Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
The role of prior warnings when cheating is easy and punishment is credible

Marc Humbert\textsuperscript{a}, Xavier Lambin\textsuperscript{b,∗}, Eric Villard\textsuperscript{a}

\textsuperscript{a} Grenoble Ecole de Management, 12 rue Pierre Semard, Grenoble F-38000, France
\textsuperscript{b} ESSEC Business School and THEMA, 3 Av. Bernard Hirsch, B.P. – 50105, Cergy 95021, France

\textbf{A R T I C L E  I N F O}

Article history:
Received 11 September 2020
Revised 13 September 2021
Accepted 26 November 2021
Available online 4 December 2021

\textbf{JEL codes:}
D8
I21
I23

\textbf{Keywords:}
Education
Online exams
Unproctored exams
Prior warnings
Nudges
Randomized control trial

\textbf{A B S T R A C T}

During the COVID-19 sanitary crisis, many exams were hastily moved to online mode. This revived a much needed debate over the privacy issues associated with online proctoring of exams, while the validity and fairness of unproctored exams were increasingly questioned. With a randomized control trial, we estimate the effectiveness of prior warnings as a means of discouraging academic dishonesty in exams. We use original, non-intrusive technologies to surreptitiously identify cheating in a series of unproctored assignments and send a targeted warning to half of the students who were identified as cheaters. We then compare their cheating behavior on the final exam with the behavior of the group of unwarned cheaters. The warning proves effective but does not completely eliminate cheating, as some students’ cheating strategies become more sophisticated following issuance of the warnings. We conclude that switching traditional exams to online mode should be accompanied by proctoring. When proctoring is not possible, credible and effective anti-cheating technologies should be deployed together with adequate warnings.

© 2021 Elsevier B.V. All rights reserved.

\textbf{1. Introduction}

Online education has experienced sustained growth in recent decades. The 2020 global sanitary crisis caused by the COVID-19 virus suddenly rendered remote learning ubiquitous and paved the way for even more extensive use in the future. Naturally, these abrupt developments stimulated active debates over the benefits of online teaching and the associated risks, in particular the issue of academic dishonesty on 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 (Becker et al., 2006; Brodowsky et al., 2020).

Following the observation that unproctored online exams result in extensive cheating (Norris, 2019; Holden et al., 2021), several strategies have been proposed. The randomization of questions (see e.g. (Li et al., 2021)), when implementable, provides satisfying results but raises issues of fairness between students facing distinct sets of questions. This process also has technical limits, because an examiner may not be able to find enough variations of a given question to avoid repetition. Online proctoring is also a popular solution (Bawarith et al., 2017; Chirumamilla and Gutterm, 2019; Cluskey et al., 2011; Idemudia et al., 2016; Mellar et al., 2018; Varble, 2014) but faces strong public opposition reflecting concerns over students’ access to the necessary technologies (such as a webcam or a stable internet connection) and, most importantly, privacy.

We took advantage of the specific conditions under which online exams were administered during the COVID-19 crisis, where cheating was predictably widespread, to develop a new strategy to discourage cheating. The strategy, based on friendly warnings sent to students, is respectful of student privacy and does not require specific equipment. In this research, we seek an answer to the following research question “can targeted warnings discourage students from cheating on exams?”. We restrict our analysis to cheating in the form of illicit sharing of information between students. To that aim, we surreptitiously analyze similarities between copies of pre-exam assignments and various trick questions

\textsuperscript{∗} Corresponding author.
E-mail address: lambin@essec.edu (X. Lambin).

\url{https://doi.org/10.1016/j.infoecopol.2021.100959}
0167-6245/© 2021 Elsevier B.V. All rights reserved.
to identify (illicit) collaborations. Students are not aware of these techniques, which makes very accurate identification of cheaters and their strategies possible. To a randomly selected subgroup of students identified as cheaters on the assignments, we send a friendly warning stating that their copies were suspicious and reminding them that cheating on the final exam is prohibited. Mean comparisons indicate that only 9% of the members of our treated group (assignment-cheaters who received a warning) cheated on the final exam, down from 40% in the control group (unwarned assignment-cheaters). Controlling for various variables, the warning sent to cheaters decreases the cheating rate by 23 percentage points. By design of the experiment, this effect is causal. A warned cheater has a probability to cheat on the final exam 4 percentage points lower than a similar student who was not identified as assignment-cheater. We conclude that warnings are effective in discouraging cheating, insofar as warned cheaters behave similarly to non-cheaters. Cheating is, however, not eliminated entirely. We observe a similar effect when we restrict the analysis to “leaders” (students who gave their exam copies to other students). This effect is, however, subject to measurement errors, and properly testing leadership status would require a larger dataset to gain significance.

To the best of our knowledge the present paper is the first to exploit observational methods to analyze the dynamics of cheating in a series of distance exams. It is the first to analyze how suspected students respond to a targeted nudge warning. The experiment was pre-registered at the AEA’s Social Science Registry. The rest of the paper is organized as follows. Section 2 reviews the literature. In Section 3 we describe our experimental setting. In Section 4 we report our results. Section 5 provides a discussion of our results and we conclude in Section 6.

2. Literature review

This paper lies at the nexus between three very active streams of the literature.

