



# e-Examinations: The Impact of Technology Problems on Student Experience

Mathew Hillier, Naveen Kumar, Nirmani Wijenayake

## ► To cite this version:

Mathew Hillier, Naveen Kumar, Nirmani Wijenayake. e-Examinations: The Impact of Technology Problems on Student Experience. Open Conference on Computers in Education (OCCE), Jan 2020, Mumbai, India. pp.35-45, 10.1007/978-3-030-59847-1\_4 . hal-03519208

**HAL Id: hal-03519208**

**<https://inria.hal.science/hal-03519208>**

Submitted on 10 Jan 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# e-Examinations: The Impact of Technology Problems on Student Experience

Mathew Hillier<sup>[0000-0001-9990-0479]</sup>, Naveen Kumar and Nirmani Wijenayake

University of New South Wales, Sydney, NSW, Australia  
m.hillier@unsw.edu.au

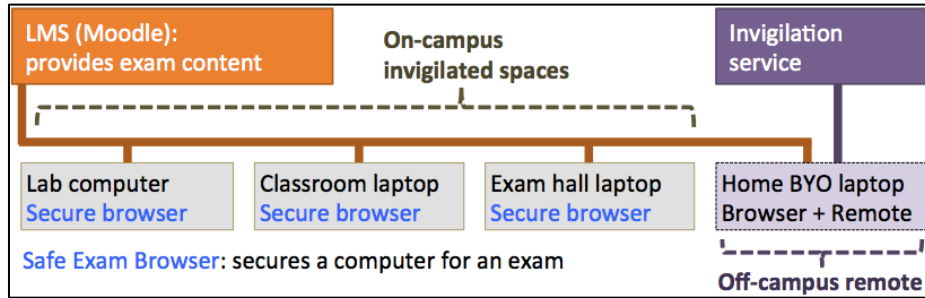
**Abstract.** This study investigated the impact of technology problems on students' perceptions of computerised examination technology and procedures. Measures included the suitability of the assessment task to computerisation, ease of use of the e-examination (e-exam) software, technical reliability, and the perceived security of the approach. A case study was conducted around the introduction of computerised tests into a second-year undergraduate biochemistry course. A series of three e-exam trial events were conducted at an Australian university in 2019 using laboratory bench computers. All students in the course were required to undertake the series of computerised examinations. Data were gathered using pre-post surveys of students' perceptions ( $n = 215$ ) that included qualitative comments and Likert items. This study focuses on the impact of a server slowdown at one of the sessions upon participants' responses to Likert survey items that included their recommendation of the e-exam approach to peers.

**Keywords:** E-exams, Assessment, Student perceptions, Acceptance.

## 1 Introduction and Background

The study extends a series of computerised examination (e-exam) studies following an Australian Government funded project on the topic [1]. The results of previous studies are available [2-5]. The Australian-based work is in company with a number of similar projects underway in other countries [6-8]. This study aimed to explore the viability of re-purposing biology teaching laboratories for use as an examination (exam) space for students undertaking STEM courses. As a result of this circumstance, the definition of an e-exam in this study was simplified in a pragmatic manner over that espoused in previous studies in the series where the focus was previously on the development of authentic assessment capabilities [9, 10] in the exam room [3, 11, 12]. The previous studies explored ideas of authentic tasks in exams and the use of bring-your-own-device (BYOD) [3, 12]. While the use of BYOD is beyond the scope of this study, prior work [13, 14] included exploration of issues relevant to the student experience of computerised testing. In this case study, we limited the e-exam to an event supervised by human invigilators undertaken on controlled, institution-owned computers with the exam content served via the quiz functionality of a learning management system (Moodle). This study was situated as the first phase of a broader approach to implementing digital exams at the host university. As such, this study sits

within the context of a broader digital exams deployment (see Fig. 1) that includes off-campus via remote invigilation, on-campus in classrooms using BYOD, in exam halls and, as in the case of this study, in computer-equipped spaces such as a laboratory where computers were already being used for formative learning tasks.



