International Journal of Communication, Science and Technology
vol. 1, no. 2 (February-March, 2025), pp.41-74

DIGITAL EXAMINATIONS IN HIGHER EDUCATION – A SYSTEMATIC LITERATURE REVIEW

Kristopher Pantani

Chair of Business Administration, esp. Corporate and Management Accounting

FernUniversität in Hagen

Universitätsstraße 41, 58084 Hagen, Germany

Niklas Lettow

Chair of Business Administration, esp. Corporate and Management Accounting

FernUniversität in Hagen

Universitätsstraße 41, 58084 Hagen, Germany

Prof. Dr. Jörn Littkemann

Chair of Business Administration, esp. Corporate and Management Accounting

FernUniversität in Hagen

Universitätsstraße 41, 58084 Hagen, Germany

ABSTRACT

The digitalization of university examinations has become a prominent issue, especially following the global COVID-19 pandemic. While digitalization offers opportunities for more efficient processes and resource conservation, it also presents challenges. This paper systematically reviews current research on digital examinations in higher education, based on an analysis of 85 studies. The review focuses on the

characteristics of digital examinations and their perceptions among different user groups. Findings show that all user groups recognize both advantages and disadvantages of digital ex- aminations. Acceptance tends to increase with experience, but concerns over security still limit widespread adoption. Clear policies and secure procedures are therefore essential to harness the full potential while mitigating associated risks.

Keywords: *Distance Learning, Higher Education, Learning Assessment, Learning Management System (LMS), Literature Review, User Acceptance*

1. INTRODUCTION

In the context of advancing digitalization and the accompanying transformation of education, the digital transformation of examination processes in academic teaching is becoming increasingly important (Bryant & Ruello, 2021). This trend, accelerated by the global COVID-19 pandemic, presents universities and colleges with numerous challenges, while also offering opportunities alongside potential risks (Dayananda et al., 2021; Dilini et al., 2021; Nurdin et al., 2021). Additionally, managing the increasing student population amid dwindling resources poses a significant challenge for university systems (Hillier, 2015). Although the digitalization holds the potential to enhance efficiency, conserve resources and standardize assessment processes, both teachers and students often view this development critically (Froehlich et al., 2023). The issue of acceptance, combined with challenges related to adapting to different disciplines and examination types, as well as adhering to restrictive legal regulations, underscores the complexity of this issue (Böhmer et al., 2018). Beyond problems with acceptance, concerns surrounding the security of examination data and the integrity of the process are also of central importance.

Although there are several studies on digital examinations, they have not been systematically summarized in a comprehensive synthesis. Therefore, the primary aim of this paper is to present the current state of the literature on digital examinations, particularly focusing on digital systems in academic teaching. Based on these objectives, the following research questions arise:

- What fundamental aspects need to be considered in digital exam creation?
- What are the characteristics of digital examination systems?
- How do students, teachers and the administration staff perceive digital examinations?
- How can the quality of digital exams be ensured?

To answer these research questions, this study uses the Systematic Literature Review (SLR) methodology. This method provides a comprehensive overview of the current research landscape on a specific topic through a targeted approach. As part of this analysis, the entire digital examination process is examined, including preparation,

procedure and post processing. The analysis and synthesis of existing studies will then serve to develop targeted recommendations and guidelines for the design of digital examinations.

2. BACKGROUND

In the academic context, the validation of learning outcomes through the successful completion of examinations is of paramount importance. Examinations that are aligned with the modules serve as proof of the attainment of knowledge, proficiencies, and competencies and are therefore essential for completing a degree (HG, 2000; UNESCO, 2012). Each assessment completed during the course of study is typically incorporated into the final grade, thereby constituting a significant component of academic success (Becker, 2022). In the traditional model, students demonstrate the knowledge they have accumulated throughout the semester in a final exam, which typically takes the form of a written test held in person at the end of the semester. They typically take place in a controlled setting where examinees and examiners do not communicate, ensuring the integrity and guaranteeing that all students are tested under the same conditions (Koh et al., 2021; Kuyoro et al., 2016). The objective of a fair assessment is to evaluate the knowledge of participants on presented topics adequately, to examine all students under equal conditions and to justify the results (Gorgani & Pak, 2020; Gorgani & Shabani, 2021).

The ongoing societal changes, particularly those accelerated by the global pandemic caused by the severe acute respiratory syndrome COVID-19 (SARS-CoV-2), have forced higher education institutions to confront new challenges and implement digital formats, allowing students to take examinations without being physically present (Dayananda et al., 2021; Dilini et al., 2021; Nurdin et al., 2021). These examinations entail the utilization of technological devices for the generation, administration, storage and/or dissemination of assessment outcomes and student feedback (Bryant & Ruello, 2021; Kuikka et al., 2014). In the majority of cases, students use their own devices to take the examinations in the room (a process known as 'Bring Your Own Device', or BYOD), or on devices that are loaned to them. During the pandemic, the majority of students took their examinations at locations they chose themselves. (Jaap et al., 2021;

Koh et al., 2021). Specialized software or digital platforms, collectively known as 'digital examination systems,' manage and deliver the content digitally (Kuyoro et al., 2016).

3. METHODOLOGY

To address the research questions, the authors chose the SLR methodology for its structured approach, transparency and repeatability (Elrod et al., 2022; Materla et al., 2019; Wetterich & Plänitz, 2021; Xiao & Watson, 2019). In accordance with the approach suggested by Xiao and Watson (2019), the first step involved developing appropriate search terms based on relevant terminology (Tranfield et al., 2003). As previously stated, the subject of investigation was digital examinations within the academic context. Accordingly, the following search terms and their synonyms were established: 'digital', 'examination' and 'university'.

Along with defining key terms, it is crucial to set specific inclusion and exclusion criteria to reduce the selection of irrelevant studies within the research area. This study conducted an extensive review of academic sources, including peer-reviewed journal articles, book chapters, and conference proceedings, in both German and English. This study explicitly excluded topics such as e-learning, digital proctoring and cheating, examination assessment and specific academic disciplines from its scope. The exclusion criteria emerged after an initial analysis of the titles and abstracts. A significant proportion of the existing studies primarily focused on digital proctoring, cheating and assessment. This indicates that extensive research has already been conducted in this area, necessitating a separate SLR. Additionally, the area of e-learning, which exclusively addresses the student learning process, is excluded as it is not conducive to answering the research questions. Finally, studies with a pronounced disciplinary focus are excluded, as their conclusions may not be universally applicable.

As stated by Siddaway et al. (2019), at least two databases relevant to the topic should be consulted. The search encompassed ten databases in the fields of education and social sciences, information technology and psychology, selected based on the scope of their studies. The search ended when no new information emerged and most studies began to show repeated results (Xiao & Watson, 2019). A total of 1,714 studies were identified, of which 610 were duplicates. In accordance with the recommendations set

forth by Siddaway et al. (2019), the titles were initially reviewed, followed by an abstract review to ascertain their eligibility for full-text reading. The exclusion of certain articles after full-text reading was based on the pre-established inclusion and exclusion criteria.

Additionally, this study did not investigate the effect of digital examinations on student performance compared to traditional formats, as this topic has already been widely explored in previous research. The extant literature offers no definitive findings on this matter. Rather, numerous authors conclude that various factors - such as the type of exam questions, students' digital competencies and other influencing variables - can lead to different performance outcomes (Blazer, 2010; Karay et al., 2015; Lim et al., 2006; Zheng & Bender, 2019). An overview of the identified studies is presented in Figure 1.

