students say that camera use conflicts with privacy rights and advocate taking exams without cameras. In other words, these students suggest to play the following simultaneous-move game.

2.2. Simultaneous-move game

Let us consider a simple simultaneous-move game between a student and a professor. The student has two actions: cheat or be honest, and the professor also has two actions: report the student for cheating or not.[9] There are four outcomes in this game, and the professor and the student rank these outcomes differently. The professor's outcomes from the best to the worst are (honest, do not report), (cheat, report), (cheat, do not report), (honest, report). The student's outcomes from the best to the worst are (cheat, do not report), (honest, do not report), (honest, report), (cheat, report). Table 1 gives an example of players' payoffs. This game has a unique mixed-strategy equilibrium, which means that the student and the professor should randomize between their two actions in equilibrium. Thus cheating as well as reporting is a part of the equilibrium.

In the mixed-strategy equilibrium, the student randomizes between her two choices in such a way to make the professor indifferent between his two choices. So, in order to determine the equilibrium probability of the student's cheating, we have to look at the professor's payoffs. To make our point clear, we restrict our attention on a simplified payoff Table 2, where we only record the professor's payoffs. Moreover, we normalize the best payoff at one and the worst payoff at zero, and $0 \leq c \leq b \leq 1$.

It is easy to find now that the equilibrium probability of the student cheating, p, is equal to

$$p = \frac{1}{1 + (b - c)}. \tag{1}$$

[9] In fact, the game can be played sequentially without the professor knowing the student's action. The normal-form of this game and our results are the same.

Table 2
Payoff Matrix for the Professor.

		Professor *Report*	*Not*
Student	*Cheat*	., b	., c
	Honest	., 0	., 1

If the professor does not feel a big difference between (cheat, report) and (cheat, do not report) outcomes, or the difference $(b-c)$ is small and close to zero, then the cheating probability is the highest, and in the extreme case, if $(b-c)=0$, this probability is equal to one, $p=1$. However, if the professor is concerned and sees a significant difference between (cheat, report) and (cheat, do not report) outcomes, then the student cheating probability goes down.

2.3. The problem and the solution

Note that the student is honest in the unique subgame perfect equilibrium in the sequential-move game, which approximates in-class exams. However, the student is supposed to cheat in the unique mixed-strategy equilibrium in the simultaneous-move game, which is a proxy for the online exams. These findings demonstrate that cheating should be higher in online tests, and these observations are intuitive, with many instructors having first hand experience with them from face-to-face and online teaching.[10]

The main problem revolving around online tests is how to prove cheating. Typically, students require to see evidence of their cheating and the professor only has indirect evidence of cheating. Having only indirect evidence makes it much harder to prove that cheating took place in case of an academic integrity referral. Thus, the professor is reluctant to report cheating in online exams, or the difference $b-c$ is close to zero in expression (1) for the equilibrium cheating probability. This in turn encourages even more cheating.[11]

How can this evidence problem be resolved? Many instructors have their own statistical evidence of students cheating. Some of these statistics are simple but very efficient. We present one such statistic – time spent per question – in the next section. However, once such a statistic is revealed to students, the instructor would not be able to use it again because students adjust accordingly. The solution is not to reveal information based on which the student was found guilty of cheating. In other words, if the professor claims that the student was cheating and this decision is final, then we indeed get the simultaneous-move game from the previous section.

We suggest to offer two options to a student. If the student buys a camera and uses it during the exam, then the sequential-move game is played, cheating is not expected (in the equilibrium), and both the student and the professor have evidence of the student's behavior on the exam. Alternatively, the student can have an exam without a camera in the privacy of their own home. In this case the simultaneous-move game is played, some cheating is expected in the equilibrium, and if the professor has evidence of cheating and claims cheating, then the student cannot request any evidence and appeal the verdict.

3. Data: evidence

Finding concrete evidence on cheating in an online examination is potentially challenging, as demonstrated in earlier studies (Watson and Sottile 2010, Fask et al. 2014). In this section, we present cases in which students were able to correctly solve several questions under thirty seconds per question.[12] The mechanism we use to identify cheating is the "Access Log" provided on the Blackboard. The Access Log provides detailed timestamps which show exactly how much time a student spends on each question. Many students appear to be unaware that the time they spend on each question is recorded although they seem to expect that the information on the "total time" they take for the exam is recorded.[13]

Our data comes from students who were enrolled in an intermediate-level course in Spring 2020 at a large public university. The course had three Midterms and had an optional Final Exam which replaced the lowest Midterm exam. The first two Midterms were held face-to-face with proctoring; the third Midterm and the Final Exam were held online asynchronously on the Blackboard following COVID-19 related campus closures. On these online exams, students received the same set of questions in a random order. The questions were all short answer questions: the student had to type in the correct answer

[10] See our discussion in Section 2.

[11] Anectodal evidence suggests that many instructors were indeed reluctant to report cheating in Spring 2020. Despite this, the number of reported cheating cases at a large public university (reported by their Academic Integrity Office) went up by almost 10% in March - June 2020 relative to March - June 2019.

[12] The questions on the exam are problem-solving questions which are arguably not-so-trivial in terms of finding the solutions. The exam is *not* multiple-choice – the student must type in the correct answer to receive credit.

[13] There were instances of students finishing their tests and waiting to submit them. The test was designed such that students could not go back and recheck their answers. Therefore, waiting could not improve or change their results. In one extreme case, a student finished the test in 11 min and waited for more than 1 h before submitting it, so that the total time spent on the test would look "normal". We believe this provides evidence on individuals involved in cheating attempting to "hide their trails" similar to what was observed in Jacob and Levitt (2003).

Table 3
Student 1 and 2's scoresheets.

