Evaluating Student Perceptions of Smart E-assessment Systems in Hong Kong Higher Education

Hon Keung YAU
City University of Hong Kong, Department of System Engineering, Kowloon Tong, Kowloon, Hong Kong.
honkyau@cityu.edu.hk

Choi Ho Man
City University of Hong Kong, Department of System Engineering, Kowloon Tong, Kowloon, Hong Kong.
homanchoi8-c@my.cityu.edu.hk

Abstract
This study explores Hong Kong higher education students' perceptions of E-assessment systems, focusing on factors shaping acceptance of E-examinations over traditional formats. Quantitative analysis of 107 respondents reveals significant positive correlations between diverse pre-exam guidance (e.g., tutorials) and key system features (e.g., usability, user friendly), and between these features and effective E-examination feedback (e.g., instants feedback) and identity monitoring techniques (e.g., webcam surveillance). E-examinations' advantages also correlate with enhanced monitoring identity. Prior experience boosts guidance value. These findings urge institutions to refine E-assessment with robust guidance, intuitive design, and reliable monitoring to enhance engagement, integrity, and satisfaction. Limited sample diversity suggests broader demographic studies. Results inform E-assessment optimization across contexts.
Keywords: E-assessment; Importance of diverse guidance; Key E-assessment system features; E-exams versus traditional exams; E-exam feedback; Monitoring techniques.

1. Introduction

Educational assessment has dramatically evolved from traditional pen-and-paper methods to technology-driven approaches, thanks to advancements in digital technology. This shift has led to more dynamic and interactive assessment formats. As institutions increasingly adopt e-assessment systems, it is essential to understand their impact on student engagement and learning. Technology integration allows for rapid, automated results and immediate feedback, significantly enhancing the evaluation process (Kiryakova, 2021).

Numerous studies have explored students' perceptions of E-assessment, highlighting a range of insights and experiences (Crisp et al., 2016; Iahad et al., 2004; Khan et al., 2021; Kumar & Owston, 2015; Kundu & Bej, 2021; Kiryakova, 2021; Pham, 2022; Rostaminezhad, 2019). However, there is a notable gap in research regarding Hong Kong higher education students' perceptions of e-assessment systems. Given that the e-learning landscape has changed significantly over the past decade, it is essential to conduct a study on E-assessment in today's higher education context.The research will address the following questions:

1. How do demographic differences and personal characteristics, such as university or institution, major, year of study, and prior experience, affect students' perceptions of E-assessment in higher education?
2. What are the students' perceptions regarding the importance of providing various types of guidance before E-examinations?
3. What are the students' perceptions regarding the essential features of an effective E-assessment system?
4. What are the students' perceptions regarding the feedback students prefer to receive from E-examinations?
5. What are the students' perceptions regarding E-examinations compare to traditional examinations?
6. What are the students' perceptions regarding techniques used for monitoring and verifying identity during E-examinations?
7. What are the students' perceptions regarding the idea of E-examinations becoming a standard format in higher education, and what features do they feel are missing from current E-assessment systems?

To narrow the research's focus, the study will also compare perceptions based on various factors, such as differences between fields of study(Major), year of study, and prior experience with e-assessments. This focused strategy will offer helpful insights into how these factors influence the way students view e-assessment systems.

2. Literature

2.1 Definition and Purpose of Educational Assessment

Educational assessment systematically collects data on students' knowledge, skills, and attitudes to enhance teaching and learning by evaluating instructional methods, identifying program weaknesses, and monitoring progress (Institute of Education Sciences, 2023). By analyzing student performance data, educators can evaluate how effectively their teaching methods meet set learning goals, enabling continuous improvement and tailored education for students. (Baleni, Z., 2015; Owan, V. et al., 2023; Osiesi, M., 2020; Hernández, R., 2012). Assessments engage students by fostering ownership, encouraging goal-setting, and promoting active participation, while regular feedback highlights strengths and areas for improvement, boosting confidence and enriching the learning experience (Amangeldina, G. & Dudovich, D., 2022; Magdalena, I. et al., 2023; Osiesi, M., 2020; Hernández, R., 2012; Zhou, M., 2023).

Technology enhances assessments with engaging tools like online quizzes, simulations, and multimedia presentations, providing immediate feedback that allows students to adjust their learning strategies in real time (Rostaminezhad, 2019; Pham, 2022; Duterte, J., 2024; Ejjami, R., 2024). Formative assessments, integrated throughout the learning process, enable teachers to adapt their methods to meet diverse student needs, as Sadler, D. (1989) emphasizes. This shift transforms assessments into dynamic tools for improving teaching and creating responsive learning environments (Sortwell, A. et al., 2024; Lee, H. et al., 2020).

2.2 Types of Assessments

Various assessments are used to achieve educational goals, each serving distinct purposes in the learning process.

1. Diagnostic Assessments

Diagnostic assessments identify students' existing skills and areas of need before teaching starts (Fan, T. et al., 2021).

2. Formative Assessments

Ongoing feedback from these assessments helps teachers and students adjust learning strategies in real-time (Cañadas, L., 2021; Lee, H. et al., 2020).

3. Interim Assessments

Interim assessments track student progress and inform teaching adjustments throughout the year (Wang, C., 2021; Perie, M. et al., 2009).

4. Summative Assessments

These final evaluations assess students' overall achievement and mastery of course objectives at the end of a term (Kibble, J., 2017; Kusumoriny, L., 2024).

Assessments serve distinct purposes for various stakeholders in education (Bulut, O. et al., 2024; Kusumoriny, L., 2024; Sri, S. et al., 2024). Students gain insights into their learning progress, identifying strengths and areas for improvement, which helps them make informed decisions about their learning plans (Bulut, O. et al., 2024; Kusumoriny, L., 2024; Sri, S. et al., 2024; Ifenthaler, D. et al., 2022; Harris, L. et al., 2022; Shin, N. et al., 2010). Educators use assessments to evaluate group and individual needs, adapt teaching strategies, and provide targeted support, fostering deeper learning (Kusumoriny, L., 2024; Constantinou, P., 2017; Graue, M., 1993).

Administrators leverage assessment data to evaluate instructional programs, allocate resources, and align school plans with student needs (Sievertsen, H., 2022; Omoeva, C. et al., 2021). At the state level, assessments measure the effectiveness of interventions and ensure students meet standards in key subjects like mathematics and literacy (Goertz, M., & Duffy, M., 2001; Institute of Education Sciences, 2023). A balanced assessment system enhances student learning and refines teaching practices (Turdieva, R., 2024; Kusumoriny, L., 2024; Din, A. et al., 2023; Muhammad, I. et al., 2024).

2.3 E-assessment: Definition and Characteristics

E-assessment refers to any type of assessment that is constructed, delivered, and marked using technology, typically through a tailor-made assessment platform. E-assessment encompasses a wide range of practices that integrate technology into various stages of the assessing process, from creation to delivery, marking, and feedback (Jordan, S., 2013; Mo, D. et al., 2022; Kundu, A. & Bej, T., 2021). In the majority of cases, E-assessment is the complete end-to-end process on a computer or digital device, but it may also be used to refer to some components, such as on-screen marking or test construction (Gibson, 2020).

E-assessment has also been referred to traditionally as computer-based testing, online assessment, or computer-based assessment, and the phrases are interchanged (Heil, J. & Ifenthaler, D., 2023). The E-assessment can be utilized both for summative and formative assessments, and therefore an effective tool for educational institutions (Mate, K. & Weidenhofer, J., 2022). Formative e-assessments provide ongoing feedback during the learning process, enabling students and educators to make timely adjustments, while summative e-assessments evaluate student learning at the conclusion of an instructional unit (Serutla, L. et al., 2024).

One significant aspect of e-assessment is its distinction from traditional assessment methods, such as pen-and-paper tests. While both approaches aim to evaluate learning outcomes, the implementation differs significantly. In e-assessment, technology facilitates not only the delivery of assessments but also the collection and analysis of data, allowing for innovative question formats and immediate feedback (Heil, J. & Ifenthaler, D., 2023). This shift in methodology can enhance the overall assessment experience for both students and educators (Mate, K. & Weidenhofer, J., 2022).