First, it sheds some light on the prevalence of cheating in higher education. All aspects of the notorious “fraud triangle” (Ramos 2003; Becker et al., 2006; King et al., 2009) are present in our experimental setting and strengthened by the strict lockdown of Spring 2020 in France. The fraud triangle outlines three components that contribute to increasing the risk of fraud. The first component is the “opportunity”. Cheating was exceptionally easy in the particular circumstances of the COVID-19 sanitary crisis. In Spring 2020, the entire country was under a severe lockdown, preventing any non-motivated trips. All institutions and in particular schools and universities were closed. Within that context all courses and exams in the institution where the experiment took place were moved to online mode, with no possibility of proctoring. Despite the physical distance between students who were all strictly confined at home, all channels of communication such as online messaging were accessible. This makes collaboration unusually easy. The second component “incentives” was also particularly salient: grades are a strong determinant of exchange opportunities in following years. These opportunities were likely to become scarcer as a result of anticipated travel restrictions, making cheating an appealing option. Third component is “rationalization”. Anecdotal feedback from students support the claim of a moral disengagement (Bandura, 2002) where cheating is rationalized by the fact that “everyone else does it”. This claim that was also widely supported by specialized and general media during the sanitary crisis. Finally, Becker et al. (2006) and McCabe and Treviño (1995) have reported that business students, which is our population, are consistently near the top of the rankings of students who are most likely to cheat.

In light of the fraud triangle, the reasons why students may cheat are quite clear. However, the mechanism through which students react to our warning is less evident. Our methodology does not allow attribute causality to a specific mechanism, but the literature described a few plausible explanations: the warning may make the threat more realistic (Bing et al., 2012), but also have a normative effect on student’s perception of rules (Cialdini, 2004), or focus their attention on the abnormality of their behavior (Cialdini, 2003).

Second, the paper offers a methodological contribution to a rich and expanding literature on curbing academic dishonesty in higher education. Academic dishonesty may take many forms and taking account of this variety goes beyond the scope of this report. A comprehensive overview of the main mechanisms of cheating may be found in (McCabe et al., 2001; McCabe and Bretag, 2016). We focus on cheating in the form of seeking outside help during an exam. Such cheating has become more prevalent and harder to detect over the past decades due to increased availability of information on the internet (Scanlon, 2003). The issue is likely to be even more acute when exams are taken at a distance without proctoring, as several studies have found that participants are less honest when they interact online than when they interact face-to-face (Rockmann and Northcraft, 2008; Van Zant and Kray, 2014). Reviews of literature from Norris (2019) and Holden et al. 2021 reveal that between 60% and 90% of students admit to have cheated on online exams. Contrary to most previous studies that are based on anonymous surveys administered after exams, we use original technologies to reveal cheating behavior. This brings about several improvements over methodologies that are commonly found in the existing literature. First, cheating is not stated (as in self-reports) but revealed, which eliminates the strong declarative biases of surveys, as reported in general settings by Sudman and Bradburn (1974) and Kerkvliet and Sigmund (1999). The bias is particularly strong in the case of cheating (Bing et al., 2012). With observational data, interesting statistical approaches at a group-level such as those found in Harmon and Lambrinos (2006), Arnold (2016), D’Souza (2017), and Fendler et al. (2018) reduce this bias but fail to explain the cheating mechanisms.

In the present paper, cheating status is attributed to specific individuals, which allows us to analyze the cheating strategies in considerable detail.

Third, our paper contributes to the literature on the effectiveness of “nudges” (Thaler and Sunstein, 2008). Nudges have been widely analysed in the context of consumer choice (Allcott and Mullainathan, 2010; Ferraro et al., 2014; Beshears et al., 2010) but also showed encouraging results in education. As an example, Zamir et al. (2017) show that adequate deadline management by professors can encourage students self-benefiting behavior such as limiting procrastination and forgetfulness. The careful design of the nudge is a key determinant of its success, as is the prior understanding of the behavioral mechanism behind cheating (Damgaard and Nielsen, 2018). Auferoth (2020) shows that the effectiveness of reminders during exam preparation highly depends on the backgrounds of the students. Most studies on the effect of nudges of the propensity to cheat on exams, concern the “code of honor”. This rich literature shows strikingly mixed results (McCabe et al., 1999; Arnold et al., 2007; Konheim-Kalkstein et al., 2008; Mazur et al., 2008), making the “honor-code” at best a partial solution. Indeed, in a review of 63 economic and psychological studies, Rosenbaum et al. (2014) conclude that there “ap-

2 The surreptitious method has the advantage of not relying on students’ candor but makes cheating difficult to prove with certainty. Further, it may seem like entrapment. For this reason, the similarity analysis was used only for the purpose of this article and did not result in consequences for students.

3 RCT ID: AEARCTR-0005696. Available at https://www.socialsciencregistry.org/trials/5696
pear[s] to be a consistent proportion of unconditional cheaters and noncheaters [ . . . ], with the honesty of the remaining individuals being susceptible to a range of variables, most notably monitoring and intrinsic lying costs” (Rosenbaum et al. 2014).

In the present paper, we reinforce the monitoring pressure on students suspect of cheating and implement an original nudge in the form of a warning prior to the exam. The warning informs some of the students that their professors suspect they cheated on the assignments and reminds them that cheating on the final exam will be penalized. Close to our work, Bing et al. (2012) and Corrigan-Gibbs et al. (2015) show that pre-exam warnings produce significantly greater reduction in cheating than “honor codes”. However, the warning is untargeted. In our study, members of the treated group receive individual warnings, which makes the threat of being caught considerably more realistic. And indeed we observe that our simple and inconsequential but targeted warning is remarkably more effective than the untargeted warnings of Bing et al. (2012) and Corrigan-Gibbs et al. (2015). We finally note that the high benefits (in the form of curbing cheating) and relatively low costs (in the form of designing our detection technology) of our strategy, is one of the main practical appeals of nudges (Madrian and Shea, 2001; Thaler and Benartzi, 2004).