Exam	Score	Letter
(a) Student 1's scoresheet		
Midterm 1	35/100	F
Midterm 2	55/100	F
Midterm 3	30/100	F
Final	95/100	A
(b) Student 2's scoresheet		
Midterm 1	50/100	F
Midterm 2	40/100	F
Midterm 3	50/100	F
Final	95/100	A

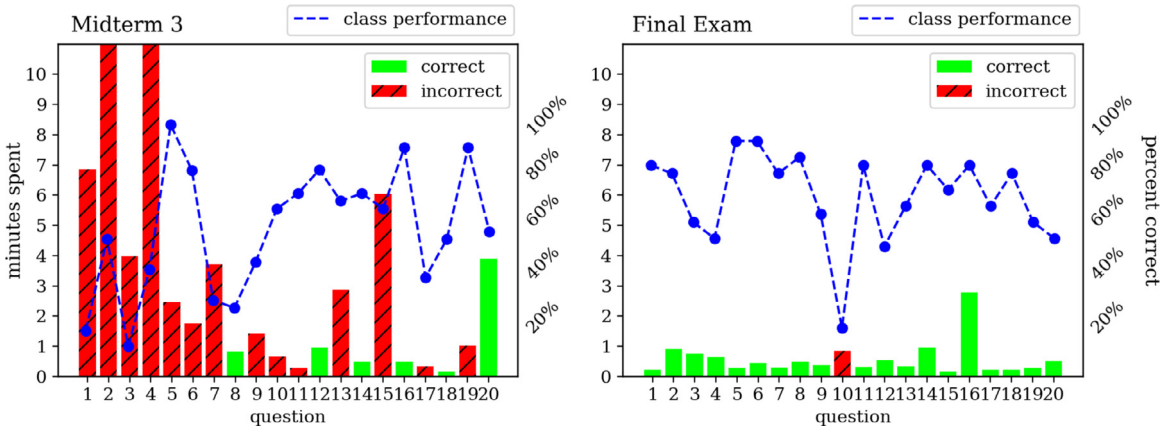


Fig. 3. Time spent per question for Student 1. *Note:* Student's Midterm 3 score is 6/20; Final score is 19/20. Total time he spent on Midterm 3 was 1 h and 10 min. Total time he spent on the Final exam was 11 min.

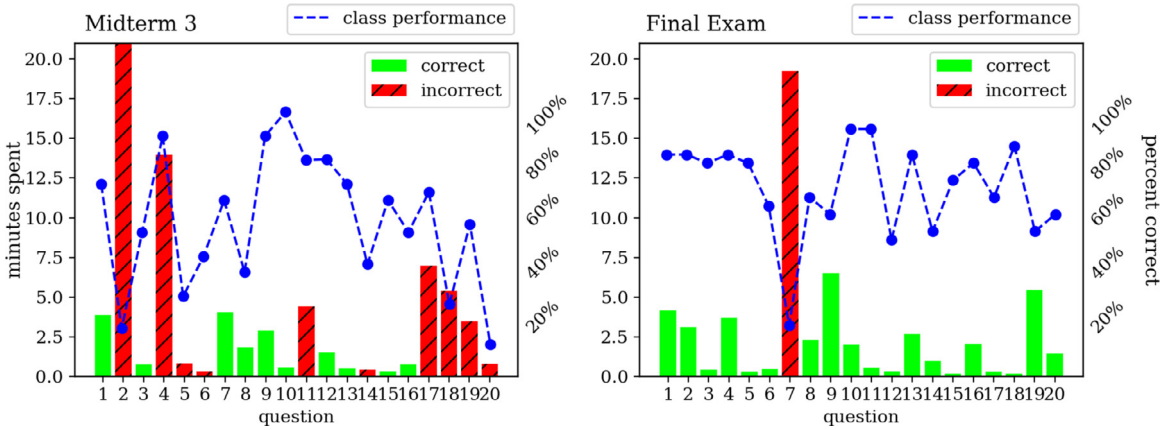


Fig. 4. Time spent per question for Student 2. *Note:* Student's Midterm 3 score is 10/20; Final score is 19/20. Total time he spent on Midterm 3 was 1 h 20 min. Total time he spent on the Final exam (excluding Question #7) was 36 min.

to receive credit with no multiple choice options given. In order to move to the next question, the student had to save and submit their answer; no moving back and forward was allowed.

Table 3 presents scoresheets for two particular students who took the course in Spring 2020. Figs. 3–4 show how much time each student spent on each question during Midterm 3 and the Final Exam with responses in Tables 4–5. On Midterm 3, students have both correct and incorrect answers and had to spend some time reading the problems and working to solve them.

Their time allocation, combined with their performance, shows no strange results for Midterm 3. However, their Access Logs reveal very peculiar information for the Final Exam.

Table 4
Student 1's answers on Midterm 3 (left) and Final (right).

Question #	Answer	Minutes
1	700	06:51
2	30	12:50
3	11.25	03:59
4	39.47	18:35
5	1	02:28
6	11.25	01:46
7	192	03:43
8	0	00:50
9	65	01:25
10	700	00:40
11	2500	00:17
12	1	00:58
13	900	02:53
14	3	00:30
15	3.18	06:03
16	2	00:30
17	3.18	00:20
18	"Elise"	00:10
19	400	01:01
20	10	03:54
		Total: 01:09:55

Question #	Answer	Minutes
1	-1	00:13
2	124	00:54
3	25	00:45
4	6534	00:38
5	25	00:17
6	-3.5	00:27
7	550	00:18
8	6	00:29
9	68	00:23
10	125	00:51
11	800	00:19
12	200	00:33
13	26	00:20
14	1440	00:58
15	2	00:10
16	1	02:47
17	500	00:14
18	82	00:14
19	11	00:17
20	4356	00:31
		Total: 00:11:49