**Fig. 1.** A holistic digital exam architecture using common core technologies across contexts.

We next explore literature related to user acceptance of technologies, with a focus on the role technical reliability can have in students' attitudes towards the use of computerised examinations.

## 2 Literature

As we have previously explored in prior work [3, 4], it is the perceptions of users of a technology that play a large part in the acceptance of that technology by individuals and organisations. We review those elements of the literature that are relevant to this particular study. The information systems literature [15, 16] has long regarded ease of use and usefulness as being key issues when it comes to people accepting new information technology. It is acknowledged that while a range of other variables can be taken into consideration [17], this study will focus on user perceptions. Prior work undertaken in the area of student acceptance of e-exams has shown that relevance, ease of use, reliability and security have been key concerns [3, 4, 13, 14, 18]. Given that students are those with the most at stake with a change to e-exams, it behoves educational institutions that their views are taken into consideration.

Researchers at the University of Bradford in the United Kingdom (UK) undertook a survey of students' opinions following their use of the online quiz-based testing tool "QuestionMark Perception" [18]. The range of issues that were canvassed served as a starting point for the survey tools used in this study. Subsequent work in Australia has included a survey [13] that sought to tap into students' preconceptions of e-exams with follow-up work at three universities also examining students' opinions after they had taken an e-exam [3, 4, 14] using similar technologies.

In the aforementioned studies, it became apparent that the issue of technical reliability had a significant place in the minds of students. The news media in Australia recently publicised a high profile example of this risk playing out [19, 20], in which a

high stakes medical board e-exam failed half way through resulting in the exam being cancelled. In the researchers' own work, it was previously reported [13] that students yet to undertake an e-exam placed the fear of technology failure high on their list of concerns. Fortunately, this concern was shown to dissipate after students had undertaken one or more e-exams that went well [3-5, 14].

In seeking to minimise outages, modern information system designers endeavour to build in technology failure mitigation measures. In terms of an e-exam system, this can include adding a frequent auto-save trigger to an online quiz system (as per Moodle quiz), adding redundant components into the network infrastructure such as hot swappable routing components, multiple network connections and placing the server on scalable cloud infrastructure. Yet the cost increases of such measures, or the lack of control over all parts of the connection chain, can mean that there is less than complete coverage for all components in the chain between the exam candidate and the exam server. Such measures may also have their limits in terms of how effective they will be in a crisis, when a highly time-sensitive and stressful event such as an exam is being run. The extent of the risk and fallout is further highlighted by lists [21] and high profile cases [22]. When client-server software systems are used, not all fail-over measures are immediate and a total severing of the connection can still result in an unscheduled end to the exam or a delay to the start of the exam [23]. The latter also occurred in the current study. However, in our series of work, it remained to be seen what the nature of the impact on students' feedback about their e-exam experience would be in the event that technical issue did occur that impacted the running of the exam event. This paper explores just this scenario.

### 3 Study Context

The study draws on ideas and techniques previously developed as part of a completed Australian national project on e-exams using bring-your-own (BYO) laptops [1-4, 12]. The same user survey instrument was used in this study to enable comparisons to the previous cited work (ibid). This study differs because it examined the use of a lock-down browser "Safe Exam Browser" (SEB) installed within university-owned computers located on biochemistry laboratory benches, instead of the alternative operating system boot method used on BYO laptops utilised in the previous studies (ibid). These laboratories are the same spaces used for scheduled practical classes for the course where formative learning activities take place. This study was undertaken at the University of New South Wales, Australia within an undergraduate second year biochemistry course (subject). The exam trials were carried out using online Moodle quizzes within the institutional learning management system (LMS). SEB provided a key capability to allow white-listed access to selected resources and software tools. This feature allowed us to go beyond a typical 'locked-down quiz' to allow controlled access to students' prepared notes in a digital format during the exam. In this case, the online LabArchives service [24] was used as the host environment for students' notes. Assessments were undertaken in-class, under supervised conditions at the mid-point of the term and at the end of the term as a final exam. The practice session did not

involve any grading, while the mid-term test and final exam were each weighted at 25% of the course. Multiple classes were split over four laboratory rooms with morning and afternoon sessions run to allow all students to use a computer. Questions in both exam sessions were in the format of selected response (multiple-choice questions) and an extended essay style response. The study was run in conjunction with the third author who was the course coordinator and lead teacher in the classes in which the e-exam trials were conducted.