3. Experimental setting

Our data consist of exam copies for 644 undergrad students in a French business school. We examine their performance in a series of five tests in a programming class in Spring 2020. These tests account for a very small share of the final grade (10%). They are used for pedagogical and participation purposes as well as in preparation for a final exam that accounts for most of the final grade (90%). The form of the test and exam are similar and consist in writing small pieces of code on an online platform. They differ only in length and topics. The exam covers the whole course and lasts 1 h and 30 min, while the tests cover course chapters and last about 60 min each. While students can do the tests any time over a 1-week time window, the final exam is synchronous (same time slot for all students). In the present paper we describe cheating to collaboration, i.e. situations when two or more students take the examination together or exchange answers. We report the communications between the professors and the students throughout the experiment: collaboration between students is very explicitly disallowed on both the tests and the exam.

We used copies of the last assignment before the exam (“Test 5”), that was very close in format to the final exam, to categorize students as assignment-cheaters or assignment-non-cheaters. We used a totally non-intrusive technology to identify collaboration: we analysed the syntax to identify suspicious similarities between submitted copies. The method, which is based on exact text matching, Gestalt Pattern Matching (for textual syntax similarity) and Syntax Analysis (for abstract code similarity) and random questions, is described in more detail in Appendix A2. The method is probabilistic but we were able, with a high degree of confidence, to identify 233 assignment-cheaters of the 644 students. Between this last assignment and the final exam, a standard email was sent to all students reminding them of the rules of the exam as well as the sanction policy for cases of cheating. In addition to this email, half of the cheaters (117 students selected randomly from among the 233 cheaters) also received a warning stating that they had been identified and put on a watch list. There was no sanction applied at this point, but the warning reminded students that such behavior at the exam would be sanctioned. This is our treated group. The other half of the assignment-cheaters were not warned and received the same standard information as any other student. This constitutes our control group. The students who were not identified as cheaters on Test 5 constitute a “baseline” group, which we use 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 applied a few minutes after the standard email. The use of several distinct identification techniques for Test 5 and the exam (where we added a few “trap questions” described in Appendix A2) made it possible to circumvent sophisticated cheating strategies (Bachore et al. 2016). The messages sent to students and more detailed regarding our identification techniques can be found in Appendices A1 and A2, respectively. Table 1 displays our descriptive statistics. Though we believe the labels are transparent, a detailed description may be found in Appendix A3.

A few observations are in order. First, we observe that some degree of Exam-cheating (14%) was observed even within the benchmark group of students who did not cheat on Test 5. This is not surprising, as the stakes of the final exam are much higher than those of the assignments and student were allotted more time to organize for collaboration. We refer to this level of cheating as the “baseline” level. Second, not all unwarned assignment-cheaters cheated on the final exam (40%). Again, this is not surprising because informational spillovers might have happened: an assignment cheater may have received a warning and told other students that cheating was credibly traced. This means that some students in the control group may have adjusted their cheating behavior in reaction to the warning sent to their classmate (e.g. with less or more subtle cheating). One may expect such informational spillovers especially within cheating group. Similarly a student who was warned might have decided to stop sharing information with their cheating group at the final exam. Some of (unwarned) members of this group might have stopped cheating too, for lack of a source of information.

The last two columns of Table 1 confirm that our control and treated groups are very similar. The main purpose of this paper is to analyze the response of the treated group to the treatment with respect to cheating behavior associated with the final exam. We can also analyze the behavior of students who passed their answers to other students in Test 5 (the “leaders”), as opposed to students who received these answers (the “followers”). A leader is defined as a student who first submitted answers in the Test 5 – which students could hand-in over a 1-week period— that were then also submitted by other students. Leaders handed-in their Test-5 on average 6.5 h earlier than followers. The secondary purpose of the study is to analyze how leaders respond to the warning. The random selection of treated students was clustered by leadership status, enabling us to treat half of the 60 leaders and keep the other half in the control group.

To best account for the binary nature of the outcome variable, we use a logistic regression model. The classic model assumes that student’s binary decision to cheat or not is governed by a latent (unobserved) variable Exam_cheat* = β0X_i + ε_i, where ε_i is an error distributed by the standard logistic distribution, and X_i is a vector of individual-specific explanatory variables. The probability of experiencing the event is in this model the probability that the latent variable exam_cheat* is larger than zero. As such, the logistic regression can be understood simply as finding the β0 parameters
4. Results

The main results of our experiment are summarized in Table 2. Because the course was taught in classes of around 40 students, we use robust standard errors clustered at the class level. Specification (1) does not include any controls. Specification (2) controls for the majors of the students, their gender and whether they handed-in the Test 5. Specification (3) also controls for class fixed effects. Specifications (4) and (5) correspond to specifications (2) and (3), respectively, but also control for leadership status. For interpretability, the table reports average marginal effects. We use “non-cheaters” as a baseline, which means all the coefficients should be interpreted as the marginal effect of the parameter relative to the cheating rate among “non-cheaters”.

Our preferred specification is Specification (3), as it controls for class fixed effects and excludes the leader/follower distinction, which lacks significance and may be used only as suggestive evidence for future work.

These regressions enable us to assess the effects of a warning on cheaters. In all specifications, the magnitude of the coefficient (in absolute terms) is greater for was_cheater_got_warning, than for was_cheater; having collaborated illicitly on assignments increases the probability of cheating on the final exam by 23% (coefficient of “Cheated on assignment”, Specification (3)). Being warned more than offsets this effect, as the marginal effect of “cheated on assignment and got warned” is −27%. This means that warned cheaters have a 4% lower probability of cheating relative to the baseline level (that of students who did not cheat on Test 5). This suggests that prior warnings are highly effective at discouraging cheating.

