A systematic review of online examinations: A pedagogical innovation for scalable authentication and integrity

Kerryn Butler-Henderson [a,*], Joseph Crawford [b]

[a] College of Health and Medicine, University of Tasmania, Locked Bag 1322, Launceston, Tasmania, 7250, Australia
[b] Academic Division, University of Tasmania, Locked Bag 1322, Launceston, Tasmania, 7250, Australia

ARTICLE INFO

ABSTRACT

Digitization and automation across all industries has resulted in improvements in efficiencies and effectiveness to systems and process, and the higher education sector is not immune. Online learning, e-learning, electronic teaching tools, and digital assessments are not innovations. However, there has been limited implementation of online invigilated examinations in many countries. This paper provides a brief background on online examinations, followed by the results of a systematic review on the topic to explore the challenges and opportunities. We follow on with an explication of results from thirty-six papers, exploring nine key themes: student perceptions, student performance, anxiety, cheating, staff perceptions, authentication and security, interface design, and technology issues. While the literature on online examinations is growing, there is still a dearth of discussion at the pedagogical and governance levels.

1. Introduction

Learning and teaching is transforming away from the conventional lecture theatre designed to seat 100 to 10,000 passive students towards more active learning environments. In our current climate, this is exacerbated by COVID-19 responses (Crawford et al., 2020), where thousands of students are involved in online adaptions of face-to-face examinations (e.g. online Zoom rooms with all microphones and videos locked on). This evolution has grown from the need to recognize that students now rarely study exclusively and have commitments that conflict with their University life (e.g. work, family, social obligations). Students have more diverse digitally capability (Margaryan et al., 2011) and higher age and gender diversity (Eagly & Sczesny, 2009; Schwalb & Sedlacek, 1990). Continual change of the demographic and profile of students creates a challenge for scholars seeking to develop a student experience that demonstrates quality and maintains financial and academic viability (Gross et al., 2013; Hainline et al., 2010).

Universities are developing extensive online offerings to grow their international loads and facilitate the massification of higher learning. These protocols, informed by growing policy targets to educate a larger quantity of graduates (e.g. Kemp, 1999; Reiko, 2001), have challenged traditional university models of fully on-campus student attendance. The development of online examination software has offered a systematic and technological alternative to the end-of-course summative examination designed for final authentication and testing of student knowledge retention, application, and extension. As a result of the COVID-19 pandemic, the initial response in higher education across many countries was to postpone examinations (Crawford et al., 2020). However, as the pandemic continued, the need to move to either an online examination format or alternative assessment became more urgent.

---

* Corresponding author.
  E-mail addresses: Kerryn.Butlerhenderson@utas.edu.au (K. Butler-Henderson), joseph.crawford@utas.edu.au (J. Crawford).

This paper is a timely exploration of the contemporary literature related to online examinations in the university setting, with the hopes to consolidate information on this relatively new pedagogy in higher education. This paper begins with a brief background of traditional examinations, as the assumptions applied in many online examination environments build on the techniques and assumptions of the traditional face-to-face gymnasium-housed invigilated examinations. This is followed by a summary of the systematic review method, including search strategy, procedure, quality review, analysis, and summary of the sample.

Print-based educational examinations designed to test knowledge have existed for hundreds of years. The New York State Education Department has "the oldest educational testing service in the United States" and has been delivering entrance examinations since 1865 (Johnson, 2009, p. 1; NYSED, 2012). In pre-Revolution Russia, it was not possible to obtain a diploma to enter university without passing a high-stakes graduation examinations (Karp, 2007). These high school examinations assessed and assured learning of students in rigid and high-security conditions. Under traditional classroom conditions, these were likely a reasonable practice to validate knowledge. The discussion of authenticating learning was not a consideration at this stage, as students were face to face *only*. For many high school jurisdictions, these are designed to strengthen the accountability of teachers and assess student performance (Mueller & Colley, 2015).

In tertiary education, the use of an end-of-course summative examination as a form of validating knowledge has been informed significantly by accreditation bodies and streamlined financially viable assessment options. The American Bar Association has required a final course examination to remain accredited (Sheppard, 1996). Law examinations typically contained brief didactic questions focused on assessing rote memory through to problem-based assessment to evaluate students' ability to apply knowledge (Sheppard, 1996). In accredited courses, there are significant parallels. Alternatives to traditional gymnasium-sized classroom paper-and-pencil invigilated examinations have been developed with educators recognizing the limitations associated with single-point summative examinations (Butt, 2018).

The objective structured clinical examinations (OSCE) incorporate multiple workstations with students performing specific practical tasks from physical examinations on mannequins to short-answer written responses to scenarios (Turner & Dankoski, 2008). The OSCE has parallels with the patient simulation examination used in some medical schools (Botezatu et al., 2010). Portfolios assess and demonstrate learning over a whole course and for extracurricular learning (Wasley, 2008).

The inclusion of online examinations, e-examinations, and bring-your-own-device models have offered alternatives to the large-scale examination rooms with paper-and-pencil invigilated examinations. Each of these offer new opportunities for the inclusion of innovative pedagogies and assessment where examinations are considered necessary. Further, some research indicates online examinations are able to discern a true pass from a true fail with a high level of accuracy (Ardid et al., 2015), yet there is no systematic consolidation of the literature. We believe this timely review is critical for the progression of the field in first stepping back and consolidating the existing practices to support dissemination and further innovation. The pursuit of such systems may be to provide formative feedback and to assess learning outcomes, but a dominant rationale for final examinations is to authenticate learning. That is, to ensure the student whose name is on the student register, is the student who is completing the assessed work. The development of digitalized examination pilot studies and case studies are becoming an expected norm with universities developing responses to a growing online curriculum offering (e.g. Al-Hakeem & Abdulrahman, 2017; Alzu'bi, 2015; Anderson et al., 2005; Fluck et al., 2009; Fluck et al., 2017; Fluck, 2019; Seow & Soong, 2014; Sindre & Vegendla, 2015; Steel et al., 2019; Wibowo et al., 2016).

As many scholars highlight, cheating is a common component of the contemporary student experience (Jordan, 2001; Rettinger & Kramer, 2009) despite that it should not be. Some are theorizing responses to the inevitability of cheating from developing student capacity for integrity (Crawford, 2015; Wright, 2011) to enhancing detection of cheating (Dawson & Sutherland-Smith, 2018, 2019) and legislation to ban contract cheating (Amigud & Dawson, 2020). We see value in the pursuit of methods that can support integrity in student assessment, including during rapid changes to the curriculum. The objective of this paper is to summarize the current evidence on online examination methods, and scholarly responses to authentication of learning and the mitigation of cheating, within the confines of assessment that enables learning and student wellbeing. We scope out preparation for examinations (e.g. Nguyen & Henderson, 2020) to enable focus on the online exam setting specifically.

2. Material and methods

2.1. Search strategy

To address the objective of this paper, a systematic literature review was undertaken, following the PRISMA approach for article selection (Moher et al., 2009). The keyword string was developed incorporating the U.S. National Library of Medicine (2019) MeSH (Medical Subject Headings) terms: [("online" OR "electronic" OR "digital") AND ("exam*" OR "test") AND ("university" OR "educat*" OR "teach" OR "school" OR "college")]. The following databases were queried: A + Education (Informit), ERIC (EBSCO), Education Database (ProQuest), Education Research Complete (EBSCO), Educational Research Abstracts Online (Taylor & Francis), Informit, and Scopus. These search phrases will enable the collection of a broad range of literature on online examinations as well as terms often used synonymously, such as e-examination/eExams and BYOD (bring-your-own-device) examinations. The eligibility criteria included peer-reviewed journal articles or full conference papers on online examinations in the university sector, published between 2009 and 2018, available in English. As other sources (e.g. dissertations) are not peer-reviewed, and we aimed to identify rigorous best practice literature, we excluded these. We subsequently conducted a general search in Google Scholar and found no additional results. All records returned from the search were extracted and imported into the Covidence® online software by the first author.