Figs. 3 and 4 reveal cases where students had to spend less than thirty seconds to solve a question. However, each of these questions requires complex problem-solving skills, demonstrating that a student would need to spend a reasonable amount of time to find each solution. Furthermore, the students had to type in their answers since the exam was not a multiple choice exam. The questions typically had non-trivial answers such as "6534" or "650" which would make it very challenging to randomly guess the correct answers. In fact, the probability of randomly guessing the correct answers on this exam is much less than the probability of winning the lottery.[14]

Furthermore, there is evidence that these two cases are connected. Students' answers for all twenty questions on the Final Exam perfectly match. Both students made only one mistake on the same question where they both submitted the same incorrect answer of "125". Fig. 5 shows the answers submitted by the rest of the class on this particular question, and it appears only three students submitted "125" while the rest of the class submitted a whole range of different numbers. Two of these three students are Students 1 and 2.[15] Lastly, timestamps from their Access Logs show that once Student 2 finished his exam, Student 1 started his immediately (in 2 min) after Student 2 finished submitting his answers. We believe the probability that these students cheated and cooperated is higher than a random statistical occurrence.

How did these students cheat? The most likely explanation is that they used online resources, where private tutors helped them solve problems. In fact, we have found evidence that the answer key for the Final Exam was distributed online in a common web platform. For a price of several dollars, students could get access to the solution key. Once the student

[14] A rough estimate on the probability of "being lucky" on the Final Exam and guessing the correct answers by randomly submitting numbers is less than 1×10^{-20}. The probability of winning the lottery is around 1×10^{-7}. This rough estimate takes into account a student's potential to "guesstimate" the range for the correct answer.

[15] This question was by far the most difficult question on the exam with a correct response rate of only 19.3%. The third student who submitted "125" scored 18/20 on the Final Exam. His other incorrect answer was "12", and the correct answer for that question was "124". It appears he simply must have made a typo.

Table 5
Student2's answers on Midterm 3 (left) and Final (right).

Question #	Answer	Minutes
1	1	03:51
2	109.2	26:42
3	1800	00:45
4	0	13:58
5	300	00:49
6	26.6	00:20
7	10	04:03
8	18	01:51
9	300	02:54
10	0	00:35
11	3	04:24
12	300	01:31
13	"Elise"	00:31
14	82.6	00:25
15	2	00:20
16	3	00:45
17	72	06:59
18	250	05:23
19	5	03:28
20	4	00:47
		Total: 01:20:31

Question #	Answer	Minutes
1	1440	04:09
2	-1	03:07
3	82	00:25
4	1	03:41
5	550	00:18
6	68	00:28
7	125	19:15
8	500	02:19
9	11	06:31
10	25	02:01
11	-3.5	00:33
12	200	00:19
13	800	02:40
14	6534	00:58
15	2	00:10
16	124	02:04
17	26	00:17
18	6	00:10
19	4356	05:28
20	25	01:26
		Total: 00:56:28

obtained the solution key, the only task they would need to complete would be to type in the answers for the questions they were presented in a random order.[16]

Why did these students cheat? They had a lot to gain. Had they not cheated, they were very likely to fail the class. Student 1 had accumulated an overall score of 47.8/100 and Student 2 had 53.5/100 before the Final Exam. Their only chance to pass the class (getting at least 70) was to perform extraordinarily well on the Final Exam. Having received 95/100 each on the Final Exam, both students would pass the class. The students' relative performances on Fig. 6 and Figs. 10 and 11 show their extraordinary performance on the Final Exam relative to the rest of the class. Furthermore, these students performed even more remarkably compared to students who took the very same course within the past 10 years. Fig. 7 shows how students - who accumulated a failing score up to the final exam - performed on the final exams given since 2010. Out of 68 such students, only 4 managed to secure a high enough score to attain a passing letter grade. These students are Students 1 and 2, as well as two other students from the same section in 2020.

Note that there is nothing that stops students to type in their answers "slowly" which would mimic a case with no cheating. It appears these two students were not aware that their Access Logs had timestamps showing how much time they spent on each question. Had they known, they would have most likely submitted their answers in a longer time period so that their Access Log would look perfectly "normal". Thus it is essential that any information on the "cheating-detection" tools the instructors possess be kept private. Once these tools are public knowledge, they become useless in detecting cheating. We have presented two specific cases with compelling cheating evidence during the Final Exam. Could it be true that cheating was not limited to these particular cases? The same version of the course was given several times in the past all with proctored in-class exams. Fig. 12 shows the students' performance across exams since 2010.

[16] It would potentially take 15–20 seconds to identify what question comes up on the screen and match it with the solution key they have at hand since the order of the questions is randomized for each student.

