# ONLINE WRITTEN EXAMS DURING COVID-19 CRISIS

Goffredo Haus, Yuri Benvenuto Pasquinelli, Daniela Scaccia and Nello Scarabottolo
*Università degli Studi di Milano, Milan, Italy*[*]

## ABSTRACT

The paper deals with the problem of carrying on online written exams in the University of Milan, suddenly closed due to the Covid-19 emergency. Main goal of the paper is to present and compare the different scenarios envisioned, depending on the number of students to be monitored in parallel to avoid cheating. For limited numbers, direct monitoring by the teacher has been suggested, with different supports depending on the type of exam proposed by the teacher (open vs. closed answer questions, presence of handwritten parts). For larger numbers, a proctoring system has been adopted, after some comparisons between two commercial solutions. Results show a very large adoption of a solution allowing teachers to monitor in real time students answering open questions on their home computers.

## KEYWORDS

Universities, Written Exams, Closed Vs. Open Answer Exams, Online Student Monitoring, Proctoring

## 1. INTRODUCTION

The lockdown imposed by the Covid-19 pandemic around the end of February 2020 forced Italian Universities to suddenly transfer online all the teaching activities normally carried on with students physically present in classrooms in a traditional university.

To better understand the context, the University of Milan is constituted by height faculties: Agricultural and Food Sciences, Humanities, Law, Medicine, Pharmacy, Political, Economic and Social Sciences, Science and Technology, Veterinary Medicine, and 2 schools: Exercise and Sport Sciences, Language Mediation and Intercultural Communication. They offer 67 bachelor degrees (3 years, 180 ECTS – European Credit Transfer System – credits) 64 master degrees (2 years after bachelor, 120 ECTS credits) and 9 single-cycle master degrees (5 or 6 years, 300 or 360 ECTS credits). More than 3000 courses are taught by 2179 staff professors and almost 2000 contract professors, supported by 1960 support people (technical and administrative staff units).

Due to urgency requirements, the implementation of online lectures has been left to personal initiative of teachers, without the possibility of significant methodological guidelines. The support offered by the university personnel mainly consisted in a set of instructions published on the web portal helping teachers in registering audio and video lectures, in connecting with students through web conference platforms, in publishing didactical materials on the proprietary LMS (Learning Management Systems).

Similar approach has been followed to allow the thesis discussion of graduating students: the web conference platforms used for lectures have been adopted also to connect these students from home with the commission of teachers evaluating their final exam.

A third aspect of online learning required on the contrary some deeper considerations: the management of exams evaluating the knowledge and the competences acquired by students at the end of each single course present in their study curriculum.

Among the various references in literature regarding online exams, it is worth mentioning (Truszkowski 2019) (Weiner & Hurtz 2017) (Wibowo 2016) (Ardid et al. 2014) and (Hillier & Fluck 2013). In the first two, some comparisons between proctored and non-proctored tests are done, clearly showing that in absence of some form of proctoring the final grades are higher, due to usage of unauthorized supports and cheating. In the third one, a comparison between online and onsite proctored exams shows on the contrary no significant differences in final grades, allowing to state that the level of student behavior control can be satisfactory both online and onsite.

---

[*] *firstname.lastname@unimi.it*

To proctor exams online, the first suggestion coming from the University governance to all teachers has been to convert them in oral form, and to interact with students through the already mentioned web conference platforms to be able to evaluate and to grade them. However, this approach presents several limitations, among which the most critical one are:

- the impossibility in an oral exam to ask students to solve problems requiring (even a short) autonomous work, perhaps to be written by hand on a piece of paper;
- the difficulties in finding a set of equally difficult questions to several students (thus the difficulties in making a fair evaluation);
- the excessive time required for courses followed by hundreds of students.