2.2. Selection procedure and quality assessment

The online Covidence® software facilitated article selection following the PRISMA approach. Each of the 1906 titles and abstracts were double-screened by the authors based on the eligibility criteria. We also excluded non-higher education examinations, given the context around student demographics is often considerably different than vocational education, primary and high schools. Where there was discordance between the authors on a title or abstract inclusion or exclusion, consensus discussions were undertaken. The screening reduced the volume of papers significantly because numerous papers related to a different education context or involved online or digital forms of medical examinations. Next, the full-text for selected abstracts were double-reviewed, with discordance managed through a consensus discussion. The papers selected following the double full-text review were accepted for this review. Each accepted paper was reviewed for quality using the MMAT system (Hong et al., 2018) and the scores were calculated as high, medium, or low quality based on the matrix (Hong et al., 2018). A summary of this assessment is presented in Table 1.

2.3. Thematic analysis

Following the process described by Braun and Clarke (2006), an inductive thematic approach was undertaken to identify common

Table 1
Summary of article characteristics.

First Author	Year	Country	Method	Participants	Theme	QAS
AbdelKarim	2016	Saudi, Jordan, Malaysia	Survey	119 students	Student perception, interface design	Medium
Abumansour	2017	Saudi	Description	NA	Authentication and security	Low
Aisyah	2018	Indonesia	Description	NA	Authentication and security	Low
Attia	2014	Saudi	Survey	34 students	Student perception, anxiety	High
Böhmer	2018	Germany	Survey	17 students	Student perception, student performance	Medium
Chao	2012	Taiwan	Survey	25 students	Authentication and security	Medium
Chebrolu	2017	India	Description	NA	Authentication and security	Low
Chen	2018	China	Exam data analysis	Not provided	Student performance	Medium
Chytrý	2018	Czech Republic	Exam data analysis	115 students	Student performance	High
Daffin	2018	USA	Exam data analysis	1694 students	Student performance	High
Dawson	2016	Australia	Description	NA	Cheating	Low
Ellis	2016	UK	Survey, exam data analysis	>120 students	Student performance	Medium
Gehringer	2013	USA	Survey	85 staff and 315 students	Cheating, administration	Medium
Gold	2009	USA	Exam data analysis	1800 students	Student performance	Medium
Guillen-Ganez	2015	Spain	Exam data analysis	70 students	Authentication and security	Medium
HearnMoore	2017	USA	Description, exam data analysis	Not provided	Cheating	Medium
Hylton	2016	Jamaica	Survey, exam data analysis	350 students	Cheating	High
Kolagari	2018	Iran	Test Anxiety Scale	39 students	Anxiety	High
Kolski	2018	USA	Test Anxiety Scale, exam data analysis, interviews	238 students	Anxiety	High
Kumar	2015	USA	Problem analysis	2 staff	Anxiety	High
Li	2015	USA	Exam data analysis	9 students	Cheating	High
Matthiasdottir	2016	Iceland	Survey	183 students	Student perceptions, anxiety	Medium
Mitra	2016	USA	Interviews, survey	5 staff; 30 students	Cheating, administration	Medium
Mohanna	2015	Saudi	Exam data analysis	127 students	Student performance, technical issues	High
Oz	2018	Turkey	Exam data analysis	97 students	Student performance	High
Pagram	2018	Australia	Interviews, survey	Interviews: 4 students, 2 staff; Survey: 6 students	Student perceptions, academic perceptions, anxiety	Medium
Park	2017	USA	Survey	37 students	Student perception	Medium
Patel	2014	Saudi	Exam data analysis	180 students	Student performance	High
Petrović	2017	Croatia	Exam data analysis	591 students	Cheating	Medium
Rios	2017	USA	Exam data analysis, survey	1126 students	Student performance, student perceptions, Authentication and security (under user friendliness)	High
Rodchua	2011	USA	Description	NA	Cheating	Low
Schmidt	2009	USA	Survey	49 students	Student performance, academic perception, student perception, anxiety, tech issues	High
Stowell	2010	USA	Test Anxiety Scale, exam data analysis	69 students	Anxiety	High
Sullivan	2016	USA	Exam data analysis, survey	178 students	Cheating	Medium
Williams	2009	Singapore	Survey	91 students	Student perception, cheating	Medium
Yong-Sheng	2015	China	Description	NA	Authentication and security	Low

QAS, quality assessment score.

themes identified in each article. This process involves six stages: data familiarization, data coding, theme searching, theme review, defining themes, and naming themes. Familiarization with the literature was achieved during the screening, full-text, and quality review process by triple exposure to works. The named authors then inductively coded half the manuscripts each. The research team consolidated the data together to identify themes. Upon final agreement of themes and their definitions, the write-up was split among the team with subsequent review and revision of ideas in themes through independent and collaborative writing and reviewing (Creswell & Miller, 2000; Lincoln & Guba, 1985). This resulted in nine final themes, each discussed in-depth during the discussion.

3. Results

There were thirty-six (36) articles identified that met the eligibility criteria and were selected following the PRISMA approach, as shown in Fig. 1.