E-assessment offers several key characteristics that enhance the assessment experience. Firstly, it provides immediate feedback to students, helping them understand their performance in real-time (Heil, J. & Ifenthaler, D., 2023; Mate, K. & Weidenhofer, J., 2022). This prompt feedback loop encourages reflection and allows learners to make necessary adjustments to their study strategies. Additionally, e-assessment facilitates efficient data collection and analysis, providing educators with valuable insights into student performance trends and learning outcomes (Heil, J. & Ifenthaler, D., 2023; Mate, K. & Weidenhofer, J., 2022).

Another important aspect of e-assessment is its scalability (Isaías, P. et al., 2023; Isaías, P. et al., 2024). Educational institutions can administer assessments to large groups of students simultaneously, streamlining logistics and resource management (Heil, J. & Ifenthaler, D., 2023; Ristov, S. et al., 2014; Masagazi, J. et al., 2024). E-assessment platforms often support diverse question formats, including multiple-choice, short-answer, and essay questions, which cater to various learning styles and assessment objectives (Isaías, P. et al., 2023; Isaías, P. et al., 2024).

Moreover, e-assessment enhances accessibility for students with disabilities or those requiring accommodations. Many platforms offer features such as text-to-speech, adjustable font sizes, and alternative input methods, ensuring equitable access to assessment opportunities (Kumar, K. & Owston, R., 2015; Mate, K., & Weidenhofer, J., 2022). Security is another significant advantage of e-assessment. With content stored digitally, the risk of loss is minimal, and access can be controlled through measures such as two-factor authentication (Alnasser, F. & Elrashidi, A., 2023; Ubah, A. et al., 2022). The ability to randomize assessment content further decreases the potential for cheating, as each student can receive different questions or variations of the same test (Miguel, J. et al, 2015; Küppers, B. et al., 2020). E-assessments are usually conducted in a locked-down environment, where all other functionalities on the computer, including internet access, are disabled during the assessment (Gibson, 2020; Küppers, B. et al., 2020; Keykha, A. et al., 2025).

2.4 E-assessment format and examples

E-assessment offers diverse formats ideal for online platforms, providing benefits like automated grading and instant feedback (Heil & Ifenthaler, 2023; Karunarathne & Wijewardene, 2021; Baleni, 2015). These include matching exercises, fill-in-the-gap tasks, and interactive elements like drag-and-drop or simulations (Mate & Weidenhofer, 2022; Heil, J. & Ifenthaler, D., 2023; Crisp et al., 2016). Traditional assessments, such as essays, case studies, and discussion forums, can be adapted online to assess skills like communication and teamwork (Bayne et al., 2022; Gunning & Collins, 2018). Live or recorded presentations further enhance engagement and support varied learning styles, fostering active learning and collaboration (Mate & Weidenhofer, 2022; Hughes et al., 2024; Chen et al., 2022; Hulton & Gapper, 2020).

2.5 Development of Hypothesis

2.5.1 Prior experience in E-assessment

Students with prior experience in E-assessment may affect their perspectives on the E-assessment system and E-examinations. Research indicates that these students place a high value on institutional support (Khan et al., 2021). Their familiarity with the challenges of E-examinations leads them to appreciate guidance even more. Moreover, students have positive perceptions of specific features like webcam proctoring and digital interfaces for answering questions when they become more familiar with E-examinations (Domínguez-Figaredo & Gil-Jaurena, 2024). This suggests that prior experience with e-assessment systems significantly influences students' views on these key features which enhance their comfort and satisfaction with the system's functionalities. Furthermore, students with prior E-assessment experience tend to have a positive overall perception of E-examinations, with a specific appreciation for immediate feedback and efficient processes (Khan et al., 2021). However, while they often feel less stressed and find E-examinations reliable, these positive perceptions can vary across different groups (Afacan Adanır et al., 2020). Thus, prior experience is essential to become an indicator for analysis in understanding student attitudes toward E-assessment. Therefore, I hypothesize:

H1: Students with prior experience in E-assessment have a greater understanding of the importance of diverse guidance before E-examinations.

H2: Students with prior experience in E-assessment hold a more positive perspective on the key features of the E-assessment system.

H3: Students with prior experience in E-assessment have a favorable view of the feedback provided during E-examinations.
H4: Prior experience with E-assessment influences students' perspectives on E-examinations compared to traditional examinations.
H5: Students with prior experience in E-assessment view monitoring techniques during E-examinations positively.

2.5.2 Major Difference in E-assessment
Each academic major has its own specific learning outcomes, which are tailored to the unique requirements and goals of the discipline. As a result, E-assessment methods may not be equally suitable for all majors. According to Saha, D., Das, S., and Acharjee, D. (2023), E-examinations may be particularly challenging for majors that emphasize practical skills and hands-on learning experiences. In these fields, such as engineering, medicine, or the arts, the reliance on E-assessment formats may limit the effectiveness of evaluating students' true competencies. Furthermore, it is likely that the administrative tools used in e-assessments will not be that good in showing students' real practical abilities, which, in the end, can affect the completeness of the assessment. Therefore, it is essential to analyze the differences in majors to determine how they affect Hong Kong students' perspectives on E-assessment. Thus, I hypothesize:
H6: Students from different Majors perceive the Identity Monitoring Techniques in E-Examinations differently.

2.5.3 Study Year Difference in E-assessment
Students from different academic years may have different points of view about E-assessment. According to Khan et al. (2021), early-year students tend to focus on monitoring issues, while later-year students prioritize practicality. Early-year students may resist monitoring due to a lack of familiarity, whereas later-year students embrace it for its efficiency. Similarly, Iahad et al. (2004) suggest that students in later years generally have a more positive perspective on E-examinations feedback, which can also extend to their views on monitoring. Therefore, it is essential to analyze the differences in study years to determine how they affect Hong Kong students' perspectives on E-assessment. Thus, I hypothesize:
H7: Students from different Study Year perceive E-Examinations Feedback differently.
H8: Students from different Study Year perceive Identity Monitoring Techniques in E-Examinations differently.

2.5.4 Diverse Guidance before E-examinations and Key Features of E-assessment system
The study demonstrates a positive relationship between the importance of diverse guidance before E-exams and the key features of E-assessment systems (Mo, Tang, Wu, & Tang, 2022). This relationship exists because guidance reduces uncertainty, develops skills, facilitates adaptation to new technology, and aligns with the digital context. Thus, I hypothesize:
H9: There is a positive relationship between Diverse Guidance and Key Features
2.5.5 Key Features of E-assessment system and E-Examinations Feedback
On the other hand, the positive relationship between Key Features of E-assessment system and E-Examinations Feedback indicates that the more effectively these features function, the more valuable students find the feedback for their learning process (Iahad, Dafoulas, Kalaitzakis, & Macaulay, 2004). Thus, I hypothesize:
H10: There is a positive relationship between Key Features and E-Examinations Feedback

2.5.6 Key Features of E-assessment system and Monitoring Techniques in E-examinations
The features enhance monitoring by improving security, efficiency, and adaptability, as demonstrated by the benefits for administrators and the outcomes of system design. This relationship exists because these features facilitate proactive oversight, reduce instances of cheating, and ensure consistent monitoring during large-scale assessments, aligning with the demands of modern digital education (Saha, Das, & Acharjee, 2023). Thus, I hypothesize:
H11: There is a positive relationship between Key Features and Monitoring Techniques
2.5.7 E-Examinations vs. Traditional Examinations and Monitoring Techniques in E-examinations
The monitoring techniques are significant factors affecting respondents' attitudes toward the comparison of E-examinations and traditional examinations. Effective monitoring enhances security, efficiency, and adaptability in the examination process, ultimately influencing students' views on the reliability and integrity of the assessment (Khan et al., 2021). Thus, I hypothesize:
H12: There is a positive relationship between E-Examinations vs. Traditional Examinations and Monitoring Techniques

3. Research methodology

3.1 Research Design

This study applies a quantitative research design to comprehensively examine the perceptions of Hong Kong higher education students regarding E-assessment systems. The research uses a structured questionnaire to measure the variability of students' attitudes, preferences, and concerns towards e-assessment. The quantitative methodology used is one of the easiest ways to obtain numerical data for the purposes of statistical analysis giving insights into trends and patterns in student perceptions. This approach is particularly effective in capturing a broad variety of opinions across different demographic groups, which strengthens the trustworthiness and accuracy of the results.