For the above reasons, the authors of this paper:
- the Deputy Rector for Digital Innovation, ICT Services, Strategic and Special Projects;
- the Director of the ICT Division;
- the Director of CTU: the Teaching and Learning Innovation and Multimedia Technology Center of the University;
- the Rector Delegate to ICT infrastructures coordination;

have been charged with the task of studying possible ways to conduct online written exams, allowing teachers to avoid the limitations of oral exams while still guaranteeing a reasonable level of reliability in their evaluations. To this purpose, main aspects to be taken into account were:

- the possibility of submitting to students open answer tests as well as closed answer quizzes;
- the possibility of asking students to submit their work (partially or totally) written on paper, for example when mathematical formulas or graphical elements are required;
- the possibility for students to use some applications (e.g., compilers, statistical software, etc.) during their exam;
- the possibility of examining groups of students ranging from a few units to a few hundreds;
- the adoption of technological solutions enforcing teacher's control over students behavior (i.e., avoiding students to copy answers, to surf the web and to interact among each other during the exam).

The following section summarizes the impact of exams in our University, in terms of numbers of students and distribution over the year. Section 3 discusses the solutions identified to monitor student behavior during their exam, on the basis of the number of participating students. Section 4 presents three different scenarios identified by the authors and proposed around the end of April 2020 to the teachers of the university, letting them to adopt the most suitable one(s) for their needs. In section 5, the usage of these solutions up to beginning of June is considered in terms of numbers of exams, numbers of examined students, etc. Section 6 draws some concluding remarks and possible future developments of this aspect of online learning.

## 2. ANALYSIS OF EXAM SESSIONS

To better understand the different scenarios described in the following sections of this paper, it is worth knowing that in our University:

- each course must propose to students at least six different exam dates every academic year;
- a student can decide her/his first exam date after the end of the course lectures, but in case of failure or unsatisfying result, she/he can reapply for the same exam in a following date, without any penalty;
- there is no penalty for students that after application do not show up at the exam date: they are simply considered absent and they can reapply whenever they want.

Thus, each teacher has to plan for at least six exam sessions every year, with a number of participating students known only a few days before each session, i.e., after closing of the student application period for each exam date.

To evaluate the impact of exams both during the lockdown period and for the subsequent months – when we expect to be forced to ensure social distancing for a long time – data related to last complete "normal" year 2019 have been taken into account. The results are shown in Tables 1 and 2, letting to the following considerations:

- the largest number of exam sessions are located at the end of the winter and summer semesters, during breaks between lecture times (i.e., January-February and June-September, with August being the traditional vacation months in our country);

- the average number of students applying for each exam session is reasonably limited, and equal to 18.6 considering the overall year, but:
- there are exam sessions characterized by huge numbers of students, over 200 and arriving to more than 540 for a couple of exams;
- the largest number of exam sessions is organized by the faculty of Medicine, far above the faculty of Science and Technology which has the largest number of examined students;
- Humanities, i.e., the faculty where the highest number of students are enrolled, usually adopts oral exams, thus the number of written sessions is quite limited.

To propose to our colleagues suitable scenarios for conducting written exams online, we had to then consider such a wide variety of situations in terms of student numbers, always keeping in mind their uneven distribution over the year.

Table 1. Written exam sessions in 2019

| Month | Exam sessions | Examined students | Average students per session | Max students per session |
|---|---|---|---|---|
| January | 2˙116 | 45˙617 | 21.6 | 309 |
| February | 2˙868 | 54˙806 | 19.1 | 353 |
| March | 471 | 8˙998 | 19.1 | 275 |
| April | 838 | 18˙255 | 21.8 | 214 |
| May | 1˙102 | 25˙677 | 23.3 | 302 |
| June | 2˙477 | 51˙301 | 20.7 | 428 |
| July | 3˙197 | 49˙074 | 15.4 | 206 |
| September | 3˙043 | 39˙083 | 12.8 | 223 |
| October | 497 | 5˙533 | 11.1 | 245 |
| November | 896 | 19˙780 | 22.1 | 545 |
| December | 827 | 23˙464 | 28.4 | 401 |
| **Total** | **18˙332** | **341˙588** | **18.6** | **545** |

Table 2. Distribution of students and sessions per faculties in 2019

| Faculty/School | Enrollments in 2018/19 | Exam sessions | Examined students |
|---|---|---|---|
| Agricultural and Food Sciences | 3˙431 | 1˙613 | 27˙462 |
| Exercise and Sport Sciences | 1˙398 | 349 | 9˙518 |
| Humanities | 14˙878 | 1˙037 | 32˙633 |
| Language Mediation and Intercultural Communication + Interfaculties | 4˙737 | 1˙115 | 32˙394 |
| Law | 6˙748 | 397 | 10˙696 |
| Medicine | 7˙682 | 5˙494 | 62˙523 |
| Pharmacy | 3˙075 | 1˙161 | 24˙286 |
| Political, Economic and Social Sciences | 7˙708 | 2˙614 | 53˙776 |
| Science and Technology | 8˙450 | 3˙823 | 75˙325 |
| Veterinary Medicine | 1˙836 | 729 | 12˙975 |
| **Total** | **59˙943** | **18˙332** | **341˙588** |