Even though there is no statistical significance, the effect seems slightly more pronounced for the cheaters who were leaders of their cheating group (specifications 4 and 5). To ascertain this, further work with a larger sample and more accurate identification of cheating clusters would however be needed.

Because our identification methods are probabilistic, we acknowledge that other specifications could have been used, as well as alternative criteria for distinguishing between cheaters and non-cheaters. Appendix A4 offers a few robustness checks that confirm that our results convey to a broad range of sensible criteria.

5. Discussion

As we have noted, our approach is based on a statistical analysis of exam responses. A natural question is the extent to which our methodology can be applied in different exam settings and formats. Indeed, our approach is facilitated by the structure of our exam, which is based on evaluating competencies such as code writing and code understanding. The questions are open-ended, which enables us to identify suspicious similarities confidently, as opposed to multiple-choice questions where the questions are by nature closed-ended and leave little room for student creativity. Because we tailored our detection strategies to our specific exam format, the methodology may not readily apply to any other exam format. Yet, the increased availability and democratization of sophisticated machine learning techniques heralds the possibil-
Table 2
Effects of warnings on cheating, dependent variable: student cheats at final exam.

|                                | (1)        | (2)        | (3)        | (4)        | (5)        |
|--------------------------------|------------|------------|------------|------------|------------|
| Cheated in assignment          | 0.188***   | 0.191***   | 0.228***   | 0.174***   | 0.201***   |
|                                | (0.037)    | (0.036)    | (0.030)    | (0.036)    | (0.031)    |
| Cheated in assignment and got warned | −0.251*** | −0.245*** | −0.265*** | −0.232*** | −0.246*** |
| and got warned                  | (0.057)    | (0.048)    | (0.042)    | (0.046)    | (0.036)    |
| Absent at test 5                | 0.117*     | 0.103*     | 0.122**    | 0.097*     |            |
|                                | (0.062)    | (0.053)    | (0.061)    | (0.053)    |            |
| Female                          | −0.017     | −0.022     | −0.023     | −0.027     |            |
|                                | (0.022)    | (0.023)    | (0.023)    | (0.023)    |            |
| Age                             | −0.018     | −0.018     | −0.022     | −0.021     |            |
|                                | (0.026)    | (0.025)    | (0.025)    | (0.025)    |            |
| Admission grade written        | 0.002      | −0.005     | −0.002     | −0.003     |            |
|                                | (0.018)    | (0.017)    | (0.018)    | (0.017)    |            |
| Admission grade oral           | 0.005      | 0.004      | 0.002      | 0.003      |            |
|                                | (0.009)    | (0.008)    | (0.008)    | (0.007)    |            |
| Grade other classes            | −0.032*    | −0.037**   | −0.031*    | −0.037**   |            |
|                                | (0.019)    | (0.017)    | (0.018)    | (0.016)    |            |
| Led cheating group in assignment | 0.097**   | 0.115**    |            |            |            |
|                                | (0.045)    | (0.046)    |            |            |            |
| Led cheating group in assignment and got warned |            |            |            |            |            |
|                                | −0.083     | −0.077     |            |            |            |
| Constant                       | −1.806     | 4.491      | 5.206      | 5.741      | 5.849      |
| Curriculum FE                  | no         | yes        | no         | no         | no         |
| Class FE                       | no         | no         | yes        | yes        | yes        |
| Nationality FE                 | no         | yes        | yes        | yes        | yes        |
| High school type FE            | no         | yes        | yes        | yes        | yes        |
| High school honors FE          | no         | yes        | yes        | yes        | yes        |
| Admission procedure FE         | no         | yes        | yes        | yes        | yes        |
| Observations                   | 644        | 644        | 644        | 644        | 644        |
| Log Likelihood                 | −281.650   | −253.494   | −243.574   | −254.111   | −241.559   |
| Akaike Inf. Crit.              | 569.300    | 590.988    | 591.149    | 592.222    | 591.118    |

Note: *p<0.1; **p<0.05; ***p<0.01.
Robust standard errors, clustered at the class level.

ity to address many more exam formats in the future. More importantly, we believe our general methodology (stage 1: identify likely cheaters in assignments; stage 2: send personalized warning to likely cheaters prior to the exam) could be used in most settings, provided that cheaters can be identified with reasonable accuracy.

Our solution is, however, only partial. We restricted our technologies to the analysis of similarities between copies. However, as Karmalin (2017) shows, cheating takes almost infinite forms, as do the detection avoidance strategies. Replicating a similar protocol in more controlled conditions would allow a more precise representation of the mechanisms at play. In particular, we acknowledge that our study may include a few “false negatives” (true cheaters identified as non-cheaters). This may occur in both Test 5 (which allows us to identify assignment-cheaters) and the final exam (where classify students as ‘exam-cheaters’ or not). First, very good programmers might have anticipated some of the principles that underlie our detection technology, and acted in such a way that they evade detection. Because the class was only an introductory course, we believe such obfuscation techniques were far beyond the reach of most of the students. Further, obfuscation would be a superfluous effort for good programmers, for whom this class was very easy. Also, there is still a possibility that cheaters casually used the copy of another student with the exact same versions of the three “trap questions”. Because there were 5 versions of each of these questions, there is a very low probability that a given student who always cheats is unidentified. Still, some students who cheated only in a few questions might have been classified as non-cheaters. Students who cheated only in standard questions would also escape detection.

Another caveat we want to highlight is that our criteria for assignment into cheating categories slightly differs between Test 5 and the exam. The reason for this is that we anticipated that some students may try to escape detection, in case they formed educated guesses about the possible techniques implemented in Test 5. This is why we also added “trap questions” in the final exam, which were arguably very difficult to anticipate. Also, the motivations for cheating may have been different for the exam where stakes were higher. Cheating might have been harder too, due to the synchronous timing. For all these reasons, one cannot compare Test-5 and exam cheating rates. This is why we focus our attention to the effect of the warning at the exam only.