3.2 Participants

The target population for this study consists of students enrolled in higher education institutions across Hong Kong, encompassing undergraduate, postgraduate, master's, and doctoral students. Specifically, the survey focuses on non-exchange students from various universities and colleges to ensure a representative sample of the local student body. The inclusion of students from diverse fields of study enriches the data and allows for a comparative analysis of perceptions based on academic discipline and level of study. A convenience sampling method was utilized to recruit participants, leveraging online forums and Canvas discussions to reach a broad audience efficiently. This method not only facilitates easier access to participants but also ensures a timely collection of data.

3.3 Questionnaire Development

The survey comprises 37 questions divided into three categories to capture participants' perspectives on e-assessment systems:

Demographic Information (4 multiple choice questions): Collects data on educational institution, field of study, study year, and e-assessment experience to analyze perception trends across groups.

Perceptions of E-assessment (31 Likert scale questions): Participants rate agreement on e-assessment aspects, including pre-exam guidance, system features, feedback, traditional exam comparisons, identity verification, and comfort, allowing nuanced insights.

Open-ended Questions (2 questions): Enables free expression on standardizing online exams and desired e-assessment features, enriching quantitative data with qualitative insights.

3.4 Data Collection

A pilot test was conducted with 10 students who were my friends from various universities and institutions in Hong Kong before the full distribution of the survey to all students. In this pilot study, the participants were asked to complete a questionnaire in order to determine if they could understand it on their own. They were then asked to provide their feedback on the questionnaire once it had been collected. It was concluded that all of them could understand it. The questionnaire's length and complexity were also moderate.

Data were collected using an online survey platform, which guarantees anonymity and confidentiality for all participants. The questionnaire was distributed through online forums for higher education students and shared among friends to ensure broad reach and effective engagement. The survey remained open for a specified period, providing ample time for responses. Participants were informed about the study's purpose and methodology, promoting transparency throughout the process. They were reassured that their responses would be kept confidential and utilized solely for research purposes, thereby fostering trust and encouraging honest feedback.

3.5 Data Analysis

Data will be analyzed using SPSS, with descriptive statistics (means, standard deviations, frequencies) summarizing demographic and Likert-scale responses to identify perception trends. Open-ended responses will be qualitatively analyzed for themes, enabling a mixed-methods approach. Factor analysis will validate scales by assessing one-dimensionality, factor loadings (>0.3, Fornell & Larcker, 1981), and corrected item-total correlations (>0.2, Kline, 2015), removing items below thresholds. Scale reliability will be tested using Cronbach's alpha (>0.7, or >0.5 for scales with <10 items; Fornell & Larcker, 1981; Pallant, 2010). Independent t-tests and one-way ANOVA will compare group means (e.g., by gender, experience, study year), and Pearson correlations will explore factor relationships.

3.6 Ethical Considerations

Informed consent was obtained from all participants, who were informed about the study's purpose, question types, and time commitment, ensuring transparency and trust. They were assured of their right to withdraw anytime without consequences. Data were anonymized to protect identities, using secure handling and confidentiality protocols to uphold ethical standards, safeguard privacy, and enhance research credibility.

3.7 Validation of Measurement Scale

Item 22, being nominal data, is excluded from Component Analysis, Factor Analysis, and Reliability Analysis due to its categorical nature, which complicates meaningful correlation calculations.

3.7.1 Factor Analysis

As mentioned, a factor loading of 0.3 is considered acceptable (Fornell & Larcker, 1981), while a corrected item-total correlation of 0.2 is used to verify the items (Kline, 2015). The values of factor loading of each Item are larger than 0.3. However, the value of Corrected Item-Total Correlation of Item27 is less than 0.2. Thus, it should be deleted. After removing Item27, all items' Corrected Item-Total Correlation values are greater than 0.2 and all items' Factor Loading values are larger than 0.3. Thus, the scale is reliable, and all items have large variances for factor analysis.

3.7.2 Component Analysis

Table 1: Component Analysis for each factor

Construct	Components Extracted	Variance Explained (Component 1, Component 2)
Diverse Guidance	1	56.33%
Key Features	2	48.31%, 14.95%
E-exam Feedback	1	50.73%
E-exams vs. Traditional exam	1	40.46%
Monitoring Techniques	2	39.61%, 17.21%

Principal Component Analysis (PCA) validated the survey's measurement scales by identifying underlying components for each construct. One component was extracted for Importance of Diverse Guidance (56.33% variance), E-Examinations Feedback (50.73%), and E-Examinations vs. Traditional Examinations (40.46%). Two components were extracted for Key Features (48.31% and 14.95% variance) and Identity Monitoring Techniques (39.61% and 17.21% variance). Component 1 was prioritized for each construct due to higher variance explained, confirming scale suitability for further analysis

3.7.3 Reliability Analysis

Table 2: Reliability Statistics for each factor

Item	Cronbach's Alpha	N of Items	Notes
Importance of Diverse Guidance	0.604	3	All correlations >0.2
Key Features of E-Assessment System	0.832	8	All correlations >0.2
E-Examinations Feedback	0.826	5 (originally 6)	Item 21 deleted (alpha improved from 0.787)
E-Exams vs. Traditional Exams	0.61	5(originally 6)	All correlations >0.2; Item 27 excluded earlier (Section 3.7.1)
Identity Monitoring Techniques	0.735	6(originally 7)	Item 29 deleted (alpha improved from 0.732)

Reliability analysis confirmed internal consistency using Cronbach's alpha (>0.5 for scales with <10 items; Pallant, 2010), with values ranging from 0.604 to 0.826 (Table X). Items 21 and 29 were removed from Feedback and Monitoring Techniques to improve alpha to 0.826 and 0.735, respectively, ensuring scale reliability for further analyses.

4. Results and Discussion

4.1 Introduction

This section discusses the data analysis results. It begins with demographic statistics of respondents, providing background on university/institutions, major, study year, and prior E-assessment experience for subgroup allocation. Next, descriptive statistics are summarized, followed by independent samples t-tests comparing factors by prior experience, one-way ANOVA assessing differences by university/institutions, major, and study year, and Pearson correlations analyzing relationships between Diverse Guidance and Key Features, Key Features and E-Examinations Feedback, Key Features and Monitoring Techniques, and E-Examinations vs. Traditional Examinations and Monitoring Techniques. The discussion follows, addressing each analysis in order.

4.2 Demographic Information of Respondents

119 questionnaires were collected, and 107 questionnaires were useful. Respondents come from various universities and institutions. There are 28 CityU students with 26.2%, 9 CUHK students with 8.4%, 3 EdUHK students with 2.8%, 8 HKBU students with 7.5%, 3 HKCC students with 2.8%, 1 HKMU student with 0.9%, 1 HKSYU student with 0.9%, 14 HKU students with 13.1%, 1 HKU SPACE student with 0.9%, 4 HKUST students with 3.7%, 1 HSU student with 0.9%, 4 IVE students with 3.7%, 4 Ling U students with 3.7%, 2 MU students with 1.9%, 2 NY students with 1.9%, 20 PolyU students with 18.7%, 1 SFU student with 0.9%, 1 YCCECE student with 0.9%. Among respondents, about 11.2% of respondents are postgraduate or master, 19.6% of respondents are year1, 16.8% of respondents are year2, 16.8% of respondents are year3, 35.5% of respondents are year4 or above. 30.8% of respondents study in Art/Social Science/Education, 16.8% of respondents study in Business, 18.7% of respondents study in Engineering, 9.3% of respondents study in Pharmacy/Medicine, 9.3% of respondents study in Creative Media/Computer Science/Information Systems, 12.1% of respondents study in Science, 2.8% of respondents study in Law. About 87.9% of respondents have experience with E-assessments. Only 12.1% of respondents do not have experience with E-assessments. The background information is demonstrated below.

