

Did they really Work? English Teachers' Attitude towards the Effectiveness of Remote Online Exams in Times of Emergencies

Salim Al-Hashmi^{1,2}

¹ Preparatory Studies Center, University of Technology and Applied Sciences, Nizwa, Oman

² Department of Educational Research, University of Lancaster, Lancaster, UK

Correspondence: Salim Al-Hashmi, Preparatory Studies Center, University of Technology and Applied Sciences, Nizwa, Oman. Department of Educational Research, University of Lancaster, Lancaster, UK.

Received: October 23, 2023

Accepted: November 24, 2023

Online Published: November 27, 2023

doi: 10.5539/elt.v16n12p27

URL: <https://doi.org/10.5539/elt.v16n12p27>

Abstract

This paper explores English teachers' perception of the effectiveness of employing remote online exams during the COVID-19 pandemic after many universities and colleges worldwide moved most of their activities online. The transition to online learning has enabled numerous higher education institutions (HEIs) to utilise options they were previously hesitant to employ due to concerns regarding how they would affect academic integrity. Eight English teachers from a university in Oman were interviewed by email to share their experiences and thoughts on the effectiveness of remote online exams (ROEs). In addition, pre-COVID-19 students' scoresheets were analysed to determine whether students' performance was consistent across face-to-face and remote online exams. The results indicate significant differences in students' scores between the two types of exams, which is attributable to factors such as invigilation issues, students' academic misconduct, and exam security. Finally, the eight interviewed teachers recommended making improvements to online exams as they believe ROEs will continue to be used in the future.

Keywords: remote online exams, academic integrity, contract cheating, COVID-19

1. Introduction

Following the declaration of COVID-19 as a global pandemic in 2020, numerous educational institutions shifted their operations to electronic formats, which brought about significant changes in various sectors, including higher education (Hodges et al., 2020). Notably, calls for the shift from traditional paper-based exams to online exams existed before the pandemic, but the nature and speed of the eventual transition from direct education to distance education led to both unanticipated challenges and opportunities for universities and colleges (Al-Hashmi, 2021; Egarter et al., 2021). One of the challenges that has received widespread attention among educators is the subject of student evaluation.

2. Literature Review

2.1 Advantages of Remote Online Exams (ROEs)

Remote online exams (ROEs) are examinations or assessments that students take remotely, typically from their own location, such as their home or another suitable environment, rather than at a physical testing centre or classroom. These exams are administered and proctored electronically using internet-connected devices, often under specific conditions, to ensure fairness, security, and integrity. ROEs during COVID-19 provided many educational institutions worldwide with an alternative to traditional invigilated paper exams. The main objective of ROEs is to assess students in a way that is non-disruptive to the educational process while simultaneously complying with the precautionary measures established by the health authorities (OECD, 2020).

The literature has identified several advantages of ROEs. For example, online exams offer instant feedback to students (Chan & Liou, 2005; Peat & Franklin, 2002), allowing them to engage in a friendly environment (Aljohani et al., 2021). Moreover, teachers can efficiently complete their marking tasks, reducing examination costs for educational institutions. Awad Ahmed et al. (2021) suggested that online exams '[enable] the provision of new kinds of questions through the use of multimedia, easy ways to gather feedback, instant support, assistance, and aids during the examination, distribution of results, easy use of data, and flexibility of examination timing.'

2.2 Challenges to ROEs

However, Awad Ahmed et al. (2021) also point out that administering ROEs poses several challenges, such as verifying the identity of exam takers and reviewing their submitted scripts for cheating or plagiarism. Several other studies have reported additional challenges, such as the availability or speed of the internet, technical issues, lack of staff training, and the expensive programs used to check for plagiarism (Egarter et al., 2021; Elsalem et al., 2021).

Perhaps the main challenge related to ROEs, though, is maintaining academic integrity. In traditional invigilated exams, academic integrity can be upheld in several ways: seating arrangements, creating different versions of an examination, and enforcing codes of professional behaviour. However, the shift to technology-driven exam administration during the COVID-19 pandemic has given rise to new methods of breaching academic integrity codes. Consequently, there is a pressing need for innovative measures to ensure academic integrity in the context of ROEs (Farland & Childs-Kean, 2021).

2.3 Effectiveness of ROEs

Although many have questioned the effectiveness of online instruction compared to face-to-face instruction, evidence has emerged that remote learning and teaching provide students with a high quality of education (Hodge et al., 2020). However, the same cannot necessarily be said of ROEs. Indeed, recent studies on the effectiveness of ROEs have yielded conflicting results. For example, Marín García et al. (2021) investigated the impact of online exams on academic integrity and students' adaptability to ROEs. The researchers used three exam modalities: an online format from home using the Respondus Lockdown Browser system (Modality 1), online in person using the Respondus Lockdown Browser system with the supervision of a teacher (Modality 2), or a paper format in person invigilated by a teacher (Modality 3). The results revealed that the students favoured Modality 1 (online at home with the Respondus Lockdown Browser system), and there were no statistically significant differences in the scores obtained by students across the three modalities.