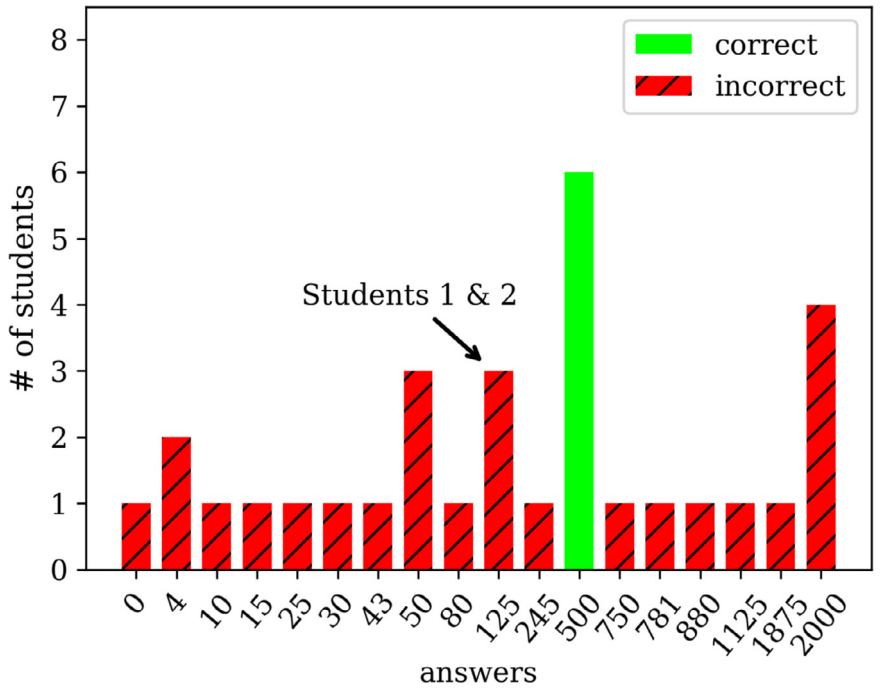


Fig. 5. All responses for the question both Students 1 & 2 made their only mistake on *Note:* Both Students 1 & 2 gave the same incorrect answer of "125". This question was by far the most difficult question on the exam with a correct response rate of only 19.3%. The third student who submitted "125" scored 18/20 on the exam. His other incorrect answer was "12", and the correct answer for that question was "124". It appears he simply must have made a typo.

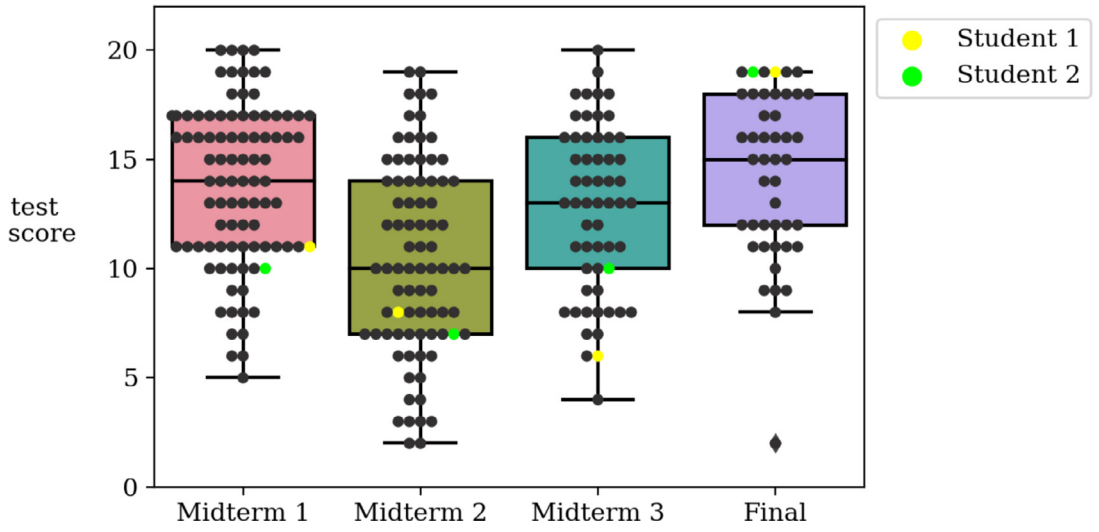


Fig. 6. Student 1 &2/s performance relative to the rest of the class. *Note:* Midterms 1 and 2 were held face-to-face with proctoring; Midterm 3 and the Final Exam were held online asynchronously without proctoring following COVID-19 related campus closures. Each dot represents a student's test score.

To compare the performance of students in Spring 2020 with past students who took the course, we use the following simple specification,

$$Score_{i,s,j} = \alpha_0 + \alpha_1 (Treat \times Midterm2)_{s,j} + \alpha_2 (Treat \times Midterm3)_{s,j} + \alpha_3 (Treat \times Final)_{s,j} + \alpha_4 Treat_s + \eta_j + \epsilon_{i,s,j} \quad (2)$$

where $Score_{i,s,j}$ is the exam score of student i in section s in exam j; $Treat_s = 1$ if the section is from Spring 2020; $Midterm2_j$, $Midterm3_j$, $Final_j$ are indicators for the corresponding exams; η_j is exam fixed-effects; $\epsilon_{i,s,j}$ is the idiosyncratic shock. Midterm 1 is taken as the baseline.

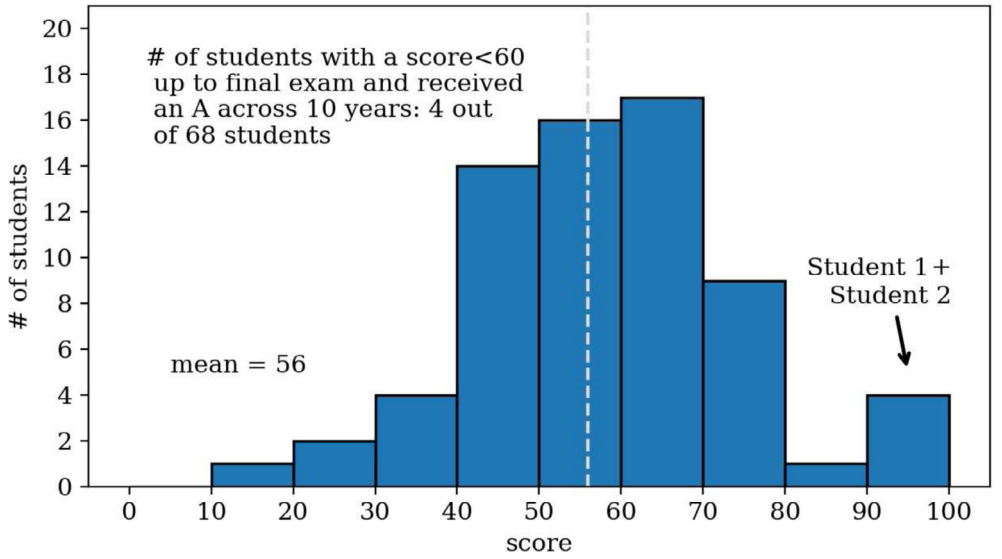


Fig. 7. Score distribution on all final exams given between 2010–2020 for students with a score <60 before the final exam. *Note:* Total # of students across 10 years with a score <60 up to final exam and manage to pass the course with a C is only 4. These students are Student 1, Student 2, Student 3 (and the fourth student from the same section.) Student 3 gave identical answer keys (with 1 exception) with Students 1 and 2.