## 4 Research Questions

Given the issues raised in the literature review and the circumstances in which the study was carried out, we focus this paper on two main research questions:

- 1) “Would the approach of using SEB for Moodle quiz-based exams undertaken within a biochemistry laboratory setting be acceptable for students in terms of perceived suitability, ease of use, reliability and security?”, and, given the events as they transpired,
- 2) “What was the impact on these measures of any technology problems that occur during an exam event?”

In seeking to explore the impact of technical problems, the severity of the problem is worth considering. A problem that is catastrophic will prevent the exam from starting or will cause a complete break in proceedings. This would negate any data collected, because the students would not have experienced a complete exam. However, problems that occur that are moderate and cause inconvenience rather than a complete failure mean that students experience the full exam procedure. This provides an opportunity to explore the extent to which such technical difficulties influence students’ perceptions of various aspects of the e-exam process.

## 5 Method

This study utilised a case study approach that sought to capture students’ perceptions via a quantitative survey. Procedures followed those used in prior work [4, 18]. The procedure used is outlined in Table 1 and was designed to be similar to that described in previous work [3, 4] in order to facilitate comparison. This study included a short series of events that comprised one practice session, then two weeks later one mid-term test and four weeks onward one final exam. The study was conducted in term 2 in 2019 involving 215 undergraduate university students.

Previous experience has shown that a zero stakes practice run was important in providing an opportunity for students to become familiar with e-exam procedures and the software environment before undertaking a real e-exam. The ungraded practice session was run in class time with all students participating. All students were asked to use a computer for the exams, and therefore this study contrasts to another study [4] where students were provided a choice to type or hand-write their exam responses. However, university special consideration (alternative assessment) was available to

students in accordance with university procedures that would have allowed an opt-out.

**Table 1.** SEB exam trial process

Stage	Activities
1. Practice session done two weeks prior to exam.	Students were able to preview the exam process by following the ‘digital exam start guide’ printed instruction sheet. Students could complete the practice questions that used the same format as those presented in the real exam. Data were collected using observation and a survey of students’ first impressions (Pre-survey). Following the session, data analysis of the surveys was carried out to detect any concerns.
2. Real exam session(s).	For the mid-term and final exams: benches were set up with a paper copy of the ‘digital exam start guide’ and post-exam survey. Each student was provided with a desktop computer equipped with a wired network connection and the Safe Exam Browser software. All students in the course were asked to use a computer to undertake the exam. <ol style="list-style-type: none"> <li>1. Students entered the room and were seated at a bench.</li> <li>2. Students could read the printed instruction sheet.</li> <li>3. Once logged into the desktop computer, students then logged into the LMS using a Chrome browser and then clicked on a link leading to an SEB setting file. This then launched Safe Exam Browser. Students were required to again login to the LMS and were then taken to the Moodle quiz start page.</li> <li>4. When all students in the room were ready to begin, the invigilator announced the start of the exam by providing the Moodle quiz password. This enabled all students to start at the same time. The Moodle quiz was set to auto-save responses each 30 seconds.</li> <li>5. Exam end: students used the Moodle quiz submission button and then exited SEB. The quiz was set to automatically submit if the quiz timer expired.</li> <li>6. Students were requested to log-off from the computer.</li> <li>7. Students completed the post-exam survey before leaving the room.</li> </ol>
3. Grading.	In the following week, the teacher did the grading. Students were given grades and feedback comments. Surveys results were analysed.