Conversely, a study conducted in Jordan by Elsalem et al. (2021), which aimed to evaluate medical students' experiences with ROEs, produced negative outcomes. The researchers found that online exams had a detrimental impact on students, leading to stress and behavioural changes as they struggled to adapt to this new mode of assessment. Another study conducted in Germany by Egarter et al. (2021) reported that even after two semesters of implementing ROEs, the exam mode resulted in unequal opportunities for students, with two main factors affecting students' ability to use technology for their exams: their willingness to use it and the reliability of their internet connection. Moreover, White (2021) reported that students engaged in misconduct in online accounting assessments during COVID-19, noting that those students 'hold complex perceptions around their attitudes towards academic integrity and rationalisations of misconduct behaviour.'

To overcome these concerns about the effectiveness of ROEs, researchers have conducted various experiments to demonstrate that such challenges are manageable. For example, a study by Vazquez et al. (2021), which investigated the differences in students' performance on exams in various environments, concluded that students' scores tend to be lower on proctored online exams than on online exams that are not proctored. Other suggestions include cancelling or postponing exams until the situation improves, redesigning examinations, and substituting examinations with other assessment forms (OECD, 2020).

3. Significance of the Study

The study comes in response to using ROEs as the most reliable tool for assessing students during the pandemic. As these circumstances are unique and unprecedented, the study offers first-hand experience of English teachers in a college in Oman who struggled to navigate the risks of examining students online without any proper training and resources. It is believed that the findings of this study will add to the ongoing discussion about the effectiveness of ROEs and help institutions prepare for the challenges in similar circumstances. The views of the interviewed teachers in this study can inform best practices and guidelines, ultimately shaping the future of assessment strategies, and improving the overall quality and credibility of assessments for remote learning.

4. Research Questions

This study addresses critical questions concerning the effectiveness of ROEs in terms of evaluating student performance, understanding the challenges and risks associated with their implementation, and exploring potential avenues for improvement. These questions attempt to fill a gap in the literature that lacks adequate evidence on the effectiveness of ROEs. The following research questions will guide this investigation:

- (1) How do the sample English teachers view the effectiveness of ROEs?

(2) What risks do these teachers identify about the use of ROEs?

(3) What recommendations do these teachers suggest for improving the use of ROEs?

5. Methodology

5.1 Participants and Context

Eight English teachers working at a university in Oman contributed to the study by providing their thoughts and feedback on using ROEs. The teachers were of different nationalities and varied in their years of teaching experience. They taught students as part of the Foundation Programme (FP), which is designed to enhance new students' proficiency in English, Maths, and IT before they embark on their credit programmes. Within the FP, the students take a placement test that determines their placement in one of the four levels. The results students achieve in Level 4 determine their eligibility for academic qualifications, such as a diploma, higher diploma, or bachelor's degree. If students aspire to pursue a higher qualification than their Level 4 results allow, they can take the IELTS exam at their own expense, with a target score of 5 required to progress to the bachelor's degree level.

5.2 Data Collection Methods

This study employed two qualitative methods to collect the necessary information to address the set objectives. Due to pandemic restrictions and teachers' annual holidays, I opted for email interviews as the first data collection method. According to James (2016), email interviews are increasingly being used to gather educational research data as they provide participants with the time and space needed to reflect on the interview questions and revise their answers. In addition, interviews by email are shifting the control of the interviews from the researcher to the participants. This power shift is valuable as it may lead to detailed and unexpected participant responses (Bowker & Tuffin, 2004).

The second qualitative research method I used for this study was document analysis. According to Bell and Waters (2014), researchers typically use document analysis in conjunction with other qualitative research methods to ensure accuracy through triangulation. In addition, Bowen (2009) argues that by studying information collected through more than one method, the researcher can validate their findings across multiple datasets and thus reduce the potential bias that might arise from relying solely on one method. Hence, I used the data obtained from the scoresheets of Level 4 students to corroborate the themes that emerged from the interviews. This approach ensured that conclusions were not drawn from a single source, contributing to the overall robustness of the research.

5.3 Data Analysis Method

The findings of the project are presented using thematic analysis (Braun & Clarke, 2012). Joffe (2012) suggests that 'thematic analysis is not tied to a particular theoretical outlook, and so can be applied when using a range of theories and epistemological approaches.' Thematic analysis offers several advantages: it is effective at capturing a diverse range of thoughts from participants and then identifying commonalities and trends across those thoughts, and its flexibility enables researchers to modify their project requirements (Braun & Clarke, 2012; Joffe, 2012).

The participants' responses were analysed to identify common themes (Braun & Clarke, 2012). These common themes were then grouped to find similarities and differences in the participants' views under each theme. For instance, one of the themes regarding the threats to ROEs had to be subdivided into several sub-themes, offering readers a more nuanced perspective on each issue.

In addition, quotations are used in the findings section to enable readers to corroborate the authenticity of the analysis and validate the findings by demonstrating the connection between data, interpretation, and conclusions (Corden & Sainsbury, 2006). Careful use of quotations ensures the protection of respondents' identities while still providing valuable insights (Eldh et al., 2020).