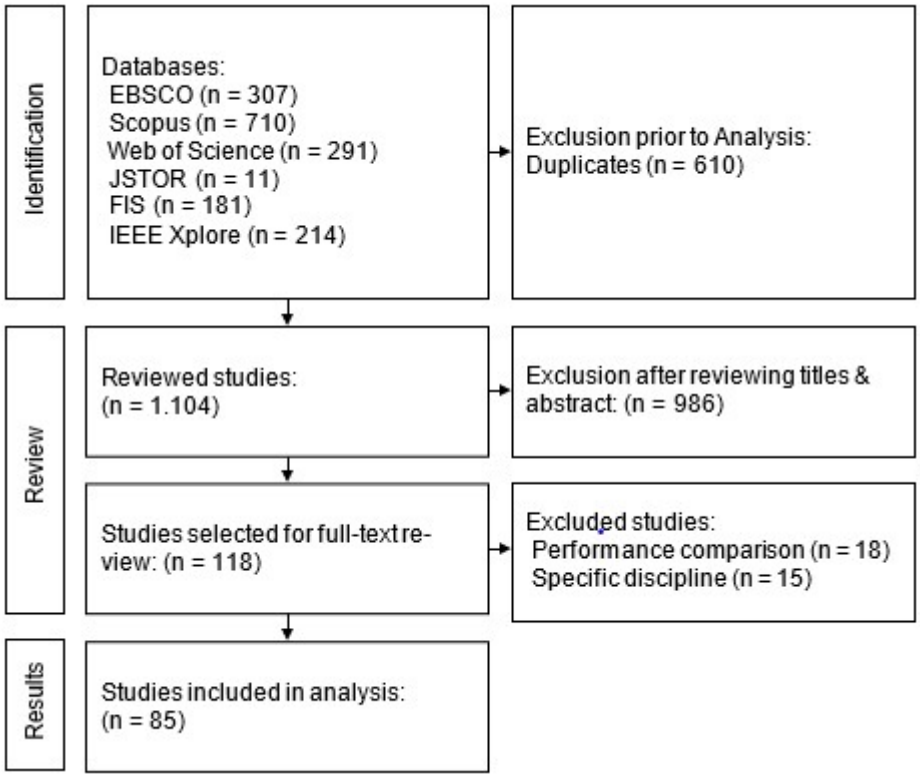

Figure 1: Approach to the Systematic Literature Review (based on Page et al., 2021)

A total of 38 journal articles, 28 book chapters and 19 conference proceedings were identified through the systematic approach. The following section will outline the findings

from the analysis of these studies, focusing on their content to uncover similarities and potential differences.

4. FINDINGS

4.1 Descriptive Findings and Classification of the Studies

The analysis encompasses a total of 85 studies, distributed across different regions of the world as follows: A total of 39 studies (46%) were conducted in Asia, while 27 studies (32%) were carried out in Europe, with the majority of these studies taking place in Germany. A total of 12 studies (14%) were conducted in North America, with the majority of these being carried out in the United States. Additionally, four studies (5%) were conducted in Australia. The remaining three studies (3%) were sourced from other regions. A further analysis of the regional data reveals that, despite often having similar national backgrounds and frequently being public institutions, universi- ties tend to pursue individual developments or third-party solutions rather than lever- aging synergies to create shared systems and standards. It is also important to con- sider that the prerequisites, such as the technical infrastructure and the legal frame- work, vary significantly between countries.

The temporal distribution of studies, as shown in Table 1, also highlights a growing focus on the digitalization of university examinations, particularly in the context of the ongoing advancements in digital technologies and the impact of the pandemic (Boeskens et al., 2023).

Time frame	Number of studies
Publication date < 2010	8 (9 %)
Publication date between 2010 and 2015	15 (18 %)
Publication date between 2015 and 2020	28 (33 %)
Publication date > 2020	34 (40%)

Table 1: Time Frame of the Studies

In general, there is a persistent tendency towards the investigation and application of digital examination techniques. In particular, the number of studies increased significantly after 2020. This reflects the heightened interest in digital transformation in assessment practices and the urgency prompted by the global pandemic of 2020, which led to an inevitable shift within the education sector (Boeskens et al., 2023; Jaap et al., 2021; Lehane et al., 2022; Starkey et al., 2021).

The regional distribution and temporal dynamics of these studies demonstrate that the digitalization of university examinations is a globally relevant issue, with varying levels of attention across different regions. This trend indicates that universities are increasingly recognizing the digital transformation of assessments as an element of contemporary educational strategies. It has the potential to enhance the efficacy, accessibility, and equity of assessment methods (Gehringer & Peddycord, 2013; Kaddoura & Gumaei, 2022; Noller, 2022).

In addition to analyzing the temporal and regional distribution, the researchers classified the studies into seven thematic areas; see Figure 2 for reference. These categories provide a structured approach to understanding the key focal points of the studies:

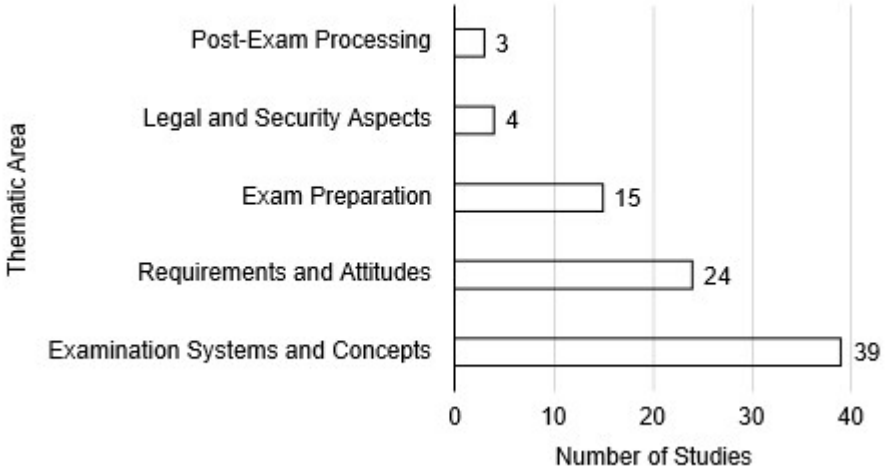

Figure 2: Classification of the Studies

A clear trend has emerged in academic discourse, as the detailed description and evaluation of digital examination concepts, structures and systems have become a prominent focus.

These are analyzed with regard to their functionality and structure, and are assessed by all involved parties, including students (n = 15), teachers (n = 5), and administrative staff (n = 4). Another area of research focuses on the attitudes of all stakeholders towards digital examination processes. While studies on preparation and post processing are less numerous, they nevertheless occupy a central position in the field. These studies address the creation of tasks and the subsequent assessment of the quality. In addition, legal and security-related frameworks for digital examinations were identified, but in detail was precluded by their country-specific nature.

In accordance with the classification and chronology of the studies, the following sections are divided into three categories: examination preparation and creation, procedure and systems and post processing. As part of the analysis, the attitudes of participants towards digital exams are examined and significant theories and techniques are analyzed and discussed with the objective of deriving recommendations for action. Finally, this review provides an overview of the methodological approaches employed in research on the digitalization of examinations in higher education.

4.2 Examination Preparation and Creation

As already explained, the analysis begins chronologically with the preparation and creation of examinations as the first step. A review of the theories used reveals that Bloom's Taxonomy and its extensions are frequently cited in the context of preparation. The classification system created by Benjamin Bloom in 1956 is an educational framework designed to classify learning objectives, streamline teaching methods and promote standardization. The taxonomy is divided into three major learning domains: the cognitive domain, the affective domain and the psychomotor domain (Ahmad et al., 2011; Bloom et al., 1956). The cognitive domain is concerned with mental abilities and comprises six hierarchical levels: knowledge, comprehension, application, analysis, synthesis and evaluation (Anderson, 2014; Bloom et al., 1956). These levels are employed to quantify the degree of cognitive complexity inherent in learning objectives,

rendering them indispensable in the context of higher education (Ahmad et al., 2011). The affective domain relates to the emotional components of learning, while the psychomotor domain involves physical skills and motor functions (Yang & Lin, 2023).

A substantial body of research has demonstrated the efficacy of employing Bloom's Taxonomy and its derivatives to categorize inquiries during the design phase. This approach has been shown to enhance learning outcomes and ensure the fairness of assessment. This method aligns with the constructive alignment theory, which asserts that learning is most effective when instructional techniques and assessment strategies are closely matched to the learning objectives (Schulze-Achatz et al., 2018; Schulze-Vorberg et al., 2016). Learning objectives offer students guidance and organization, increasing the chances of achieving successful learning outcomes (Blumberg, 2009). This suggests that learning objectives should be clearly defined and that examination questions should be derived from and categorized according to these objectives. The aim of this approach is to improve teaching outcomes and, consequently, students' performance.