A sub-set of the survey items selected for analysis is displayed in Table 2. These questions directly asked for students’ perceptions of the issues detailed in the research question. These items were Likert-style items asking for agreement ranging from 1 equating to strongly disagree to 5 strongly agree with 3 being neutral. The analysis procedure followed that of a previous study on e-exams [4]. The responses were analysed using SPSS v24 with an alpha level of .05. Likert items pertaining to students’ perceptions were treated as non-parametric [25]. This stance was also used in another study [18] when they analysed students’ perceptions of a quiz-based e-exam system. Mann and Whitney’s U test [26] was used to test the variance between groups (morn-

ing versus afternoon) on Likert items. When comparing paired Likert items (Post 1 and Post 2), a Sign Test was used [27].

It is important to note that specific conditions prevailed during this pilot. As such, the results are only descriptive of this group and do not represent a generalised view of e-exams by university students. Like Dermo [18], we take the position that statistical tests serve as a tool to summarise the body of opinion from this particular group and, as such, we do not present this as a search for an objective truth regarding e-exams.

## 6 Findings

This study involved 215 undergraduate students, 60% were female and 40% were male. The total number of students varied at each stage by about 2% because not all students participated in each event or responded to each question. Approximately 65% of the students were undertaking a computerised exam for the first time.

Following each exam, the respondents were asked to reflect on the technology approach used for the exam. The results for agreement items pertaining to perceived suitability, ease of use, reliability and security are displayed in Table 2. Students were also asked if they would recommend the approach to others (also shown in Table 2). The majority of items received positive agreement, most with mean agreement ratings at 3 or greater out of 5 (strongly agree). The sentiment within this group overall was relatively uniform as evidenced by the small standard deviations (Table 2), although some divergence is addressed later. Parametric statistics are shown here to provide clarity to the reader in terms of what the responses were for each item at each stage.

**Table 2.** Post-exam survey responses regarding the exam system

Question	Pre			Post 1			Post 2		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
I felt this particular exam suited the use of computers.	n/a	-	-	216	3.9	1.0	216	3.8	1.1
I felt the exam software was easy to use.	215	4.3	0.8	216	4.2	0.8	210	3.9	1.0
I felt the exam software was reliable against technical failures.	214	3.5	1.0	216	3.5	1.0	211	3.0	1.2
I felt the exam software was secure against cheating.	215	4.1	0.9	216	4.0	0.8	208	3.9	0.8
I would recommend this approach to doing exams to others.	215	3.6	1.0	216	3.6	1.0	210	3.6	1.1

The impact of a technical issue in the form of a substantial drop in the performance (i.e. a slow-down in server response times) of the LMS resulted in disruption to the start of the exams held during the last event (Post 2). Exam candidates were asked: “Did you experience any technical difficulties during this exam?” Those who indicated ‘yes’ during the mid-term exam were 12% (27), while following the final exam this increased to 43% (88). It should be noted that all those who undertook the exams

were able to successfully submit the exam. The disruption mainly impacted the ability for candidates to enter into the exam quiz, but once candidates had begun the quiz, further impacts were not noticeable.

The differences in students' opinions between the mid-term and final assessment events also demonstrate the impact of the technical issue. The results of a Sign Test are shown in Table 3. Note the requirement of a normal distribution of differences was not met for a Wilcoxon Signed Ranks Test [28] to be used because Shapiro-Wilk Tests [29] for each pair were all  $< .000$ . The items related to ease of use, reliability and security were significantly marked down by students in the second exam event.

**Table 3.** Survey responses comparing two exam events using a Sign Test

Post 1 versus Post 2 <b>Question</b>	<i>Z</i>	<i>Sig</i>
I felt this exam suited the use of computers.	-1.812	0.07
I felt the exam software was easy to use.	-3.816	<b>0.00</b>
I felt the exam software was reliable against technical failures.	-3.350	<b>0.00</b>
I felt the exam software was secure against cheating.	-1.925	<b>0.05</b>
I would recommend this approach to doing exams to others.	-0.297	0.77

It is worth noting that the infrastructure problems impacted the morning and afternoon sessions of the final exam to a different extent. A comparison of the morning and afternoon groups was made using a Mann-Whitney Test (Table 4). Note, means are shown for reader clarity as to the direction of differences. Results show a statistically significant divergence of opinions when looking at the final (Post 2) event. The differences between morning and afternoon sessions at the practice (pre-) event and mid-term event (Post 1) were generally not significant, although the afternoon group tended to provide lower ratings across all events.