## 3. MONITORING STUDENT BEHAVIOR DURING EXAMS

A first discriminating aspect considered in defining the different scenarios has been the possibility for the teacher to monitor in real time the behavior of the whole set of examined students or of a limited number of student groups through some web conference platforms.

After some tests, it has been seen that a reasonable number of students that can be monitored by a single person is in the range of 20-30, and that it was not worth to ask teachers to split students in more than 4 to 5 groups, to be monitored in parallel (with the help of some collaborators) or one after each other. Then, we decided to consider 100 students applying for the same exam date as the limit for exam sessions monitored under teacher's responsibility and exams sessions requiring external support.

Table 3 subdivides the number of sessions already given in Table 1 using this threshold to discriminate between sessions up to 99 applying students and sessions with 100 or more applying students. As it can be seen, the percentage of large sessions is definitely limited, even if they must deal with a number of applying students corresponding to 20% of the total number of students registered to exams in year 2019.

Table 3. Exam sessions in 2019 having less or at least 100 applying students

| Month | Total # of sessions | Sessions with up to 99 students | Sessions with 100 or more students | % of large sessions | Total # of students in large sessions |
|---|---|---|---|---|---|
| January | 2˙116 | 2˙043 | 73 | 3.45% | 11˙269 |
| February | 2˙868 | 2˙819 | 49 | 1.71% | 7˙299 |
| March | 471 | 458 | 13 | 2.76% | 2˙166 |
| April | 838 | 804 | 34 | 4.06% | 5˙056 |
| May | 1˙102 | 1˙056 | 46 | 4.17% | 7˙443 |
| June | 2˙477 | 2˙413 | 64 | 2.58% | 9˙865 |
| July | 3˙197 | 3˙163 | 34 | 1.06% | 4˙383 |
| September | 3˙043 | 3˙021 | 22 | 0.72% | 2˙995 |
| October | 497 | 494 | 3 | 0.60% | 545 |
| November | 896 | 856 | 40 | 4.46% | 7˙752 |
| December | 827 | 768 | 59 | 7.13% | 10˙245 |
| **Total** | **18˙332** | **17˙895** | **437** | **2.38%** | **69˙018** |

## 3.1 Direct Monitoring

Monitoring of a group of 20-30 students by means of a web conference platform (e.g., Microsoft Teams, Zoom, etc.) has been proposed in the following way:
- the web conference is established between the computer of the teacher and the smartphone of each student, who is requested to install the app of the web conference adopted by the teacher;
- students are admitted to the web conference planned by the teacher one after each other;
- after admission to the web conference, the student has to show an identity document to the teacher (to prove her/his right to participate to the exam;
- the teacher can ask the student to use her/his smartphone to show the room where she/he will take the exam, to demonstrate that no other people nor other support device (e.g., a second computer) is present;
- then, the student must position her/his smartphone – suitably powered to avoid battery exhaustion – as shown in Figure 1, in order to frame her/his work place (table and computer desktop).



Figure 1. Student smartphone framing student work place

Student monitoring through her/his smartphone instead of through her/his computer proved to have several advantages:
- the teacher can control that no forbidden material (e.g., books, written notes, other smartphone, etc.) is used by the student during the exam;
- by zooming on each student window in the web conference, the teacher can look at the desktop and see if the student is operating correctly (i.e., using only the allowed applications);

- the audio of the student smartphone is always kept on, allowing to immediately identify unexpected noises, other people's voices, etc.;
- students have no idea about what is the web conference window examined in each moment by the teacher, who can randomly browse among them. This has a psychological effect on student behavior, since they must expect to be deeply controlled in any moment.

Of course, this monitoring relies from one side on the reliability of the connection between student smartphone and web conference platform, and on the other side on the ability of the teacher to properly handle the web conference.