To effectively design examinations using Bloom's Taxonomy, a number of prerequisites must be met. Firstly, curriculum objectives must be addressed. Secondly, examinations must include a variety of question formats and difficulty levels. Thirdly, cognitive levels must be considered. Finally, points must be allocated appropriately across the questions (Amria et al., 2018; Bardesi & Razek, 2014; Kale & Kiwelekar, 2013). In order to implement Bloom's framework in digital creation, Amria et al. (2018) put forth three potential approaches:

- The utilization of a question bank, from which examiners may select questions;
- A random selection of questions is employed for the generation of examinations through the use of simple randomization;
- The generation of questions is facilitated through the implementation of algorithms and artificial intelligence (Amria et al., 2018).

A fundamental necessity for all scenarios is the availability of a structured question bank that enables both examiners and automated processes to select questions and facilitate the exchange of questions across departments and universities (Borromeo, 2013;

Elkhatat, 2022; Imran et al., 2019; Ware et al., 2014). Furthermore, Elkhatat (2022) proposes the creation of multiple sub-pools within the comprehensive database to minimize the prevalence of repetitive questions and guarantee an equitable distribution. Furthermore, a decision must be made regarding the use of randomization and if so, the type of randomization to employ. In order to determine the optimal approach, it is essential to consider the various scenarios that may be employed, including the absence of randomization, the implementation of individual randomization for each examination, the utilization of A/B testing and its associated variations, or the randomization of the order of examination tasks (Bardesi & Razek, 2014). For example, the Fisher-Yates algorithm can be used to avoid repetition and is particularly well suited for randomization processes (Febriani et al., 2021).

In addition to straightforward randomization, the third approach - generation via algorithms/AI - is especially noteworthy because it enables the creation based on a multitude of criteria. A common methodology employed in research is natural language processing (NLP), a subfield of artificial intelligence. Natural language processing (NLP) can be employed to categorize examination questions in accordance with Bloom's Taxonomy by identifying pertinent keywords (Jayakodi et al., 2016; see, for instance, Khedr et al., 2022). Alternatively, the entire examination can be subjected to this process. NLP can be employed to generate examinations based on pre-established criteria, which are then randomized (for an example see Amria et al., 2018; Ferreyra & Backhoff-Escudero, 2016).

In addition to the specific criteria for creation, general requirements, such as location (e.g. at the university, at home, etc.), time (e.g. fixed dates, 24/7-examination availability, etc.) and equipment (e.g. university devices, your own device, etc.) specifications, must be considered in all scenarios. Pagram et al. (2018) found in their study that students are often reluctant to use their personal devices due to concerns about privacy and place greater trust in the functionality of university-provided devices.

In summary, the following findings can be derived from the study on preparation:

- Curricula and learning objectives are explicitly delineated and accurately reflected in the construction of examinations (Amria et al., 2018; Bardesi & Razek, 2014; Kale & Kiwelekar, 2013);
- An important prerequisite for the creation is the development of a question bank (Elkhatat, 2022; Imran et al., 2019; Ware et al., 2014), which includes different types of tasks, difficulty levels and classification according to Bloom's taxonomy;
- In addition to manual creation, examinations can be generated by simple randomization or using complex algorithms and artificial intelligence (Amria et al., 2018);
- Finally, general criteria such as location, time and equipment use must be clearly defined.

After discussing studies on preparation, the focus now shifts to the procedure and the characteristics of the systems used.

4.3 Examination procedure and systems

Digital systems, whether web-based or program-based, must support the specific needs of different user groups. Three primary roles can be identified in an exam system: students, teachers and administrators (Bardesi & Razek, 2014). An overview of these user groups, along with their main functions, is shown in Figure 3 and will be discussed in more detail below.

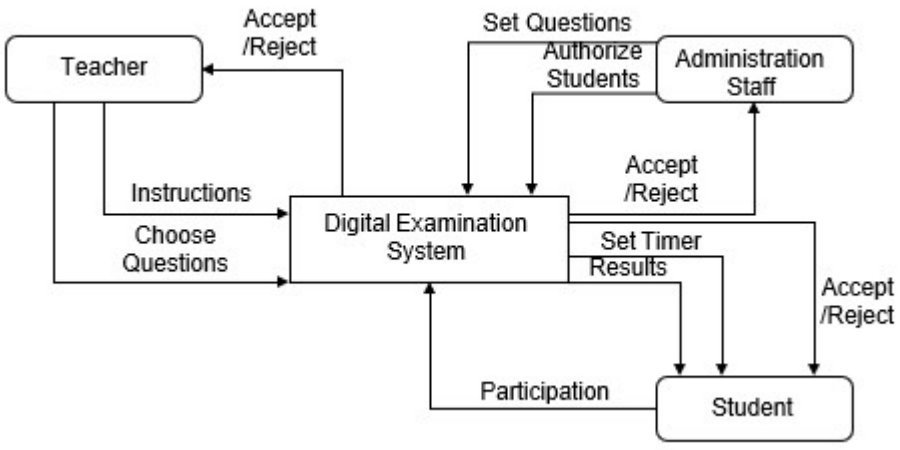

Figure 3: Elements of Digital Examination Systems (based on Febriani et al., 2021; Kuyoro et al., 2016)

Students authenticate to the system, take examinations and access their grades through the system. Teachers are responsible for creating and grading questions, as well as setting parameters such as duration or time per question. They can also set the level of question randomization, maintain the question bank with sample solutions and review results. Administrators manage the server infrastructure, enroll both teachers and students in the system, implement security protocols such as authentication and anti-fraud measures and ensure examination integrity. In addition, the administration staff defines the organizational parameters (Al-Hakeem & Salim Abdulrahman, 2017; Bardesi & Razek, 2014; Bella et al., 2011; Kuyoro et al., 2016). The system itself must be able to authenticate both student and teachers' credentials and create sessions for each user. It should also allow students to review or modify their answers, randomly generate questions and prevent repetition of previously presented questions. Grading can be fully automated by the system, performed in a hybrid format where the examiner reviews the results before finalizing them, or performed entirely manually. At the end of the examination, the system should calculate the student's score based on their answers and communicate the results to the student (Kuyoro et al., 2016).

Following the overview of the functions of the individual user groups, a detailed examination of the requirements and attitudes of the user groups towards digital examination systems is presented. The studies analyzed, as shown in Figure 2, specifically capture the requirements of the students. Particular focus is given to student attitudes, as evidenced by the extensive number of studies conducted in this field. The results of these studies are summarized in Table 2 for students and in Table 3 for teachers and administrators:

Discussion and Results	Authors
Faster writing skills and uncomplicated editing and revision of solutions	Hillier, 2015
Perceived performance enhancement	Hamsatu et al., 2016

Increased acceptance through: Increased perceived user-friendliness and accessibility; flexibility, freedom and transparency, particularly through independence of time and place; reduced test anxiety and stress; increased experience; faster assessment.	Adanir & Çinar, 2021; Adanir et al., 2020; Froehlich et al., 2023; Jaap et al., 2021; Miettunen, 2006; Ngafif, 2018; Yilmaz & Hebebci, 2022; Zheng & Bender, 2019
Reduced acceptance and increased test anxiety due to: Academic dishonesty; a lack of fairness and the use of generic examinations; a lack of experience and information and privacy; technical difficulties and failures; and time constraints.	Hartmann et al., 2021; Hillier, 2015; Joshi & Brastad, 2019; Keijzer-de Ruijter & Draaijer, 2019; Ocak & Karakuş, 2021; Romaniuk & Łukasiewicz-Wieleba, 2021; Ziehfreund et al., 2022

Table 2: Students' Attitudes towards Digital Exams

The analysis of the studies reveals that students have varying perceptions of digital exams, which can affect their level of acceptance. The success largely depends on students' perceptions of their experience, ease of use, system reliability and fairness (Adanir & Çinar, 2021; Borisov et al., 2020; Froehlich et al., 2023; Hillier, 2015). Examination-related anxiety decreases as students gain initial experience with the system, with initial uncertainty being replaced by familiarity. This highlights the importance of introductory programs and supportive measures (Hamsatu et al., 2016; Hartmann et al., 2021; Romaniuk & Łukasiewicz-Wieleba, 2021; Ziehfreund et al., 2022).

