Dealing with Cheating in Online Exams: A Systematic Review of Proctored and Non-Proctored Exams

Bilal Oncul, bilal.oncul@erzurum.edu.tr, Erzurum Technical University, Turkey, https://orcid.org/0000-0003-3970-0824

SUMMARY
With the effect of the pandemic, higher education is maintained all over the world with the help of online education methods. How to make online exams, which are an important part of distance education processes, in a way to prevent cheating and illegal behavior is an important issue that brings with it discussions. This research, in which the solutions to this issue were discussed theoretically, evaluated in terms of educational sciences, and the studies carried out in this context were examined, was carried out in the form of a systematic review. Within the scope of the research, the data of 15 studies were evaluated. The findings of the research were evaluated in a way that will reveal the trends in this subject, the results of the researches and compare the proctored exams and unproctored exams. Based on the findings, it has been determined that the proctored exam application, in which various technology-based authentication and supervision methods are used in online education processes, is beneficial to use these methods with the same technological opportunities to prevent cheating. It can be said that these efforts can be effective for assessment and evaluation processes, and unproctored applications can be used in online education in various ways.

Keywords: Proctored online exam, non-proctored online exam, online education, systematic review.

INTRODUCTION
The COVID-19 epidemic, which broke out in the world in 2019, is a social phenomenon that affects all social life, health systems, transportation and trade. To cope with this situation, people have completely reorganized their lives. In addition to activities such as socializing activities, travels and trade, extraordinary measures were taken in education as well. After the pandemic process started, schools at all education levels were either suspended or education continued with distance education methods. Especially at higher education level, in educational activities carried out in the form of online or blended learning, problems related to digital access (inequality among students) and unethical behaviors in online exams are among the top issues to be dealt with.

In line with the recommendations of the World Health Organization globally, it has implemented preventive measures such as social distancing (Mbuuge, 2020), quarantine, self-isolation (Williams et al. 2020; Suppawittaya, Yiempfat & Yasri, 2020) and face masking (Zhou et al. 2020) among people to mitigate the spread of COVID-19. The emerging need for social distance, closure of country borders and radical control measures such as the closure of institutions have caused people to review their life routines, which directly affect their habits (Suppawittaya, Yiempfat & Yasri, 2020). In terms of studying and teaching at the height of the COVID-19 pandemic, more than 1.5 billion students, or 91.3 percent of new enrollments globally, were directly affected by school closures (UNESCO, 2020; Crawford et al. 2020; Muftahu, 2020). In addition, the already existing economic, cultural and geographical inequality of opportunity in the world has begun to be felt more deeply (Corlatean, 2020). Online education (used with names such as e-learning, distance education), which is an important way to cope with all these negativities, is explained as the fact that the interaction between the learner and the learning environment cannot be realized due to time or geographical differences, unlike face-to-face education (King et al. 2001). Although it seems to have made the functioning of schools dramatically under pandemic conditions, online or distance education was already popular in paid or free universities globally (Palvia, et al. 2018; Chaney, Chaney & Eddy, 2010). Thanks to the development of online language translator applications, the facilitation of video-based training, and the emergence of online management systems, distance education has become more widespread in recent years and has started to become popular as a training method used to provide individuals with skills (Li & Lalani, 2020; Corlatean, 2020).

In distance or face-to-face education, it is critical to measure what the learner wants to teach and how much he has learned, and to determine whether he is successful as a result of the process. The traditionally used pencil-paper exams have also changed over time. Studies conducted to compare paper-and-pencil exams reveal the advantages of online exams in higher education such as quick response, rich content, easy readability (automatic), flexibility, and creating individual exams in terms of assessment and evaluation (Bayazit & Askar, 2012; Ilgaz & Adanur, 2020; Jeong, 2014). On the other hand, different results may emerge in the assessment and evaluation process regarding the competence/familiarity of individuals to use computers (McClelland & Cuevas, 2020). With the widespread use of technology in education, one of the most important question marks in online applications in
achievement tests or performance tests is the existence of cheating or unethical student behaviors (Carpenter, 2004; Harper, 2006; Kelley & Dooley, 2014). Various studies explain why students exhibit cheating or unethical behaviors due to lack of knowledge or misconceptions (Ersoy, 2014; Çakmak, 2015, Greenberger et al. 2016), academic dishonesty that become easier with the widespread use of technology (Peytcheva-Forsyth, Aleksieva & Yovkova, 2018; Blau & Eshet-Alkalai, 2017; Ravasco, 2012; Curran, Middleton & Doherty, 2011; Bracey, 2005). In some studies, it is stated that the attitudes and behaviors of the instructors in the lessons or in the exam applications (insufficient feedback, difference between the course content and the exam, insufficient effort to prevent it etc.) lead students to academic dishonesty (Özden, Özden & Biçer, 2015; Murdock, Beauchamp & Hinton, 2008; Bertram Gallant, 2017). Parallel to this situation, failure to detect academic dishonesties and impunity also sets a bad example for honest students (Haney & Clarke, 2007; Peytcheva-Forsyth, Aleksieva & Yovkova, 2018). Irregular behaviors commonly encountered in online exams can be listed as follows (King, Guyette Jr & Piotrowski, 2009):