As regards the first aspect, the student may have a low quality internet connection (maybe due to the area where her/his home is located), her/his smartphone may have heating troubles after too much time of video and audio connection, or she/he may have other problem resulting in unexpected end of her/his presence in the web conference. In these cases, it is up to the teacher to decide how to proceed: our suggestion is to let the student to try recovering the connection quickly; in case of failure, she/he must have a second chance in the following days, either as part of a successive student group, or as a single examined student, perhaps in oral form.

As regards the second aspect, we published a set of guidelines on the University web portal to help teachers to train themselves before undergoing their exams. Up to now the only critical situation happened for a teacher not realizing in time that a group of students in a larger set was not monitored at all, and he decided to cancel the exams of that group, obviously raising student complaints that left the University governance to force the teacher to re-admit them.

## 3.2 Proctoring

When 100 or more students applied for a single exam session, it is not worth asking the teacher to monitor many student groups, either in parallel (too many collaborators needed) or in sequence (too much time required).

We then started considering different proctoring proposals available on the market, i.e., systems offering different kinds of monitoring supports having these general characteristics:
- no control in real time by the teacher or other collaborators is required: the systems record the behavior of the student during the exam through the webcam of her/his computer, and/or a second device like a smartphone;
- after the end of the exam, all recordings are processed by suitable Artificial Intelligence algorithms, that mark in red any "suspect" behavior of the student (e.g., eyes or head movements, noises, appearance of other people, etc.);
- the marked tracks are then sent to the teacher together with all recordings, to allow her/him to analyze the suspect behaviors and decide accordingly how to manage them.

Among the various offerings, we concentrated on two options – namely: Respondus and Proctorio – and we started some comparative tests in early May, first on simulated exams involving staff and then on real exams using volunteer teachers.

### 3.2.1 Respondus

As stated in (Respondus 2020) Respondus use its LockDown Browser to prevent cheating. With such a browser, assessments are displayed full-screen and cannot be minimized, almost all menu and toolbar options are removed, access to other applications including messaging, screen-sharing, virtual machines, and remote desktops is forbidden, as well as copying and pasting to or from an assessment.

Once activated, Lockdown Browser closes all other applications on the student computer, asks the student to show her/him identity document and her/his room and work place, tests audio and video devices on the student computer and finally let the student start accessing exam question on a suitably connected LMS.

Main pros and cons emerging from the testing phase were the following:
- Respondus has a very aggressive and attractive pricing policy for newcomers, especially related to the Covid-19 emergency, when all exams were to be planned online;
- setup of the student computer was not always straightforward, especially in MacOS environments, where the Lockdown Browser installation and/or activation required sometimes the intervention of a call center to let the student take the exam;

- sometimes, the system was not able to process the recording of some students, implying exam repetition;
- the AI algorithms revealed some limits in putting reds on the tracks: in some cases, the appearance of a second person during the exam was not spotted.

As a final result, the teacher had to examine in a different way around 10% of the students present at the session, either for setup or for post processing troubles.

### 3.2.2 Proctorio

As stated in (Proctorio 2020) Proctorio does not use a specific lockdown browser to prevent cheating, but a simple add-on for browsers like Chrome that creates a secure exam environment by restricting internet navigation and computer functionality. All internet browsing can be blocked or specific websites may be whitelisted to allow third-party quiz applications or blacklisted to prevent the use of unauthorized resources during exams.

Main pros and cons emerging from the testing phase where the following:
- Proctorio does not have any specific pricing policy related to the Covid-19 emergency, thus resulting in being far more expensive than Respondus especially when all exams have to be planned online; however:
- setup of the student computer was definitely simpler due to the absence of a pervasive tool like the Lockdown Browser;
- student behavior monitoring is more accurate, since Proctorio records the webcam stream and also the desktop of the student computer;
- the browser add-on sends only some video frames instead of a continuous streaming, thus significantly reducing the network bandwidth requirements (and facilitating monitoring of students with poor internet connections);
- AI algorithms can be tuned by the teacher in terms of sensitivity of the different kinds of suspected behaviors after receiving the recorded exams; this allows the teacher to emphasize the aspects considered most dangerous and/or more common.