User Group	Discussion and Results	Authors
	Time and resource savings through more efficient and flexible testing processes Customized testing Integration of multimedia elements	Ferreyra & Backhoff-Escudero, 2016; Hamsatu et al., 2016; Rjoub et al., 2009; Yilmaz & Hebebci, 2022
	Necessity of fixed design patterns E-	Joshi & Brastad, 2019

Teachers	tests	
	Increased risk due to: academic dishonesty (especially from external sources); interactions with each other; loss of control; loss of validity and reliability	Keijzer-de Ruijter & Draaijer, 2019; Miettunen, 2006; Romaniuk & Łukasiewicz-Wieleba, 2021; Yilmaz & Hebebci, 2022
Administration Staff	Integration of the inspection system into the overall system leads to optimized processes	Bryant & Ruello, 2021
	Uncomplicated archiving	Miettunen, 2006; Yilmaz & Hebebci, 2022
	Lack of common pedagogical and technological basis for faculty-wide examination systems	Chirumamilla & Sindre, 2021; Yilmaz & Hebebci, 2022
	Technical and infrastructural problems due to: Long preparation; security issues; interruptions due to network problems.	

Table 3: Teachers' and Administrators' Attitudes towards Digital Examinations

The studies suggest that teachers and administrators, similar to students, perceive both benefits and challenges of digital examinations. New formats expand the possibilities of examination design, particularly through personalization and the integration of multimedia elements (Ferreyra & Backhoff-Escudero, 2016; Rjoub et al., 2006; Romaniuk & Łukasiewicz-Wieleba, 2021). However, digital examinations also carry in- creased risks, particularly related to academic dishonesty, loss of control over the process and challenges related to validity and reliability (Gurung et al., 2012; Kim & Choi, 2020; Stadler et al., 2021).

Another key issue in the discussion of digital examinations is cost considerations,

which are viewed differently. Savings in travel, space and materials costs due to the lack of physical presence requirements, as well as the reduction of effort due to optimized and accelerated examination processes are notable advantages (Miettunen, 2006; Yilmaz & Hebebci, 2022). Others highlight the high investment costs for IT infra-structure, arguing that due to additional expenses such as servers and maintenance, the long-term costs may be comparable to traditional cost structures (Keijzer-de Ruijter & Draaijer, 2019). Overall, there is no consensus on the costs.

The following results can be derived from the studies of the procedure:

- Identification of the following user groups: students, teachers and administrators (Bardesi & Razek, 2014);
- Involving all stakeholders in the implementation process and increasing familiarity with the system can improve student adoption and performance. Proactive approaches such as user training should be prioritized (Froehlich et al., 20-23; Hillier, 2015; Joshi & Brastad, 2019; Schulze-Vorberg et al., 2016; Wallace & Clariana, 2005);
- System security is essential for all user groups, especially in terms of authentication, external protection and fraud prevention (Elkhatat, 2022);
- Despite the need for initial investment, institutions can benefit from optimized pro- cesses (Keijzer-de Ruijter & Draaijer, 2019; Miettunen, 2006; Yilmaz & Hebebci, 2022).

Overall, it is evident that successful implementation demands thorough attention to the needs of all stakeholders, with a strong emphasis on security and user engagement being central to these efforts (Joshi & Brastad, 2019).

4.4 Post-Examination Processing

The review reveals an increasing focus on the evaluation of examinations and test items. In this context, Item Response Theory (IRT) is the most frequently mentioned approach (n = 7). IRT is a set of statistical models used to analyze test and questionnaire items (Brennan, 2006). It is commonly used in educational or psychological testing, where data from questionnaires or standardized measurement instruments are modeled. A key feature of IRT models is the use of latent variables to represent the constructs being assessed. In the educational context, 'ability' is referred

to as the la- tent variable in IRT that describes an individual's competence in solving test items. The model correlates participant characteristics with item characteristics in order to fore- cast the probability of a correct response and the examinee's aptitude (Cai et al., 2016; Jumailiyah, 2017). For individuals with low ability, the chance of answering correctly is almost zero, while for those with high ability, it nears 100% (Baker, 2002).

In the 3-parameter model of the IRT, three item parameters are required in addition to participant ability to calculate the probability of a correct answer: difficulty, discrimination and guessing behavior (Baker, 2002; Jumailiyah, 2017). The difficulty parameter reflects the challenge of a task, while the discrimination parameter indicates how sensitive an item is to differences in participants' abilities (Baker & Kim, 2017). This means that items are more likely to be answered correctly by individuals with higher abilities. Finally, the guessing parameter describes the likelihood that individuals with lower ability can answer a question correctly by guessing. These parameters are estimated using the maximum likelihood method (Baker & Kim, 2017; Cai et al., 2016).

By analyzing different items, the quality can be improved. Ferreyra and Backhoff-Escudero (2016) propose continuous validation of examination, especially in the case of automated test generation. The IRT item parameters-difficulty, discrimination and guessing behavior-contribute to improving the fairness when these factors are considered and incorporated into the generation of question banks. This approach allows for better classification of questions, which ultimately leads to fairer assessments (Zhuang et al., 2020).

Overall, the following trends are emerging in post-examination processing:

- Ongoing evaluation analysis of test items using IRT can help improve fairness;
- Incorporating IRT parameters into a question bank can improve quality (Ferreyra & Backhoff-Escudero, 2016; Zhuang et al., 2020).

5. IMPLICATIONS

The following section begins with recommendations for the content of digital examinations and concludes with implications for the design. The content structure should start with the definition of learning objectives based on Bloom's taxonomy to

derive appropriate questions (Schulze-Vorberg et al., 2016). These questions are stored and classified in databases, where the IRT item parameters (difficulty, discrimination and guessing behavior) should be considered to facilitate comprehensive and structured preparation as well as balanced assessments (Ahmad et al., 2011; Amria et al., 2018; Fowler et al., 2022). In summary, the digital examination process can be divided into four phases, as shown in Figure 4.

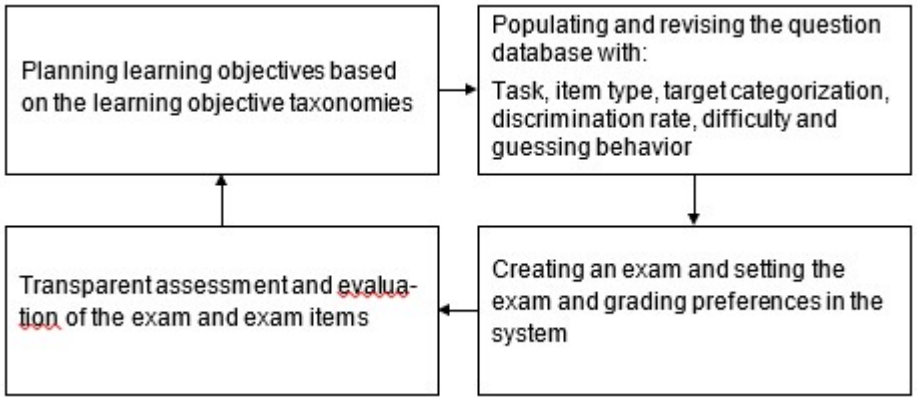

Figure 4: Digital Examination Process

In the preparation phase, interpersonal aspects are of great importance. Lack of experience can lead to test anxiety and the perception that students with prior knowledge have an advantage during the testing process (Hartmann et al., 2021). Therefore, in- sufficient experience should be addressed through various offerings such as consulta- tions, training programs and practice modes (Alkhateeb et al., 2022; Hamsatu et al., 2016; Romaniuk & Łukasiewicz-Wieleba, 2021; Ziehfreund et al., 2022) to reduce un- certainty and increase student acceptance (Froehlich et al., 2023; Hillier, 2015; Schulze-Vorberg et al., 2016). Creative development approaches such as design thinking, as well as engaging and communicating with students during the implementation and adaptation of the user interface, can simplify usability and increase adoption (Ocak & Karakuş, 2021; Rowan et al., 2024; Ziehfreund et al., 2022).