3.1. Characteristics of selected articles

The selected articles are from a wide range of discipline areas and countries. Table 1 summarizes the characteristics of the selected articles. The United States of America held a vast majority (14, 38.9%) of the publications on online examinations, followed by Saudi Arabia (4, 11.1%), China (2, 5.6%), and Australia (2, 5.6%). When aggregated at the region-level, there was an equality of papers from North America and Asia (14, 38.9% each), with Europe (6, 16.7%) and Oceania (2, 5.6%) least represented in the selection of articles. There has been considerable growth in publications in the past five years, concerning online examinations. Publications between the years 2009 and 2015 represented a third (12, 33.3%) of the total number of selected papers. The majority (24, 66.7%) of papers were
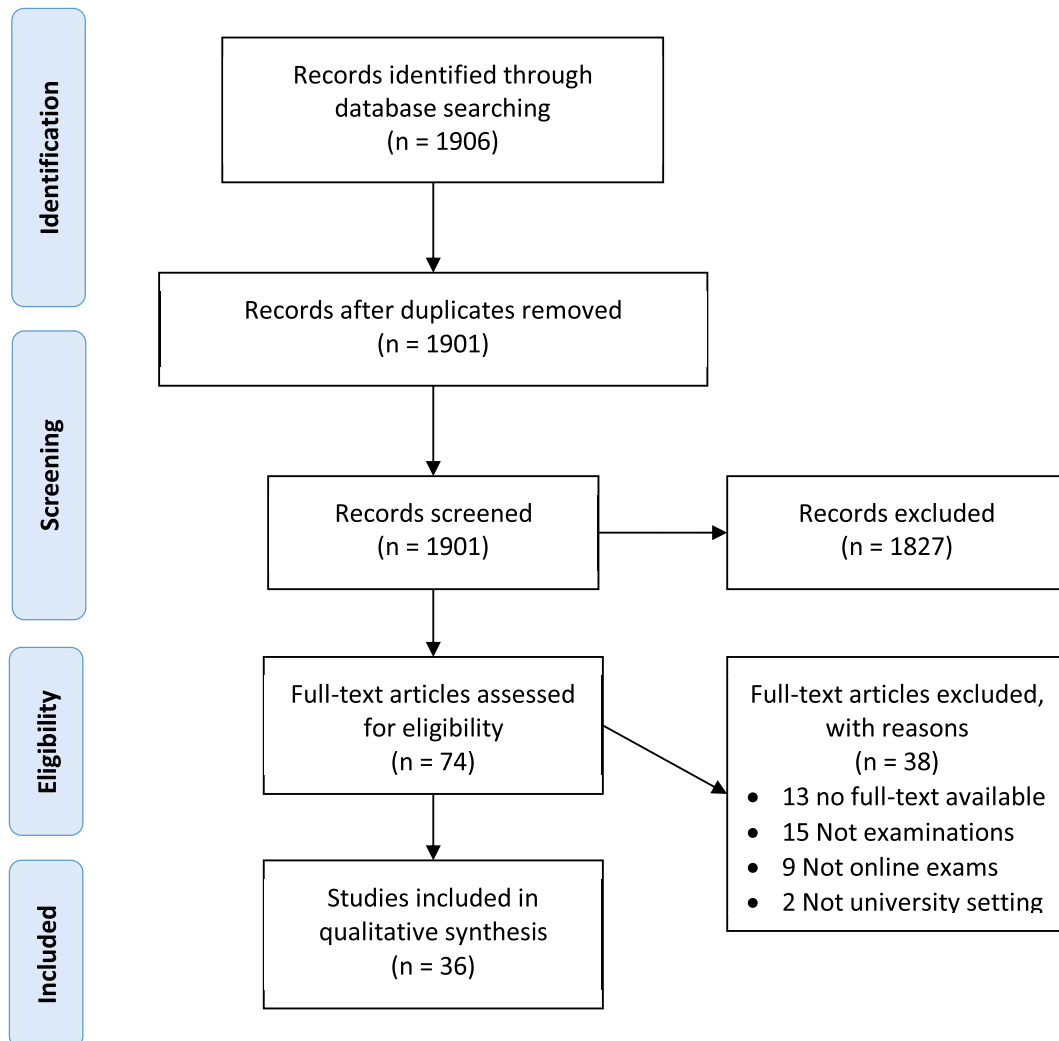

Fig. 1. PRISMA results.

published in the last three years. Papers that described a system but did not include empirical evidence scored a low-quality rank as they did not meet many of the criteria that relate to the evaluation of a system.

When examining the types of papers, the majority (30, 83.3%) were empirical research, with the remainder commentary papers (6, 16.7%). Of the empirical research papers, three-quarters of the paper reported a quantitative study design (32, 88.9%) compared to two (5.6%) qualitative study designs and two (5.6%) that used a mixed method. For quantitative studies, there was a range between nine and 1800 student participants ($\bar{x} = 291.62$) across 26 studies, and a range between two and 85 staff participants ($\bar{x} = 30.67$) in one study. The most common quantitative methods were self-administered surveys and analysis of numerical examination student grades (38% each). Qualitative and mixed methods studies only adopted interviews (6%). Only one qualitative study reported a sample of students ($n = 4$), with two qualitative studies reporting a sample of staff ($n = 2$, $n = 5$).

3.2. Student perceptions

Today's students prefer online examinations compared to paper exams ([68.75% preference of online over paper-based examinations: Attia, 2014; 56–62.5%: Böhmer et al., 2018; no percentage: (Schmidt, Ralph & Buskirk, 2009); 92%: Matthíasdóttir & Arnalds, 2016; no percentage: Pagram et al., 2018; 51%: Park, 2017; 84%: Schmidt, Ralph & Williams & Wong, 2009). Two reasons provided for the preference is the increased speed and ease of editing responses (Pagram et al., 2018), with one study finding two-thirds (67%) of students reported a positive experience in online examination environment (Matthíasdóttir & Arnalds, 2016). Students believe online examinations allows a more authentic assessment experience (Williams & Wong, 2009), with 78 percent of students reporting consistencies between the online environment and their future real-world environment (Matthíasdóttir & Arnalds, 2016).

Students perceive the online examinations saves time (75.0% of students surveyed) and is more economical (87.5%) than paper examinations (Attia, 2014). It provides greater flexibility for completing examinations (Schmidt et al., 2009) with faster access to remote student papers (87.5%) and students trust the result of online over paper-based examinations (78.1%: Attia, 2014). The majority of students (59.4%: Attia, 2014; 55.5%: Pagram et al., 2018) perceive that the online examination environment makes it easier to cheat. More than half (56.25%) of students believe that a lack of information communication and technology (ICT) skill do not adversely affect performance in online examinations (Attia, 2014). Nearly a quarter (23%) of students reported (Abdel Karim & Shukur, 2016) the most preferred font face (type) was Arial, a font also recommended by Vision Australia (2014) in their guidelines for online and print inclusive design and legibility considerations. Nearly all (87%) students preferred black text color on a white background color (87%). With regards to onscreen time counters, a countdown counter was the most preferred option (42%) compared to a traditional analogue clock (30%) or an ascending counter (22%). Many systems allow students to set their preferred remaining time reminder or alert, including 15 min remaining (35% students preferred), 5 min remaining (26%), mid-examination (15%) or 30 min remaining (13%).

3.3. Student performance

Several studies in the sample referred to a lack of score variation between the results of examination across different administration methods. For example, student performance did not have significant difference in final examination scores across online and traditional examination modalities (Gold & Mozes-Carmel, 2017). This is reinforced by a test of validity and reliability of computer-based and paper-based assessment that demonstrated no significant difference (Oz & Ozturan, 2018), and equality of grades identified across the two modalities (Stowell & Bennett, 2010).

When considering student perceptions, of the studies documented in our sample, there tended to be favorable ratings of online examinations. In a small sample of 34 postgraduate students, the respondents had positive perceptions towards online learning assessments (67.4%). The students also believed it contributed to improved learning and feedback (67.4%), and 77 percent had favorable attitudes towards online assessment (Attia, 2014). In a pre-examination survey, students indicated they preferred to type than to write, felt more confident about the examination, and had limited issues with software and hardware (Pagram, 2018). With the same sample in a post-examination survey, within the design and technology examination, students felt the software and hardware were simple to use, yet many students did not feel at ease from their use of an e-examination.

Rios and Liu (2017) compared proctored and non-proctored online examinations across several aspects, including test-taking behavior. Their study did not identify any difference in the test-taking behavior of students between the two environments. There was no significant difference between omitted items and not-reached items. Furthermore, with regards to rapid guessing, there was no significant difference. A negligible difference existed for students aged older than thirty-five years, yet gender was a nonsignificant factor.

3.4. Anxiety

Scholars have an increasing awareness of the role that test anxiety has in reducing student success in online learning environments (Kolski & Weible, 2018). The manuscripts identified by the literature scan, identified inconsistencies of results for the effect that examination modalities have on student test anxiety. A study of 69 psychology undergraduates identified that students who typically experienced high anxiety in traditional test environments had lower anxiety levels when completing an online examination (Stowell & Bennett, 2010). In a quasi-experimental study ($n = 38$ nursing students), when baseline anxiety is controlled, students in computer-based examinations had higher degrees of test anxiety.