6. Findings

This section presents the results of this study, drawing from a combination of email interviews and analysis of pre-COVID-19 exam results. The aim was to address the research questions regarding the effectiveness of ROEs while also considering the associated risks during times of emergency.

6.1 The Inevitability of ROEs

All eight participants agreed that conducting ROEs was the only option available after the closure of their respective institutions. Participant 2 made the following observation:

The immediate and urgent shift to remote assessment during the pandemic was strategically the most practicable solution and alternative adopted to maintain momentum and to guarantee the assurance of the coverage and deployment of assessment as the most crucial feature of students' learning experience.

The participants unanimously agreed that ROEs were the optimal solution, but they had varying opinions regarding their experience of such exams. Some believed that ROEs were wholly ineffective and did not accurately reflect their students' actual performance. In contrast, others viewed it as a learning experience with potential for improvement, citing significant improvements in administering online exams and reporting that the technical problems they faced at the beginning of the academic year had begun to disappear by the third semester.

6.2 Face-to-Face Exams and Online Exams

The majority of the participants agreed that their students may not have attained the same grades they scored on their ROEs if called upon to take identical exams on campus. Participant 4, for instance, speculated that such a discrepancy could occur as a result of students using translation applications to write the required texts during ROEs. In addition, Participant 3 suggested that cheating was the main reason for the possible differences in students' scores between ROEs and face-to-face exams.

However, while most of the participants agreed with the above explanations, two individuals argued that some students might score the same results in either setting. They felt that some students generally performed well in their studies and were therefore likely to replicate their success online. Participant 2 wrote the following:

Students would yield either higher or lower results if different. Generally, it is not an all-inclusive nor exclusive static principle, and this view can be divided and expansive. Students' performances are dependent on students' differing knowledge, skills, ability, etc. and some other essential and consequential factors such as environment, gender, teacher, infrastructure, connectivity, anxiety level, time, and types of tests, among so many others.

Some participants also observed that students might experience frustration during ROEs due to technical issues, such as poor internet connectivity. This frustration, in turn, might have lowered students' performance if they had taken the exams on campus. For instance, Participant 7 mentioned that one of her students was always anxious before exams. This student took her exams in a relative's house due to the poor internet connection in her own area. Participant 7 also talked about her student's embarrassment when her younger sibling entered the room and started shouting and playing.

Moreover, Participant 1 speculated that students' poor understanding of digital technology might have impacted their exam performance. She explained that some students needed help adapting to the new learning system, which included taking ROEs. As a result, in some instances, the required exams had to be rescheduled, cancelled, or postponed.

Tables 1 and 2 offer an alternative perspective on predicting whether students would attain similar grades in face-to-face exams. The figures indicate that the number of students who scored 71 and above (bachelor's degree) rose significantly in the 2020/2021 academic year when ROEs were introduced following the university's closure due to the pandemic. For example, in the second semester of the 2020/2021 academic year, approximately 94% of students obtained a score that qualified them to study for a bachelor's degree, compared to only 20% in the previous year (before the pandemic).

Similarly, Tables 1 and 2 show a substantial decline in the number of students who achieved the diploma level (50–63) in the 2020/2021 academic year, constituting 1% of students in Semester 2 and 5% in Semester 3, whereas the number of students placed at the diploma level surpassed the combined count of those at the bachelor's degree level and the advanced diploma level in Semesters 1 and 2 of 2019/2020.

Another important takeaway from the two tables is the variation in the percentage of students who did not pass. For instance, no students failed in Semester 2 of 2020/2021, whereas approximately 24% failed to pass Level 4 in Semester 2 of 2019/2020. In addition, around 2% of students failed in Level 4 in Semester 3 of 2020/2021 compared to 11% in the same semester of 2018/2019.

Table 1. Degree Distribution Based on Level 4 Results (Online Exams During COVID-19)

Semester	Bachelor's (71–100)	Degree (64–70)	Advanced (50–63)	Diploma (50–63)	Failed (below 50)
April (2021)	71.90%	21.30%	5%	1.90%	1.90%
Jan (2021)	93.90%	5.00%	1.10%	0.00%	0.00%

Table 2. Degree Distribution Based on Level 4 Results (Paper-based Exams Before COVID-19)

Semester	Bachelor's (71–100)	Degree (64–70)	Diploma (50–63)	Failed (below 50)
Sep (2020)	35.50%	28%	13.5%	23%
Jan (2020)	20.30%	15.10%	40.70%	23.90%
Sep (2019)	11.30%	24.40%	53.80%	10.50%
April (2019)	30.90%	27.90%	30%	11%

6.3 Threats to Exam Reliability

The study participants identified several threats that impacted the reliability of ROEs. While most of them agreed that some of these threats diminished with their growing experience of designing and conducting evaluations, it nevertheless remains imperative to address the volume of these threats.