The structure must be adapted according to the modality to ensure fairness (Romaniuk & Łukasiewicz-Wieleba, 2021). Keijzer-de Ruijter and Draaijer (2019) outline guide- lines for the layout and design of digital examination structures, emphasizing that scrolling between questions should be avoided, line width should be limited to a

maximum of 600 pixels, spaces should be placed between paragraphs and a sans-serif font should be used. Borisov et al. (2020) also specify minimum and maximum font sizes (12-26), highlight the usefulness of a countdown timer and an autosave feature (Pagram et al., 2018). The ability to return to previous answers and make multiple attempts is essential to correct potential mistakes (Keijzer-de Ruijter & Draaijer, 2019; Ocak & Karakuş, 2021). Additionally, Agarwala et al. (2021) call for a minimum broad- band of 4G and an uninterrupted power supply during examinations. Islam et al. (2021) discuss the possibility of extending time based on bandwidth to minimize disad- vantages due to poor internet connections. In case of technical issues, logs should be reviewed and a retake offered (Adanir et al., 2020).

Another element in reducing student frustration with creation is the specification and personalization of examination (Khedr et al., 2022). Researchers refer to this as the smart-testing approach, an intelligent model that enables the creation of individualized and targeted examinations based on student behavioral and knowledge data (Qaffas et al., 2023). Offering the option to take a second examination to improve grades is also a strategy to improve student learning and performance. However, it is important not to shorten the second test, as this could result in fewer concepts being assessed (Fernandez, 2021; Fowler et al., 2022; Morphew et al., 2020).

Security is a major concern for administrators (Yilmaz & Hebebci, 2022; Ziehfreund et al., 2022). The implementation of sophisticated artificial authentication techniques (such as fingerprint scanning, facial recognition, iris scanning, keystroke and signature matching), is of paramount importance to guarantee the authenticity of the examinee and to deter any potential cheating attempts (Al-Fayoumi & Aboud, 2017; Çoban Budak et al., 2023; Romaniuk & Łukasiewicz-Wieleba, 2021). It is important to note that students are generally more willing to share sensitive data with their university than with third parties (Levy et al., 2011). Measures such as locking programs, disabling copy functions and shortcuts, using open-ended questions, randomizing questions to personalize and version tests, requiring cameras and implementing thoughtful time management, including considerations of time pressure and the duration of tasks, are all critical elements in preventing cheating (Ahn et al., 2014; Duric & Mahmutovic, 2021; Frankl et al., 2012; Gehringer & Peddycord, 2013; Koh et al., 2021; Lee, 2022).

It should be noted, however, that time pressure may increase the difficulty (Stadler et al., 2021). Comprehensive education on academic honesty is essential to increase awareness of moral and ethical behavior (Böhmer et al., 2018; Fask et al., 2014).

In summary, the design must be carefully planned and structured to ensure fairness and optimal use. The testing process should be supported by a clear definition of learn- ing objectives and the careful creation and management of question banks that incorporate various parameters. Technical reliability, flexible structures and a high level of security are essential, as is the involvement and support of all stakeholders.

6. CONCLUSION

The purpose of this study was to provide a comprehensive review of the current literature on digital examinations in higher education. The focus was on the areas of creation, procedure and post-examination evaluation. A SLR was chosen as the method- ology to provide a focused and thorough understanding of the current research in this area (Elrod et al., 2022; Materla et al., 2019; Wetterich & Plänitz, 2021; Xiao & Watson, 2019).

A total of 85 studies were identified, providing important insights into the field of creation. Three methods for selecting questions were identified: examiner selection, simple randomization and algorithmic/AI randomization. Consideration of Bloom's Taxonomy and IRT parameters is critical to ensuring the quality of items. Students, teachers and administrators have different roles and concerns, with a strong emphasis on security issues related to academic dishonesty. This underscores the need to be fully secure against external interference and for universities to implement preventative measures against cheating. Increasingly, responsibility for the environment is being shifted to students, who should be supported through practice and training (Al-Mashaqbeh & Al Hamad, 2010). After the examination, the quality of items should be assessed using testing theories such as IRT to ensure fairness and quality.

Future studies should explore what prevents universities from working together to develop more secure testing environments. Such collaborative efforts could leverage synergies and lead to cost savings (Miettunen, 2006; Yilmaz & Hebebci, 2022). While Hillier (2015) examined the use of portable computers in his study, the integration of mobile devices into systems remains a largely unexplored area that future research

should address.

In summary, one major limitation of this literature review is the choice of studies. There is a risk that not all relevant contributions were captured, which may limit the findings. Given the dynamic development of the field and the increasing number of studies in recent years, it is also possible that further relevant research will be published in the coming years that was not considered in this work. Future research should adopt a more focused selection of studies, including a more precise delimitation into the areas of preparation, administration and post evaluation, to provide a comprehensive view of the state of research in each subfield.

REFERENCES

Adanir, G. A., & Çinar, M. (2021). The Acceptance and Use of an Online Exam System by Online Learners: Implementation of the UTAUT Model. *Sakarya University Journal of Education*, *11*(3), 412–430. https://doi.org/10.19126/suje.830529

Adanir, G. A., Ismailova, R., Omuraliev, A., & Muhametjanova, G. (2020). Learners' Perceptions of Online Exams: A Comparative Study in Turkey and Kyrgyzstan. *International Review of Research in Open and Distributed Learning*, *21*(3), 1–17. https://doi.org/10.19173/irrodl.v21i3.4679

Agarwala, P., Phadke, S., & Tilak, P. (2021). Online Exams during COVID-19: Teacher's Perspective. *Turkish Online Journal of Qualitative Inquiry*, *12*(6), 5050–5056.

Ahmad, N. D., Adnan, W. A. W., Abdul Aziz, M., & Yusaimir Yusof, M. (2011). Automating Preparation of Exam Questions: Exam Question Classification System (EQCS). In IEEE Computer Society (Ed.), *2011 International Conference on Research and Innovation in Information Systems* (pp. 1–6). IEEE. https://doi.org/10.1109/ICRIIS.2011.6125715

Ahn, J. Y., Han, K. S., Choi, S. H., & Mun, G. S. (2014). Designing a Personalized Exam System to Enhance Students' Understanding. *ICIC Express Letters*, *8*(2), 349–355.

Al-Fayoumi, M., & Aboud, S. J. (2017). An Efficient E-Exam Scheme. *International Journal of Emerging Technologies in Learning*, *12*(4), 153–162. https://doi.org/10.3991/ijet.v12i04.6719

Al-Hakeem, M., & Salim Abdulrahman, M. (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), 9–16. https://doi.org/10.5815/ijmecs.2017.05.02

Alkhateeb, N. E., Ahmed, B. S., Al-Tawil, N. G., & Al-Dabbagh, A. A. (2022). Students and Examiners' Perception on Virtual Medical Graduation Exam during

the COVID-19 Quarantine Period: A Cross-Sectional Study. *PloS One*, *17*(8), 1-12. https://doi.org/10.1371/journal.pone.0272927

Al-Mashaqbeh, I. F., & Al Hamad, A. (2010). Student's Perception of an Online Exam within the Decision Support System Course at Al al Bayt University. In IEEE Computer Society (Ed.), *Proceedings of the 2010 Second International Conference on Computer Research and Development* (pp. 131–135). IEEE. https://doi.org/10.1109/ICCRD.2010.15

Amria, A., Ewais, A., & Hodrob, R. (2018). A Framework for Automatic Exam Generation Based on Intended Learning Outcomes. In B. M. McLaren, R. Reilly, S. Zvacek, & J. Uhomoibhi (Eds.), *Proceedings of the 10th International Conference on Computer Supported Education* (2nd ed., pp. 474–480). Science and Technology Publications. https://doi.org/10.5220/0006795104740480

Anderson, L. W. (Ed.). (2014). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's* (Pearson New International Edition). Pearson Education.