Table 6
Exam score differences comparing online vs. face-to-face delivery from past 10 years.

outcome: exam score	(1) Section 1	(2) Section 2
Midterm 2	-7.183***	-7.183***
	(1.732)	(1.732)
Midterm 3	-3.382*	-3.382*
	(1.617)	(1.617)
Final	-8.033***	-8.033***
	(2.116)	(2.116)
Treat	-1.670	0.009
	(2.019)	(2.019)
Treat×Midterm 2	-10.416***	-5.172**
	(1.732)	(1.732)
Treat×Midterm 3	-10.795***	-7.509***
	(1.617)	(1.617)
Treat×Final	5.913**	-1.640
	(2.116)	(2.116)
N	1674	1682

Notes: Section 1 2020 students and Section 2 2020 students are compared with past 10-year students separately in columns (1)-(2). Baseline is Midterm 1, which was held face-to-face at the beginning of Spring 2020 before the COVID-19 related campus closures. Midterm 3 and the Final exam were held online. Students 1 and 2 were enrolled in Section 1. Treat is an indicator for the corresponding online section in 2020. Exam scores are out of 100 points. Clustered standard errors (clustered by section) are shown in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

It appears that more than two students may have cheated on the Final Exam. Table 6 shows how two sections from Spring 2020 performed across exams compared to the previous students who took the very same course. Column 1 compares 2020 Section 1 students to the past students. Similarly, column 2 compares 2020 Section 2 students to past students. Both sections performed worse on Midterms 2 and 3 compared to past students.[17] However on the Final Exam, Section 1 outperformed the past students by approximately 4.3 points while Section 2 still performed slightly worse. This means that on average, a student in Section 1 received almost one higher letter grade on the Final Exam than their past counterparts.

[17] The exams were designed such that the difficulty of each exam goes up moving from the first Midterm to the Final Exam. In addition, all past students received multiple choice questions while the students from Spring 2020 received a similar set of questions with no multiple choice options, but instead were asked to type in their answers.

These findings suggest that combined with incentives and peer-effects in cheating, there may also be a learning process in cheating. There appears to be no evidence of cheating on Midterm 3 - the first online exam - but evidence of cheating exists on the Final Exam, where gains are much more salient to students for any potential improvement in their grades.

4. Conclusion

> *"I am talking about cheating. Unfortunately, no one can be trusted, except maybe for the top players for whom their reputation is the key asset."*
>
> - Arkady Dvorkovich, FIDE President, April 2020.
>
> *(In response) "Wow. So sad. The biggest insult ever to all chess players just surfaced online."*
>
> - Jovan Retronic, International Chess Master, April 2020.

Like doping, cheating cannot be completely eliminated. There always was, is, and will be cheating in face-to-face and online examinations. However, we can (try to) keep it at an expected equilibrium level. In this paper, we first looked at two simple models of face-to-face and online examinations. The theory suggests that cheating should be expected online. Then, we presented evidence of cheating that took place in an online examination in Spring 2020 under COVID-19 lockdowns and made suggestions on how to mitigate cheating based on the experience accumulated by online chess communities in the last two decades.

COVID-19 made online chess much more popular since March 2020, and there is a growing number of online chess tournaments with substantial monetary prizes. This online chess experience is very similar to the experience of many academic instructors. The recent evidence suggests that the problem is not only there, but it is getting worse. In the intermediate Section B (1401–1700) of the recent European Online Chess Championship, 5 out of the top 6 players have been banned for cheating. The comment of International Grandmaster Nigel Short, FIDE Vice-President, on May 25, 2020 is revealing: *"This scourge will not stop until people are criminally prosecuted for fraud."* Whether people indeed be prosecuted or not, one thing is clear: there is no chance to win a prize in an online chess event without proctoring if you do not cheat because you expect that everybody else will cheat and this belief is fulfilled in a bad equilibrium where everybody cheats. Of course, some people are disqualified, but not all.

What does this mean for online exams? If instructors curve their grades, then they create a competition among students similar to what is seen in chess tournaments. Now, each student has more incentives to cheat because if they believe that the rest of the group is cheating, then they must demonstrate better than at least average class performance in order to pass the class. Thus, a student's chance to pass the class without cheating would be very slim. Of course in this case, the cost of cheating goes down because the alternative to not cheating is failing the class. Therefore, any grade curving should not be used for online teaching.

If universities decide to implement online exams with no proctoring in the upcoming semesters, we expect that there will be widespread cheating among students, who will not be penalized since it is almost impossible to present definitive evidence of cheating in an online exam. Unlike in online chess platforms, it will be difficult to implement our second suggestion for public universities that the instructors should be allowed to make the final decision for students who refuse to use a camera. Therefore, universities should implement a uniform online exam policy where a camera capturing each student's computer screen and room is a requirement. A camera will also help to check a student's ID and eliminate the possibility of another person taking the test.[18] For instructors, in addition to not curving any grades, we also suggest to give students less time but easier questions to increase the value of time, making it more costly to cheat.

Fig. 8 shows a rise in cheating in the world's largest online chess platform since shutdowns due to COVID-19 began. Today, several online chess tournaments have a policy that requires participants to use a camera to live-stream and record themselves during the tournament.[19] Fig. 9 shows such setup used in a recent online chess tournament.[20] Players live-stream from a side-angle camera showing their screen and surroundings with their microphone enabled. Even though this method

[18] See a discussion on this issue https://www.michigandaily.com/section/academics/university-faculty-and-students-discuss-academic-integrity-digital-classroom.

[19] See the regulations for a recent online chess event held in Spring 2020 and the regulations for an upcoming online chess event to be held in Fall 2020.