As a conclusion, we decided to proceed for the coming summer period of exams, i.e., June-September 2020, using Proctorio for monitoring exam sessions with at least 100 students applying.

## 4. TYPES OF SUPPORTED EXAMS

Besides identifying the most suitable ways usable to monitor student behavior, we had to take into account the nature of the exam each teacher may decide to use; as already stated in the introduction, we had then to consider:
- the possibility of submitting to students open answer tests as well as closed answer quizzes;
- the possibility of asking students to submit their work (partially or totally) written on paper;
- the possibility for students to use some applications (e.g., compilers, statistical software, etc.) during their exam.

The resulting scenarios identified and proposed to our University teachers are described in the following subsections.

## 4.1 Open Answer Tests

For limited numbers of students – i.e., groups that can be monitored by the teacher – a proposed scenario is based on the exam.net platform (Exam.net 2020) implemented by the Swedish company Teachiq AB and free to use outside Sweden during 2020. Main characteristics of this platform are:
- a very easy teacher interface, greatly facilitating creation and test of exams;
- real time monitoring of student work, since the teacher can browse among students and see what each of them already wrote;
- chat support, allowing the teacher to interact with every student without disturbing the overall group;
- download of student work as a pdf file for correction and grading;
- possibility for students to include pictures of hand written work, taken through their smartphone in a controlled way (QR code generated by the platform to allow each student to take and submit her/his pictures).

Moreover, since the visibility of the student desktop through her/his smartphone (as described in section 3.1) can be limited by environment lighting, non ideal smartphone positioning, etc., the scenario adopts the safer approach proposed by exam.net and based on the usage of SEB (Safe Exam Browser). As stated in (SEB 2020) SEB is a software that turns any computer temporarily into a secure workstation. It controls access to resources like system functions, other websites and applications and prevents unauthorized resources being used during an exam. SEB cannot be considered totally safe: some ways to bypass it are claimed to exist, though not so easy to implement (usage of virtual machines, access through a second computer, etc.). However, the appearance of the student desktop when SEB is running and the back monitoring through student smartphone facilitate the teacher work in identifying such bypasses.

For larger numbers of students, requiring proctoring, Proctorio has been integrated with the Moodle LMS already adopted by CTU, where several types of exam questions can be defined and submitted to students. The only unavailable option is the possibility for students to submit handwritten contributions, since the activities students must scan and submit them and are not compatible with the AI monitoring approach of all proctoring systems.

## 4.2 Closed Answer Quizzes

The exam.net platform cannot be easily used for this kind of exams: it has no native support for quizzes and it does not integrate with an LMS like Moodle; it just gives the possibility to indicate some URLs accessible through SEB during the exam. However, there is no guarantee that from these URLs it is not possible to start surfing the web; moreover, we had several troubles in integrating exam.net with the Single Sign On feature of our University, used by students to authenticate themselves.

For the above reasons, closed answer quizzes for limited numbers of students (as well as any other types of exams that can be defined in Moodle) are implemented by directly integrating SEB with the Moodle LMS, without passing through exam.net.

For large numbers of students, quizzes are obviously handled through proctoring, as outlined at the end of section 4.2.

## 4.3 Exams Requiring Usage of External Applications

A reliable control of the actual usage of computer programs like compilers, spreadsheets, statistical software, and so forth, made by students during the exam cannot be easily ensured, especially for groups of students directly monitored by the teacher. In fact, even when the student desktop can be clearly seen by the teacher, it is almost impossible to ensure that no other unadmitted programs or websites are accessed during the exam.

The only possible approach, allowed e.g. by Proctorio, is the adoption of proctoring for student monitoring, but this would result in a far larger number of exam sessions requiring such a solution. In fact, since even for small student groups it would become necessary to set up a proctoring session, this would lead to extra costs but above all to unaffordable workload for the technicians managing them.

For these reasons, teachers have been requested to convert this kind of exams into an oral format.

## 5. FIRST EVALUATIONS

After the first month of application of the scenarios above described, we are able to show some very preliminary figures. As already stated, the hugest numbers of exam sessions just started beginning of June 2020, with the peak expected between the second part of June and the first part of July.