Baker, F. B. (2002). *The Basics of Item Response Theory* (2nd ed.). ERIC Clearinghouse on Assessment and Evaluation.

Baker, F. B., & Kim, S.-H. (2017). *The Basics of Item Response Theory Using R*. *Statistics for Social and Behavioral Sciences*. Springer International Publishing. https://doi.org/10.1007/978-3-319-54205-8

Bardesi, H. J. A., & Razek, M. A. (2014). Learning Outcome E-Exam System. In IEEE Computer Society (Ed.), *2014 Sixth International Conference on Computational Intelligence, Communication Systems and Networks* (pp. 77–82).

IEEE. https://doi.org/10.1109/CICSyN.2014.29

Becker, M. (2022). *Universitäre Bildung in Deutschland: Ideengeschichtliche Perspektiven und aktuelle Herausforderungen in NRW* (1st ed.). *Nomos eLibrary Open Access: Vol. 29*. Nomos Verlagsgesellschaft.

Bella, G., Costantino, G., Coles-Kemp, L., & Riccobene, S. (2011). Remote Management of Face-to-Face Written Authenticated though Anonymous Exams. In A. Verbraeck, M. Helfert, J. Cordeiro, & B. Shishkov (Eds.), *Proceedings of the 3rd International Conference on Computer Supported Education* (pp. 431–437). SciTePress.

Blazer, C. (2010). Computer-Based Assessments. *Research Services, Miami-Dade County Public Schools*, *0918*, 1–18.

Bloom, B. S., Krathwohl, D. R., & Masia, B. B. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals*. Longmans, Green & Co.

Blumberg, P. (2009). Maximizing Learning through Course Alignment and Experience with Different Types of Knowledge. *Innovative Higher Education*, *34*(2), 93–103. https://doi.org/10.1007/s10755-009-9095-2

Boeskens, L., Meyer, K., & Minea-Pic, A. (2023). Building a Digital Education Policy Ecosystem for Quality, Equity and Efficiency. In Organisation for Economic Co-operation and Development (Ed.), *Shaping Digital Education: Enabling Factors for Quality, Equity and Efficiency.* OECD Publishing.

Böhmer, C., Feldmann, N., & Ibsen, M. (2018). E-Exams in Engineering Education — Online Testing of Engineering Competencies: Experiences and lessons learned. In IEEE Computer Society (Ed.), *Proceedings of 2018 IEEE Global Engineering Education Conference* (pp. 571–576). IEEE. https://doi.org/10.1109/EDUCON.2018.8363281

Borisov, E., Petrenko, E., Zatsarinnaya, E., Romanov, A., Cherkina, V., & Sotnikova, L. (2020). Organization of Taking Exams of University Students in Online Format: Problems and Opportunities. *E3S Web of Conferences*, *217*(8), Article 08005, 1–11. https://doi.org/10.1051/e3sconf/202021708005

Borromeo, R. M. H. (2013). Online Exam for Distance Educators Using Moodle. In D. Tan & L. Fang (Eds.), *Proceedings of the 2013 IEEE 63rd Annual Conference International Council for Education Media* (pp. 1–4). IEEE. https://doi.org/10.1109/CICEM.2013.6820155

Brennan, R. L. (2006). *Educational measurement*. *ACE Praeger series on higher education*. Praeger.

Bryant, P., & Ruello, J. (2021). One System to Examine Them All: Defining the Complexities of Implementing an Institution Wide Online Exam Model. *ASCILITE Publications*, 370–374. https://doi.org/10.14742/apubs.2019.290

Cai, L., Choi, K., Hansen, M., & Harrell, L. (2016). Item Response Theory. *Annual Review of Statistics and Its Application*, *3*(1), 297–321. https://doi.org/10.1146/annurev-statistics-041715-033702

Chirumamilla, A., & Sindre, G. (2021). E-Exams in Norwegian Higher Education: Vendors and Managers Views on Requirements in a Digital Ecosystem Perspective. *Computers & Education*, *172*, 104263–104282. https://doi.org/10.1016/j.compedu.2021.104263

Çoban Budak, E., Yurtay, N., Budak, Y., & Geçer, A. K. (2023). Voice-Assisted Online Exam Management and System Usability Analysis with Visually Impaired Students. *Interactive Learning Environments*, *31*(9), 5508–5522. https://doi.org/10.1080/10494820.2021.2010219

Dayananda, D. P., Chathumini, K. G., & Vasanthapriyan, S. (2021). A Novel Framework for Online Exams during the Pandemic of COVID-19: Evaluation Methods, Students' Priorities and Academic Dishonesty in Online Exams. In IEEE Computer Society (Ed.), *2021 IEEE 1st International Conference on Advanced Learning Technologies on Education & Research* (pp. 1–4). IEEE. https://doi.org/10.1109/ICALTER54105.2021.9675092

Dilini, N., Senaratne, A., Yasarathna, T., Warnajith, N., & Seneviratne, L. (2021). Cheating Detection in Browser-based Online Exams through Eye Gaze Tracking. In *2021 6th International Conference on Information Technology Research (ICITR)* (pp. 1–8). IEEE. https://doi.org/10.1109/ICITR54349.2021.9657277

Duric, J., & Mahmutovic, A. (2021). Software for Writing Online Exam with Video and Audio Surveillance - Cheatless. In IEEE Computer Society (Ed.), *2021 44th International Convention on Information, Communication and Electronic Tech-*

nology (pp. 654–659). IEEE. https://doi.org/10.23919/MIPRO52101.2021.9596723

Elkhatat, A. M. (2022). Practical Randomly Selected Question Exam Design to Address Replicated and Sequential Questions in Online Examinations. *International Journal for Educational Integrity*, *18*(1), 1–18. https://doi.org/10.1007/s40979-022-00103-2

Elrod, C. C., Stanley, S. M., Cudney, E. A., Hilgers, M. G., & Graham, C. (2022). Management Information Systems Education: A Systematic Review. *Journal of Information Systems Education*, *33*(4), 357–370.

Fask, A., Englander, F., & Wang, Z. (2014). Do Online Exams Facilitate Cheating? An Experiment Designed to Separate Possible Cheating from the Effect of the Online Test Taking Environment. *Journal of Academic Ethics*, *12*(2), 101–112. https://doi.org/10.1007/s10805-014-9207-1

Febriani, I., Ekawati, R., Supriadi, U., & Abdullah, M. I. (2021). Fisher-Yates Shuffle Algorithm for Randomization Math Exam on Computer Based-Test. In *AIP Conference Proceedings, Transforming Research and Education of Science and Mathematics in the Digital Age* (060015-1-060015-7). AIP Publishing. https://doi.org/10.1063/5.0042534

Fernandez, O. E. (2021). Second Chance Grading: An Equitable, Meaningful, and Easy-to-Implement Grading System that Synergizes the Research on Testing for Learning, Mastery Grading, and Growth Mindsets. *PRIMUS*, *31*(8), 855–868. https://doi.org/10.1080/10511970.2020.1772915

Ferreyra, M. F., & Backhoff-Escudero, E. (2016). Validez del Generador Automático de Ítems del Examen de Competencias Básicas (Excoba). *Revista Electrónica De Investigación Y Evaluación Educativa*, *22*(1), 1–17. https://doi.org/10.7203/relieve.22.1.8048

Fowler, M., Smith, D. H., Emeka, C., West, M., & Zilles, C. (2022). Are We Fair? In L. Merkle, M. Doyle, J. Sheard, L.-K. Soh, & B. Dorn (Eds.), *Proceedings of the*

53rd ACM Technical Symposium on Computer Science Education (pp. 647–653). ACM. https://doi.org/10.1145/3478431.3499388

Frankl, G., Schartner, P., & Zebedin, G. (2012). Secure Online Exams Using Students' Devices. In IEEE Computer Society (Ed.), *Proceedings of the 2012 IEEE Global Engineering Education Conference* (pp. 1–7). IEEE Computer Society. https://doi.org/10.1109/EDUCON.2012.6201111