In 34 postgraduate student interviews, only three opposed online assessment based on perceived lack of technical skill (e.g. typing;

Attia, 2014). Around two-thirds of participants identified some form of fear-based on internet disconnection, electricity, slow typing, or family disturbances at home. A 37 participant Community College study used proximal indicators (e.g. lip licking and biting, furrowed eyebrows, and seat squirming) to assess the rate of test anxiety in webcam-based examination proctoring (Kolski & Weible, 2018). Teacher strategies to reduce anxiety in their students include enabling students to consider, review, and acknowledge their anxieties (Kolski & Weible, 2018). Responses such as students writing of their anxiety, or responding to multiple-choice questionnaire on test anxiety, reduced anxiety. Students in the test group and provided anxiety items or expressive writing exercises, performed better (Kumar, 2014).

3.5. Cheating

Cheating was the most prevalent area among all the themes identified. Cheating in asynchronous, objective, and online assessments is argued by some to be at unconscionable levels (Sullivan, 2016). In one survey, 73.6 percent of students felt it was easier to cheat on online examinations than regular examinations (Aisyah et al., 2018). This is perhaps because students are monitored in paper and pencil examinations, compared to online examinations where greater control of variables is required to mitigate cheating. Some instructors have used randomized examination batteries to minimize cheating potential through peer-to-peer sharing (Schmidt et al., 2009).

Scholars identify various methods for mitigating cheating. Identifying the test taker, preventing examination theft, unauthorized use of textbook/notes, preparing a set-up for online examination, unauthorized student access to a test bank, preventing the use of devices (e.g. phone, Bluetooth, and calculators), limiting access to other people during the examination, equitable access to equipment, identifying computer crashes, inconsistency of method for proctoring (Hearn Moore et al., 2017). In another, the issue for solving cheating is social as well as technological. While technology is considered the current norm for reducing cheating, these tools have been mostly ineffective (Sullivan, 2016). Access to multiple question banks through effective quiz design and delivery is a mechanism to reduce the propensity to cheat, by reducing the stakes through multiple delivery attempts (Sullivan, 2016). Question and answer randomization, continuous question development, multiple examination versions, open book options, time stamps, and diversity in question formats, sequences, types, and frequency are used to manage the perception and potential for cheating. In the study with MBA students, perception of the ability to cheat seemed to be critical for the development of a safe online examination environment (Sullivan, 2016).

Dawson (2016) in a review of bring-your-own-device examinations including:

- Copying contents of USB to a hard drive to make a copy of the digital examination available to others,
- Use of a virtual machine to maintain access to standard applications on their device,
- USB keyboard hacks to allow easy access to other documents (e.g. personal notes),
- Modifying software to maintain complete control of their own device, and
- A cold boot attack to maintain a copy of the examination.

The research on cheating has focused mainly on technical challenges (e.g. hardware to support cheating), rather than ethical and social issues (e.g. behavioral development to curb future cheating behaviors). The latter has been researched in more depth in traditional assessment methods (e.g. Wright, 2015). In a study on Massive Open Online Courses (MOOCs), motivations for students to engage in optional learning stemmed from knowledge, work, convenience, and personal interest (Shapiro et al., 2017). This provides possible opportunities for future research to consider behavioral elements for responding to cheating, rather than institutional punitive arrangements.

3.6. Staff perception

Schmidt et al. (2009) also examined the perceptions of academics with regards to online examination. Academics reported that their biggest concern with using online examinations is the potential for cheating. There was a perception that students may get assistance during an examination. The reliability of the technology is the second more critical concern of academic staff. This includes concerns about internet connectivity as well as computer or software issues. The third concern is related to ease of use, both for the academic and for students. Academics want a system that is easy and quick to create, manage and mark examinations, and students can use with proficient ICT skills (Schmidt et al., 2009). Furthermore, staff reported in a different study that marking digital work was easier and preferred it over paper examinations because of the reduction in paper (Pagram et al., 2018). They believe preference should be given to using university machines instead of the student using their computer, mainly due to issues around operating system compatibility and data loss.

3.7. Authentication and security

Authentication was recognized as a significant issue for examination. Some scholars indicate that the primary reason for requiring physical attendance to proctored examinations is to validate and authenticate the student taking the assessment (Chao et al., 2012). Importantly, the validity of online proctored examination administration procedures is argued as lower than proctored on-campus examinations (Rios & Liu, 2017). Most responses to online examinations use bring-your-own-device models where laptops are brought to traditional lecture theatres, use of software on personal devices in any location desired, or use of prescribed devices in a

classroom setting. The primary goal of each is to balance the authentication of students and maintain the integrity and value of achieving learning outcomes.

In a review of current authentication options (AbuMansoor, 2017), the use of fingerprint reading, streaming media, and follow-up identifications were used to authenticate small cohorts of students. Some learning management systems (LMS) have developed subsidiary products (e.g. Weaver within Moodle) to support authentication processes. Some biometric software uses different levels to authenticate keystrokes for motor controls, stylometry for linguistics, application behavior for semantics, capture to physical or behavioral samples, extraction of unique data, comparison of distance measures, and recording decision-making. Development of online examinations should be oriented towards the same theory of open book examinations.

A series of models are proposed in our literature sample. AbuMansoor (2017) propose to use a series of processes into place to develop examinations that minimize cheating (e.g. question batteries), deploying authentication techniques (e.g. keystrokes and fingerprints), and conduct posthoc assessments to search for cheating. The Aisyah et al. (2018) model identifies two perspectives to conceptualize authentication systems: examinee and admin. From the examinee perspective, points of authentication at the pre-, intra-, and post-examination periods. From the administrative perspective, accessing photographic authentication from pre- and intra-examination periods can be used to validate the examinee. The open book open web (OBOW: Mohanna & Patel, 2016) model uses the application of authentic assessment to place the learner in the role of a decision-maker and expert witness, with validation by avoiding any question that could have a generic answer.

The Smart Authenticated Fast Exams (SAFE: Chebrolu et al., 2017) model uses application focus (e.g. continuously tracking focus of examinee), logging (phone state, phone identification, and Wi-Fi status), visual password (a password that is visually presented but not easily communicated without photograph), Bluetooth neighborhood logging (to check for nearby devices), ID checks, digitally signed application, random device swap, and the avoidance of 'bring your own device' models. The online comprehensive examination (OCE) was used in a National Board Dental Examination to test knowledge in a home environment with 200 multiple choice questions, and the ability to take the test multiple times for formative knowledge development.

Some scholars recommend online synchronous assessments as an alternative to traditional proctored examinations while maintaining the ability to manually authenticate (Chao et al., 2012). In these assessments: quizzes are designed to test factual knowledge, practice for procedural, essay for conceptual, and oral for metacognitive knowledge. A 'cyber face-to-face' element is required to enable the validation of students.

3.8. Interface design

The interface of a system will impact on whether a student perceives the environment to be an enabler or barrier for online examinations. Abdel Karim and Shukur (2016) summarized the potential interface design features that emerged from a systematic review of the literature on this topic, as shown in Table 2. The incorporation of navigation tools has also been identified by students and staff as an essential design feature (Rios & Liu, 2017), as is an auto-save functionality (Pagram et al., 2018).

3.9. Technology issues

None of the studies that included technological problems in its design reported any issues (Böhmer et al., 2018; Matthíasdóttir & Arnalds, 2016; Schmidt et al., 2009). One study stated that 5 percent of students reported some problem ranging from a slow system through to the system not working well with the computer operating system, however, the authors stated no technical problems that resulted in the inability to complete the examination were reported (Matthíasdóttir & Arnalds, 2016). In a separate study, students reported that they would prefer to use university technology to complete the examination due to distrust of the system working with their home computer or laptop operating system or the fear of losing data during the examination (Pagram et al., 2018). While the study did not report any problems loading on desktop machines, some student laptops from their workplace had firewalls, and as such

Table 2
Potential interface design features (Abdel Karim & Shukur, 2016).