Table 3: Percentage of Respondents' Studying University/Institution

University/Institutions

	Frequency	Percent	Valid Percent	Cumulative Percent
CityU	28	26.2	26.2	26.2
CUHK	9	8.4	8.4	34.6
EdUHK	3	2.8	2.8	37.4
HKBU	8	7.5	7.5	44.9
HKCC	3	2.8	2.8	47.7
HKMU	1	.9	.9	48.6
HKSYU	1	.9	.9	49.5
HKU	14	13.1	13.1	62.6
HKU SPACE	1	.9	.9	63.6
HKUST	4	3.7	3.7	67.3
HSU	1	.9	.9	68.2
IVE/VTC	4	3.7	3.7	72.0
Ling U	4	3.7	3.7	75.7
MU	2	1.9	1.9	77.6
NY	2	1.9	1.9	79.4
PolyU	20	18.7	18.7	98.1
SFU	1	.9	.9	99.1
YCCECE	1	.9	.9	100.0
Total	107	100.0	100.0	

Table 4: Percentage of Respondents' Studying Major

Major

	Frequency	Percent	Valid Percent	Cumulative Percent
Art/Social Science/Education	33	30.8	30.8	30.8
Business	18	16.8	16.8	47.7
Engineering	20	18.7	18.7	66.4
Pharmacy/Medicine	10	9.3	9.3	75.7

	Creative Media/Computer Science/Information Systems	10	9.3	9.3	85.0
	Science	13	12.1	12.1	97.2
	Law	3	2.8	2.8	100.0
	Total	107	100.0	100.0	

Table 5: Percentage of Respondents' Study Year

Study_Year

		Frequency	Percent	Valid Percent	Cumulative Percent
	Postgraduate/Master	12	11.2	11.2	11.2
	Year1	21	19.6	19.6	30.8
	Year2	18	16.8	16.8	47.7
	Year3	18	16.8	16.8	64.5
	Year4 or above	38	35.5	35.5	100.0
	Total	107	100.0	100.0	

Table 6: Percentage of Respondents' prior experience of E-assessments

Prior experience of E-assessments

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	No	13	12.1	12.1	12.1
	Yes	94	87.9	87.9	100.0
	Total	107	100.0	100.0	

4.3 Descriptive Statistics

The items are 5-point Likert-scale questions (1=Strongly agree to 5=Strongly disagree).

4.3.1 The significance of offering various types of guidance before the E-examinations

The overall mean of Importance of Diverse Guidance Before E-Examinations is around 2.35. It means that the respondents rated this factor positively. From the five-point scale, point 2 denoted "Agree", meaning the respondents tend to agree that giving diverse guidance before the E-examinations is important.

4.3.2 The essential features of an E-assessment system

The overall mean of Key Features of an E-Assessment System is around 1.68. It means that the respondents rated this factor positively. From the five-point scale, point 2 denoted "Agree", meaning the respondents tend to agree that some features mentioned in the questionnaire are important.

4.3.3 The feedback from the E-examinations

The overall mean of E-Examinations Feedback is around 2.21. It means that the respondents rated this factor positively. From the five-point scale, point 2 denoted "Agree", meaning the respondents tend to agree that feedback after examination is important. The mean of item 17 is the lowest and close to 2, which is "I believe that students should receive a detailed explanation of the incorrect answers". This revealed that higher education students prefer having solutions or model answers after examinations.

4.3.4 E-examinations compared to traditional examinations

The overall mean of E-Examinations vs. Traditional Examinations is around 2.87. This means that the respondents rated this factor slightly positively. From the five-point scale, point 3 denoted "Neutral", meaning the respondents slightly agree that the performance on E-examinations. The mean of item 28 is the lowest and close to 2, which is "I believe that students should receive a detailed explanation of the incorrect answers". This revealed that the duration of E-examinations is appropriate, and higher education students can manage E-examinations similarly to traditional examinations.

Item 22 is “How do you rate the comfort level of E-examinations compared to traditional examinations” with five answers “E-exams are far less comfortable than traditional exams”, “E-exams are less comfortable than traditional exams”, “Both are similar in comfort”, “E-exams are more comfortable than traditional exams”, “E-exams are far more comfortable than traditional exams”. The majority of respondents selected option 3, which states that 'Both are similar in comfort.' The mean value for Item 22 is also 3, suggesting that most higher education students do not have a preference for either E-examinations or traditional examinations based on comfort, as they find them to be equally comfortable.

4.3.5 Techniques for monitoring and verifying identity in E-examinations

According to Table 43, the overall mean of Identity Monitoring Techniques in E-Examinations is around 2.73. This means that the respondents rated this factor slightly positively. From the five-point scale, point 3 denoted “Neutral”, meaning the respondents slightly agree on the methods of identity monitoring in E-examinations.

4.3.6 Open-ended questions

The questionnaire includes two open-ended questions aimed at capturing higher education students' perspectives on E-assessment. Item 36 seeks to determine whether students support the idea of E-examinations replacing traditional exams, while Item 37 explores the features that students believe would be beneficial in an E-assessment system.

The responses to Item 36 reveal a noticeable split among students. Out of 107 total replies, 30 students were in favor of E-examinations, appreciating the flexibility, convenience, and cost savings they offer, especially for those who might struggle to attend in-person assessments. On the other hand, 61 students expressed concerns, citing issues like the potential for cheating, technical glitches, and the challenges of effectively monitoring students during online tests. Many also felt that certain subjects require in-person exams due to their hands-on nature. Furthermore, 16 respondents took a neutral stance, suggesting that a combination of online and traditional exams could work well, depending on personal preferences and the specific needs of different courses.

From the Item37 results, respondents shared important ideas for improving future E-assessment systems. Many expressed a desire for more flexible exam timing, allowing students to choose when to take their tests. There was also a call for better communication tools to enable monitored interactions during exams. Improved monitoring features, like screen observation and behavior checks, were seen as essential for fairness. Quick feedback on answers, insights into classmates' responses, and clear explanations would greatly help students. Suggestions include using technology like handwriting and voice recognition, as well as AI for grading longer answers. To prevent cheating, automated checks and screen-locking at the end of exams were considered vital. Meanwhile, support for special learning needs and a question bank for revision were emphasized.

4.4 Independent Sample T-test

The Independent Samples T-test assesses significant differences between two groups, using Levene’s Test for Equality of Variances and the t-test for Equality of Means. Levene’s Test indicates equal variances if the significance level exceeds 0.05 (use equal variance results) or unequal variances if ≤0.05 (use unequal variance results). The t-test shows a significant mean difference if the 2-tailed significance is ≤0.05; otherwise, no difference exists.

4.4.1 Comparison of Importance of Diverse Guidance Before E-Examinations between prior experience

The significance value from Levene’s Test is 0.366, which is greater than 0.05. This suggests that the variances between respondents who have experienced E-assessment and those who have not are equal. As a result, the assumption of equal variances will proceed. The significance value (1-tailed) from the 't-test for Equality of Means' is 0.049, also below 0.05. This means there is statistically significant difference in the perceived importance of diverse guidance before E-examinations based on prior experience. Based on the results, the hypothesis **H1: Students with prior experience in E-assessment have a greater understanding of the importance of diverse guidance before E-examinations** is *supported*.

4.4.2 Comparison of Key Features of an E-Assessment System between prior experience

The significance value from Levene’s Test is 0.597, which is greater than 0.05. This suggests that the variances between respondents who have experienced E-assessment and those who have not are equal. As a result, assumption of equal variances will be proceed. The significance value (1-tailed) from the 't-test for Equality of Means' is 0.261, also above 0.05. This means there is no statistically significant difference in the perceived key features of an E-assessment system based on prior experience. Based on the results, the hypothesis **H2: Students with prior experience in E-assessment hold a more positive perspective on the key features of the E-assessment system** is *rejected*.