**Table 4.** Survey responses comparing morning and afternoon exam sessions

Moring versus afternoon groups <b>Question (abbreviated)</b>	<i>AM</i>		<i>PM</i>		<i>Statistics</i>	
	<i>N</i>	<i>M*</i>	<i>N</i>	<i>M*</i>	<i>U</i>	<i>Sig</i>
[pre-] ... easy to use.	108	4.4	107	4.3	5275	0.22
[pre-] ... reliable against technical failures.	107	3.5	107	3.5	5437	0.50
[pre-] ... secure against cheating.	108	4.0	107	4.2	5218	0.19
[pre-] ... recommend this approach... to others.	108	3.7	107	3.6	5522	0.56
[Post1] ... this exam suited the use of computers.	109	4.0	107	3.8	5368	0.29
[Post1] ... was easy to use.	109	4.3	107	4.1	5007	<b>0.05</b>
[Post1] ... was reliable against technical failures.	109	3.5	107	3.5	5750.5	0.85
[Post1] ... was secure against cheating.	109	4.1	107	4.0	5650.5	0.66
[Post1] ... recommend this approach... to others.	109	3.8	107	3.4	4797	<b>0.02</b>
[Post2] ... this exam suited the use of computers.	106	4.1	105	3.5	3813.5	<b>0.00</b>
[Post2] ... was easy to use.	106	4.3	104	3.4	2617.5	<b>0.00</b>
[Post2] ... was reliable against technical failures.	106	3.6	105	2.4	2481.5	<b>0.00</b>
[Post2] ... was secure against cheating.	106	4.0	102	3.7	4311.5	<b>0.01</b>
[Post2] ... recommend this approach... to others.	106	3.9	104	3.3	3626	<b>0.00</b>



\* Means shown for reader clarity.

The implications for practice in dealing with technical issues during e-exams are considered in the following section.

## 7 Discussion and Conclusion

Despite the technology issues encountered in the final exam session, the introduction of e-exams was a success in the context of the use of bench-top computers in a biochemistry course. Generally, positive responses (i.e. above 3 on the 5-point scales) were received from students across the survey questions. What did become apparent was that the technical issue that was encountered during the final exam event did have a statistically significant impact on student ratings of reliability, ease of use and to a lesser extent on their perception of the system security. Their recommendation of the computerised exam approach was not impacted when comparing Post 1 and Post 2 events (Table 3). However, when focusing on the Post 2 event (Table 4), it was found that there were significant differences between the opinion ratings of those in the morning and afternoon groups, with the afternoon session experiencing more severe technical problems than the morning.

The extent of the impact is worth reflecting upon, given the relatively minor extent of the technology problem. This was no more than an inconvenience, causing a delay to the start of the exam. The interruption did not prevent the students from undertaking and submitting the exam, yet it still had a statistically measurable impact on students' perceptions. It is likely that the stress of an exam event means that students are extra sensitive to anything that is not perfect on the day.

When considering the possibility of mediating factors in the findings, it is possible students may have regarded the final exam as being higher stakes than the mid-term and that may have also contributed to an elevated level of pre-existing stress in the second event. However, the study design was such that both exam events (Post 1 and Post 2) were run in a near-identical manner, sessions were led by the same team members, with the same groups participating in each. Yet, the greater impact of technology problems in the afternoon session appears to coincide with the statically significant drop in that group's ratings on the survey. Therefore, this suggests that the technology problems did influence the changes seen in the students' responses to the survey following the second event.