We also have to 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 show that targeted warnings make prevention significantly more effective than traditional warnings. However, prevention may not be enough. Repression in the form of sanctions is often necessary because sanctions increase confidence in the results of exams and limit the feeling of injustice in those who do not cheat. Our statistical analysis provides an indication only of the probability of fraud but rarely provides sufficient evidence to trigger a sanction, and we acknowledge that the proof mechanisms should be improved. The “trap random questions” that we used and present in Appendix A2 represent a promising avenue.

An important issue in research on academic dishonesty is the difficulty of estimating the prevalence of cheating. With the increasing prevalence of online exams, we believe the need to examine methods that allow for creativity and at the same time are robust to cheating will only increase in the coming years. This paper proposes solutions to enhance this robustness. Because of a relatively small sample size and inaccurate identification of cheating groups inherent to our probabilistic technique, we leave the more detailed study of the leader/followers dynamics for future research. We believe this could uncover many of the fundamental mechanisms behind cheating and, in turn, options to curb this behavior.
6. Conclusion

The COVID-19 crisis gave digital technologies a unique opportunity to showcase how they can contribute to our education systems. There is overwhelming evidence that these technologies will occupy an ever-growing place in education not only in times of crisis but also as a new teaching standard. The crisis has also revealed, however, some of the limits of online education, not the least of which involves the viability of online exams. Digital tools meant to address these limits, such as online exam platforms, were also questioned, as they support exam-taking but at the same time facilitate cheating.

In this paper we develop an original and effective method that can be used to identify illicit collaborations in a probabilistic way. We first show that cheating can be massive when exams are given at a distance without proctoring. We then demonstrate with a randomized experiment that a credible and effective mechanism could be deployed to discipline students and restore the validity and fairness of exams. Prior warnings are particularly effective in inducing honest behavior. In light of these results, we argue that the wise use of warnings constitutes a promising alternative to proctoring, when supervision is impossible for practical or ethical reasons.

Our work could be extended fruitfully in at least two directions. First, a larger sample size could allow to uncover some interesting mechanisms behind group formation and disbandment, and in particular the relative effectiveness of warnings sent to specific students (leaders or not), or warnings sent to all members of a group relative to only part of it. Second, the ability to control intra-group and inter-group communication between students between the treatment and the final exam would allow to better understand the mechanism through which the warning disciplines students.

Author statement

None

Appendix

A1. Messages sent to students

The students received two generic emails. The first was sent after Test 4 and before Test 5. It stated that Test 5 and the Exam were individual and reminded students of the consequences of verified cheating. It also notified the students that Test 5 would be a bit different from the previous tests, so as to better match the format of the final exam.
The second message was sent between Test 5 and the exam. It stated that some cheating had been observed to have occurred on Test 5 and again reminded students of the rules of the exam.

In addition to these generic emails, we sent the half of the students we identified as cheaters a warning email.

The preamble and first question on both Test 5 and the exam require students to adhere to a no-cheating “honor code”:
A2. Plagiarism detection technologies: details

Cheaters: We used three distinct detection technologies. The first two technologies were used to identify cheaters on Test 5. To defeat potential hacking (students understanding our technologies and intentionally avoiding them during the exam), we used the third technology only on exam results.

The first two technologies consist in textual analysis of copies. Because some simple hacking techniques are well known, we did more than use exact matching of copies (technology 1). Previous analysis of tests 1 through 4 showed that students quickly learn to add spaces, change variable names, etc., to avoid matching exam copies perfectly.

Following the spirit of many plagiarism-detection algorithms such as the one used by Karnalim (2017), our second technology analyses text similarity between copies. To account for the fact that our texts are typically very short (consisting of a few lines as opposed to several pages for most plagiarism-detection services), we use Gestalt Pattern Matching (Ratcliff and Obershelp, 1983). This enables us to identify copies that “look similar” to each other even if they are not perfect matches. Crucially, we complement Gestalt Pattern Matching with analysis of the abstract syntax of the codes. A given copy’s abstract syntax tree is extracted and compared with all other abstract syntax trees. For both concrete and abstract syntax analysis, a distance matrix between all copies is computed.

Copies are then clustered with a DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm. Copies that are closely packed together are considered clusters of cheaters. The members who constitute such a cluster are labelled as “cheaters.” It is useful at this stage to note that this strategy yields only a probabilistic estimation of cheating behavior. It is therefore an effective prevention tool but is of little help when it comes to applying sanctions (see discussion); hence the need for a third technology.

The third technology consists in a classical approach, the use of random questions, to which we added an original twist. For each specific question (displayed to all students as, say, “question B”), 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 answer that is incorrect but corresponds to the correct answer of another version of the question. We designed our versions such that it was extremely unlikely that a student could have genuinely come up with the answer to another version of a question. Further, differences between versions are visible only to particularly attentive eyes or to students expecting such a strategy to be deployed. Given that there is no precedent for such “trap random questions” in the institution, this eventuality is very unlikely to materialize. A cheater will be classified as such if s/he gives an answer that corresponds to the correct answer of another version of the question. We use this third technology for exam copies but not for the assignments, so students cannot learn to avoid detection.