4.4.3 Comparison of E-Examinations Feedback between prior experience

The significance value from Levene's Test is 0.646, which is greater than 0.05. This suggests that the variances between respondents who have experienced E-assessment and those who have not are equal. As a result, assumption of equal variances will be proceed. The significance value (1-tailed) from the 't-test for Equality of Means' is 0.369, also above 0.05. This means there is no statistically significant difference in the perceived E-Examinations Feedback based on prior experience. Based on the results, the hypothesis **H3: Students with prior experience in E-assessment have a favorable view of the feedback provided during E-examinations** is *rejected*.

4.4.4 Comparison of E-Examinations vs. Traditional Examinations between prior experience

The significance value from Levene's Test is 0.262, which is greater than 0.05. This suggests that the variances between respondents who have experienced E-assessment and those who have not are equal. As a result, the assumption of equal variances will proceed. The significance value (2-tailed) from the 't-test for Equality of Means' is 0.566, also above 0.05. This means there is no statistically significant difference in the perceived E-Examinations vs. Traditional Examinations based on prior experience. Based on the results, the hypothesis **H4: Prior experience with E-assessment influences students' perspectives on E-examinations compared to traditional examinations** is *rejected*.

4.4.5 Comparison of Identity Monitoring Techniques in E-Examinations between prior experience

The significance value from Levene's Test is 0.614, which is greater than 0.05. This suggests that the variances between respondents who have experienced E-assessment and those who have not are equal. As a result, assumption of equal variances will be proceed. The significance value (1-tailed) from the 't-test for Equality of Means,' it is 0.098, also above 0.05. This means there is no statistically significant difference in the perceived E-Examinations vs. Traditional Examinations based on prior experience. Based on the results, the hypothesis **H5: Students with prior experience in E-assessment view monitoring techniques during E-examinations positively** is *rejected*.

4.5 One-way ANOVA

One-way ANOVA tests for significant mean differences among two or more independent groups. A significance level ≤ 0.05 indicates differences, while >0.05 shows none. Post-hoc tests identify specific group differences, with pairwise significance ≤ 0.05 indicating a difference and >0.05 showing no difference.

4.5.1 Comparison of Students' Perceptions on Identity Monitoring Techniques in E-Examinations between Major

The significance level for ANOVA is 0.375 ($F = 1.088$, $p > 0.05$). This denotes that there are no significant differences in the selection of Identity Monitoring Techniques in E-Examinations among the Majors. Furthermore, Table 58 shows that all significance levels from the post-hoc test exceed 0.05, further confirming that there are no significant differences between Majors regarding the Identity Monitoring Techniques in E-Examinations. Based on the results, the hypothesis **H6: Students from different Majors perceive the Identity Monitoring Techniques in E-Examinations differently** is *rejected*.

4.5.2 Comparison of Students' Perceptions on E-Examinations Feedback between Study Year

Table 60 demonstrates that the significance level for ANOVA is 0.113 ($F = 1.921$, $p > 0.05$). This denotes that there are no significant differences in the selection of E-Examinations Feedback among the Study Year. Furthermore, Table 61 shows that all significance levels from the post-hoc test exceed 0.05, further confirming that there are no significant differences between Study Year regarding the E-Examinations Feedback. Based on the results, the hypothesis **H7: Students from different Study Year perceive E-Examinations Feedback differently** is *rejected*.

4.5.3 Comparison of Students' Perceptions on Identity Monitoring Techniques in E-Examinations between Study Year

Table 63 demonstrates that the significance level for ANOVA is 0.881 ($F = 0.295$, $p > 0.05$). This denotes that there are no significant differences in the selection of Identity Monitoring Techniques in E-Examinations among the Study Year. Furthermore, Table 64 shows that all significance levels from the post-hoc test exceed 0.05, further confirming that there are no significant differences between Study Year regarding the Identity Monitoring Techniques in E-Examinations. Based on the results, the hypothesis **H8: Students from different Study Year perceive Identity Monitoring Techniques in E-Examinations differently** is *rejected*.

4.6 Pearson Correlation

Pearson Correlation measures the direction and strength of the relationship between two variables, with a positive coefficient (r) indicating both increase together and a negative (r) indicating one increases as the other decreases. The coefficient ranges from -1 to +1, with values near ±1 showing a strong relationship and near 0 a weak one. A 2-tailed significance <0.05 indicates a statistically significant correlation; >0.05 suggests none.

4.6.1 Relationship between Diverse Guidance and Key Features

The relationship between Diverse Guidance and Key Features, showing a Pearson correlation coefficient of 0.435 ($p = 0.001$, which is below 0.01). This result points to a positive and significant influence of Diverse Guidance on Key Features.

A coefficient of +0.435 indicates a meaningful connection between these two variables. As noted by Field (2017), a correlation value between 0.3 and 0.5 reflects a medium to large correlation. This suggests that there is a moderate positive correlation between Diverse Guidance and Key Features. Therefore, our hypothesis **H9: There is a positive relationship between Diverse Guidance and Key Features**, is *supported*.

4.6.2 Relationship between Key Features and E-Examinations Feedback

Table 66 highlights the relationship between Key Features and E-Examinations Feedback, showing a Pearson correlation coefficient of 0.390 ($p = 0.001$, which is below 0.01). This result points to a positive and significant influence of Key Features on E-Examinations Feedback.

A coefficient of +0.390 indicates a meaningful connection between these two variables. As noted by Field (2017), a correlation value between 0.3 and 0.5 reflects a medium to large correlation. This suggests that there's a moderate positive correlation between Key Features and E-Examinations Feedback. Therefore, our hypothesis **H10: There is a positive relationship between Key Features and E-Examinations Feedback**, is *supported*.

4.6.3 Relationship between Key Features and Monitoring Techniques

Table 67 highlights the relationship between Key Features and Monitoring Techniques, showing a Pearson correlation coefficient of 0.383 ($p = 0.001$, which is below 0.01). This result points to a positive and significant influence of Key Features on Monitoring Techniques.

A coefficient of +0.383 indicates a meaningful connection between these two variables. As noted by Field (2017), a correlation value between 0.3 and 0.5 reflects a medium to large correlation. This suggests that there is a moderate positive correlation between Key Features and Monitoring Techniques. Therefore, our hypothesis **H11: There is a positive relationship between Key Features and Monitoring Techniques**, is *supported*.

4.6.4 Relationship between E-Examinations vs. Traditional Examinations and Monitoring Techniques

Table 68 highlights the relationship between E-Examinations vs. Traditional Examinations and Monitoring Techniques, showing a Pearson correlation coefficient of 0.323 ($p = 0.001$, which is below 0.01). This result points to a positive and significant influence of E-Examinations vs. Traditional Examinations on Monitoring Techniques. A coefficient of +0.323 indicates a meaningful connection between these two variables. As noted by Field (2017), a correlation value between 0.3 and 0.5 reflects a medium to large correlation. This suggests that there is a moderate positive correlation between E-Examinations vs. Traditional Examinations and Monitoring Techniques. Therefore, our hypothesis **H12: There is a positive relationship between E-Examinations vs. Traditional Examinations and Monitoring Techniques**, is *supported*.

4.7 Discussion

4.7.1 Discussion of Prior Experience Comparison

The t-test p-value in Section 4.4.1 ($p=0.049$) indicates a significant difference between students with and without prior E-examination experience in rating diverse guidance, supporting hypothesis H1. Khan et al. (2021) suggest that experienced students recognize E-examination challenges, preferring guidance, while novices may underestimate difficulties, explaining the preference for varied support among the former.

Section 4.4.2 shows a t-test p-value of 0.261 (>0.05), indicating no significant difference in key features ratings between students with and without prior E-examination experience, rejecting hypothesis H2. Although previous research notes experienced students often view features like webcam proctoring positively (Domínguez-Figaredo & Gil-Jaurena, 2024), this study's broader "key features" scope may explain the lack of difference, suggesting prior experience has little impact on perceptions.