The particular technical incident occurred the day following an overnight upgrade that saw additional modules added to the LMS. Overall, the experience suggested that the mitigation measures in place for the infrastructure verified it was capable of hosting an e-exam event even when server performance degradation was being experienced.

The findings from this study show that technical performance of an e-exam system does impact exam candidates' perceptions of the system and therefore their acceptance of an e-exam approach. This highlights that test administrators and teachers need to have open communication with the technical support department and with

students about how problems will be handled in order to minimise any additional stress for students.

## References

1. Transforming Exams across Australia. Australian Government Department of Education and Training, Grant ID15-4747, <http://transformingexams.com>, last accessed 2020/05/24.
2. Hillier, M., Lyon, N.: Writing e-Exams in Pre-University College. In: Passey D., Bottino R., Lewin C., Sanchez E. (eds.) *Empowering Learners for Life in the Digital Age*, AICT, Volume 524, 264-274. Springer, Cham, Switzerland (2019).
3. Hillier, M., Grant, S., Coleman, M.: Towards authentic e-Exams at scale: robust net-worked Moodle. In: *Proceedings of the Australasian Society for Computers in Learning in Tertiary Education Conference*, Geelong, VIC, <http://transformingexams.com/files/hillier-grant-coleman-Towards-authentic-e-Exams-at-scale-ASCILITE-2018.pdf>, last accessed 2020/05/24.
4. Hillier, M., Lyon, N.: Student experiences with bring your own laptop e-Exams in pre-university college. In: Passey, D., Bottino, R., Lewin, C., Sanchez, E. (eds.) *Empowering Learners for Life in the Digital Age*, AICT, Volume 524, pp. 253-263. Springer, Heidelberg, Germany (2019).
5. Hillier, M., Grant, S.: Do-it-yourself e-Exams, In: *Proceedings of the Australasian Society for Computers in Learning in Tertiary Education Conference*, Geelong, VIC, <http://transformingexams.com/files/hillier-grant-diy-eexams-ASCILITE-2018.pdf>, last accessed 2020/05/24.
6. Fluck, A., Hillier, M.: eExams: Strength in Diversity. In: Tatnall, A., Webb, M. (eds.) *Tomorrow's Learning: Involving Everyone*, AICT, Volume 515, 409-417. Springer, Cham, Switzerland (2017).
7. Fluck, A., Pálsson, H., Coleman, M., Hillier, M., Schneider, D., Frankl, G., Uolia, K.: eExam symposium: design decisions and implementation experience. Paper presented at the IFIP World Conference on Computers in Education, 3-6 July, Dublin, Ireland, [http://transformingexams.com/files/Fluck\\_etal\\_2017.pdf](http://transformingexams.com/files/Fluck_etal_2017.pdf), last accessed 2020/05/24.
8. Küppers, B., Schroeder, U.: Students' Perceptions of e-Assessment: A Case Study from Germany. In: Passey, D., Bottino, R., Lewin, C., Sanchez, E. (eds.) *Empowering Learners for Life in the Digital Age*, AICT, Volume 524, 275-284. Springer, Heidelberg, Germany (2019).
9. Crisp, G.: Towards Authentic e-Assessment Tasks In: *Proceedings of the EdMedia: World Conference on Educational Media and Technology*, Honolulu, HI, <http://www.editlib.org/p/31689/>, last accessed 2020/05/24.
10. Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., Herrera-Seda, C.: Authentic assessment: creating a blueprint for course design. *Assessment & Evaluation in Higher Education* 43(5), 840-854 (2018).
11. Fluck, A., Hillier, M.: Innovative assessment with eExams. In: *Proceedings of the Australian Council for Computers in Education Conference*, Brisbane, QLD, <http://conference.acce.edu.au/index.php/acce/acce2016/paper/view/34/27>, last accessed 2020/05/24.
12. Hillier, M., Fluck, A.: Transforming Exams - How IT works for BYOD e-Exams. In: *Proceedings of the Australasian Society for Computers in Learning in Tertiary Education Conference*, pp. 100-105, <http://2017conference.ascilite.org/wp-content/uploads/2017/11/Concise-HILLIER.pdf>, last accessed 2020/05/24.