1. Use of books or online source material,
2. Taking the exam for someone else/taking the exam for someone else,
3. Getting help from third parties in the exam,
4. Receiving or disseminating questions,
5. Copying and selling the exam questions to other students or using them in the following years,
6. Violation of the exam period with various excuses,
7. Use of mobile devices for copy communication

In order to cope with academic dishonesties, the view that it may be effective to reorganize the educational environment or the entire instructional design rather than correcting the student is becoming more and more common (Murdock, Beauchamp & Hinton, 2008; Palazzo et al. 2010). Yet overcoming academic dishonesties are a challenging task. Especially the development of technology has brought the conflict between academic dishonesty and honesty to a different dimension (Peytcheva-Forsyth, Aleksieva & Yovkova, 2018). While face-to-face exams require students to physically go to a testing center to take the exam at a certain time (Larson & Sung, 2009; Shapley, 2000), it is possible to conduct virtual/remote proctoring with the help of an application or learning management system in online exams. As a concept, unlike face-to-face exams, the execution of the exams without the need for constant supervision of an official, without a control mechanism, is called unproctored exam (Lievens & Burke, 2011).

The use of techniques that can verify the identity of students in online exams is one of the most important strategic moves. In this context, biometric data such as fingerprint, handwriting, electronic signature, voice recognition, face recognition (comparison with the photograph registered in the system) are checked at the beginning of the exam, as the identity verification strategies used (Online exam control procedures) (Alotaibi, 2010; Sukmandhani & Sutedja, 2019; Wisher, Curnow & Belanich, 2005; Lee-Post & Hapke, 2017; Valera, Valera & Gelogo, 2015; Rodchua, Yiadom-Boakye & Woolsey, 2011). In addition, user authentication methods are used in the form of requesting a password from the participant, password typing speed, and checking with mobile search (Bailie & Jortberg, 2009; Ullah, Xiao & Lilley, 2012; Kang & Kim, 2015). The main purpose of all these strategies is to determine whether the person taking the exam is a real student.

Various procedures have been defined in the literature during the online exam application process after the students have entered the online system thanks to the verification procedures. These procedures include efforts designed to prevent cheating that may occur during the online exam administration process. These technology-assisted efforts are listed as preventing remote assistance applications before the exam, recording student IP addresses, removing students using different IPs from the exam, scheduling the exam at the same time for different countries and time zones, and using the computer only for the exam (Ruipererez-Valiente et al. 2017; Gao, 2012; Amigud et al. 2016; Cluskey Jr, Ehlen & Raiborn, 2011). At the time of the exam, software solutions such as forcing different tabs and browsers to be closed or taking the exam in full screen, taking the exam again with the approval of the instructor (in synchronous exams) are applied (Jung & Yeom, 2009; Kitahara, Westfall & Mankelwicz, 2011; Sinha & Yadav, 2020). In addition, solutions such as not prolonging the exam period unnecessarily, preparing questions and options specifically for each student, and presenting the questions to the student one by one, have been included in the literature in terms of measurement and evaluation science (Bailie & Jortberg, 2009). According to McMurtry (2001) written instructions on exam duration should be consistent with reference materials and allowable forms of communication between students regarding exam questions.

There are suitable solutions to prevent cheating by students who take high-risk exams, both face-to-face and online. However, the concept of proctoring/unproctoring exams in synchronous or asynchronous applications is a concept discussed in the literature to eliminate cheating and irregularities in online exams (Cluskey Jr, Ehlen & Raiborn, 2011; Hollister & Berenson, 2009). In the literature, online proctoring system types are listed with live, recorded
or computer-assisted proctoring application based on automation. Online proctoring types and basic features are given in Figure 1 (Hussein et al. 2020).