Section 4.4.3 shows a t-test p-value of 0.369 (>0.05), indicating no significant difference in E-examinations feedback ratings between students with and without prior experience, rejecting hypothesis H3. Despite previous

research noting experienced students value immediate feedback (Khan et al., 2021), this study finds no difference, possibly due to standardized feedback formats in Hong Kong, where students practice past exams and receive consistent solutions, and a universal emphasis on immediate feedback.

Section 4.4.4 shows a t-test p-value of 0.566 (>0.05), indicating no significant difference in perceptions of E-examinations versus traditional exams between students with and without prior experience, rejecting hypothesis H4. Despite previous research noting reduced stress with experience (Afacan Adanır et al., 2020), similar exam structures in Hong Kong and fading novelty may explain the uniform acceptance across groups.

Section 4.4.5 shows a t-test p-value of 0.098 (>0.05), indicating no significant difference in monitoring techniques ratings between students with and without prior E-examination experience, rejecting hypothesis H5. Despite previous research noting positive views on webcam proctoring (Domínguez-Figaredo & Gil-Jaurena, 2024), this study finds uniform perceptions, likely due to increased familiarity post-COVID-19, emphasizing fairness and reliability over individual experience.

4.7.2 Discussion of Major Comparison

Section 4.5.1 shows a one-way ANOVA significance value of 0.375 ($F = 1.088$, $p > 0.05$), indicating no significant difference in perceptions of identity monitoring techniques in E-examinations across majors, rejecting hypothesis H6. Although previous research notes that E-assessment often fails to evaluate practical skills (Saha, D., Das, S., and Acharjee, D., 2023), this study finds uniform perceptions, likely due to standardized monitoring techniques in Hong Kong's paper-based E-examinations, where practical skills are assessed separately via assignments, leading to similar student attitudes across fields.

4.7.3 Discussion of Study Year Comparison

Section 4.5.2 shows a one-way ANOVA significance value of 0.113 ($F = 1.921$, $p > 0.05$), indicating no significant difference in E-examination feedback perceptions across study years, rejecting hypothesis H7. Although previous research suggests later-year students view feedback more positively (Iahad et al., 2004), this study finds uniform perceptions, likely due to consistent exposure to E-examinations and improved feedback quality, aligning student views across academic levels.

Section 4.5.3 shows a one-way ANOVA significance value of 0.881 ($F = 0.295$, $p > 0.05$), indicating no significant difference in perceptions of monitoring techniques across study years, rejecting hypothesis H8. Although previous researches suggest later-year students view monitoring more positively (Iahad et al., 2004; Khan et al., 2021), this study finds uniform perceptions, likely due to standardized techniques post-COVID-19 and shared trust in their fairness, aligning views across academic levels.

4.7.4 Discussion of Relationship between Diverse Guidance and Key Features

Section 4.6.1 shows a Pearson correlation coefficient of 0.435 ($p < 0.001$), indicating a moderate positive correlation between Diverse Guidance and Key Features, supporting hypothesis H9. This aligns with Mo et al. (2022), who emphasize varied guidance's role in preparing students for E-examinations by reducing anxiety and enhancing familiarity with digital interfaces. Catering to diverse learning styles and providing clear instructions on system features (e.g., time management, feedback tools) boosts engagement and performance. Inadequate guidance may lead to underutilized resources, lowering scores, while effective support empowers students, underscoring its importance in E-examination success.

4.7.5 Discussion of Relationship between Key Features and E-Examinations Feedback

Section 4.6.2 reveals a Pearson correlation coefficient of 0.390 ($p < 0.001$), indicating a moderate positive correlation between Key Features and E-Examinations Feedback, supporting hypothesis H10. This aligns with Iahad et al. (2004), showing that user-friendly E-assessment systems enhance feedback accessibility, improving communication between educators and students. Khan et al. (2021) note that intuitive features increase interaction, making feedback more valued, while Domínguez-Figaredo and Gil-Jaurena (2024) highlight that streamlined design helps students focus on feedback interpretation, boosting engagement and learning outcomes by addressing strengths and weaknesses effectively.

4.7.6 Discussion of Relationship between Key Features and Monitoring Techniques

Section 4.6.3 shows a Pearson correlation coefficient of 0.383 ($p < 0.001$), indicating a moderate positive correlation between Key Features and Monitoring Techniques, supporting hypothesis H11. This aligns with Saha et al. (2023), emphasizing that user-friendly platforms enhance monitoring efficiency, reducing implementation effort for teachers and students (Domínguez-Figaredo & Gil-Jaurena, 2024; Khan et al., 2021). Minimized

technical issues improve dependability and accessibility, making the system a reliable tool for tracking student progress (Domínguez-Figaredo & Gil-Jaurena, 2024).

4.7.7 Discussion of Relationship between E-Examinations vs. Traditional Examinations and Monitoring Techniques

Section 4.6.4 shows a Pearson correlation coefficient of 0.323 ($p < 0.001$), indicating a moderate positive correlation between E-Examinations vs. Traditional Examinations and Monitoring Techniques, supporting hypothesis H12. Aligning with Khan et al. (2021), the remote nature of E-Examinations requires robust Monitoring Techniques, including authentication processes, to ensure academic integrity. Advanced tools like webcam monitoring, which do not rely on in-person invigilators (Domínguez-Figaredo & Gil-Jaurena, 2024), enhance efficiency and fairness while addressing higher cheating risks in E-Examinations (Khan et al., 2021). These innovations boost trust, emphasizing the critical role of Monitoring Techniques in E-Examinations compared to Traditional Examinations.

5. Conclusion

The questionnaire results were in line with the hypotheses H1, H9 to H12, but did not accept the hypotheses H2 to H8. The table below displays the outcome of the hypotheses.

Table 7: Results of the Hypothesizes

Hypothesis	**Results**
H1: Students with prior experience in E-assessment have a greater understanding of the importance of diverse guidance before E-examinations	*supported*
H2: Students with prior experience in E-assessment hold a more positive perspective on the key features of the E-assessment system	*rejected*
H3: Students with prior experience in E-assessment have a favorable view of the feedback provided during E-examinations	*rejected*
H4: Prior experience with E-assessment influences students' perspectives on E-examinations compared to traditional examinations	*rejected*
H5: Students with prior experience in E-assessment view monitoring techniques during E-examinations positively	*rejected*
H6: Students from different Majors perceive the Identity Monitoring Techniques in E-Examinations differently	*rejected*
H7: Students from different Study Year perceive E-Examinations Feedback differently	*rejected*
H8: Students from different Study Year perceive Identity Monitoring Techniques in E-Examinations differently	*rejected*
H9: There is a positive relationship between Diverse Guidance and Key Features	*supported*
H10: There is a positive relationship between Key Features and E-Examinations Feedback	*supported*
H11: There is a positive relationship between Key Features and Monitoring Techniques	*supported*
H12: There is a positive relationship between E-Examinations vs. Traditional Examinations and Monitoring Techniques	*supported*

5.1 Summary

This study examines Hong Kong higher education students' perceptions of e-assessment, focusing on factors influencing their preference for e-examinations versus traditional paper-based formats. A survey of 107 students assessed pre-exam guidance (e.g., tutorials, practice tests), system usability, instant feedback, and identity monitoring (e.g., webcam surveillance), comparing e-examinations to traditional assessments.

The study found that prior experience only had an impact on importance of diverse guidance before. Demographic background, such as field of study or study year, do not affect students' views in Identity Monitoring Techniques in E-Examinations and E-Examinations Feedback. There is a moderate positive correlation between Diverse Guidance and Key Features, Key Features and E-Examinations Feedback, Key Features and Monitoring Techniques, E-Examinations vs. Traditional Examinations and Monitoring Techniques. These insights guide optimizing e-assessment systems, enhancing their integration into higher education and supporting innovation.