Interface design features	Recommended values	Description
Font size	10, 12, 14, 18, 22, and 26 points	Font size has a significant effect on objective and subjective readability and comprehensibility.
Font face (type)	Andale Mono, Arial, Arial Black, Comic Sans Ms, Courier New, Georgia, Impact, Times New Roman, Trebuchet Ms, Verdana, and Tahoma	Reading efficiency and reading time are important aspects related to the font type and size.
Font style	Regular, Italic, Bold, and Bold Italic	
Text and background colour	Either: - Black text on white background (Standard) - Black text on grey background (Reduced) - White text on black background (Reversed)	Text and background colour affect text readability and colours, with greater contrast ratio generally lead to greater readability.
Time counter	Countdown timer, ascending counter and traditional clock	Online examination systems should display the time counter on the screen until the examination time has ended.
Alert	5 min (M) remain, 15 M remain, 30 M remain, Mid-exam and No alert	An alert can be used to give attention about remaining examination time.

had to load the system from a USB.

4. Discussion

This systematic literature review sought to assess the current state of literature concerning online examinations and its equivalents. For most students, online learning environments created a system more supportive of their wellbeing, personal lives, and learning performance. Staff preferred online examinations for their workload implications and ease of completion, and basic evaluation of print-based examination logistics could identify some substantial ongoing cost savings. Not all staff and students preferred the idea of online test environments, yet studies that considered age and gender identified only negligible differences (Rios & Liu, 2017).

While the literature on online examinations is growing, there is still a dearth of discussion at the pedagogical and governance levels. Our review and new familiarity with papers led us to point researchers in two principal directions: accreditation and authenticity. We acknowledge that there are many possible pathways to consider, with reference to the consistency of application, the validity and reliability of online examinations, and whether online examinations enable better measurement and greater student success. There are also opportunities to synthesize online examination literature with other innovative digital pedagogical devices. For example, immersive learning environments (Herrington et al., 2007), mobile technologies (Jahnke & Liebscher, 2020); social media (Giannikas, 2020), and web 2.0 technologies (Bennett et al., 2012). The literature examined acknowledges key elements of the underlying needs for online examinations from student, academic, and technical perspectives. This has included the need for online examinations need to accessible, need to be able to distinguish a true pass from a true fail, secure, minimize opportunities for cheating, accurately authenticates the student, reduce marking time, and designed to be agile in software or technological failure.

We turn attention now to areas of need in future research, and focus on accreditation and authenticity over these alternates given there is a real need for more research prior to synthesis of knowledge on the latter pathways.

4.1. The accreditation question

The influence of external accreditation bodies was named frequently and ominously among the sample group, but lacked clarity surrounding exact parameters and expectations. Rios (2017, p. 231) identified a specific measure was used "for accreditation purposes". Hylton et al. (2016, p. 54) specified that the US Department of Education requires "appropriate procedures or technology are implemented" to authentic distance students. Gehringer and Peddycord (2013) empirically found that online/open-web examinations provided more significant data for accreditation. Underlying university decisions to use face-to-face invigilated examination settings is to enable authentication of learning – a requirement of many governing bodies globally. The continual refinement of rules has enabled a degree of assurance that students are who they say they are.

Nevertheless, sophisticated networks have been established globally to support direct student cheating from completing quick assessments and calculators with secret search engine capability through to full completion of a course inclusive of attending on-campus invigilated examinations. The authentication process in invigilated examinations does not typically account for distance students who have a forged student identification card to enable a contract service to complete their examinations. Under the requirement assure authentication of learning, invigilated examinations will require revision to meet contemporary environments. The inclusion of a broader range of big data from keystroke patterns, linguistics analysis, and whole-of-student analytics over a student lifecycle is necessary to identify areas of risk from the institutional perspective. Where a student has a significantly different method of typing or sentence structure, it is necessary to review.

An experimental study on the detection of cheating in a psychology unit found teachers could detect cheating 62 percent of the time (Dawson & Sutherland-Smith, 2017). Automated algorithms could be used to support the pre-identification of this process, given lecturers and professors are unlikely to be explicitly coding for cheating propensity when grading multiple hundreds of papers on the same topic. Future scholars should be considering the innate differences that exist among test-taking behaviors that could be codified to create pattern recognition software. Even in traditional invigilated examinations, the use of linguistics and handwriting evaluations could be used for cheating identification.

4.2. Authentic assessments and examinations

The literature identified in the sample discussed with limited depth the role of authentic assessment in examinations. The evolution of pedagogy and teaching principles (e.g. constructive alignment; Biggs, 1996) have paved the way for revised approaches to assessment and student learning. In the case of invigilated examinations, universities have been far slower to progress innovative solutions despite growing evidence that students prefer the flexibility and opportunities afforded by digitalizing exams. University commitments to the development of authentic assessment environments will require a radical revision of current examination practice to incorporate real-life learning processes and unstructured problem-solving (Williams & Wong, 2009). While traditional examinations may be influenced by financial efficacy, accreditation, and authentication pressures, there are upward pressures from student demand, student success, and student wellbeing to create more authentic learning opportunities.

The online examination setting offers greater connectivity to the kinds of environments graduates will be expected to engage in on a regular basis. The development of time management skills to plan times to complete a fixed time examination is reflected in the business student's need to pitch and present at certain times of the day to corporate stakeholders, or a dentist maintaining a specific time allotment for the extraction of a tooth. The completion of a self-regulated task online with tangible performance outcomes is reflected in many roles from lawyer briefs on time-sensitive court cases to high school teacher completions of student reports at the end

of a calendar year. Future practitioner implementation and evaluation should be focused on embedding authenticity into the examination setting, and future researchers should seek to understand better the parameters by which online examinations can create authentic learning experiences for students. In some cases, the inclusion of examinations may not be appropriate; and in these cases, they should be progressively extracted from the curriculum.

4.3. Where to next?

As institutions begin to provide higher learning flexibility to students with digital and blended offerings, there is scholarly need to consider the efficacy of the examination environment associated with these settings. Home computers and high-speed internet are becoming commonplace (Rainie & Horrigan, 2005), recognizing that such an assumption has implications for student equity. As Warschauer (2007, p. 41) puts it, "the future of learning is digital". Our ability as educators will be in seeking to understand how we can create high impact learning opportunities while responding to an era of digitalization. Research considering digital fluency in students will be pivotal (Crawford & Butler-Henderson, 2020). Important too, is the scholarly imperative to examine the implementation barriers and successes associated with online examinations in higher education institutions given the lack of clear cross-institutional case studies. There is also a symbiotic question that requires addressing by scholars in our field, beginning with understanding how online examinations can enable higher education, and likewise how higher education can shape and inform the implementation and delivery of online examinations.