[20] Must be on Zoom (use real name) to be eligible for prizes (side/rear camera angle).

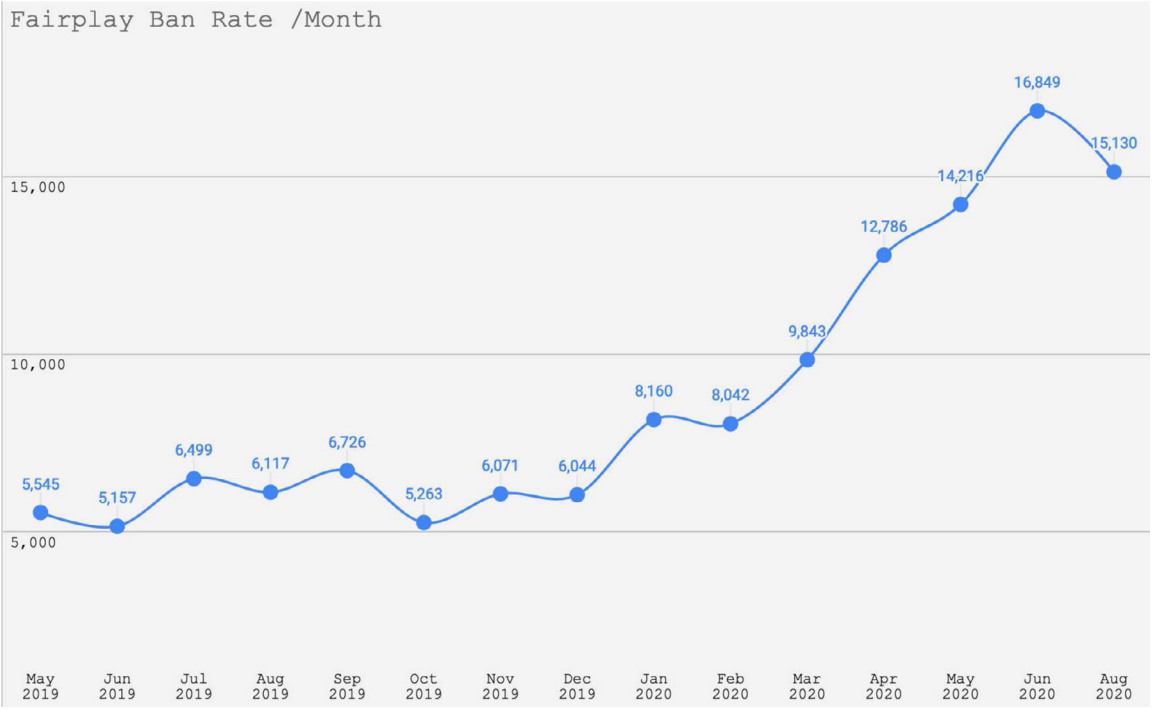
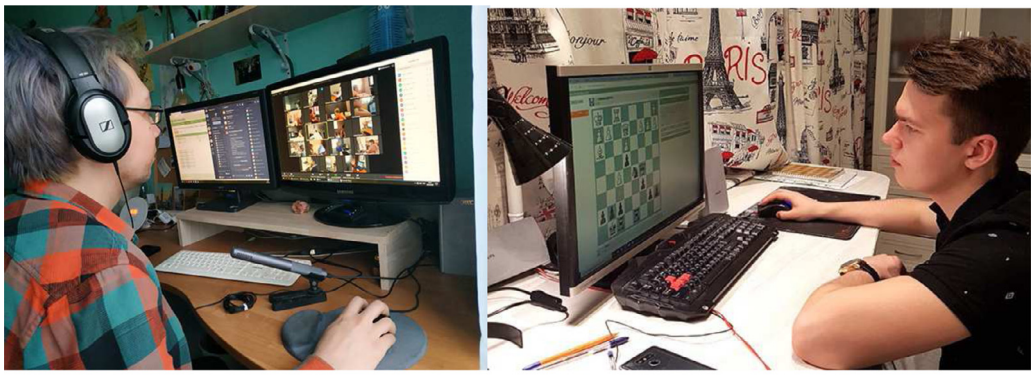


Fig. 8. Account closures on Chess.com since shutdowns due to COVID-19 began *Note:* Source: https://www.chess.com/article/view/online-chess-cheating#false-positives.



(a) An arbiter monitoring chess players (b) A chess player

Fig. 9. Playing chess on Chessking.com. *Note:* Source: https://ruchess.ru/news/report/sokhranyaya_distantsiyu.

cannot eliminate all cheating, we believe it is a great balance between having no proctoring and using online proctoring services.[21]

Of course, the problem is much bigger, as was noted by Peter Heine Nielsen, Coach of World Chess Champion Magnus Carlsen, on May 25, 2020: *"The same could be said about corruption, pre-arranged games, buying of votes, jobs going to friends or political allies instead of an open recruitment procedure based on merits etc. These are big issues for the chess world, not a 1400–1700 online event."*

[21] Online proctoring services are often associated with privacy concerns related to the use and storage of personal data. These services are also criticized for their use of the "AI" and created petitions against using them. See among many articles https://www.cnn.com/2020/08/29/tech/online-school-test-surveillance/index.html, https://www.dailymail.co.uk/news/article-8243637/Creepy-software-used-stop-university-students-cheating-online-exams-amid-coronavirus.html and https://www.vox.com/recode/2020/5/4/21241062/schools-cheating-proctorio-artificial-intelligence.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