6.3.1 Invigilation

Several participants noted that invigilating students was one of the main threats to the reliability of ROEs. There were several reasons for this. For example, one participant complained about the poor quality of cameras and internet connection, as these made it challenging for invigilators to verify students' identities. She wrote the following:

There were many times when I could not see the face or the ID of the student because of the blurry image. I reported the issue to the management, and some came to check. But as you know, having all the other problems, such an issue might be overlooked.

Another participant explained that, even if invigilators could see the students taking the exam, their view of what the students were actually doing may still have been obscured. He commented:

For administrative, cultural, and technical reasons, the invigilator could see only the student (usually to the chest level) on the screen, which was not helpful to check if there was any cheating/malpractice. In addition, the camera was not showing the monitor nor the adjacent places to check if a student was getting any help.

6.3.2 External Help

Most respondents discussed instances where it was evident that some students received assistance from external sources during ROEs, particularly in the writing sections of the exams. One of the primary sources of help was translation applications. Given that many of the students were fluent in Arabic, some were able to use these applications to translate both questions and answers, often without attracting the invigilator's attention. Participant 3 elaborated on this issue:

There was a possibility of using Google Translate or other translation apps. Students could type their essays/answers in their mother tongue and translate them into English using such apps. These applications are so advanced now that you can't question the quality of the language.

Participant 7 echoed this observation, emphasising that students could evade detection because plagiarism detector programs like Turnitin are unable to identify this type of plagiarism. In addition, Participant 4 commented on the differences in performance in writing and speaking:

I have assessed students' speaking, and I have found that some needed to improve in English. They hardly say any complete sentences. But I was surprised when I looked at their writing, and I'm sure that this is because they use Google Translate, something they couldn't do in speaking. After all, we watched them during the speaking assessment.

The second source of external help that students may have utilised was websites. Some participants mentioned cases where the writing submitted by students was clearly not their own work.

Some students have definitely got help from the internet. Therefore, you could easily say that the level of English is not that of a student studying in the foundation programme. We could identify some essays that have been taken from the internet and awarded them zero.

Participant 5 clarified that some students managed to evade detection when engaging in plagiarism. She explained that these students used websites that paraphrased the original work from the internet and then submitted the essays without leaving a trace.

Students learn new tricks every time we set rules. For example, after we awarded zero to students for submitting plagiarised work, some of them started looking for ways to change a word here or modify the structure there with the help of AI paraphrasing tools. Unfortunately, using these tools made it difficult for our plagiarism detector to find.

In addition to the two sources of external help mentioned above, two of the participants raised their concern that some students might have solicited the help of freelancers. One of them expressed his surprise when marking several essays with perfect English. He said the following:

If you think about it, there is a high possibility that this happened. We used MS Teams to invigilate the students while they used Moodle to answer the questions. So, a student may have given his credentials to a freelancer who could access Moodle and answer the exams instead of the student. I have heard about this in other institutions and believe it happened in ours.

The biggest concern expressed by all the participants was the inflation in the listening and reading scores. For instance, they believed that some students could share their answers through WhatsApp. Participant 6 voiced this concern:

I can't explain how most students got high marks in listening and reading. Yes, some of them were really good. But, when most students get these high marks in these two skills, you start thinking they might share the answers through their WhatsApp groups or other means.

6.3.3 Exam Design and Specification

Several participants assumed that the design and specifications of online exams could have improved their reliability. One participant stated the following:

When online exams were introduced, I thought the format would be different to tackle the anticipated issues like malpractice. But, to my surprise, the e-exam format was the same as the paper-based exams used before the pandemic. So, the format was one of the reasons for the inflation of marks.

Another participant suggested that a leading cause of the inflation of marks was the question types used in the reading and listening section. She explained the following:

Most listening and reading questions were multiple-choice and true/false questions, which makes me think that even weak students can get easy marks if they tick answers randomly. Unfortunately, these types of questions are also easy to cheat on when students take exams from their homes.

6.3.4 Marking

Despite the importance of marking and the crucial role of markers in evaluating students, only two participants stated that marking was a risk to the reliability of ROEs, particularly in writing. Participant 7 explained that some markers became frustrated by the number of students whose work contained phrases and sentence structures that exceeded their actual ability levels. On the one hand, these markers ignored the issue and graded students' work as per the guidelines without reporting any malpractices, fearing that they would be asked to reevaluate the exams again. On the other hand, Participant 1 noted that there is a prevailing belief among several markers that being stringent in grading students' writing would have minimal impact. The students engaging in these unethical practices, they argue, would pass regardless, as their scores in other skills are very high. Therefore, the consequences were deemed to ultimately be insignificant, even if these students were graded zero in their writing skills.

To pass, a student has to score 50. However, I haven't seen any work done to compare the scores in the four different skills to validate the achieved grades. For example, some colleagues say that a student might be good at listening and weak at writing. But I do not favour this explanation because the differences are sometimes bigger than expected.

Participant 5, meanwhile, raised the issue of fatigue when grading students' online work, which might negatively impact the consistency of markers. He explained that 'marking is challenging when it comes to online exams. Double marking the writing scripts on-screen strains eyes and causes physical discomfort.'