4.4. Limitations

This study adopted a rigorous PRISMA method for preliminary identification of papers for inclusion, the MMAT protocol for identifying the quality of papers, and an inductive thematic analysis for analyzing papers included. These processes respond directly to limitations of subjectivity and assurance of breadth and depth of literature. However, the systematic literature review method limits the papers included by the search criteria used. While we opted for a broad set of terms, it is possible we missed papers that would typically have been identified in other manual and critical identification processes. The lack of research published provided a substantial opportunity to develop a systematic literature review to summarize the state of the evidence, but the availability of data limits each comment. A meta-analysis on quantitative research in this area of study would be complicated because of the lack of replication. Indeed, our ability to unpack which institutions currently use online examinations (and variants thereof) relied on scholars publishing on such implementations; many of which have not. The findings of this systematic literature review are also limited by the lack of replication in this infant field. The systematic literature review was, in our opinions, the most appropriate method to summarize the current state of literature despite the above limitations and provides a strong foundation for an evidence-based future of online examinations. We also acknowledge the deep connection that this research may have in relation to the contemporary COVID-19 climate in higher education, with many universities opting for online forms of examinations to support physically distanced education and emergency remote teaching. There were 138 publications on broad learning and teaching topics during the first half of 2020 (Butler-Henderson et al., 2020). Future research may consider how this has changed or influenced the nature of rapid innovation for online examinations.

5. Conclusion

This systematic literature review considered the contemporary literature on online examinations and their equivalents. We discussed student, staff, and technological research as it was identified in our sample. The dominant focus of the literature is still oriented on preliminary evaluations of implementation. These include what processes changed at a technological level, and how students and staff rated their preferences. There were some early attempts to explore the effect of online examinations on student wellbeing and student performance, along with how the changes affect the ability for staff to achieve.

Higher education needs this succinct summary of the literature on online examinations to understand the barriers and how they can be overcome, encouraging greater uptake of online examinations in tertiary education. One of the largest barriers is perceptions of using online examinations. Once students have experienced online examinations, there is a preference for this format due to its ease of use. The literature reported student performance did not have significant difference in final examination scores across online and traditional examination modalities. Student anxiety decreased once they had used the online examination software. This information needs to be provided to students to change students' perceptions and decrease anxiety when implementing an online examination system. Similarly, the information summarized in this paper needs to be provided to staff, such as the data related to cheating, reliability of the technology, ease of use, and reduction in time for establishing and marking examinations. When selecting a system, institutions should seek one that includes biometrics with a high level of precision, such as user authentication, and movement, sound, and keystroke monitoring (reporting deviations so the recording can be reviewed). These features reduce the need for online examinations to be invigilated. Other system features should include locking the system or browser, cloud-based technology so local updates are not required, and an interface design that makes using the online examination intuitive. Institutions should also consider how it will address technological failures and digital disparities, such as literacy and access to technology.

We recognize the need for substantially more evidence surrounding the post-implementation stages of online examinations. The current use of online examinations across disciplines, institutions, and countries needs to be examined to understand the successes and gaps. Beyond questions of 'do students prefer online or on-campus exams', serious questions of how student mental wellbeing, employability, and achievement of learning outcomes can be improved as a result of an online examination pedagogy is critical. In

conjunction is the need to break down the facets and types of digitally enhanced examinations (e.g. online, e-examination, BYOD examinations, and similar) and compare each of these for their respective efficacy in enabling student success against institutional implications. While this paper was only able to capture the literature that does exist, we believe the next stage of literature needs to consider broader implications than immediate student perceptions toward the achievement of institutional strategic imperatives that may include student wellbeing, student success, student retention, financial viability, staff enrichment, and student employability.

Author statement

Both authors Kerryn Butler-Henderson and Joseph Crawford contributed to the design of this study, literature searches, data abstraction and cleaning, data analysis, and development of this manuscript. All contributions were equal.

Funding

References

Abdel Karim, N., & Shukur, Z. (2016). Proposed features of an online examination interface design and its optimal values. *Computers in Human Behavior, 64*, 414–422. https://doi.org/10.1016/j.chb.2016.07.013

AbuMansour, H. (2017). Proposed bio-authentication system for question bank in learning management systems. In *2017 IEEE/ACS 14th international conference on computer systems and applications (AICCSA)* (pp. 489–494). https://doi.org/10.1109/aiccsa.2017.215

Aisyah, S., Bandung, Y., & Subekti, L. B. (2018). Development of continuous authentication system on android-based online exam application. In *2018 international conference on information technology systems and innovation (ICITSI)* (pp. 171–176). https://doi.org/10.1109/icitsi.2018.8695954

Al-Hakeem, M. S., & Abdulrahman, M. S. (2017). Developing a new e-exam platform to enhance the university academic examinations: The case of Lebanese French University. *International Journal of Modern Education and Computer Science, 9*(5), 9. https://doi.org/10.5815/ijmecs.2017.05.02

Alzu'bi, M. (2015). The effect of using electronic exams on students' achievement and test takers' motivation in an English 101 course. In *Proceedings of conference of the international journal of arts & sciences* (pp. 207–215).

Amigud, A., & Dawson, P. (2020). The law and the outlaw: is legal prohibition a viable solution to the contract cheating problem? *Assessment & Evaluation in Higher Education, 45*(1), 98–108. https://doi.org/10.1080/02602938.2019.1612851

Anderson, H. M., Cain, J., & Bird, E. (2005). Online course evaluations: Review of literature and a pilot study. *American Journal of Pharmaceutical Education, 69*(1), 34–43. https://doi.org/10.5688/aj690105

Ardid, M., Gómez-Tejedor, J. A., Meseguer-Dueñas, J. M., Riera, J., & Vidaurre, A. (2015). Online exams for blended assessment. Study of different application methodologies. *Computers & Education, 81*, 296–303. https://doi.org/10.1016/j.compedu.2014.10.010

Attia, M. (2014). Postgraduate students' perceptions toward online assessment: The case of the faculty of education, Umm Al-Qura university. In A. Wiseman, N. Alromi, & S. Alshumrani (Eds.), *Education for a knowledge society in Arabian Gulf countries* (pp. 151–173). Bingley, United Kingdom: Emerald Group Publishing Limited. https://doi.org/10.1108/s1479-3679_2014_0000024015.

Bennett, S., Bishop, A., Dalgarno, B., Waycott, J., & Kennedy, G. (2012). Implementing web 2.0 technologies in higher education: A collective case study. *Computers & Education, 59*(2), 524–534.

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*(3), 347–364. https://doi.org/10.1007/bf00138871

Böhmer, C., Feldmann, N., & Ibsen, M. (2018). E-exams in engineering education—online testing of engineering competencies: Experiences and lessons learned. In *2018 IEEE global engineering education conference (EDUCON)* (pp. 571–576). https://doi.org/10.1109/educon.2018.8363281

Botezatu, M., Hult, H., Tessma, M. K., & Fors, U. G. (2010). Virtual patient simulation for learning and assessment: Superior results in comparison with regular course exams. *Medical Teacher, 32*(10), 845–850. https://doi.org/10.3109/01421591003695287

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Butler-Henderson, K., Crawford, J., Rudolph, J., Lalani, K., & Sabu, K. M. (2020). COVID-19 in Higher Education Literature Database (CHELD V1): An open access systematic literature review database with coding rules. *Journal of Applied Learning and Teaching, 3*(2). https://doi.org/10.37074/jalt.2020.3.2.11. Advanced Online Publication.

Butt, A. (2018). Quantification of influences on student perceptions of group work. *Journal of University Teaching and Learning Practice, 15*(5).