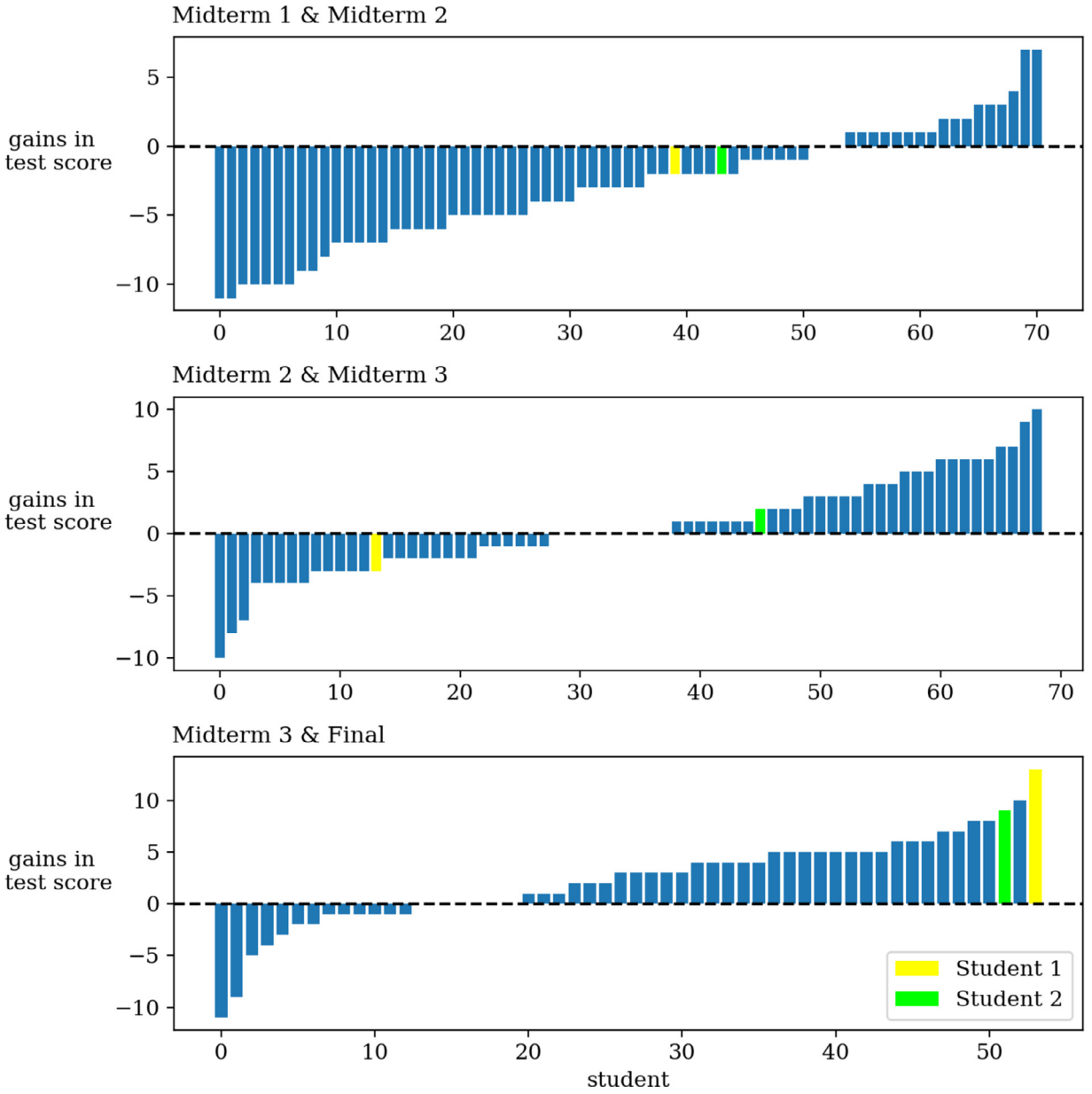


Fig. 10. Comparison of gains in scores for students between exams. *Note:* Midterms 1 and 2 were held face-to-face with proctoring; Midterm 3 and the Final Exam were held online asynchronously without proctoring following COVID-19 related campus closures. Each bar represents a student's test score gains between corresponding exams. Each exam is worth 20 points.

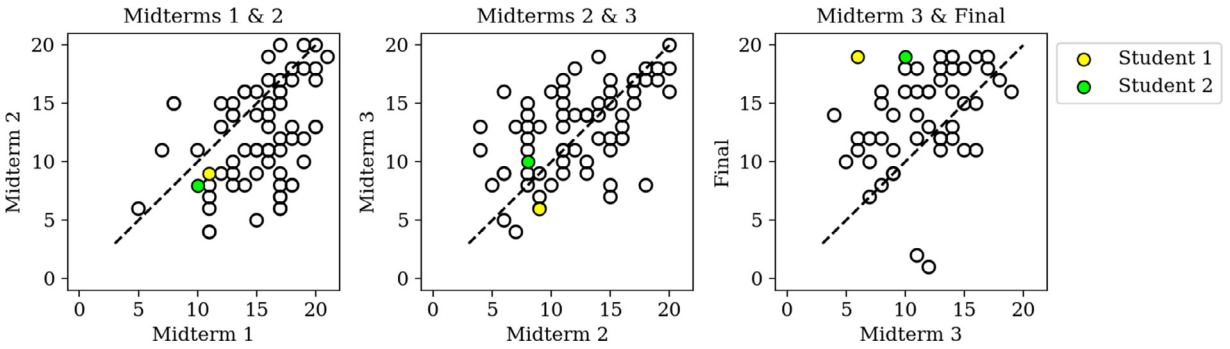


Fig. 11. Comparison of gains in scores for students across exams. *Note:* Midterms 1 and 2 were held face-to-face with proctoring; Midterm 3 and the Final Exam were held online asynchronously without proctoring following COVID-19 related campus closures. Each dot represents a student's performance in corresponding exams. The dashed line is the 45 degree line.