6.3.5 E-exam Security

Another participant highlighted the issue of ROE security. She explained that the security of exams was always at stake in the institution:

Students were allowed to use two devices at our university to ensure that their cameras were on one device and a second to take the English e-exam. It was a testing time for the invigilators as they could not guarantee security.

Identity verification and monitoring the examinees during the conduct of the exam was a tedious task for the examiners.

To support this suggestion, Tables 3 and 4 below demonstrate the average writing, reading, listening, and speaking scores of the online version of the Level 4 Exit Exam during COVID-19 and the paper-based version before COVID-19, respectively. Table 3 covers two semesters of online exams from the university under study, while Table 4 shows the average score for six semesters before COVID-19. Each of these four components contributes 25 points to the total mark.

A closer look at the two tables indicates that the differences in the average scores in writing and speaking were marginal. For example, the average writing score was almost one point higher than the nearest average score (Semester 1 of 2019/2020) before COVID-19. Likewise, the average speaking score in Semester 2 during ROEs was almost the same as the average speaking score during the six semesters before the pandemic.

However, the differences in average scores in listening and reading are more pronounced for ROEs than for paper exams. For example, the data shows that students scored an average of nearly 22 points in the listening part of the online English Level 4 Exit Exam compared to an average of between 11.5 and 12.3 in the six semesters before shifting to ROEs. Likewise, students scored around 23 in reading in the off-campus online exam, while in the previous six semesters, students averaged between approximately 13 and 16 points.

Table 3. Average of Students' Scores in Reading, Writing, Listening, and Speaking (Online Exams During COVID-19)

Semester	Average Score	Writing	Average Score	Listening	Average Score	Reading	Average Score	Speaking
April (2021)	16.81		21.8		22.09		16.98	
Jan (2021)	17		22		23		16	

Table 4. Average of Students' Scores in Reading, Writing, Listening, and Speaking (Paper-based Exams Before COVID-19)

Semester	Average Score	Writing	Average Score	Listening	Average Score	Reading	Average Score	Speaking
Sep (2020)	14.3		11.5		12.7		16.7	
Jan (2020)	14.4		11.9		13		16.9	
Sep (2019)	15.5		12.3		12.9		16.7	
April (2019)	14.8		11.6		16.3		17.2	
Jan (2019)	14.9		14.7		14.5		17.2	
Sep (2018)	14.6		12.6		13.3		17.2	

6.4 The Management of Online Exams

Most participants found the management of online exams to be acceptable, but they did identify some areas for improvement. Participant 6, for example, commented on the effectiveness of the management of online exams while also noting the potential for unreliability:

Indeed, throughout the year, we were used to the process. We could minimise the invigilation problems and build effective marking protocols; however, this doesn't mean our exams were reliable.

In addition, participants praised the rapid improvement in marking and welcomed the changes made by the administration concerning suspected malpractices, particularly in writing:

I was happy when the administration asked teachers to report all the suspected malpractices regardless of their number. Students were informed of the new penalties for writing more than the targeted number of words and translating from Arabic to English. We had seen an improvement.

Another participant raised the issue of students' attitudes towards academic integrity. She cited her experience with changing the format of the listening and reading sections in the Level 4 Exit Exam:

I was one of the teachers who protested against the exam format because I thought it was the reason for the inflation of marks in listening and reading. So, the design was changed, but the scores remained high. It made me feel that some students would always find a way to cheat in ROEs.

6.5 Suggestions for Improvement of ROEs

Given the evaluative nature of this study, it was essential to gather suggestions from the interviewed staff on how to improve ROEs. Their recommendations included the following:

- (1) Use safe browsers that prevent students from using other applications during online exams.
- (2) Introduce fingerprint devices to verify students' identities. These devices will reduce the likelihood of individuals other than the students taking the exam.
- (3) Design exams that assess language skills rather than knowledge-based skills.
- (4) Create multiple versions of exams, particularly in reading and listening.
- (5) Shift the focus from achievement assessments to formative assessments, prioritising the evaluation of the learning process over outcomes.
- (6) Introduce and clearly communicate new penalties for cheating and using translation devices, such as apps, websites, and freelancers.
- (7) Conduct face-to-face interviews with students who score high marks in listening and reading online but low marks in writing.

7. Discussion

The findings of this study provide valuable insights into English teachers' perceptions of ROEs during the COVID-19 pandemic in Oman. These insights align with the existing literature on remote assessment, educational technology, and the challenges associated with maintaining the integrity of online examinations.

7.1 Effectiveness of ROEs in Emergency Situations

The participants in this study unanimously recognised the necessity of ROEs as a response to the abrupt closure of their institution during the pandemic. This aligns with existing literature that highlights the vital role of online assessments in ensuring the continuity of education during emergencies (Hodges et al., 2020). Such assessments are crucial to academic progress when emergencies disrupt face-to-face education (Elsalem et al., 2021).