Chao, K. J., Hung, I. C., & Chen, N. S. (2012). On the design of online synchronous assessments in a synchronous cyber classroom. *Journal of Computer Assisted Learning, 28*(4), 379–395. https://doi.org/10.1111/j.1365-2729.2011.00463.x

Chebrolu, K., Raman, B., Dommeti, V. C., Boddu, A. V., Zacharia, K., Babu, A., & Chandan, P. (2017). Safe: Smart authenticated Fast exams for student evaluation in classrooms. In *Proceedings of the 2017 ACM SIGCSE technical symposium on computer science education* (pp. 117–122). https://doi.org/10.1145/3017680.3017754

Chen, Q. (2018). An application of online exam in discrete mathematics course. In *Proceedings of ACM turing celebration conference-China* (pp. 91–95). https://doi.org/10.1145/3210713.3210734

Chytrý, V., Nováková, A., Rícan, J., & Simonová, I. (2018). Comparative analysis of online and printed form of testing in scientific reasoning and metacognitive monitoring. In *2018 international symposium on educational technology (ISET)* (pp. 13–17). https://doi.org/10.1109/iset.2018.00013

Crawford, J. (2015). *Authentic leadership in student leaders: An empirical study in an Australian university.* University of Tasmania, Australia: Honours Dissertation.

Crawford, J., & Butler-Henderson, K. (2020). Digitally empowered workers and authentic leaders: The capabilities required for digital services. In K. Sandhu (Ed.), *Leadership, management, and adoption techniques for digital service innovation* (pp. 103–124). Hershey, Pennsylvania: IGI Global. https://doi.org/10.4018/978-1-7998-2799-3.ch006.

Crawford, J., Butler-Henderson, K., Rudolph, J., Malkawi, B., Glowatz, M., Burton, R., Magni, P., & Lam, S. (2020). COVID-19: 20 countries' higher education intra-period digital pedagogy responses. *Journal of Applied Teaching and Learning, 3*(1), 9–28. https://doi.org/10.37074/jalt.2020.3.1.7

Creswell, J., & Miller, D. (2000). Determining validity in qualitative inquiry. *Theory into Practice, 39*(3), 124–130. https://doi.org/10.1207/s15430421tip3903_2

Daffin, L., Jr., & Jones, A. (2018). Comparing student performance on proctored and non-proctored exams in online psychology courses. *Online Learning, 22*(1), 131–145. https://doi.org/10.24059/olj.v22i1.1079

Dawson, P. (2016). Five ways to hack and cheat with bring-your-own-device electronic examinations. *British Journal of Educational Technology, 47*(4), 592–600. https://doi.org/10.1111/bjet.12246

Dawson, P., & Sutherland-Smith, W. (2018). Can markers detect contract cheating? Results from a pilot study. *Assessment & Evaluation in Higher Education, 43*(2), 286–293. https://doi.org/10.1080/02602938.2017.1336746

Dawson, P., & Sutherland-Smith, W. (2019). Can training improve marker accuracy at detecting contract cheating? A multi-disciplinary pre-post study. *Assessment & Evaluation in Higher Education, 44*(5), 715–725. https://doi.org/10.1080/02602938.2018.1531109

Eagly, A., & Sczesny, S. (2009). Stereotypes about women, men, and leaders: Have times changed? In M. Barreto, M. K. Ryan, & M. T. Schmitt (Eds.), *Psychology of women book series. The glass ceiling in the 21st century: Understanding barriers to gender equality* (pp. 21–47). American Psychological Association. https://doi.org/10.1037/11863-002.

Ellis, S., & Barber, J. (2016). Expanding and personalizing feedback in online assessment: A case study in a school of pharmacy. *Practitioner Research in Higher Education, 10*(1), 121–129.

Fluck, A. (2019). An international review of eExam technologies and impact. *Computers & Education, 132*, 1–15. https://doi.org/10.1016/j.compedu.2018.12.008

Fluck, A., Adebayo, O. S., & Abdulhamid, S. I. M. (2017). Secure e-examination systems compared: Case studies from two countries. *Journal of Information Technology Education: Innovations in Practice, 16*, 107–125. https://doi.org/10.28945/3705

Fluck, A., Pullen, D., & Harper, C. (2009). Case study of a computer based examination system. *Australasian Journal of Educational Technology, 25*(4), 509–533. https://doi.org/10.14742/ajet.1126

Gehringer, E., & Peddycord, B., III (2013). Experience with online and open-web exams. *Journal of Instructional Research, 2*, 10–18. https://doi.org/10.9743/jir.2013.2.12

Giannikas, C. (2020). Facebook in tertiary education: The impact of social media in e-learning. *Journal of University Teaching and Learning Practice, 17*(1), 3.

Gold, S. S., & Mozes-Carmel, A. (2009). A comparison of online vs. proctored final exams in online classes. *Journal of Educational Technology, 6*(1), 76–81. https://doi.org/10.26634/jet.6.1.212

Gross, J., Torres, V., & Zerquera, D. (2013). Financial aid and attainment among students in a state with changing demographics. *Research in Higher Education, 54*(4), 383–406. https://doi.org/10.1007/s11162-012-9276-1

Guillén-Gámez, F. D., García-Magariño, I., Bravo, J., & Plaza, I. (2015). Exploring the influence of facial verification software on student academic performance in online learning environments. *International Journal of Engineering Education, 31*(6A), 1622–1628.

Hainline, L., Gaines, M., Feather, C. L., Padilla, E., & Terry, E. (2010). Changing students, faculty, and institutions in the twenty-first century. *Peer Review, 12*(3), 7–10.

Hearn Moore, P., Head, J. D., & Griffin, R. B. (2017). Impeding students' efforts to cheat in online classes. *Journal of Learning in Higher Education, 13*(1), 9–23.

Herrington, J., Reeves, T. C., & Oliver, R. (2007). Immersive learning technologies: Realism and online authentic learning. *Journal of Computing in Higher Education, 19*(1), 80–99.

Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M. P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M. C., Vedel, I., & Pluye, P. (2018). The mixed methods appraisal tool (MMAT) version 2018 for information professionals and researchers. *Education for Information, 34*(4), 285–291. https://doi.org/10.3233/EFI-180221

Hylton, K., Levy, Y., & Dringus, L. P. (2016). Utilizing webcam-based proctoring to deter misconduct in online exams. *Computers & Education, 92*, 53–63. https://doi.org/10.1016/j.compedu.2015.10.002

Jahnke, I., & Liebscher, J. (2020). Three types of integrated course designs for using mobile technologies to support creativity in higher education. *Computers & Education, 146*. https://doi.org/10.1016/j.compedu.2019.103782. Advanced Online Publication.

Johnson, C. (2009). *History of New York state regents exams. Unpublished manuscript*.

Jordan, A. (2001). College student cheating: The role of motivation, perceived norms, attitudes, and knowledge of institutional policy. *Ethics & Behavior, 11*(3), 233–247. https://doi.org/10.1207/s15327019eb1103_1

Karp, A. (2007). Exams in algebra in Russia: Toward a history of high stakes testing. *International Journal for the History of Mathematics Education, 2*(1), 39–57.

Kemp, D. (1999). *Knowledge and innovation: A policy statement on research and research training*. Canberra: Australian Government Printing Service.

Kolagari, S., Modanloo, M., Rahmati, R., Sabzi, Z., & Ataee, A. J. (2018). The effect of computer-based tests on nursing students' test anxiety: A quasi-experimental study. *Acta Informatica Medica, 26*(2), 115. https://doi.org/10.5455/aim.2018.26.115-118

Kolski, T., & Weible, J. (2018). Examining the relationship between student test anxiety and webcam based exam proctoring. *Online Journal of Distance Learning Administration, 21*(3), 1–15.