Figure 1. Types of Online Proctoring

As seen above, there is a real-time control in the live proctoring application. However, the person who will follow this system is expected to be competent in technology. Live proctoring can be used to detect academic dishonesties such as cheating, along with input operations such as eye movements and face recognition. Recorded proctoring is carried out at the end of the exam by recording the exam environment with the help of various tools and keeping log records. This method requires human intervention and is disadvantageous in terms of time and cost. The automatic proctoring application, on the other hand, has the duty of controlling only with the help of the established system. For this reason, although the labor cost decreases, professional support is needed for the installation of the system.

Discussions about the necessity of human control in online exams show themselves precisely in this area. Within the framework of these discussions, this unsupervised situation that may arise in distance education exams has been tried to be resolved with a hybrid approach in the form of "remote supervision". Various procedures are put into effect for the inspections carried out before the exam and during the exam process in the exam applications where remote inspection is carried out. Figure 2 of the classification related to proctoring in exams.

Figure 2. Proctoring Mechanism Classification in Exams Types of Online Proctoring

The online education approaches of educators are formed depending on the content, the policies of the institution and the general conditions of the classroom. The hybrid audit in Figure 2 usually requires the support of an application or an organization. It is difficult for an instructor to do such an application alone. On the other hand, using the required software or system-based control mechanisms has a monetary cost per exam (54).

Through learning management systems, it is possible for the educator to test students asynchronously and provide immediate corrective feedback with online grade results. This feedback can be done in a number of different test
formats, including multiple choice, true-false, and short answer. Combined with automatic grading, asynchronous testing or practice can give students immediate feedback on their performance, enhancing learning while questions and course content are still fresh. However, the number of studies examining the duration of administration of supervised and unsupervised tests and their impact on learning in an online teaching environment is quite limited. In this context, the aim of this research is to make a systematic analysis of the researches comparing supervised exams and unsupervised exams used in online exam applications.

METHOD

This research was designed as a systematic review study. Systematic reviews, based on a comprehensive review of studies conducted in a particular field, are made to answer a question that the researcher is investigating. Some inclusion and exclusion criteria are used in determining the studies to be systematically reviewed, and the findings in the studies that meet the criteria are synthesized (55). In this context, studies in the literature comparing unproctored and proctored exams have been examined. The purpose of this inclusion criterion is to reveal the advantages and disadvantages with the help of studies related to unsupervised and supervised exams in online education. The findings of the studies examined were evaluated in this respect.

Data Collection and Analysis

The main data of the research consists of research findings that compare proctored and non-proctored exams in terms of features such as academic achievement scores. The findings of the studies were examined in this context and systematic inferences were formed.

To address the objective of this paper, a systematic literature review was undertaken, following the PRISMA approach for article selection (Moher et al., 2009). Nine stages of the systematic review created by Gough (2007) and simplified by Bearman (2012) were applied. Accordingly, detailed inclusion/exclusion criteria as well as articulated search and search strategies should be used prior to database searches. In this context, potential studies to be included in the research were primarily searched by keywords in five databases (Educational database, ERIC, Scopus and Web of Science), and were identified by following the references cited in these studies (both forward and backward). Finally, the process was completed with a scan on Google Scholar.

While searching the databases, “non-proctored* OR proctored* OR unproctored* (Title)” were used in the relevant databases. Search was carried out on the presence of keywords in the titles of the articles. The number of articles accessed in the first stage and the databases accessed are given in Table 1.

<table>
<thead>
<tr>
<th>Database</th>
<th>Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERIC</td>
<td>109</td>
</tr>
<tr>
<td>Scopus</td>
<td>64</td>
</tr>
<tr>
<td>Web of Science</td>
<td>80</td>
</tr>
<tr>
<td>Ebsco</td>
<td>45</td>
</tr>
<tr>
<td>Google Scholar</td>
<td>569</td>
</tr>
</tbody>
</table>

Table 1. Article Hits and Access Databases

From Table 1 we observe that the number of articles with the phrase non-proctored and proctored in their titles, among the most used databases in the field of education. In order to obtain articles comparing two online exam applications, the algorithm used in this study is given in Figure 3.

![Figure 3. PRISMA results.](image-url)
As seen above, the studies that have been reached in the field of education were first identified by using the titles. Because multiple databases are used, an article is likely to be screened in multiple searches. After the pre-screening process, some actions have been taken to prevent the same article from being displayed multiple times (duplicate screening). The articles that will serve the purpose of the study were separated by reading the abstracts, and some articles were accessed with the help of "snowball sampling". Then, the elimination of the articles continued with the determination of the titles in which both words were used. Finally, in this study, in accordance with the purpose of examining the studies comparing two online auditing exam strategies, articles containing "comparison" were selected, and studies to be systematically compiled in the research were obtained. In this context, 15 articles were included in the study, and the findings of the studies were studied.