The participants' differing opinions on the effectiveness of ROEs are reflective of larger debates about the quality and authenticity of online assessment. While some educators view online exams as a practical solution with room for improvement (Khan et al., 2021), others express concerns about their ability to measure student learning accurately (Farland & Childs-Kean, 2021). This tension between the necessity of remote assessments in emergencies and their limitations in replicating in-person experiences is a recurring theme in the literature (Elsalem et al., 2021; Marín García et al., 2021; OECD, 2020).

The significant increase in high scores observed in this study during the academic year in which ROEs were implemented raises questions about the comparability of online and offline assessments. This finding resonates with existing evidence of potential grade inflation in online environments in other contexts, such as Turkey (Karadag, 2021) and Saudi Arabia (Al-Jarf, 2022). This discrepancy between student performance in ROEs and traditional exams underscores the need for rigorous assessment design and monitoring to ensure fairness and validity (Awad Ahmed et al., 2021; Mate & Weidenhofer, 2022).

7.2 Risks and Threats to Exam Reliability in ROEs

The current study uncovered various threats to the reliability of ROEs, including invigilation challenges, external assistance, exam design, marking issues, and e-exam security concerns. These themes align with existing literature on the vulnerabilities of online assessments.

7.2.1 Invigilation Challenges

The challenges identified by the participants regarding invigilation, such as blurred images and limited visibility, are consistent with prior research highlighting the complexities of proctoring in online environments (Mutawa & Sruthi, 2021; Raman et al., 2021). Several studies suggest that proctoring technology and techniques require ongoing development to address these challenges effectively (Al-Jarf, 2022; Mate & Weidenhofer, 2022; Mutawa & Sruthi, 2021; Raman et al., 2021).

7.2.2 External Assistance

The use of external aids, including translation applications and websites, to manipulate exam results, as discussed by this study's participants, has been a long-standing concern with online assessments, as corroborated by other studies (Momeni, 2022; Naghdipour, 2022). In addition, studies have reported a rise in contract cheating

or ghost-writing (Erguvan, 2021a). The ability of students to exploit technology to facilitate academic dishonesty underscores the need for innovative anti-cheating measures (Erguvan, 2021b; Farland & Childs-Kean, 2021).

7.2.3 Exam Design

The questioning methods employed in exam design, particularly multiple-choice and true/false questions, can be ineffective. Hence, this study, along with previous studies, calls for careful consideration of assessment formats and the alignment of assessment methods with learning objectives (Butler, 2018; Whisenhunt et al., 2022). Research suggests that a well-designed assessment can enhance student engagement and mitigate cheating (Erguvan, 2021b; Naghdipour, 2022; Raman et al., 2021).

7.2.4 Marking Issues

While not extensively discussed in our study, we highlight potential issues related to the consistency and fairness of marking in ROEs. This theme aligns with research emphasising the importance of clear and reliable grading rubrics and processes in online assessments (Gaytan & McEwen, 2007).

7.2.5 E-exam Security

The concerns raised by our participants regarding the security of ROEs, especially identity verification and monitoring, resonate with broader discussions on maintaining the integrity of online assessments (Ghizlane et al., 2019). Innovations in secure online testing systems are crucial to addressing these concerns (Baseer et al., 2022; Erguvan, 2021b; Mitra & Gofman, 2016).

7.2.6 Management of Online Exams

The participants generally found the management of online exams to be acceptable but still identified areas for improvement. Their comments align with the literature on best practices for online exam administration, emphasising the need for clear guidelines, faculty development, and continuous improvement of exam protocols (Ghanbari & Nowroozi, 2021).

7.3 Suggestions for Improvement of ROEs

The suggestions provided by the participants to enhance the reliability and fairness of ROEs align with recommendations in the existing literature. These include the use of secure browsers (Slusky, 2020), biometric authentication (Traoré et al., 2017; Vegendla & Sindre, 2019), diversification of assessment formats (Abood & Maizer, 2022), a shift towards formative assessment (Mohamadi, 2018), and the implementation of clear academic integrity policies (Khan et al., 2021; Mitra & Gofman, 2016; Traoré et al., 2017; Whisenhunt et al., 2022).

In conclusion, the findings of this study contribute to the ongoing discourse on remote online assessments by contextualising the experiences and perceptions of English teachers in Oman. These findings underscore the importance of addressing the multifaceted challenges and vulnerabilities associated with ROEs during emergencies, aligning with existing literature on online assessment practices. Implementing the suggested improvements and utilising existing research can help institutions effectively navigate the evolving landscape of remote assessment. To achieve this, it will be necessary to engage in greater collaboration and to undertake more research actions to refine online assessment practices in the face of future uncertainties.

Acknowledgement

This paper originated as an assignment submitted as part of the author's PhD studies and has undergone substantial revisions to meet the standards and requirements of journal publications. The author acknowledges the support of academic staff on the Doctoral Programme in Higher Education Research, Evaluation, and Enhancement at Lancaster University.