Kumar, A. (2014). Test anxiety and online testing: A study. In *2014 IEEE Frontiers in Education Conference (FIE) Proceedings* (pp. 1–6). https://doi.org/10.1109/fie.2014.7044376

Li, X., Chang, K. M., Yuan, Y., & Hauptmann, A. (2015). Massive open online proctor: Protecting the credibility of MOOCs certificates. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing* (pp. 1129–1137). https://doi.org/10.1145/2675133.2675245

Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. California: Sage Publications.

Margaryan, A., Littlejohn, A., & Vojt, G. (2011). Are digital natives a myth or reality? University students' use of digital technologies. *Computers & Education, 56*(2), 429–440. https://doi.org/10.1016/j.compedu.2010.09.004

Matthíasdóttir, Á., & Arnalds, H. (2016). e-assessment: students' point of view. In *Proceedings of the 17th international conference on computer systems and technologies 2016* (pp. 369–374). https://doi.org/10.1145/2983468.2983497

Mitra, S., & Gofman, M. (2016). Towards greater integrity in online exams. In *Proceedings of the twenty-second americas conference on information systems (28)*.

Mohanna, K., & Patel, A. (2015). Overview of open book-open web exam over blackboard under e-Learning system. In *2015 fifth international conference on e-learning* (pp. 396–402). https://doi.org/10.1109/econf.2015.81

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151*(4). https://doi.org/10.7326/0003-4819-151-4-200908180-00135, 264-249.

Mueller, R. G., & Colley, L. M. (2015). An evaluation of the impact of end-of-course exams and ACT-QualityCore on US history instruction in a Kentucky high school. *Journal of Social Studies Research, 39*(2), 95–106. https://doi.org/10.1016/j.jssr.2014.07.002

Nguyen, H., & Henderson, A. (2020). Can the reading load Be engaging? Connecting the instrumental, critical and aesthetic in academic reading for student learning. *Journal of University Teaching and Learning Practice, 17*(2), 6.

NYSED. (2012). *History of regent examinations: 1865 – 1987. Office of state assessment*. http://www.p12.nysed.gov/assessment/hsgen/archive/rehistory.htm. (Accessed 7 January 2020).

Oz, H., & Ozturan, T. (2018). Computer-based and paper-based testing: Does the test administration mode influence the reliability and validity of achievement tests? *Journal of Language and Linguistic Studies, 14*(1), 67.

Pagram, J., Cooper, M., Jin, H., & Campbell, A. (2018). Tales from the exam room: Trialing an e-exam system for computer education and design and technology students. *Education Sciences, 8*(4), 188. https://doi.org/10.3390/educsci8040188

Park, S. (2017). Online exams as a formative learning tool in health science education. In *Proceedings of the 21st world multi-conference on systemics, cybernetics and informatics* (pp. 281–282). WMSCI 2017.

Patel, A. A., Amanullah, M., Mohanna, K., & Afaq, S. (2014l). E-exams under e-learning system: Evaluation of onscreen distraction by first year medical students in relation to on-paper exams. In *Third international conference on e-technologies and networks for development* (pp. 116–126). ICeND2014. https://doi.org/10.1109/icend.2014.6991364.

Petrović, J., Vitas, D., & Pale, P. (2017). Experiences with supervised vs. unsupervised online knowledge assessments in formal education. In *2017 international symposium ELMAR* (pp. 255–258). https://doi.org/10.23919/elmar.2017.8124480

Rainie, L., & Horrigan, J. (2005). *A decade of adoption: How the internet has woven itself into American life*. Washington, DC: Pew Internet & American Life Project.

Reiko, Y. (2001). University reform in the post-massification era in Japan: Analysis of government education policy for the 21st century. *Higher Education Policy, 14*(4), 277–291. https://doi.org/10.1016/s0952-8733(01)00022-8

Rettinger, D. A., & Kramer, Y. (2009). Situational and personal causes of student cheating. *Research in Higher Education, 50*(3), 293–313. https://doi.org/10.1007/s11162-008-9116-5

Rios, J. A., & Liu, O. L. (2017). Online proctored versus unproctored low-stakes internet test administration: Is there differential test-taking behavior and performance? *American Journal of Distance Education, 31*(4), 226–241. https://doi.org/10.1080/08923647.2017.1258628

Rodchua, S., Yiadom-Boakye, G., & Woolsey, R. (2011). Student verification system for online assessments: Bolstering quality and integrity of distance learning. *Journal of Industrial Technology, 27*(3).

Schmidt, S. M., Ralph, D. L., & Buskirk, B. (2009). Utilizing online exams: A case study. *Journal of College Teaching & Learning (TLC), 6*(8). https://doi.org/10.19030/tlc.v6i8.1108

Schwalb, S. J., & Sedlacek, W. E. (1990). Have college students' attitudes toward older people changed. *Journal of College Student Development, 31*(2), 125–132.

Seow, T., & Soong, S. (2014). Students' perceptions of BYOD open-book examinations in a large class: A pilot study. In *Proceedings of the australasian society for computers in learning in tertiary education, Dunedin* (pp. 604–608).

Sheppard, S. (1996). An informal history of how law schools evaluate students, with a predictable emphasis on law school final exams. *UMKC Law Review, 65,* 657.

Sindre, G., & Vegendla, A. (2015, November). E-exams and exam process improvement. In *NIK: Norsk Informatikkonferanse (n.p.)*.

Steel, A., Moses, L. B., Laurens, J., & Brady, C. (2019). Use of e-exams in high stakes law school examinations: Student and staff reactions. *Legal Education Review, 29* (1), 1.

Stowell, J. R., & Bennett, D. (2010). Effects of online testing on student exam performance and test anxiety. *Journal of Educational Computing Research, 42*(2), 161–171. https://doi.org/10.2190/ec.42.2.b

Sullivan, D. P. (2016). An integrated approach to preempt cheating on asynchronous, objective, online assessments in graduate business classes. *Online Learning, 20* (3), 195–209. https://doi.org/10.24059/olj.v20i3.650

Turner, J. L., & Dankoski, M. E. (2008). Objective structured clinical exams: A critical review. *Family Medicine, 40*(8), 574–578.

US National Library of Medicine. (2019). *Medical subject headings.* https://www.nlm.nih.gov/mesh/meshhome.html. (Accessed 28 March 2019).

Vision Australia. (2014). *Online and print inclusive design and legibility considerations. Vision Australia.* https://www.visionaustralia.org/services/digital-access/blog/12-03-2014/online-and-print-inclusive-design-and-legibility-considerations. (Accessed 12 November 2019).

Warschauer, M. (2007). The paradoxical future of digital learning. *Learning Inquiry, 1*(1), 41–49. https://doi.org/10.1007/s11519-007-0001-5

Wibowo, S., Grandhi, S., Chugh, R., & Sawir, E. (2016). A pilot study of an electronic exam system at an Australian University. *Journal of Educational Technology Systems, 45*(1), 5–33. https://doi.org/10.1177/0047239516646746

Williams, J. B., & Wong, A. (2009). The efficacy of final examinations: A comparative study of closed-book, invigilated exams and open-book, open-web exams. *British Journal of Educational Technology, 40*(2), 227–236. https://doi.org/10.1111/j.1467-8535.2008.00929.x

Wright, T. A. (2015). Distinguished Scholar Invited Essay: Reflections on the role of character in business education and student leadership development. *Journal of Leadership & Organizational Studies, 22*(3), 253–264. https://doi.org/10.1177/1548051815578950

Yong-Sheng, Z., Xiu-Mei, F., & Ai-Qin, B. (2015). The research and design of online examination system. In *2015 7th international conference on information technology in medicine and education (ITME)* (pp. 687–691). https://doi.org/10.1109/ITME.2015.96