5.2 Limitation of study

This study's findings have limitations affecting their generalizability and reliability. The sample size of 107 responses may be insufficient to detect significant effects or represent the broader population. Uneven distribution, with only CityU, HKU, and PolyU contributing over ten participants each, raises concerns about applicability. Limited demographic diversity further restricts the findings' relevance across different groups.

5.3 Recommendations for Future Research

This study on e-examination performance and monitoring techniques sets the stage for deeper e-assessment research. Future studies could analyze specific factors in greater detail. While this study examined students' perceptions based on year, major, and prior e-assessment experience, additional variables like gender, age, and GPA should be explored to assess their impact on perceptions. Furthermore, this study on students' perceptions of e-assessment in higher education could extend to secondary schools or international contexts to better understand its impact across educational settings. Additionally, given respondents' interest in AI integration, future research should explore AI's role in enhancing e-assessment effectiveness, accessibility, and personalization.

Reference

Afacan Adanır, G., İsmailova, 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.

Alnasser, F., & Elrashidi, A. (2023). Improving The Security of E-Exam Systems. *2023 International Conference on IT Innovation and Knowledge Discovery (ITIKD)*, 1-7. https://doi.org/10.1109/ITIKD56332.2023.10100104.

Amangeldina, G., & Dudovich, D. (2022). The impact of pedagogical assessment on increasing motivation to learn and the quality of education. *Focus on Language Education and Research*. https://doi.org/10.35213/2686-7516-2022-3-1-11-23.

Baleni, Z. (2015). Online Formative Assessment in Higher Education: Its Pros and Cons.. *Electronic Journal of e-Learning*, 13, 228-236.

Bayne, L., Birt, J., Hancock, P., Schonfeldt, N., & Agrawal, P. (2022). Best practices for group assessment tasks. *Journal of Accounting Education*. https://doi.org/10.1016/j.jaccedu.2022.100770.

Bulut, O., Gorgun, G., & Yildirim-Erbasli, S. (2024). The impact of frequency and stakes of formative assessment on student achievement in higher education: A learning analytics study. *Journal of Computer Assisted Learning*. https://doi.org/10.1111/jcal.13087.

Cañadas, L. (2021). Contribution of formative assessment for developing teaching competences in teacher education. *European Journal of Teacher Education*, 46, 516 - 532. https://doi.org/10.1080/02619768.2021.1950684.

Chen, J., Lai, P., Chan, A., Man, V., & Chan, C. (2022). AI-Assisted Enhancement of Student Presentation Skills: Challenges and Opportunities. *Sustainability*. https://doi.org/10.3390/su15010196.

Constantinou, P. (2017). Instructional Assessment Strategies for Health and Physical Education. *Strategies*, 30, 3 - 9. https://doi.org/10.1080/08924562.2017.1297747.

Crisp, G., Guàrdia, L., & Hillier, M. (2016). Using e-Assessment to enhance student learning and evidence learning outcomes. *International Journal of Educational Technology in Higher Education*, 13. https://doi.org/10.1186/s41239-016-0020-3.

Din, A., Zabidin, M., Tahawi, A., Din, O., & Aawi, F. (2023). Educational Assessment: Its Types and Outcomes: A Bibliometric Analytical Study. *International Journal of Academic Research in Business and Social Sciences*. https://doi.org/10.6007/ijarbss/v13-i11/19516.

Domínguez-Figaredo, D., & Gil-Jaurena, I. (2024). Effects of familiarity with digital assessment in online education. *Distance Education*, 1-16.

Duterte, J. (2024). INNOVATIONS IN DISTANCE EDUCATION PRACTICES: A COMPREHENSIVE REVIEW. EPRA International Journal of Multidisciplinary Research (IJMR). https://doi.org/10.36713/epra18210.

Ejjami, R. (2024). The Future of Learning: AI-Based Curriculum Development. International Journal For Multidisciplinary Research. https://doi.org/10.36948/ijfmr.2024.v06i04.24441.

Fan, T., Song, J., & Guan, Z. (2021). Integrating diagnostic assessment into curriculum: a theoretical framework and teaching practices. *Language Testing in Asia*, 11, 1-23. https://doi.org/10.1186/s40468-020-00117-y.

Field, A. (2017). *Discovering statistics using IBM SPSS statistics* (PDF). Retrieved from http://repo.darmajaya.ac.id/5678/1/Discovering%20Statistics%20Using%20IBM%20SPSS%20Statistics%20%28%20PDFDrive%20%29.pdf

Fornell, C. and Larcker, D (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. Journal of marketing research 15, pp.282-388.

Gibson, M. (2020). GuideNo17: Introduction to e-assessment. WATT WORKS quick guides. Heriot-Watt University. Retrieved from https://lta.hw.ac.uk/wp-content/uploads/GuideNo17_Introduction-to-e-assessment.pdf

Goertz, M., & Duffy, M. (2001). Assessment and Accountability Systems in the 50 States, 1999-2000. CPRE Research Report Series.. . https://doi.org/10.1037/e383742004-001.

Graue, M. (1993). Integrating Theory and Practice Through Instructional Assessment. *Educational Assessment*, 1, 283-309. https://doi.org/10.1207/S15326977EA0104_1.

Gunning, T., & Collins, P. (2018). Fostering teamwork skills across the School of Engineering using online self and peer assessment. *ASCILITE Publications*. https://doi.org/10.14742/apubs.2018.1962.
Harris, L., Adie, L., & Wyatt-Smith, C. (2022). Learning Progression–Based Assessments: A Systematic Review of Student and Teacher Uses. *Review of Educational Research*, 92, 996 - 1040. https://doi.org/10.3102/00346543221081552.
Heil, J., & Ifenthaler, D. (2023). Online Assessment in Higher Education: A Systematic Review. *Online Learning*, *27*(1), 187-218.
Hernández, R. (2012). Does continuous assessment in higher education support student learning?. *Higher Education*, 64, 489-502. https://doi.org/10.1007/S10734-012-9506-7.
Hughes, H., Padgett, R., & Donald, W. (2024). Preparing students for hybrid working: The place and case for authentic assessment via synchronous online presentations. *Industry and Higher Education*. https://doi.org/10.1177/09504222241266163.
Hulton, A., & Gapper, K. (2020). An Innovative Presentation Tool as an Alternative to Traditional Methods for Student Assessments. , 273-291. https://doi.org/10.1007/978-3-030-35396-4_17.
Iahad, N., Dafoulas, G. A., Kalaitzakis, E., & Macaulay, L. A. (2004, January). Evaluation of online assessment: The role of feedback in learner-centered e-learning. In *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the* (pp. 10-pp). IEEE.
Ifenthaler, D., Schumacher, C., & Kuzilek, J. (2022). Investigating students' use of self-assessments in higher education using learning analytics. *J. Comput. Assist. Learn.*, 39, 255-268. https://doi.org/10.1111/jcal.12744.
Institute of Education Sciences. (2023). *Educational assessment: A factsheet for Maine* (REL Northeast & Islands). U.S. Department of Education. https://ies.ed.gov/ncee/rel/regions/northeast/pdf/RELNEI_FactsheetME.pdf
Isaías, P., Miranda, P., & Pífano, S. (2023). E-Assessment Systems: An Evaluation Framework from the Perspective of Higher Education Experts. *2023 International Symposium on Computers in Education (SIIE)*, 1-6. https://doi.org/10.1109/SIIE59826.2023.10423677.
Isaías, P., Miranda, P., & Pífano, S. (2024). Framework for the analysis and comparison of e-assessment systems. *ASCILITE Publications*. https://doi.org/10.14742/apubs.2017.786.
Jordan, S. (2013). E-assessment: Past, present and future. *New Directions*, *9*(1), 87-106.
Karunarathne, G., & Wijewardene, L. (2021). Online Assessments: Their Importance in Higher Education. *Advances in Social Sciences Research Journal*. https://doi.org/10.14738/assrj.812.11448.
Keykha, A., Imanipour, M., Shahrokhi, J., & Amiri, M. (2025). The Advantages and Challenges of Electronic Exams: A Qualitative Research based on Shannon Entropy Technique. *Journal of Advances in Medical Education & Professionalism*, 13, 1 - 11. https://doi.org/10.30476/jamp.2024.102951.1987.
Khan, M. A., Vivek, V., Khojah, M., Nabi, M. K., Paul, M., & Minhaj, S. M. (2021). Learners' Perspective towards E-Exams during COVID-19 Outbreak: Evidence from Higher Educational Institutions of India and Saudi Arabia. *International Journal of Environmental Research and Public Health*, *18*(12), 6534. https://doi.org/10.3390/ijerph18126534
Kibble, J. (2017). Best practices in summative assessment.. *Advances in physiology education*, 41 1, 110-119 . https://doi.org/10.1152/advan.00116.2016.
Kiryakova, G. (2021). E-assessment-beyond the traditional assessment in digital environment. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1031, No. 1, p. 012063). IOP Publishing.
Kline, P. (2015). *A handbook of test construction (psychology revivals): introduction to psychometric design*. Routledge.
Kumar, K., & Owston, R. (2015). Evaluating e-learning accessibility by automated and student-centered methods. *Educational Technology Research and Development*, 64, 263 - 283. https://doi.org/10.1007/s11423-015-9413-6.
Kundu, A., & Bej, T. (2021). Experiencing e-assessment during COVID-19: an analysis of Indian students' perception. *Higher Education Evaluation and Development*, *15*(2), 114-134.
Küppers, B., Eifert, T., Zameitat, R., & Schroeder, U. (2020). EA and BYOD: Threat Model and Comparison to Paper-based Examinations. , 495-502. https://doi.org/10.5220/0009578004950502.
Kusumoriny, L. (2024). Enhancing Learning Through Assessment and Evaluation : Strategies for Student Progress. .
Lee, H., Chung, H., Zhang, Y., Abedi, J., & Warschauer, M. (2020). The Effectiveness and Features of Formative Assessment in US K-12 Education: A Systematic Review. *Applied Measurement in Education*, 33, 124 - 140. https://doi.org/10.1080/08957347.2020.1732383.
Magdalena, I., Andreani, M., Nurhasanah, S., & Ushaybiah, Z. (2023). DAMPAK PENILAIAN UNTUK PEMBELAJARAN TERHADAP MOTIVASI DAN KETERLIBATAN SISWA. *JURNAL RISET PENDIDIKAN DAN PENGAJARAN*. https://doi.org/10.55047/jrpp.v2i1.450.

