25.

DISTANCE EXAMS: CAN TARGETED WARNINGS DISCOURAGE CHEATING?

Marc Humbert and Xavier Lambin

Abstract

During the COVID-19 sanitary crisis of 2020, many exams were hastily moved to online mode. This revived a much-needed debate on the privacy issues of online proctoring of exams, while the validity and fairness of unproctored exams were increasingly questioned. In a randomized control trial, we send a targeted warning to half of the students who were identified as cheaters in previous exams. We then compare their cheating behavior at the final exam to the group of unwarned cheaters. Preliminary results show that the warning proves effective but does not completely annihilates cheating as the cheating strategies of some students become more sophisticated. We conclude that switching traditional exams to online mode should come with proctoring. When proctoring is not possible, credible and effective anti-cheating technologies should be deployed, together with adequate warnings.[*]

1. Introduction

Online education has experienced sustained growth over the past decades. The 2020 global public health crisis suddenly made it ubiquitous and paved the way for even more extensive use in the future. Naturally, these sudden developments stimulated active debates about the benefits of online teaching and the associated risks, in particular the issue of academic dishonesty in distance exams. The stakes go beyond the already crucial issue of fairness in education, as several authors have noted a strong correlation between academic and professional dishonesty.[645]

Following the observation that unproctored online exams result in extensive cheating, several strategies have been proposed.[646] The randomization of questions, when possible, provides satisfactory results but raises the issue of fairness between students facing different sets of questions.[647] It also has technical limits, because an exam designer may not be able to find enough variations of a similar question. Online proctoring is also a popular solution but faces strong public opposition due to concerns over students' access to the necessary technologies (e.g. a webcam or a stable Internet connection) and, most importantly, privacy.[648]

---

[645] D. Becker and others, 'Using the Business Fraud Triangle to Predict Academic Dishonesty among Business Students', *Academy of Educational Leadership Journal*, 10(1) (2006), 37-54; G. H. Brodowsky and others, 'Tolerance for Cheating from the Classroom to the Boardroom: A Study of Underlying Personal and Cultural Drivers', *Journal of Marketing Education*, 42(1) (2019), 23-36.

[646] M. Norris, 'University Online Cheating—How to Mitigate the Damage', *Research in Higher Education Journal*, 37 (2019).

[647] N. I. Nizam and others, 'Scheme for Cheating Prevention in Online Exams during Social Distancing', *Preprints*, (2020), Article 2020040327.

[648] R. Bawarith and others, 'E-Exam Cheating Detection System', *International Journal of Advanced Computer Science and Applications*, 8(4) (2017), Article 4.

We propose another strategy to discourage cheating. Targeting a randomly selected subgroup of the students identified as cheaters on previous assignments (with a probabilistic, algorithmic method), we send a friendly warning stating that their copies were suspicious and reminding them that cheating on the final exam is prohibited. We observed that warnings are effective in reducing cheating, in that warned cheaters behave similarly to non-cheaters. Cheating is not, however, eliminated.

There is a rich and growing literature on academic dishonesty in higher education. A comprehensive overview of the latest developments may be found in the work of McCabe.[649] We are particularly interested in academic dishonesty when exams are taken without proctoring, at a distance. Recent reviews of the literature reveal that between 60% and 90% of students admit to having cheated on such exams.[650] In contrast to most previous work, which is based on anonymous post-exam surveys, we use original technologies to reveal cheating behavior. With our research protocol, cheating is not reported (as in self-reports) but revealed, which eliminates the strong reporting biases of surveys.[651] Furthermore, statistical approaches to the phenomenon may reveal cheating behavior accurately but fail to explain the mechanisms of cheating.[652] In our research, the status of cheater is attributed to specific

---

[649] D. McCabe, L. Treviño, and K. Butterfield, 'Cheating in Academic Institutions: A Decade of Research', *Ethics and Behavior*, 11 (2001), 219-32; D. McCabe, 'Cheating and Honor: Lessons from a Long-Term Research Project', in *Handbook of Academic Integrity,* ed. by T. Bretag (Singapore: Springer Singapore, 2016), pp. 187-98.

[650] Norris, 'University Online Cheating'.

[651] S. Sudman and N. Bradburn, *Response Effects in Surveys: A Review and Synthesis* (Chicago: Aldine, 1974); J. Kerkvliet and C. L. Sigmund, 'Can We Control Cheating in the Classroom?', *The Journal of Economic Education*, 30(4) (1999), 331-43.