Educational level was not considered in the analysis process, and the field and geographical preferences of the research were not reflected in the scanning of the study. In addition, calculations such as the effect size of the journals in which the studies were published were not made, and the number of citations (Google scholar) received by the study so far was included in the analysis without any sorting purposes. Table 2 contains information such as the author, publication date and number of citations of the articles accessed for analysis.

<table>
<thead>
<tr>
<th>First Author</th>
<th>Method</th>
<th>Participant</th>
<th>Theme</th>
<th>Country</th>
<th>Cites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wellman, G. S. (2004)</td>
<td>Exam Scores</td>
<td>120 College Students</td>
<td>Medical Terminology</td>
<td>USA</td>
<td>50</td>
</tr>
<tr>
<td>Hollister, K. K. (2009)</td>
<td>Exam Scores</td>
<td>173 College Students</td>
<td>General Student Performance</td>
<td>USA</td>
<td>106</td>
</tr>
<tr>
<td>Prince, D. J. (2009)</td>
<td>Exam Scores</td>
<td>76 College Students</td>
<td>Finance, Marketing And Management</td>
<td>USA</td>
<td>56</td>
</tr>
<tr>
<td>Rios, J. A. (2017)</td>
<td>Exam Scores</td>
<td>1,126 Undergraduate Students</td>
<td>Critical Thinking, Reading, Writing, Mathematics</td>
<td>USA</td>
<td>28</td>
</tr>
<tr>
<td>Daffin Jr, L. W. (2018)</td>
<td>Exam Scores</td>
<td>1,694 College students</td>
<td>Various Courses</td>
<td>USA</td>
<td>47</td>
</tr>
<tr>
<td>Dominguez, C. (2019)</td>
<td>Test Score</td>
<td>1,900 College students</td>
<td>Mathematics</td>
<td>USA</td>
<td>4</td>
</tr>
<tr>
<td>Truszkowski, D. (2019)</td>
<td>Test Score</td>
<td>510 College students</td>
<td>Introductory Programming</td>
<td>USA</td>
<td>7</td>
</tr>
<tr>
<td>Chen, B. (2020)</td>
<td>Exam Scores</td>
<td>86 College Students</td>
<td>introductory financial accounting</td>
<td>USA</td>
<td>5</td>
</tr>
<tr>
<td>Goedl, P. A. (2020)</td>
<td>Exam Scores</td>
<td>73 College Students</td>
<td>Student GPA Scores</td>
<td>USA</td>
<td>12</td>
</tr>
<tr>
<td>Reisenwitz, T. H. (2020)</td>
<td>Exam Scores</td>
<td>400</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Included Articles and Their Features (Listed by Publication Year).
As seen in the table above, 15 articles were selected for the analysis process. The number of citations of the articles examined during the analysis process and the date order of the authors are given in the table. While the total number of citations of the studies examined was 400 at the time of the analysis, it is increasing gradually due to the distance education trend that gained weight due to the pandemic. Studies are generally presented in the form of comparison of test scores. Only one study was conducted in Spain and other studies were conducted in the USA. All the studies reviewed were at the undergraduate level.

Thematic analysis method was adopted for the analysis of the research. The common themes of the articles studied in thematic analysis is an inductive approach (Braun and Clarke, 2006). In this context, a six-stage analysis was conducted. In order to identify the data, the findings of all studies were examined, and the proctored and non-proctored applications used in online exams were examined in the context of statistical differences. In addition, qualitative findings in the articles were also included in the study, unlike the meta-analysis. The findings were tried to be combined in certain themes, and the emerging similarities (patterns) were named. In order to make a review based on the literature, the epistemological structure of the study was tried to be understood by reading the full texts at least three times. The emerging themes were tried to be combined by putting the codes on paper. Finally, the researcher checked the themes and sub-themes reached by an independent researcher and created the final version (Creswell and Miller, 2000; Lincoln, 2007). At the end of this research, 6 final themes emerged.

FINDINGS AND CONCLUSION

In this research, which aims to compare the proctored and non-proctored exam applications used in online exams by using studies on the same subjects, 15 studies were included in the systematic analysis process. This systematic literature review sought to assess the current state of literature concerning comparison of proctored and unproctored exams.