As it can be seen from Table 4, the scenario based on the exam.net platform (i.e., open answer questions or quizzes with few questions, for groups of students directly monitored by the teacher) shows a large number of sessions and a significant number of examined students, but a definitely smaller average number of students per session. Besides some bias due to session tests made by teachers (that we cannot evaluate), it is clear that such a way of examining students has been greatly appreciated. In fact, it allows teachers not familiar with Moodle to submit written exams also to very small groups of students, exploiting all the facilities and friendliness of that platform.

Moodle with SEB has been introduced a couple of weeks after exam.net, and Moodle is presently used only by few teachers of the University, thus numbers of sessions and students are definitely smaller. However, the average number of students per session is four times the one of exam.net. This is due from one side to the fact that teachers learn to use Moodle when useful, i.e., when they have to deal with larger classes. From the other side, to the adoption of Moodle for an exam evaluating the ability of students from several faculties in using basic computer utilities, monitored by several CTU staff technicians in parallel on large student groups.

Proctoring has been used initially – when testing Respondus and Proctorio – for smaller exam sessions. Then, once decided to adopt Proctorio, proctoring has been limited to exam sessions with at least 100 applied students, as said before. This explains why in Table 4 the average number of students per proctored session is very low for Respondus (used almost only during test) while for Proctorio it is rapidly increasing towards 100 and more.

Table 4. Summary of first month of online written exams

| Scenario | # of sessions | % of sessions | Average # of students per session | # of examined students | % of examined students |
|---|---|---|---|---|---|
| Exam.net | 1´333 | 93.7% | 9.5 | 12´673 | 74.0% |
| Moodle with SEB | 55 | 3.9% | 32.6 | 1´793 | 10.5% |
| Proctoring using Respondus | 9 | 0.6% | 39.3 | 354 | 2.1% |
| Proctoring using Proctorio | 26 | 1.8% | 89.0 | 2´314 | 13.5% |
| **Total** | **1´423** | **100.0%** | **12.0** | **17´134** | **100.0%** |

## 6. CONCLUSIONS

After evaluation, and interaction with other Universities facing similar problems, three different scenarios have been identified for carrying on written exams online during the Covid-19 emergency period.

Cost and setup complexity of proctored exam sessions forced to limit this last scenario to very large groups of students (≥100) to be evaluated in parallel: of course, this limitation has been considered excessive by some teachers, reluctant to perform by themselves a direct monitoring.

However, the identification of two scenarios allowing reasonably affordable handling of groups of 20-30 students by a single teacher or collaborator allowed the implementation of a lot of written exams, definitely less time consuming than the oral exams initially proposed as the only possible solution. However, further studies are required to assess the efficacy of online exams on much larger cohort of students. Also, the results of this study may not be replicable to other institutions as this was based on one case study. Hence, other universities can assess the exams systems this study explored.

## REFERENCES

Ardid, M, Gómez-Tejedor, JA, Meseguer-Dueñas, JM, Riera, J & Vidaurre, A 2014, 'Online exams for blended assessment. Study of different application methodologies', Computers and Education, vol. 81, pp. 296-303.

Exam.net 2020, A robust, easy-to-use and secure exam platform, viewed June 10th 2020, < https://exam.net/>.

Hillier, M & Fluck, A 2013, 'Arguing again for e-exams in high stakes examinations', 30th ascilite Conference 2013 Proceedings, pp. 385-396.

Proctorio 2020, A Comprehensive Learning Integrity Platform, viewed June 10th 2020, < https://web.proctorio.com/>.

Respondus 2020, Assessment Tools for Learning Systems, viewed June 10th 2020, < https://web.respondus.com/>.

SEB 2020, Safe Exam Browser, viewed June 10th 2020, < https://safeexambrowser.org/>.

Truszkowski, D 2019, 'Proctored Versus Non-Proctored Testing: A Study for Online Classes', Dissertation of the Doctoral Program of the American College of Education.

Weiner, JA & Hurtz, GM 2017, 'A Comparative Study of Online Remote Proctored versus Onsite Proctored High-Stakes Exams', Journal of Applied Testing Technology, vol. 18, no. 1, pp. 13-20.

Wibowo, S, Grandhi, S, Chugh, R & Sawir, E 2016, 'A Pilot Study of an Electronic Exam System at an Australian University', Journal of Educational Technology Systems, Vol 45, Issue 1, 2016.