[652] I. J. M. Arnold, 'Cheating at Online Formative Tests: Does It Pay Off?', *The Internet and Higher Education,* 29 (2016), 98-106; R. J. Fendler, M. Yates, and

individuals, which allows for a very detailed analysis of cheating strategies.

Our study also contributes to an active literature on the effectiveness of 'nudges'.[653] Nudges have been widely analyzed in the context of consumer choices, but also in the field of education.[654] Damgaard and Nielsen described various initiatives and showed that a necessary condition for nudges to be effective is that the architect has a sufficient understanding of the behavioral mechanism underlying cheating.[655] In this chapter, we study the effect of a simple, inconsequential warning on cheating behavior in subsequent exams. The warning only informs some students that their professors suspect that they cheated on the preparatory test and reminds them that cheating will be penalized on the final exam. Contrary to the studies by Bing and others and Corrigan-Gibbs and others, the treated group receives individual warnings, which makes the threat of being identified as a cheater more realistic.[656]

J. Godbey, 'Observing and Deterring Social Cheating on College Exams', *International Journal for the Scholarship of Teaching and Learning*, 12(1) (2018), Article 4.

[653] R. H. Thaler and C. R. Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness* (Revised and expanded edition) (New York: Penguin Books, 2009).

[654] H. Allcott and S. Mullainathan, 'Behavior and Energy Policy', *Science*, 327(5970) (2010), 1204-05; P. J. Ferraro, J. J. Miranda, and M. K. Price, 'The Persistence of Treatment Effects with Norm-Based Policy Instruments: Evidence from a Randomized Environmental Policy Experiment', *American Economic Review*, 101(3) (2011), 318-22; J. Beshears and others, 'The Effect of Providing Peer Information on Retirement Savings Decisions', *The Journal of Finance*, 70(3) (2015), 1161-201.

[655] M. T. Damgaard and H. S. Nielsen, 'Nudging in Education', *Economics of Education Review*, 64 (2018), 313-42.

[656] M. N. Bing and others, 'An Experimental Investigation of an Interactive Model of Academic Cheating Among Business School Students', *Academy of Management Learning and Education*, 11(1) (2012), 28-48; H. Corrigan-Gibbs

2. Experimental setting

Our experimental setting was ideally suited to the study of cheating. All aspects of the notorious 'fraud triangle' were present and reinforced by the 2020 lockdown.[657] In 2006, Becker and others reported that business students are consistently at the top of the list of students most likely to cheat.[658] That was our population. But the conditions of our exam were exceptional. First, cheating was exceptionally easy because of the particular circumstances of the COVID-19 public health crisis. The entire country was under a strict lockdown preventing any unnecessary travel. All institutions, especially schools and universities, were closed. In this context, all courses and exams at the institution where the study was conducted were moved to online mode, without any possibility of monitoring. Despite the physical distance between students, communication channels such as online messaging were available to them, raising fears of wide-ranging collaboration. Second, there were strong incentives to cheat: grades in the first year are a key determinant of access to Erasmus-type programs in subsequent years. Finally, travel restrictions made cheating attractive. Rumors on the lines of 'everyone cheats, let's do it' were likely to spread...

Our database consists of examination papers from 644 undergraduate students at a French business school. We examined their performance in a series of five tests in a programming class in spring 2020. Together, these tests accounted for a very small proportion of the final grade (10%). They were used for pedagogical and participation purposes, as well as to prepare for a final exam that accounted for the bulk of the

---

and others, 'Deterring Cheating in Online Environments', *ACM Transactions on Computer-Human Interaction*, 22(6) (2015), 28:1-23.

[657] M. Ramos, 'Auditors' Responsibility for Fraud Detection', *Journal of Accountancy,* 195 (2003), 28-36; Becker and others, 'Using the Business Fraud Triangle'.
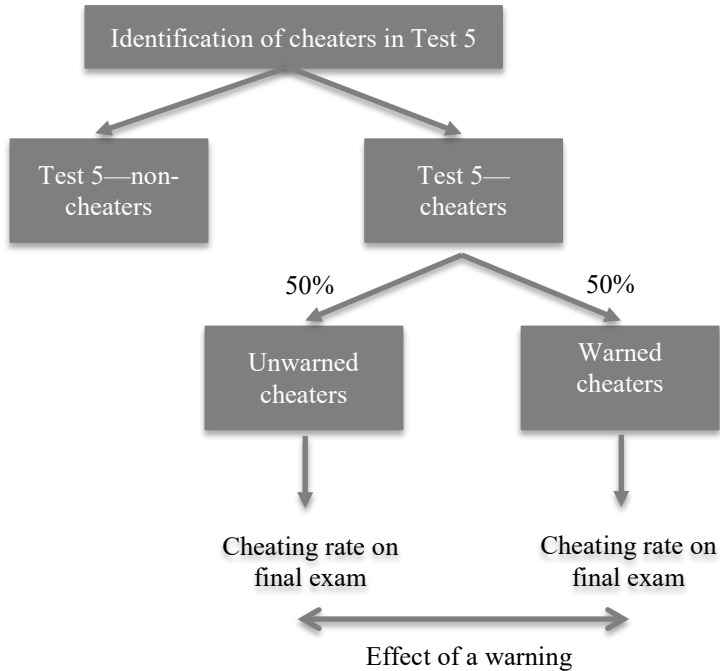
[658] Becker and others, 'Using the Business Fraud Triangle'.