As we discussed, the criteria that discriminates cheaters from non-cheaters is different in Test 5 and the Exam. This is due to the third technology (random questions) that was kept aside for the exam, and the fact the Exam was longer and harder. We note however, that this difference in criteria does not impact the validity of our randomized control trial: all students of the study group (test 5-cheaters) have been assigned to an assignment-cheating status according to the first criteria. Assignment to test and control groups (warned and unwarned cheaters) was random, and the identification of exam cheating was done according to the same criteria in both of these groups. We provide below the detail of our criteria for the assignment of the cheating status.

Test 5: questions 11 to 14 allow for significant variety in the answers. As such they can be used to identify anomalous similarities.

We classify a student as a cheater is (s)he meets at least two of these criteria:

- Exact same answer to Q11 as another copy
- Exact same answer to Q12 as another copy
- Exact same answer to Q13 as another copy
- Exact same answer to Q14 as another copy
- Gestalt Pattern Matching applied to Q11–Q14 shows similarity with another copy
- Abstract syntax tree applied to Q11–Q14 shows similarity with another copy

Exam: questions 12 to 18 allow for significant variety in the answers. We use a similar criteria to that of Test 5. We classify a student as a cheater is (s)he meets at least three of these criteria:

- Exact same answer to Q12 (Q13…Q18) as another copy
- Gestalt Pattern Matching applied to Q12-Q18 shows similarity with another copy
- Abstract syntax tree applied to Q11–Q18 shows similarity with another copy
- The student gave the correct answer to another version of either of the trap questions Q9, Q10, Q13.

Leaders: We have access to the exact time at which Test 5 was handed in by each student. We define a “leader” as a student who belongs to a cheating cluster and was the first to hand in her/his copy.

A3. Details descriptive statistics

We provide here some details on the variables displayed in Table 1. Table 3 complements these descriptive statistics, with further detail on the leader/follower distinction.

- Exam cheater: the student has been identified as a cheater at the final exam
- Test5 cheater: the student has been identified as a cheater at the last assignment before the exam.
- Test5 leader: the student has been identified as a leader of a cheating group at the last assignment before the exam. That student would also be labeled as a “Test5 cheater”
- NB missed assignments: Number of assignments, out of 5, that the student did not hand in.
- Absent at test5: Dummy variable that takes value 1 if the student did not hand in the last assignment.
- Grade test1, Grade test2, Grade test3, Grade test4, Grade test5: grade received at the various pre-exam assignments
- Female: takes value 1 if the student is a female, 0 otherwise.
- Admission grade (written): grade received at the competitive exam for admission at the school (written examination).
- Admission grade (oral): grade received at the competitive exam for admission at the school (written examination)
- Grade in other classes: average grade received in the other classes in 1st year at the school
- Did classe prépa: takes value 1 if the student followed “Classes préparatoires aux grandes écoles” (classes to prepare students for the entrance examinations to the top-ranking higher education establishments, 0 otherwise.
- Scientific high school: takes value 1 if the student went to a scientific high school (as opposed to Economic, literary or technical), 0 otherwise
- High school honors: takes value 1 if the student got “high honors” (“très bien”) or better at the French “Baccalauréat.”
A4. Robustness checks

One key technical difficulty of the present paper is to classify students into cheating and non-cheating categories. Indeed, the cheating status is by definition concealed by the students and, therefore, difficult to observe. To achieve this, we designed a few original identification techniques described in Appendix A2. None of them is perfect and the criteria for classification into cheating categories has by nature an arbitrary component.

Table 4 displays a few robustness checks related to the classification in the cheating categories. The set of controls is the one of the preferred specification of the main text (specification 3 in Table 2). The seven specifications represented in Table 4 differ by the dependent variable. The dependent variables of specifications (1) to (4) are the result of either of our detection strategies:

- Specification 1: Do the answers show strong textual similarity with at least another copy? (Gestalt Pattern Matching). (1 if yes, 0 otherwise)
- Specification 2: Does the abstract syntax of the answers show strong similarity with at least another copy? (DBSCAN). (1 if yes, 0 otherwise)
- Specification 3: Did the student give the correct answer to at another version of at least one question? (trap questions) (1 if yes, 0 otherwise)

Table 4
Robustness checks.