Masagazi, J., Lugemwa, P., & Mirembe, E. (2024). Pioneering Online Assessment Solutions: Empirical Experiences from Educational Practitioners. *International Journal on Integrating Technology in Education*. https://doi.org/10.5121/ijite.2024.13201.

Mate K, Weidenhofer J. (2022) Considerations and strategies for effective online assessment with a focus on the biomedical sciences. *FASEB BioAdvances*. 4(1), 9-21. https://doi.org/10.1096/fba.2021-00075

Miguel, J., Caballé, S., Xhafa, F., & Prieto, J. (2015). Security in online web learning assessment. *World Wide Web*, 18, 1655-1676. https://doi.org/10.1007/s11280-014-0320-2.

Mo, D. Y., Tang, Y. M., Wu, E. Y., & Tang, V. (2022). Theoretical model of investigating determinants for a successful Electronic Assessment System (EAS) in higher education. Education and Information Technologies, 27(9), 12543-12566.

Muhammad, I., Tomalá, M., Coello, C., Silva, S., & Del Rocío Cerón Silva, C. (2024). Assessment methods and their impact on learning outcomes in education. *LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades*. https://doi.org/10.56712/latam.v5i5.2865.

Omoeva, C., Cunha, N., & Moussa, W. (2021). Measuring equity of education resource allocation: An output-based approach. *International Journal of Educational Development*. https://doi.org/10.1016/j.ijedudev.2021.102492.

Osiesi, M. (2020). Educational Evaluation: Functions, Essence and Applications in Primary Schools' Teaching and Learning. *Society & Sustainability*. https://doi.org/10.38157/society_sustainability.v2i2.134.

Owan, V., Abang, K., Idika, D., Etta, E., & Bassey, B. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *Eurasia Journal of Mathematics, Science and Technology Education*. https://doi.org/10.29333/ejmste/13428.

Pallant, J. (2010). *SPSS survival manual: A step by step guide to data analysis using SPSS*. Open University Press/McGraw-Hill.

Perie, M., Marion, S., & Gong, B. (2009). Moving toward a Comprehensive Assessment System: A Framework for Considering Interim Assessments.. *Educational Measurement: Issues and Practice*, 28, 5-13. https://doi.org/10.1111/J.1745-3992.2009.00149.X.

Pham, A. T. (2022). University Students' Attitudes towards the Application of Quizizz in Learning English as a Foreign Language. *International Journal of Emerging Technologies in Learning, 17(19).*

Ristov, S., Gusev, M., Armenski, G., & Velkoski, G. (2014). Scalable and elastic e-Assessment cloud solution. *2014 IEEE Global Engineering Education Conference (EDUCON)*, 762-769. https://doi.org/10.1109/EDUCON.2014.6826180.

Rostaminezhad, M. A. (2019). STUDENTS' PERCEPTIONS OF THE STRENGTHS AND LIMITATIONS OF ELECTRONIC TESTS FOCUSING ON INSTANT FEEDBACK. *Journal of Information Technology Education, 18.*

Sadler, D. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144. https://doi.org/10.1007/BF00117714.

Saha, D., Das, S., & Acharjee, D. (2023). Exploring the Efficacy and Implementation of an Online Examination System: A Comprehensive Study on eXamPro in Modern Education.

Serutla, L., Mwanza, A., & Celik, T. (2024). Online assessments in a changing education landscape. In *Reimagining Education-The Role of E-Learning, Creativity, and Technology in the Post-Pandemic Era*. IntechOpen.

Shin, N., Stevens, S., & Krajcik, J. (2010). Designing assessments to track student progress. , 580-587.

Sievertsen, H. H. (2022). Assessments in Education. *arXiv preprint arXiv:2208.05826*.

Sortwell, A., Trimble, K., Ferraz, R., Geelan, D., Hine, G., Ramirez-Campillo, R., Carter-Thuiller, B., Gkintoni, E., & Xuan, Q. (2024). A Systematic Review of Meta-Analyses on the Impact of Formative Assessment on K-12 Students' Learning: Toward Sustainable Quality Education. *Sustainability*. https://doi.org/10.3390/su16177826.

Sri, S., Jannah, F., Widyanti, E., Tinggi, S., & Sangatta, A. (2024). Assessment of Knowledge Competency Achievement. *Journal of Scientific Research, Education, and Technology (JSRET)*. https://doi.org/10.58526/jsret.v3i2.429.

Turdieva, R. (2024). TECHNIQUES FOR ASSESSING THE PERFORMANCE OF ESL STUDENTS. *Modern approaches and new trends in teaching foreign languages*. https://doi.org/10.52773/tsuull.conf.teach.foreign.lang.2024.8.5/pyfl8084.

Ubah, A., Onakpojeruo, E., Ajamu, J., Mangai, T., Isa, A., Ayansina, N., & Al-Turjman, F. (2022). Biometrics Authentication Techniques in E-Learning Assessment. *2022 International Conference on Artificial Intelligence of Things and Crowdsensing (AIoTCs)*, 126-132. https://doi.org/10.1109/AIoTCs58181.2022.00105.

Wang, C. (2021). On Interim Cognitive Diagnostic Computerized Adaptive Testing in Learning Context. *Applied Psychological Measurement*, 45, 235 - 252. https://doi.org/10.1177/0146621621990755.

Zhou, M. (2023). Significance of Assessment in Learning: The Role of Educational Assessment Tools. Science Insights Education Frontiers. https://doi.org/10.15354/sief.23.co215.