References

- Abood, H., & Maizer, M. A. (2022). Strategies to Address Cheating in Online Exams. *International Journal of Technology in Education (IJTE)*, 5(4), 608-620. <https://doi.org/10.46328/ijte.256>
- Al-Hashmi, S. (2021). A Study on the Impact of the Sudden Change to Online Education on the Motivation of Higher Education Students. *Higher Education Studies*, 11(3), 78. <https://doi.org/10.5539/hes.v11n3p78>
- Al-Jarf, R. (2022). Grade Inflation at Saudi Universities Before, During and After the Pandemic: A Comparative Study. *Journal of Humanities and Social Sciences Studies*, 4(4), 111-125. <https://doi.org/10.32996/jhsss.2022.4.4.15>

- Aljohani, A., Aloreafy, M., Alzaidi, S., & Meccawy, Z. (2021). Saudi EFL Students' Perceptions of Online English Achievement Exams in the Era of COVID-19. *International Journal of Linguistics, Literature and Translation*, 4(5), 118-125. <https://doi.org/10.32996/ijllt.2021.4.5.11>
- Awad Ahmed, F. R., Ahmed, T. E., Saeed, R. A., Alhumyani, H., Abdel-Khalek, S., & Abu-Zinadah, H. (2021). Analysis and challenges of robust E-exams performance under COVID-19. *Results in Physics*, 23, 103987. <https://doi.org/10.1016/j.rinp.2021.103987>
- Baseer, K. K., Jahir Pasha, M., Rama, A. V., Reddy, K., Rekha, K., & Begum, M. S. (2022). Smart Online Examination Monitoring System. *JOURNAL OF ALGEBRAIC STATISTICS*, 13(3), 559-570. <https://publishoa.com>
- Bell, J., & Waters, S. (2014). *Doing your research project: A Guide for First-time Researchers* (6th ed.). McGraw Hill Education.
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27-40. <https://doi.org/10.3316/QRJ0902027>
- Bowker, N., & Tuffin, K. (2004). Using the Online Medium for Discursive Research About People With Disabilities. *Social Science Computer Review*, 22(2), 228-241. <https://doi.org/10.1177/0894439303262561>
- Braun, V., & Clarke, V. (2012). Thematic analysis. In APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological. (pp. 57–71). *American Psychological Association*. <https://doi.org/10.1037/13620-004>
- Butler, A. C. (2018). Multiple-choice testing in education: Are the best practices for assessment also good for learning? *Journal of Applied Research in Memory and Cognition*, 7(3), 323-331. <https://doi.org/10.1016/j.jarmac.2018.07.002>
- Chan, T., & Liou, H.-C. (2005). Effects of Web-based Concordancing Instruction on EFL Students' Learning of Verb–Noun Collocations. *Computer Assisted Language Learning*, 18(3), 231-251. <https://doi.org/10.1080/09588220500185769>
- Corden, A., & Sainsbury, R. (2006). Exploring 'Quality': Research Participants' Perspectives on Verbatim Quotations. *International Journal of Social Research Methodology*, 9(2), 97-110. <https://doi.org/10.1080/13645570600595264>
- Egarter, S., Mutschler, A., & Brass, K. (2021). Impact of COVID-19 on digital medical education: compatibility of digital teaching and examinations with integrity and ethical principles. *International Journal for Educational Integrity*, 17(1), 18. <https://doi.org/10.1007/s40979-021-00084-8>
- Eldh, A. C., Årestedt, L., & Berterö, C. (2020). Quotations in Qualitative Studies: Reflections on Constituents, Custom, and Purpose. *International Journal of Qualitative Methods*, 19, 160940692096926. <https://doi.org/10.1177/1609406920969268>
- Elsalem, L., Al-Azzam, N., Jum'ah, A. A., & Obeidat, N. (2021). Remote E-exams during Covid-19 pandemic: A cross-sectional study of students' preferences and academic dishonesty in faculties of medical sciences. *Annals of Medicine and Surgery*, 62, 326-333. <https://doi.org/10.1016/j.amsu.2021.01.054>
- Erguvan, I. D. (2021). The rise of contract cheating during the COVID-19 pandemic: a qualitative study through the eyes of academics in Kuwait. *Language Testing in Asia*, 11(1), 34. <https://doi.org/10.1186/s40468-021-00149-y>
- Farland, M. Z., & Childs-Kean, L. M. (2021). Stop tempting your students to cheat. *Currents in Pharmacy Teaching and Learning*, 13(6), 588-590. <https://doi.org/10.1016/j.cptl.2021.01.035>
- Gaytan, J., & McEwen, B. C. (2007). Effective Online Instructional and Assessment Strategies. *American Journal of Distance Education*, 21(3), 117-132. <https://doi.org/10.1080/08923640701341653>
- Ghanbari, N., & Nowroozi, S. (2021). The practice of online assessment in an EFL context amidst COVID-19 pandemic: views from teachers. *Language Testing in Asia*, 11(1), 27. <https://doi.org/10.1186/s40468-021-00143-4>
- Ghizlane, M., Hicham, B., & Reda, F. H. (2019). 2019 International Conference on Systems of Collaboration Big Data, Internet of Things & Security (SysCoBioTS). *A New Model of Automatic and Continuous Online Exam Monitoring*, 1-5. <https://doi.org/10.1109/SysCoBioTS48768.2019.9028027>

- Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). *The Difference Between Emergency Remote Teaching and Online Learning*. <https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning>
- James, N. (2016). Using email interviews in qualitative educational research: creating space to think and time to talk. *International Journal of Qualitative Studies in Education*, 29(2), 150-163. <https://doi.org/10.1080/09518398.2015.1017848>
- Joffe, H. (2012). Thematic Analysis. In D. Harper & A. Thompson (Eds.), *Qualitative Research Methods in Mental Health and Psychotherapy: A Guide for Students and Practitioners* (pp. 209–223). Wiley-Blackwell. <https://doi.org/10.1002/9781119973249.ch15>
- Karadag, E. (2021). Effect of COVID-19 pandemic on grade inflation in higher education in Turkey. *PLOS ONE*, 16(8), e0256688. <https://doi.org/10.1371/journal.pone.0256688>
- Khan, Z. R., Sivasubramaniam, S., Anand, P., & Hysaj, A. (2021). ‘e’-thinking teaching and assessment to uphold academic integrity: lessons learned from emergency distance learning. *International Journal for Educational Integrity*, 17(1), 17. <https://doi.org/10.1007/s40979-021-00079-5>
- Marín García, P. J., Arnau, A., & Llobat, L. (2021). Preferences and scores of different types of exams during COVID-19 pandemic in faculty of veterinary medicine in Spain: A cross-sectional study of paper and E-exams. *Education Sciences*, 11(8). <https://doi.org/10.3390/educsci11080386>
- Mate, K., & Weidenhofer, J. (2022). Considerations and strategies for effective online assessment with a focus on the biomedical sciences. *FASEB BioAdvances*, 4(1), 9-21. <https://doi.org/10.1096/fba.2021-00075>
- Mitra, S., & Gofman, M. I. (2016). *Towards Greater Integrity in Online Exams*, 1-5.
- Mohamadi, Z. (2018). Comparative effect of online summative and formative assessment on EFL student writing ability. *Studies in Educational Evaluation*, 59, 29-40. <https://doi.org/10.1016/j.stueduc.2018.02.003>
- Momeni, A. (2022). Online Assessment in Times of COVID-19 Lockdown: Iranian EFL Teachers’ Perceptions. *International Journal of Language Testing*, 12(2).
- Mutawa, A. M., & Sruthi, S. (2021). Students’ Perspective Towards Online Proctoring in Exams During Covid-19. *Journal of Engineering Research*. <https://doi.org/10.36909/jer.14749>
- Naghdipour, B. (2022). ICT-enabled informal learning in EFL writing. *Journal of Second Language Writing*, 56, 100893. <https://doi.org/10.1016/j.jslw.2022.100893>
- OECD. (2020). *Remote Online Exams in Higher Education During the COVID-19 Crisis*.
- Peat, M., & Franklin, S. (2002). Supporting student learning: the use of computer-based formative assessment modules. *British Journal of Educational Technology*, 33(5), 515-523. <https://doi.org/10.1111/1467-8535.00288>
- Raman, R., B, S., G, V., Vachharajani, H., & Nedungadi, P. (2021). Adoption of online proctored examinations by university students during COVID-19: Innovation diffusion study. *Education and Information Technologies*, 26(6), 7339-7358. <https://doi.org/10.1007/s10639-021-10581-5>
- Slusky, L. (2020). Cybersecurity of Online Proctoring Systems. *Journal of International Technology and Information Management*, 29(1), 56-83. <https://doi.org/10.58729/1941-6679.1445>
- Traoré, I., Nakkabi, Y., Saad, S., Sayed, B., Ardigo, J. D., & de Faria Quinan, P. M. (2017). *Ensuring Online Exam Integrity Through Continuous Biometric Authentication*. In *Information Security Practices* (pp. 73–81). Springer International Publishing. https://doi.org/10.1007/978-3-319-48947-6_6
- Vazquez, J. J., Chiang, E. P., & Sarmiento-Barbieri, I. (2021). Can we stay one step ahead of cheaters? A field experiment in proctoring online open book exams. *Journal of Behavioral and Experimental Economics*, 90, 101653. <https://doi.org/10.1016/j.socec.2020.101653>
- Vegendla, A., & Sindre, G. (2019). Mitigation of Cheating in Online Exams: Strengths and Limitations of Biometric Authentication. In A. V. S. Kumar (Ed.), *Biometric Authentication in Online Learning* (pp. 47–68). IGI Global. <https://doi.org/10.4018/978-1-5225-7724-9.ch003>

- Whisenhunt, B. L., Cathey, C. L., Hudson, D. L., & Needy, L. M. (2022). Maximizing learning while minimizing cheating: New evidence and advice for online multiple-choice exams. *Scholarship of Teaching and Learning in Psychology*, 8(2), 140-153. <https://doi.org/10.1037/stl0000242>
- White, A. (2021). May you live in interesting times: a reflection on academic integrity and accounting assessment during COVID19 and online learning. *Accounting Research Journal*, 34(3), 304-312. <https://doi.org/10.1108/ARJ-09-2020-0317>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).