| Dependent variable: | Text similarity | Abstract similarity | Trap questions | Number of exact matches | Cheat score ≥ 2 | Cheat score ≥ 4 | Cheat score ≥ 6 |
|---------------------|-----------------|---------------------|---------------|------------------------|----------------|----------------|----------------|
|                     | logistic        | logistic            | logistic      | OLS                    | logistic       | logistic       | OLS            |
| Was cheater         | 0.93***         | 0.97***             | 0.52**        | 0.37**                 | 0.93***        | 0.93***        | 0.93***        |
| (0.024)             | (0.019)         | (0.130)             | (0.091)       | (0.081)                | (0.016)        | (0.205)        | (0.205)        |
| Was cheater         | −0.104***       | −0.057***           | −0.035**      | −0.607***              | −0.215**       | −0.144***      | −0.997***      |
| and got warning     | (0.025)         | (0.031)             | (0.161)       | (0.093)                | (0.034)        | (0.247)        |                |
| Female              | −0.032*         | 0.027               | 0.010         | −0.128*                | −0.018         | −0.049**       | −0.132         |
| Absent at Test 5    | −0.023          | 0.034               | −0.012        | 0.264                  | 0.112**        | −0.034         | 0.344          |
| (0.069)             | (0.037)         | (0.038)             | (0.222)       | (0.051)                | (0.030)        | (0.258)        |                |
| Age                 | −0.003          | −0.023              | −0.005        | 0.014                  | 0.011          | −0.002         | 0.0002         |
| (0.015)             | (0.023)         | (0.012)             | (0.065)       | (0.029)                | (0.018)        | (0.102)        |                |
| Admission grade     | −0.010          | 0.011               | −0.009        | 0.060                  | −0.010         | 0.011          | 0.026          |
| written             | (0.012)         | (0.014)             | (0.010)       | (0.047)                | (0.022)        | (0.007)        | (0.058)        |
| Admission grade     | −0.004          | 0.001               | −0.001        | −0.002                 | 0.006          | −0.001         | −0.003         |
| oral                | (0.006)         | (0.004)             | (0.003)       | (0.021)                | (0.008)        | (0.005)        | (0.028)        |
| Grade other classes | −0.028**        | −0.002              | −0.026*       | −0.186***              | −0.071***      | −0.028*        | −0.246***      |
| (0.011)             | (0.012)         | (0.016)             | (0.052)       | (0.020)                | (0.015)        | (0.078)        | (0.078)        |
| Constant            | 6.502           | −1.416              | −7.322        | 3.754                  | 36.318         | 3.820          | 5.281          |
| Curriculum FE       | no              | no                  | no            | no                     | no             | no             | no             |
| Class fixed FE      | yes             | yes                 | yes           | yes                    | yes            | yes            | yes            |
| Nationality FE      | yes             | yes                 | yes           | yes                    | yes            | yes            | yes            |
| High school type FE | yes             | yes                 | yes           | yes                    | yes            | yes            | yes            |
| Admission procedure FE | yes      | yes                 | yes           | yes                    | yes            | yes            | yes            |
| Observations        | 644             | 644                 | 644           | 644                    | 644            | 644            | 644            |
| R²                  | 0.153           | 0.135               | 0.135         | 0.098                  | 0.150          | 0.150          | 0.150          |
| Log Likelihood      | −145.054        | −201.407            | −98.623       | −356.546               | −356.546       | −112.383       | 817.092        | 328.765        |

Note: *p<0.1; **p<0.05; ***p<0.01.
• Specification 4: How many of the answered questions exactly match the set of answers of at least another copy? (exact matching/numeric variable)

We observe that all of these criteria result in the same qualitative insight: the nudge has a strong effect on students propensity to cheat. Yet, none of these detection techniques is perfect. Some students might for example avoid exact matching by changing variable names or marginally re-ordering the code, which would generate false negatives. These techniques can also generate “false positives” as discussed in Section 5.

This is why we built a custom numeric variable, the “cheat score”, which is the sum of all tests (assuming a positive test equals 1, 0 otherwise). In the main text we classify as cheater any student whose cheat score has a cheat score greater or equal to 3. As a robustness check, Table 4 uses thresholds of 2 (Specification 5) and 4 (Specification 6). Specification (7) simply takes the cheat score as a dependent variable. Again we verify that our results are qualitatively maintained. Expectedly, when the criteria is too narrow (Specifications 1 to 4) or too strict (Specification 6), our results are attenuated. We note finally that contrary to the specifications of Table 2, Specifications (4) to (7) in Table 4 use standard optimum least squares, owing to the numeric nature of the dependent variable.