Froehlich, L., Sassenberg, K., Jonkmann, K., Scheiter, K., & Stürmer, S. (2023). Student Diversity and E-Exam Acceptance in Higher Education. *Journal of Computer Assisted Learning*, *39*(4), 1196–1210. https://doi.org/10.1111/jcal.12794

Gehringer, E. F., & Peddycord, B. W. (2013). Experience with Online and Open-Web Exams. *Journal of Instructional Research*, *2*, 10–18. https://doi.org/10.9743/JIR.2013.2.12

Gorgani, H. H., & Pak, A. J. (2020). A New Method for Assessment of Engineering Drawing Answer Scripts Using Fuzzy Logic. *Journal of Computational Applied Mechanics*, *51*(1), 170–183. https://doi.org/10.22059/JCAMECH.2019.265225.325

Gorgani, H. H., & Shabani, S. (2021). Online Exams and the COVID-19 Pandemic: A Hybrid Modified FMEA, QFD, and K-Means Approach to Enhance Fairness. *SN Applied Sciences*, *3*(10), 1–18. https://doi.org/10.1007/s42452-021-04805-z

Gurung, R. A. R., Wilhelm, T. M., & Filz, T. (2012). Optimizing Honor Codes for Online Exam Administration. *Ethics & Behavior*, *22*(2), 158–162. https://doi.org/10.1080/10508422.2011.641836

Hamsatu, P., Yusufu, G., & Mohammed, H. A. (2016). Teachers' Perceptions and Undergraduate Students' Experience in E-Exam in Higher Institution in Nigeria. *Journal of Education and Practice*, *7*(23), 158–166.

Hartmann, P., Hobert, S., & Schumann, M. (2021). The Intention to Participate in Online Exams – The Students Perspective. In Curran Associates Inc (Chair), *27th Americas Conference on Information Systems,* Remote.

Hillier, M. (2015). E-Exams with Student Owned Devices: Student Voices. In D. Churchill, T. K. F. Chiu, & N. J. Gu (Chairs), *International Mobile Learning Festival.* Symposium conducted at the meeting of The University of Hong Kong, Hong Kong.

Imran, J. B., Madni, T. D., Taveras, L. R., Clark, A. T., Ritchie, C., Cunningham, H. B., Christie, A., Abdelfattah, K. R., & Farr, D. (2019). Assessment of General Surgery Resident Study Habits and Use of the TrueLearn Question Bank for American Board of Surgery In-Training Exam Preparation. *American Journal of Surgery*, *218*(3), 653–657. https://doi.org/10.1016/j.amjsurg.2019.02.031

Islam, M. N., Hoque, A. M. T., & Habib, M. A. (2021). Design and Implementation of Bandwidth Based Time Varying Online Exam. In IEEE Computer Society (Ed.), *2021 3rd International Conference on Electrical & Electronic Engineering* (pp. 101–104). IEEE Computer Society. https://doi.org/10.1109/ICEEE54059.2021.9718990

Jaap, A., Dewar, A., Duncan, C., Fairhurst, K., Hope, D., & Kluth, D. (2021). Effect of Remote Online Exam Delivery on Student Experience and Performance in Applied Knowledge Tests. *BMC Medical Education*, *21*(1), 86. https://doi.org/10.1186/s12909-021-02521-1

Jayakodi, K., Bandara, M., & Meedeniya, D. (2016). An Automatic Classifier for Exam Questions with WordNet and Cosine Similarity. In IEEE Computer Society (Ed.), *2016 Moratuwa Engineering Research Conference* (pp. 12–17). IEEE Computer Society. https://doi.org/10.1109/MERCon.2016.7480108

Joshi, S. G., & Brastad, L. Å. (2019). Comparing Three Input Devices for Sketching Assignments in E-Exams in Computer Science. In *Proceedings of the International Conference on e-Learning 2019* (pp. 105–115). IADIS Press. https://doi.org/10.33965/el2019_201909f014

Jumailiyah, M. (2017). Item Response Theory: A Basic Concept. *Educational Research and Reviews*, *12*(5), 258–266. https://doi.org/10.5897/ERR2017.3147

Kaddoura, S., & Gumaei, A. (2022). Towards Effective and Efficient Online Exam Systems Using Deep Learning-Based Cheating Detection Approach. *Intelligent Systems with Applications*, *16*(200153), 1–12. https://doi.org/10.1016/j.iswa.2022.200153

Kale, V. M., & Kiwelekar, A. W. (2013). An Algorithm for Question Paper Template Generation in Question Paper Generation System. In *2013 The International Conference on Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE)*.

Karay, Y., Schauber, S. K., Stosch, C., & Schüttpelz-Brauns, K. (2015). Computer Versus Paper - Does It Make any Difference in Test Performance? *Teaching and Learning in Medicine*, *27*(1), 57–62. https://doi.org/10.1080/10401334.2014.979175

Keijzer-de Ruijter, M., & Draaijer, S. (2019). Digital Exams in Engineering Education. In S. Draaijer, D. Joosten-ten Brinke, & E. Ras (Eds.), *Communications in Computer and Information Science. Technology Enhanced Assessment* (Vol. 1014, pp. 140–164). Springer International Publishing. https://doi.org/10.1007/978-3-030-25264-9_10

Khedr, A. E., Almazroi, A. A., & Idrees, A. M. (2022). Intelligent Framework for Enhancing the Quality of Online Exams Based on Students' Personalization. *International Journal of Advanced Computer Science and Applications*, *13*(7), 605–614. https://doi.org/10.14569/IJACSA.2022.0130772

Kim, H., & Choi, U. (2020). Learner Perception of an Online L2-Course Summative Exam. *Multimedia-Assisted Language Learning*, *23*(3), 258–279.

Koh, L. L. A., Tan, M. X., Pee, G.-Y. M., Lee, C. H., Colla, M., & Kwan, W. L. (2021). Exploring Fair and Effective Online Electronic Exam in place of In-Person Examinations during Remote Learning. In IEEE Computer Society (Ed.), *2021 IEEE International Conference on Engineering, Technology & Education (TALE)* (pp. 1–7). IEEE. https://doi.org/10.1109/TALE52509.2021.9678669

Kuikka, M., Kitola, M., & Laakso, M.-J. (2014). Challenges when Introducing Elec-

tronic Exam. *Research in Learning Technology*, *22*(0), 1–17. https://doi.org/10.3402/rlt.v22.22817

Kuyoro, S. O., Maminor, G. U., Kanu, R. U., & Akande, O. (2016). The Design and Implementation of a Computer Based Testing System. *Journal of Applied Computation*, *1*, 1–7.

Hochschulgesetz, March 14, 2000.

Lee, H. (2022). A Framework for Synchronous Remote Online Exams. *IEICE Transactions on Information and Systems*, *E105-D*(7), 1343–1347. https://doi.org/10.1587/transinf.2022EDL8009

Lehane, P., Scully, D., & O'Leary, M. (2022). 'Time to Figure Out What to Do': Understanding the Nature of Irish Post-Primary Students' Interactions with Computer-Based Exams (CBEs) that Use Multimedia Stimuli. *Irish Educational Studies*, *41*(1), 5–25. https://doi.org/10.1080/03323315.2021.2022517

Levy, Y., Ramim, M. M., Furnell, S. M., & Clarke, N. L. (2011). Comparing Intentions to Use University-Provided vs Vendor-Provided Multibiometric Authentication in Online Exams. *Campus-Wide Information Systems*, *28*(2), 102–113. https://doi.org/10.1108/10650741111117806

Lim, E., Ong, B., Wilder-Smith, E., & Seet, R. (2006). Computer-based Versus Pen-and-paper Testing: Students' Perception. *Annals of the Academy of Medicine*, *35*(9), 599–603. https://doi.org/10.47102/annals-acadmedsg.v35n9p599

Materla, T., Cudney, E. A., & Antony, J. (2019). The Application of Kano Model in the Healthcare Industry: A Systematic Literature Review. *Total Quality Management & Business Excellence*, *30*(6), 660–681. https://doi.org/10.1080/14783363.2017.1328980

Miettunen, J. (2006). Students Prefer Online Exams. In International Association for Development of the Information Society (Chair), *The International Conference WWW/Internet 2006,* Murcia.