13. Hillier, M.: The very idea of e-Exams: student (pre)conceptions. In: Proceedings of the 31st Australasian Society for Computers in Learning in Tertiary Education Conference, Dunedin, New Zealand, <http://ascilite.org/conferences/dunedin2014/files/fullpapers/91-Hillier.pdf>, last accessed 2020/05/24.
14. Hillier, M.: e-Exams with student owned devices: Student voices. In: Proceedings of the International Mobile Learning Festival: Mobile Learning, MOOCs and 21st Century learning, Hong Kong, China, 582–608, [http://transformingexams.com/files/Hillier\\_IMLF2015\\_full\\_paper\\_formatting\\_fixed.pdf](http://transformingexams.com/files/Hillier_IMLF2015_full_paper_formatting_fixed.pdf), last accessed 2020/05/24.
15. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 13(3), 319–340 (1989).
16. Farzin, S.: Attitude of Students Towards E-Examination System: an Application of E-Learning. *Science Journal of Education* 4(6), 222–227 (2017).
17. Lunceford, B.: Reconsidering Technology Adoption and Resistance: Observations of a Semi-Luddite. *Explorations in Media Ecology* 8(1), 29–47 (2009).
18. Dermo, J.: E-assessment and the student learning experience: A survey of student perceptions of e-assessment. *British Journal of Educational Technology* 40, 203–214 (2009).
19. Aubusson, K., Noyes, J.: Burnt-out doctors deeply distressed by botched high-stakes exam. *The Sydney Morning Herald* (February 20), <https://www.smh.com.au/national/nsw/burnt-out-doctors-deeply-distressed-by-botched-high-stakes-exam-20180220-p4z117.html>, last accessed 2020/05/24.
20. Aubusson, J., Noyes, K.: Company behind botched medical exam has track record of failure. *The Sydney Morning Herald* (February 21), <https://www.smh.com.au/national/company-behind-botched-medical-exam-has-track-record-of-failure-20180221-p4z15a.html>, last accessed 2020/05/24.
21. Strauss, V.: Pearson’s history of testing problems — a list. *Washington Post* (April 21), <https://www.washingtonpost.com/news/answer-sheet/wp/2016/04/21/pearsons-history-of-testing-problems-a-list/>, last accessed 2020/05/24.
22. Donovan, D.: Wake Forest law grad sues ExamSoft over bar exam debacle. *North Carolina Lawyers Weekly* (August 18), <https://nclawyersweekly.com/2014/08/18/wake-forest-law-grad-sues-examsoft-over-bar-exam-debacle/>, last accessed 2020/05/24.
23. Peregoodoff, R.: Large-Scale BYOD Wireless LMS Midterm and Final Exams. Seminar presented at the EDUCAUSE Annual Conference, <https://events.educause.edu/annual-conference/2015/proceedings/large-scale-byod-wireless-lms-midterm-and-final-exams>, last accessed 2020/05/24.
24. LabArchives: Electronic Lab Notebook (software as a service product), <https://www.labarchives.com>, last accessed 2020/05/24.
25. Jamieson, S.: Likert scales: how to (ab)use them. *Medical Education* 38(12), 1217–1218 (2004).
26. Mann, H.B., Whitney, D.R.: On a Test of Whether One of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics* 18(1), 50–60 (1947).
27. Roberson, P.K., Shema, S.J., Mundfrom, D.J., Holmes, T.M.: Analysis of paired Likert data: how to evaluate change and preference questions. *Family Medicine* 27(10), 671–675 (1995).
28. Wilcoxon, F.: Individual comparisons by ranking methods. *Biometrics Bulletin* 1(6), 80–83 (1945).
29. Shapiro, S.S., Wilk, M.B.: An analysis of variance test for normality (complete samples), *Biometrika* 52 (3–4), 591–611 (1965).