final grade (90%). The form of the tests and the final exam was similar: it consisted in writing small pieces of code on an online platform and understanding written code. They differed only in their duration and subject matter. The final exam covered the whole course and lasted one and a half hours, while the tests covered chapters of the course and lasted about forty-five minutes. For the first four tests, we told students that they could collaborate if it helped them to learn more, but for the last assignment ('Test 5'), we strongly encouraged them to work on their own to prepare for the final exam, which was to be written individually. In this chapter, we equate cheating with collaboration (two or more students taking the test together or exchanging answers), as such collaboration was explicitly forbidden in Test 5 and the final exam.

We therefore used the results of Test 5 to classify students into two categories: those who cheated and those who did not cheat. We used a completely unintrusive technology to identify collaboration: we analyzed syntax (both textual and algorithmic) to identify suspicious similarities between submitted papers. The method is probabilistic, but following Test 5, we were able to identify 230 assignment cheaters out of the 644 students with a high degree of confidence. Between this last assignment and the final exam, a standard email was sent to all students reminding them of the rules of the exam and the penalty policy for cheating. In addition to this general email, half of the cheating sample also received a warning stating that they had been identified and placed on a watch list. There was no sanction, but the warning was a reminder that similar behavior during the exam would be penalized. This was the test group. The other half of the cheating group was not warned and received only the information sent to all students. This was the control group. The students who were not identified as cheaters in Test 5 constituted a 'reference' group, which we used as a benchmark. The standard email was sent two days after Test 5 and five days before the

final exam. The treatment (sending the warning to half of the cheaters) was done a few minutes after the standard email (see Figure 1).
Figure 1: The experimental protocol.

```
              ┌─────────────────────────────────────┐
              │  Identification of cheaters in Test 5 │
              └─────────────────────────────────────┘
                    ↙                        ↘
        ┌──────────────┐              ┌──────────────┐
        │ Test 5—non-  │              │  Test 5—     │
        │  cheaters    │              │  cheaters    │
        └──────────────┘              └──────────────┘
                            50%        ↙        ↘   50%
                    ┌──────────────┐      ┌──────────────┐
                    │  Unwarned    │      │   Warned     │
                    │  cheaters    │      │  cheaters    │
                    └──────────────┘      └──────────────┘
                           ↓                     ↓
                   Cheating rate on      Cheating rate on
                     final exam            final exam
                           ←───────────────────→
                            Effect of a warning
```

The main purpose of this report is to present a preliminary analysis of the treatment group's response to the treatment, in terms of cheating behavior on the final examination.

3. Plagiarism detection methods

We used two different types of detection methods: textual comparisons of students' answers and 'trick questions'. The second type was not used for test 5, but we kept it for the final exam because exam cheating is a learning game between students and professors. When professors introduce new technologies, students quickly learn how to avoid them.

The analysis of general textual answers has been intensively studied in the literature and has generated many anti-plagiarism solutions. They are not perfect because in the 'learning game' students find ways around these systems.

Our case was more specific because textual answers did not correspond to a natural language but to a programming language with its own particular syntax. This specific case has been already studied by professors of computer science but, in order to avoid attacks by students (avoidance strategies) and tailor our detection strategy to the specific nature of our exams, we developed our own text-based probabilistic algorithms.[659]

Papers that showed a high degree of similarity were considered to indicate a cluster of cheaters. Its constituent members were therefore labeled as 'cheaters'. It is useful at this point to note that this strategy only gives a probabilistic estimate of cheating behavior. It is therefore an effective prevention tool, but is of little use in terms of sanctions (see the discussion), whence the need for another method.