Morphew, J. W., Silva, M., Herman, G., & West, M. (2020). Frequent Mastery Testing with Second-Chance Exams Leads to Enhanced Student Learning in Under-

graduate Engineering. *Applied Cognitive Psychology*, *34*(1), 168–181. https://doi.org/10.1002/acp.3605

Ngafif, A. (2018). The Use of Web-Based Exam (WBE) to Optimize Students Testing Result. *English Review: Journal of English Education*, *6*(2), 41–50. https://doi.org/10.25134/erjee.v6i2.1241

Noller, J. (2022). Challenges and Requirements in Hybrid Written Exams Settings. In *Proceedings of the 27th European Conference on Pattern Languages of Programs* (pp. 1–5). ACM. https://doi.org/10.1145/3551902.3551975

Nurdin, E., Zubaidah Amir, M. Z., Nufus, H., Thakur, D., Shah, F. A., Dube, R., & Tommy Tanu Wijaya (2021). Let your Students Cheat on Mathematics Online Exams: Students' Perspectives. *Malikussaleh Journal of Mathematics Learning*, *4*(2), 131–136. https://doi.org/10.29103/mjml.v4i2.3286

Ocak, G., & Karakuş, G. (2021). Undergraduate students' Views of and Difficulties in Online Exams during the COVID-19 Pandemic. *Themes in E-Learning*, *14*, 13–30.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., Moher, D. (2021). The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *PLoS Medicine*, *18*(3), S. 1–15. https://doi.org/10.1371/journal.pmed.1003583

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), 1–11. https://doi.org/10.3390/educsci8040188

Qaffas, A. A., Idrees, A. M., Khedr, A. E., & Kholeif, S. A. (2023). A Smart Testing Model Based on Mining Semantic Relations. *IEEE Access*, *11*, 30237–30246. https://doi.org/10.1109/ACCESS.2023.3260406

Rjoub, A., Eyadat, Y., Ghazawi, A., Tall, B., Sharou, N., & Mardeeni, L. (2009). A

Multi-form Multiple Choice Editor Exam Tool Based on HTML Website and Artificial Intelligence Techniques. *Journal of Computer Science*, *5*(6), 405–412. https://doi.org/10.3844/jcssp.2009.405.412

Rjoub, A., Tall, B., Sharou, N., & Mardeeni, L. (2006). A Novel Multi-Forms Multiple Choice Editor Exam Tool Based on HTML Website. In *2006 7th International Conference on Information Technology Based Higher Education and Training* (pp. 854–869). IEEE. https://doi.org/10.1109/ITHET.2006.339710

Romaniuk, M. W., & Łukasiewicz-Wieleba, J. (2021). Academic Lecturers Towards the Students' Examining. Similarities and Differences of Stationary and Remote Exams in the Pandemic Era. *International Journal of Electronics and Telecommunications*, 63–68. https://doi.org/10.24425/ijet.2022.139849

Rowan, W., McCarthy, S., Mebrahtu, S., Cauche, C., O'Reilly, K., & Odili, D. (2024). Teaching Tip Embedding Sustainability in Information Systems Design Education. *Journal of Information Systems Education*, *35*(2), 122–137. https://doi.org/10.62273/HBHX1382

Schulze-Achatz, S., Pengel, N., Pachtmann, K., Franken, O., Köhler, T., Schlenker, L., & Wollersheim, H.-W. (2018). TASKtrain – Kompetenzorientierte Qualifizierung von Hochschullehrenden zur Konzeption und Erstellung von E-Prüfungsaufgaben. In H. Fischer & T. Köhler (Eds.), *Medien in der Wissenschaft: Vol. 73. Postgraduale Bildung mit digitalen Medien: Problemlagen und Handlungsansätze aus Sicht der Beteiligten.* Waxmann.

Schulze-Vorberg, L., Fabriz, S., Beckmann, N., Niemeyer, J., Tillmann, A., Kebschull, U., Krömker, D., & Horz, H. (2016). Die Potenziale von E-Prüfungen nutzen: Ein Konzept zur Unterstützung von Hochschullehrenden bei der Einführung von elektronischen Prüfungsformaten. In B. Berendt (Ed.), *Neues Handbuch Hochschullehre* (Vol. 3, pp. 127–144). DUZ Verlags- und Medienhaus.

Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to Do a Systematic Review: A Best Practice Guide for Conducting and Reporting Narrative Reviews,

Meta-Analyses, and Meta-Syntheses. *Annual Review of Psychology*, *70*, 747–770. https://doi.org/10.1146/annurev-psych-010418-102803

Stadler, M., Kolb, N., & Sailer, M. (2021). The Right Amount of Pressure: Implementing Time Pressure in Online Exams. *Distance Education*, *42*(2), 219–230. https://doi.org/10.1080/01587919.2021.1911629

Starkey, L., Shonfeld, M., Prestridge, S., & Cervera, M. G. (2021). Special issue: Covid-19 and the Role of Technology and Pedagogy on School Education during a Pandemic. *Technology, Pedagogy and Education*, *30*(1), 1–5. https://doi.org/10.1080/1475939X.2021.1866838

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management*, *14*(3), 207–222. https://doi.org/10.1111/1467-8551.00375

UNESCO. (2012). *International Standard Classification of Education ISCED 2011*. UNESCO.

Wallace, P., & Clariana, R. B. (2005). Test Mode Familiarity and Performance - Gender and Race Comparisons of Test Scores among Computer-Literate Students in Advanced Information Systems Courses. *Journal of Information Systems Education*, *16*(2), 177–182.

Ware, J., Kattan, T., Mohammed, A., & Siddiqui, I. (2014). The Perfect MCQ exam. *Journal of Health Specialties*, *2*(3), 94–99. https://doi.org/10.4103/1658-600x.137880

Wetterich, C., & Plänitz, E. (2021). *Systematische Literaturanalysen in den Sozialwissenschaften: Eine praxisorientierte Einführung*. Verlag Barbara Budrich.

Xiao, Y., & Watson, M. (2019). Guidance on Conducting a Systematic Literature Review. *Journal of Planning Education and Research*, *39*(1), 93–112. https://doi.org/10.1177/0739456X17723971

Yang, H.-H., & Lin, J.-Y. (2023). Students' Persistence Intention in MOOCs in the

Psychomotor Domain: An extended 3P Model of the Teaching and Learning Perspective. *Frontiers in Psychology*, *14*, 1–10. https://doi.org/10.3389/fpsyg.2023.1094138

Yilmaz, O., & Hebebci, M. T. (2022). Examining the Opinions of Faculty Members on Online Exams with SWOT Analysis. In M. Shelley, T. Ozturk, & H. Akcay (Chairs), *International Conference on Research in Education and Science,* Antalya.

Zheng, M., & Bender, D. (2019). Evaluating Outcomes of Computer-Based Classroom Testing: Student Acceptance and Impact on Learning and Exam Performance. *Medical Teacher*, *41*(1), 75–82. https://doi.org/10.1080/0142159X.2018.1441984

Zhuang, Z.-Y., Ho, C.-K., Tan, P. J. B., Ying, J.-M., & Chen, J.-H. (2020). The Optimal Setting of A/B Exam Papers without Item Pools: A Hybrid Approach of IRT and BGP. *Mathematics*, *8*(8), Article 1290, 1–29. https://doi.org/10.3390/MATH8081290

Ziehfreund, S., Reifenrath, J., Wijnen-Meijer, M., Welzel, J., Sauter, F., Wecker, H., Biedermann, T., & Zink, A. (2022). Considering Medical Students' Perception, Concerns and Needs for E-Exam during COVID-19: A Promising Approach to Improve Subject Specific E-Exams. *Medical Education Online*, *27*(1), Article 2114131, 1–11. https://doi.org/10.1080/10872981.2022.2114131