The first results that emerged as a result of the analysis are about the statistical comparison of the test scores of the proctored and non-proctored exams. According to the comparative test results, in some studies examined, unproctored exams caused a significant difference in students' test scores compared to the equivalent group proctored exam (articles 1, 4, 5, 9, 11, 13, 14 and 15). In some studies, no significant difference was found between the two equivalent groups (articles 2, 3, 6, 7, 8, and 12). The explanations about the absence of a significant difference were associated with the compelling nature of the way unproctored tests were administered. However, in one of the studies in which no difference was observed (no: 12), although no difference was observed in some of the students' exam scores, a significant difference was observed in the course success. In another study (no: 7) it was determined that there were differences in students' perspectives on exam practices. Another study (no: 2) in which no significant difference was found includes a comparison with the scores obtained by students in previous years. Therefore, it cannot be shown as a suitable proof for Internet-based proctored systems applied today. However, there are some studies in the literature with similar results (Salinas et al. 2015).

In studies where there was a significant difference, it was observed that the exam scores of the students in the unproctored exams were high. Only one study (no: 1) stated that students' learning levels were high in proctored exams. However, in some studies, significant differences were determined in the average student scores, while in some studies, there were differences in the test scores between the groups. In the last meta-analysis study (no: 10), it was emphasized that unproctored aptitude tests could cause cheating. In the meta-analysis study in question, the importance of unproctored measurements including tasks that cannot be found easily on the internet was mentioned.

While examining the studies in the systematic analysis process, the qualifications of the proctored and unproctored tests were revealed by the authors. The common view on this issue; is that unproctored measurements can cause copy behavior. The literature also reveals similar views (Harmon & Lambrinos, 2008; Karim, Kaminsky & Behrend, 2014). On the other hand, the high cost of online proctored exams, the difficulties in their applicability (technological expertise, hardware requirement, user-friendly environment, exam anxiety, etc.), putting the student under pressure on test-taking behavior have been revealed as negative. For example, as reported in the meta-analysis study, students need online guides much more in online proctored exams. Another study cited the use of random questions as a “best practice” in online education as a proctored exam with additional identity checks. However, it is stated that unproctored applications can be used in quizzes or less important exams. Some studies have stated that time can be seen as a factor in comparing unproctored exams with proctored exams. For example, in articles 9 and 14, it is stated that the exam time, which is strictly applied in unproctored exams, can be abused by students when given the opportunity.

Another issue encountered in the analysis process is some difficulties encountered in the application of proctoring exams. In a study on the subject, some behaviors that are considered suspicious when students are using proctoring software were defined in the system. Suspicous such as switching/closing a different page from the exam browser page (instant or long-term), using keyboard events (basic key logger) in password entry, changing the Internet connection status of the user (voluntarily), multiple mouse clicks (or different sounds), behaviors are discussed (Andreou et al. 2021; Schoenmakers & Wens, 2021). There are decisions that the instructor/institute should take
as to whether these situations are normal or related to cheating. During the systematic review process, opinions were expressed in some articles on this subject. For example, in a study (no: 1), it was stated that students experienced problems in terms of privacy and exam stress in remote proctored exams. In the same study, there is an opinion that the fact that students do not express significant negative opinions may be due to the fact that they are online learners. In another study (no: 9) it was stated that students completed the non-proctored exams in a longer time and that their existing exam anxiety might increase in the proctored exams. In the same study, it is recommended to use code-of-conduct or honor statements, which are not as costly as the proctored application, in order to cope with the academic dishonesty behavior of the students.

As a result, the studies and literature analysis show that the number of articles and the interest in the subject have increased in online education, which has increased recently due to the pandemic. In addition, it has been observed that remote proctoring applications, which are used to prevent cheating behaviors in online exams, can be effective in student performance measurements in some cases. The literature review points out that proctored exams have some question marks about cost, student concerns, and privacy.

LIMITATIONS AND IMPLICATIONS FOR FUTURE RESEARCH

Some limitations to the present systematic review must be noted. The scientific articles that constitute the data of this research are only the full text of the researcher. The researcher included only comparative articles in accordance with the research purpose, and the inferences were limited to these articles. Articles published after the date of the search can be found. In this context, more studies should emerge on the application of unproctored or proctored exams, especially in the context of online education, which has increased due to the Covid-19 pandemic. In terms of researchers and practitioners, it is recommended within the scope of this research to disseminate experimental studies and to make various comparisons about the use of internet applications specialized in proctoring in education.

REFERENCES


Li, C., & Lalani, F. (2020, April). The COVID-19 pandemic has changed education forever. This is how. In World economic forum (Vol. 29).


Salinas, I., & Opton, L. L. Exam Integrity: The Webcam Effect. [http://hdl.handle.net/10755/602858](http://hdl.handle.net/10755/602858)


SYSTEMATIC REVIEW ARTICLES