The second method, which we used only for the final exam, consisted of a classic approach, random questions, to which we added an original touch. For each specific question (displayed to all students as, say, 'question 8'), we randomly assigned a slightly different version of the question to each student (student A gets question 8A, student B gets question 8B, etc.). A cheater will therefore give an incorrect answer, but one that matches the correct answer to another version of the question. We designed our versions such that it is extremely unlikely that a person would give the right answer to another version without external help. Moreover, the difference between versions was only visible to particularly attentive eyes, or students who expected such a strategy to

---

[659] O. Karnalim, 'Python Source Code Plagiarism Attacks on Introductory Programming Course Assignments', *Themes in Science and Technology Education*, 10 (2017), 17-29.

be deployed. Since there was no precedent of this type of 'trick question' in the educational institution, this was highly unlikely. Cheaters were classified as such if they gave an answer that corresponded to the correct answer to another version of the question. Because we used this method only for exams and not for assignments, students could not learn to avoid detection.[660]

4. Results

Our preliminary results suggest that there was some degree of cheating (about 14%) even among the reference group of students who had not been identified as cheaters on Test 5. This is not surprising, as the stakes on an exam are much higher than those for assignments and students had more time to organize their collaboration. We will call this level of cheating the 'baseline' level. More importantly, the preliminary regressions allow us to assess the effect of a warning on cheaters. Having cheated in assignments increases the probability of cheating on the final exam by 25% to 30%, while being warned more than offsets this effect and results in a 3% to 5% decrease in the probability of cheating compared to the baseline. This suggests that warnings are very effective at curbing cheating.

5. Discussion

A central issue in research on academic dishonesty is the estimation of the prevalence of cheating. With the increasing prevalence of online exams, we believe the need for examination methods that allow for creativity while being resistant to cheating will continue to increase in the coming years. This paper proposes solutions to increase their

---

[660] Note that the details of our statistical analyses are presented in a longer document than this chapter and are available from the authors upon request.

robustness. As we mentioned above, our approach is based on a statistical analysis of exam responses. This approach is facilitated by the structure of our exam, which is based on the assessment of skills such as writing or understanding code. The questions are open-ended, which allows suspicious similarities to be confidently identified, as opposed to multiple-choice questions, which are by nature closed-ended and leave little room for student creativity.

However, our solution is only partial. We must keep in mind that the main objective is to eliminate or at least limit cheating. The two main levers of action are prevention and repression. We have shown that targeted warnings make prevention much more effective than traditional warnings. However, we recognize that prevention may not be sufficient. Repression (sanctions) may be necessary to reinforce preventive actions. However, although our statistical analysis provides clear indications of the probability of fraud, it does not provide sufficient evidence to trigger a sanction.

This paper examines the effect of a single treatment (a warning to cheaters). This choice was made to ensure the statistical significance of our results, given the expected effect sizes and our sample size. Our warnings proved effective for a large proportion of the cheating group, but we do not know whether the effect is sustainable or whether, if such warnings were repeated too often, they would remain credible over the long term. One avenue of research would be to determine how often messages are needed to develop honest behavior during examinations. On the other hand, the repetitive effect of messages can also be tiring and lead to an opposite effect to the desired one. Further work with researchers from the psychosocial sciences would be useful in this regard.[661] In future research, we also aim to test other treatments such as

---

[661] D. Courbet, I. Milhabet, and D. Priolo, 'Communication persuasive: Effets de la vividité et de la répétition des messages sur l'optimisme comparatif et sur

training students on exam integrity and anti-cheating strategies. Indeed, behavioral change can be achieved through the use of commitment or persuasion.[662]

6. Conclusion

The COVID-19 crisis has given digital technologies a unique opportunity to highlight how they can contribute to our education systems. These technologies will become increasingly important in education not only in times of crisis, but also as a new standard for teaching. However, the crisis also revealed some of the limits of online education, including the viability of online exams. It is essential to increase confidence in the results of exams, and therefore degrees, and to limit the sense of injustice students who do not cheat are likely to feel.

The preliminary results of our randomized control experiment suggest that a credible, effective mechanism could be deployed to discipline students and restore the validity and fairness of exams. Advance warnings are particularly effective in inducing honest behavior. In light of these preliminary results, we argue that the judicious use of warnings represents a promising alternative to proctoring, especially when proctoring is not possible for either practical or ethical reasons.

We consider it necessary to develop this new line of research, as it is essential to put safeguards in place to avoid the temptation to commit massive fraud. This does not prevent the development of a reflection process on the teaching of integrity to students—in fact, quite the contrary.

---

l'intention comportementale', *Revue Internationale de Psychologie Sociale*, 14 (2001), 163-75.