References

Allcott, A., Mullainathan, S., 2010. Behavior and energy policy. Science 327, 1204–1205.
Arnold, Ivo, 2006. Cheating at online formative tests: Does it pay off? The Internet and Higher Education 29, 96–106. doi:10.1016/j.iheduc.2016.02.001.
Arnold, R., Martin, B., Jinks, M., Bigby, L., 2007. Is there a relationship between honor codes and academic dishonesty? J. Coll. Charact. 8, 1–20.
Auferth, F., 2020. Why benefits from nudges for exam preparation? An experiment. Behavioral & Experimental Economics eJournal.
Bachore, M.M., 2016. The nature, causes and practices of academic dishonesty/cheating in higher education: the case of Hawassa University. J. Educ. Pract. 7 (19), 14–20.
Bandura, A., 2002. Selective moral disengagement in the exercise of moral agency. Journal of Moral Education 31, 101–119.
Bawarith, R., Basuhail, A., Fattouch, A., Gamalel-Din, S., 2017. E-exam Cheating Detection System. Int. J. Adv. Comput. Sci. Appl. (IJACSA) 8 (4) 2017.
Becker, D., Connolly, J., Lentz, P., Morrison, J., 2006. Using the business fraud triangle to predict academic dishonesty among business students. Acad. Educ. Leadersh. 1, 37–54.
Beshears, John, JamesChoi, J., Laibson, David, BrigitteMadrian, C., KatherineMilkman, L., 2010. The Effect of Providing Peer Information on Retirement Savings Decisions. RAND Corporation, Santa Monica, CA https://www.rand.org/pubs/working_papers/W8003.html.
Bing, M., Davison, H., Vitell, S., Ammeter, A., Garner, B., Novicic, M., 2012. An experimental investigation of an interactive model of academic cheating among business school students. Acad. Educ. Leadersh. 11 (1), 28–48. Retrieved July 8, 2020, from www.jstor.org/stable/23304455.
Brodowsky, G.H., Tari, E., Ho, E.N., Sciglimpaglia, D., 2020. Tolerance for cheating from the classroom to the boardroom: a study of underlying personal and cultural drivers. J. Market. Educ. 42, 23–36.
Cialdini, R.B., Goldstein, N.J., 2004. Social influence: compliance and conformity. Annu. Rev. Psychol. 55, 591–621.
Chirumamilla, Aparna, Guttorm, Sindre, 2019. E-Assessment in Programming Courses: Towards a Digital Ecosystem Supporting Diverse Needs? 18th Conference on e-Business, e-Services and e-Society (IEE).
Cialdini, R.B., 2003. Crafting normative messages to protect the environment. Curr. Dir. Psychol. Sci. 12, 105–109.
Cluskey, Eileen, C.R., Raibora, M.H., 2011. Thwarting online exam cheating without proctor supervision. Journal of Academic and Business Ethics 4 (1), 1–7.
Corrigian-Gibbs, H., Gupta, N., Northcott, C., Cuthrell, E., Thies, W., 2015. Detering cheating in online environments. ACM Trans. Comput. Hum. Interact. 22 (6). doi:10.1145/2810239.
D’Souza, K.A., Siegfried, D.V., 2017. A conceptual framework for detecting cheating in online and take-home exams. Decis. Sci. J. Innov. Educ. 15 (4), 370–391.
Damgaard, M.T., Nielsen, H.S., 2018. Nudging in education. Econ. Educ. Rev. doi:10.1016/j.econedrev.2018.03.008.
Ferraro, P., Miranda, Juan Jose, Price, Michael K., 2014. The persistence of treatment effects with norm-based policy instruments: evidence from a randomized environmental policy experiment the persistence of treatment Effects with norm-based policy instruments: evidence from a randomized environmental policy.
Am. Econ. Rev. 101, 318–322. doi:10.1257/aer.101.3.318, Papers Proceedings-sissn 0002-8282.
Fendler, R.J., Yates, M., Goddby, J., 2018. Observing and deterring social cheating on college exams. Int. J. Scholarsh. Teach. Learn. doi:10.20429/ijstl.2018.120104.
Green, S., 2004. Cheating. Law Polis. 23, 137–185.
Harmon, Oiskar, Lambrinos, James, 2006. Online Format vs. Live Mode of Instruction: Do Human Capital Differences or Differences in Human Capital Explain the Differences in Outcomes? Working papers 2006-07, University of Connecticut, Department of Economics.
Holden, O., Kuhlmeier, V.A., Norris, M., 2021. Academic integrity in online testing: a research review. Frontiers in Education doi:10.3389/feduc.2021.639814.
Idemudia, Samson, Rohani, Foad, Ohman, Sti, 2016. An improvement of student examination assessment through online (e-exam) by considering psychological distress factors. International Research Association of Computer Science, & Technologies ( IRACST) 2016.
Karnalim, O., 2017. Python source code plagiarism attacks on introductory programming course assignments. Themes Sci. Technol. Educ. 10, 1729.
Kerkvliet, J., Sigmund, C.L., 1999. Can we control cheating in the classroom? J. Econ. Educ. doi:10.1080/002209909569060.
Konheim-Kالkatth, Y., Stellnack, M., Shilkey, M., 2008. Comparison of honor code and non-honor code classrooms at a non-honor code Univ. J. Coll. Charact. (3) VOLUME IXNO.
Li, M., Luo, L., Sikkdar, S., et al., 2021. Optimized collusion prevention for online exams during social distancing. npj Sci. Learn. 5, 6. doi:10.1038/s41539-020-00083-3.
Madrian, B.C., Shea, D.F., 2001. The power of suggestion: inertia in 401(k) participation and savings behavior. Q. J. Econ. 116 1149-1187.
Mazar, N., Oh, A., Asley, D., 2008. The dishonesty of honest people: a theory of self-concept maintenance. J. Market. Res. 45 (6), 633–644.
McCabe, D., 2016. Cheating and honor: lessons from a long -term research project. In: Bretag, T. (Ed.), Handbook of Academic Integrity. Springer, pp. 187–196. doi:10.1007/978-981-287-098-8_22.
McCabe, D., Treviño, L., 1995. Cheating among business students: a challenge for business leaders and educators. J. Manag. Educ. 19, 205–218.
McCabe, D.L., Trevino, L.K., Butterfield, K.D., 1999. Qualitative Investigation. The Journal of Higher Education 70 (2), 211–234. https://doi.org/10.3200/HEBO.70.2.211-234.
McCabe, D.L., Treviño, L.K., Butterfield, K.D., 2001. Cheating in academic institutions: a decade of research. Ethics Behav. 11 (3), 219–232.
Mellar, Harvey, Peycheva-Forsyth, Roumiana, Kocdar, Serpih, Karademir, Abdulkadir, Y., Yovcheva, Blagoveva, 2018. Addressing cheating in e-assessment using student authentication and authorship checking systems: teachers’ perspectives. International Journal for Educational Integrity 14 (1). doi:10.1007/s40979-018-0025-x.
Norrish, Mark, 2019. University online cheating – How to mitigate the damage [2019]. Res. High Educ. J. 37, 20 eVno 2019.
Rickmann, K., Northcraft, G., 2008. To be or not to be trusted: the influence of media richness on detection and deception. Org. Behav. Hum. Decis. Process. 107 (2), 106–122.
Scallon, Patrick, 2003. Student online plagiarism: How do we respond? College Teaching 51 (4), 161–165.
Sudman, S., Bradburn, N., 1974. Response Effects in surveys: A review and Synthesis. Aldine, Chicago, IL.
Thaler, R., Sunstein, C., 2008. Nudge: Improving Decisions About Health, Wealth, and Happiness. Penguin Books, New York.
Thaler, R.H., Benartzi, S., 2004. Save More Tomorrow: using Behavioral Economics to increase Employee Saving. J. Pol. Econ. 112, 164–187.
Vanz A., Kray, L., 2014. I can’t lie to your face”: minimal face-to-face interaction promotes honesty. J. Exp. Soc. Psychol. 55 (1), 234–238.
Varble, D., 2014. Reducing cheating opportunities in online test. Atlantic Marketing Journal 3 (3).
Zamir, E., Lewisohn-Zamir, D., Ritov, L., 2017. It’s now or never! using deadlines as nudges. Law Soc. Invq. 42, 769–803. doi:10.1111/hsi.12199.