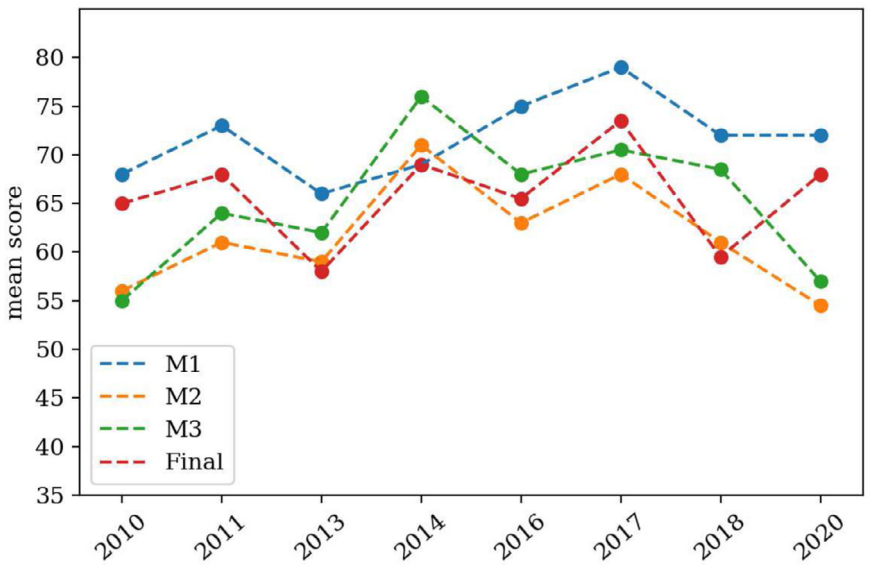


Fig. 12. Performance on exams across years (for students who took the final exam) *Notes:* All exams before 2020 were held face-to-face with in-class proctoring. Midterms 1 and 2 in 2020 were held face-to-face at the beginning of Spring 2020 before the COVID-19 related campus closures. Midterm 3 and the Final exams in 2020 were held online.

References

Alan, S., Ertac, S., Gumren, M., 2020. Cheating and incentives in a performance context: evidence from a field experiment on children. J. Econ. Behav. Organ. 179, 681–701. doi:10.1016/j.jebo.2019.03.015.
Becker, G.S., 1968. Crime and punishment: an economic approach. J. Polit. Econ. 76 (2), 169–217.
Charness, G., Blanco-Jimenez, C., Ezquerra, L., Rodriguez-Lara, I., 2019. Cheating, incentives, and money manipulation. Exp. Econ. 22 (1), 155–177. doi:10.1007/s10683-018-9584-1.
Dee, T.S., Jacob, B.A., 2012. Rational ignorance in education: a field experiment in student plagiarism. J. Hum. Resour. 47 (2), 397–434.
Diedenhofen, B., Musch, J., 2017. Pagefocus: using paradata to detect and prevent cheating on online achievement tests. Behav. Res. Methods 49 (4), 1444–1459. doi:10.3758/s13428-016-0800-7.
Duggan, M., Levitt, S.D., 2002. Winning isn't everything: corruption in sumo wrestling. Am. Econ. Rev. 92 (5), 1594–1605. doi:10.1257/000282802762024665.
Erat, S., Gneezy, U., 2012. White lies. Manage Sci 58 (4), 723–733. doi:10.1287/mnsc.1110.1449.
Fask, A., Englander, F., Wang, Z., 2014. Do online exams facilitate cheating? An experiment designed to separate possible cheating from the effect of the online test taking environment. J. Acad. Ethics 12 (2), 101–112. doi:10.1007/s10805-014-9207-1.
Fischbacher, U., Föllmi-Heusi, F., 2013. Lies in disguise-an experimental study on cheating. J. Eur. Econ. Assoc. 11 (3), 525–547. doi:10.1111/jeea.12014.
Gächter, S., Schulz, J.F., 2016. Intrinsic honesty and the prevalence of rule violations across societies. Nature 531 (7595), 496–499. doi:10.1038/nature17160.
Gneezy, U., 2005. Deception: the role of consequences. Am. Econ. Rev. 95 (1), 384–394. doi:10.1257/0002828053828662.
Jacob, B.A., Levitt, S.D., 2003. Rotten apples: an investigation of the prevalence and predictors of teacher cheating. Q. J. Econ. 118 (3), 843–877. doi:10.1162/00335530360698441.
Karim, M.N., Kaminsky, S.E., Behrend, T.S., 2014. Cheating, reactions, and performance in remotely proctored testing: an exploratory experimental study. J. Bus. Psychol. 29 (4), 555–572. doi:10.1007/s10869-014-9343-z.
Lederman, D., 2020. Best way to stop cheating in online courses? each better. Inside Higher Ed. https://www.insidehighered.com/digital-learning/article/2020/07/22/technology-best-way-stop-online-cheating-no-experts-say-better
Levitt, S.D., Dubner, S.J., 2005. Freakonomics: A Rogue Economist Explores the Hidden Side of Everything. William Morrow & Co, New York, NY, US.

Maggioni, M.A., Rossignoli, D., 2020. Clever little lies: math performance and cheating in primary schools in congo. J. Econ. Behav. Organ. 172, 380–400. doi:10.1016/j.jebo.2019.12.021.
Martinelli, C., Parker, S.W., Pérez-Gea, A.C., Rodrigo, R., 2018. Cheating and incentives: learning from a policy experiment. Am. Econ. J.: Econ. Policy 10 (1), 298–325. doi:10.1257/pol.20150066.
Newton, D., 2020. Another problem with shifting education online: a rise in cheating. Washington Post. https://www.washingtonpost.com/local/education/another-problem-with-shifting-education-online-a-rise-in-cheating/2020/08/07/1284c9f6-d762-11ea-aff6-220dd3a14741_story.html.
Vanberg, C., 2017. Who never tells a lie? Exp. Econ. 20 (2), 448–459. doi:10.1007/s10683-016-9491-2.
Watson, G., Sottile, J., 2010. Cheating in the digital age: do students cheat more in online courses? Online J. Distance Learn. Admin. 13 (1).