[662] C. A. Kiesler, *Psychology of Commitment: Experiments Linking Behavior to Belief* (New York: Academic Press, 1971).

Bibliography

Allcott, H., and S. Mullainathan, 'Behavior and Energy Policy', *Science*, 327(5970) (2010). https://doi.org/10.1126/science.1180775

Arnold, I. J. M., 'Cheating at Online Formative Tests: Does It Pay Off?', *The Internet and Higher Education,* 29 (2016).

Bawarith, R., and others, 'E-Exam Cheating Detection System', *International Journal of Advanced Computer Science and Applications*, 8(4) (2017). https://doi.org/10.14569/IJACSA.2017.080425

Becker, D., and others, 'Using the Business Fraud Triangle to Predict Academic Dishonesty among Business Students', *Academy of Educational Leadership Journal*, 10(1) (2006). https://search.proquest.com/openview/3d7dae5af62b010d3f002 d2de4dc4841/1?pq-origsite=gscholar&cbl=38741

Beshears, J., and others, 'The Effect of Providing Peer Information on Retirement Savings Decisions', *The Journal of Finance*, 70(3) (2015). https://doi.org/10.1111/jofi.12258

Bing, M. N., and others, 'An Experimental Investigation of an Interactive Model of Academic Cheating Among Business School Students', *Academy of Management Learning and Education*, 11(1) (2012). https://doi.org/10.5465/amle.2010.0057

Brodowsky, G. H., and others, 'Tolerance for Cheating from the Classroom to the Boardroom: A Study of Underlying Personal and Cultural Drivers', *Journal of Marketing Education*, 42(1) (2019). https://doi.org/10.1177/0273475319878810

Corrigan-Gibbs, H., and others, 'Deterring Cheating in Online Environments', *ACM Transactions on Computer-Human Interaction*, 22(6) (2015). https://doi.org/10.1145/2810239

Courbet, D., I. Milhabet, and D. Priolo, 'Communication persuasive: Effets de la vividité et de la répétition des messages sur l'optimisme comparatif et sur l'intention comportementale', *Revue Internationale de Psychologie Sociale*, 14 (2001).

Damgaard, M. T., and H. S. Nielsen, 'Nudging in Education', *Economics of Education Review*, 64 (2018). https://doi.org/10.1016/j.econedurev.2018.03.008

Fendler, R. J., M. Yates, and J. Godbey, 'Observing and Deterring Social Cheating on College Exams', *International Journal for the Scholarship of Teaching and Learning*, 12(1) (2018). https://doi.org/10.20429/ijsotl.2018.120104

Ferraro, P. J., J. J. Miranda, and M. K. Price, 'The Persistence of Treatment Effects with Norm-Based Policy Instruments: Evidence from a Randomized Environmental Policy Experiment', *American Economic Review*, 101(3) (2011). https://doi.org/10.1257/aer.101.3.318

Karnalim, O., 'Python Source Code Plagiarism Attacks on Introductory Programming Course Assignments', *Themes in Science and Technology Education*, 10 (2017).

Kerkvliet, J., and C. L. Sigmund, 'Can We Control Cheating in the Classroom?', *The Journal of Economic Education*, 30(4) (1999). https://doi.org/10.2307/1182947

Kiesler, C. A., *Psychology of Commitment: Experiments Linking Behavior to Belief* (New York: Academic Press, 1971).

McCabe, D., 'Cheating and Honor: Lessons from a Long-Term Research Project', in *Handbook of Academic Integrity,* ed. by T. Bretag (Singapore: Springer Singapore, 2016), pp. 187-98. https://doi.org/10.1007/978-981-287-079-7_35-1

McCabe, D., L. Treviño, and K. Butterfield, 'Cheating in Academic Institutions: A Decade of Research', *Ethics and Behavior*, 11 (2001). https://doi.org/10.1207/S15327019EB1103_2

Nizam, N. I., and others, 'Scheme for Cheating Prevention in Online Exams during Social Distancing', *Preprints*, (2020). https://doi.org/10.20944/preprints202004.0327.v1

Norris, M., 'University Online Cheating—How to Mitigate the Damage', *Research in Higher Education Journal*, 37 (2019). https://eric.ed.gov/?id=EJ1233121

Ramos, M., 'Auditors' Responsibility for Fraud Detection', *Journal of Accountancy,* 195 (2003).

Sudman, S., and N. Bradburn, *Response Effects in Surveys: A Review and Synthesis* (Chicago: Aldine, 1974).

Thaler, R. H., and C. R. Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness* (Revised and expanded edition) (New York: Penguin Books, 2009).