education sciences

MDPI

Article

University Teachers' Views on the Adoption and Integration of Generative AI Tools for Student Assessment in Higher Education

Zuheir N. Khlaif [1,*], Abedalkarim Ayyoub [2], Bilal Hamamra [3], Elias Bensalem [3,4], Mohamed A. A. Mitwally [5], Ahmad Ayyoub [6], Muayad K. Hattab [7] and Fadi Shadid [5,7]

1 Faculty of Humanities and Educational Sciences, Educational Sciences Department, An Najah National University, Nablus P.O. Box 7, Palestine
2 Faculty of Humanities and Educational Sciences, Psychology and Counseling Department, An Najah National University, Nablus P.O. Box 7, Palestine; ayyoub@najah.edu
3 Faculty of Humanities and Educational Sciences, Department of English, An Najah National University, Nablus P.O. Box 7, Palestine; bilalhamamra@najah.edu (B.H.); ebsalem@gmail.com (E.B.)
4 Department of Languages and Translation, Northern Border University, Arar 91431, Saudi Arabia
5 Open Distance Learning, University of South Africa, Pretoria 0003, South Africa; mmitwa@unisa.ac.za (M.A.A.M.); fshadid@najah.edu (F.S.)
6 Faculty of Arts, English Department, Bethlehem University, Bethlehem P1520468, Palestine; aayyoub@bethlehem.edu
7 Faculty of Law and Political Science, Department of Law, An Najah National University, Nablus P.O. Box 7, Palestine; mhattab@najah.edu
* Correspondence: zkhlaif@najah.edu

Abstract: This study examines the factors that may impact the adoption of generative artificial intelligence (Gen AI) tools for students' assessment in tertiary education from the perspective of early-adopter instructors in the Middle East. It utilized a self-administered online survey and the Unified Theory of Acceptance and Use of Technology (UTAUT) model to collect data from 358 faculty members from different countries in the Middle East. The Smart PLS software 4 was used to analyze the data. The findings of this study revealed that educators developed new strategies to integrate Gen AI into assessment and used a systematic approach to develop assignments. Moreover, the study demonstrated the importance of developing institutional policies for the integration of Gen AI in education, as a driver factor influencing the use of Gen AI in assessments. Additionally, the research identified significant factors, namely performance expectancy, effort expectancy, social influences, and hedonic motivation, shaping educators' behavioral intentions and actual use of Gen AI tools to assess students' performance. The findings reveal both the potential advantages of Gen AI, namely enhanced student engagement and reduced instructor workloads, and challenges, including concerns over academic integrity and the possible negative impact on students' writing and thinking skills. This study emphasizes the significance of targeted professional development and ethical criteria for the proper integration of Gen AI in educational assessment.

Keywords: assessment; AI literacy; AI in education; higher education

check for updates

Citation: Khlaif, Z.N.; Ayyoub, A.; Hamamra, B.; Bensalem, E.; Mitwally, M.A.A.; Ayyoub, A.; Hattab, M.K.; Shadid, F. University Teachers' Views on the Adoption and Integration of Generative AI Tools for Student Assessment in Higher Education. *Educ. Sci.* **2024**, *14*, 1090. https://doi.org/10.3390/educsci14101090

Academic Editor: Xinqiao Liu

Received: 25 July 2024
Revised: 30 September 2024
Accepted: 4 October 2024
Published: 6 October 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Educators' interest in generative artificial intelligence (Gen AI) has grown significantly [1], particularly since OpenAI released ChatGPT-3 in 2022. This is primarily due to these tools' capacity to generate writing that resembles that produced by humans, to the extent that texts generated by AI are challenging for even specialists to recognize [2]. Furthermore, students have rapidly adopted Gen AI tools in their writing (Moorhouse et al., 2023).

Opponents of the integration of Gen AI in education have expressed their concerns regarding the validity of assessments, particularly with respect to academic integrity and plagiarism [3,4]. Students' use of AI tools to complete their assignments [5] could compromise academic integrity and eventually undermine the reputations of academic

institutions. Furthermore, academics contend that if students rely too much on GAI, it might hinder the development of their writing and critical thinking abilities [4,6] and negatively impact the quality of instruction and learning outcomes for students [7].

Nevertheless, some academics argue that Gen AI can have positive effects on schooling [8,9]. For example, [10] points out that Gen AI can boost students' participation in assignments, provide prompt feedback, facilitate collaboration, and make learning more accessible. One major advantage is the capacity to offer prompt and insightful feedback through automated marking. Similarly, [11] found that Gen AI tools could assist educators in identifying and meeting their students' needs, as well as in automatic essay scoring and instantaneous student feedback, thereby reducing instructors' workloads. Zhai and Nehm (2023) argue that the discussion about Gen AI should focus on how to use these tools, rather than whether to use them. This is because many teachers are already using AI to aid formative assessment in various educational settings [12,13].

To persuade instructors and professors who doubt the validity of Gen AI in student assessment in higher education, it is essential to examine the factors that may influence their decisions. Therefore, the purpose of this study is to explore the factors that influence higher education instructors' use of Gen AI in assessing students from their perspectives.

2. Research Problems

The rapid advancement of Gen AI technologies has sparked considerable interest among educators regarding their potential applications in tertiary education, especially in student assessment. However, the adoption of Gen AI tools in educational assessment is fraught with challenges and controversies. Critics worry that the use of artificial intelligence (AI) in assessment may jeopardize academic integrity, leading to increased instances of plagiarism and reducing the reliability of assessment outcomes [3,5]. Furthermore, there are concerns that the reliance on Gen AI could impair students' writing and critical thinking skills and degrade the quality of instruction and learning outcomes [6,7].

Despite these concerns, proponents highlight the benefits of Gen AI, such as enhanced student engagement, timely feedback, and reduced instructor workloads, which can make learning more accessible and personalized [10,11]. Given these opposing viewpoints, it is vital to explore the factors affecting the acceptability and usage of Gen AI tools in higher education from the perspective of educators. This study seeks to fill this gap by investigating the constructs that influence the acceptance of Gen AI for student evaluation, providing insights that might assist in defining policies and approaches for the effective implementation of AI in schools and universities.

2.1. Research Purpose

The goal of this study is to investigate the constructs that may influence the use of Gen AI in assessing students in university settings, as seen through the viewpoints of early-adopter educators in higher education institutions. Moreover, this study provides a structural relationship between these constructs using the PLS-SEM analysis of the retrieved data from the participants.

2.2. Contribution of the Study

This study makes a significant contribution to the field of educational technology by empirically examining the factors influencing the adoption of Gen AI tools for students' assessment in higher education. Utilizing the Unified Theory of Acceptance and Use of Technology (UTAUT) model, this research validates the critical roles of performance expectancy, effort expectancy, social influences, and hedonic motivation in shaping educators' behavioral intentions (BI) and actual usage of Gen AI. By highlighting the moderating effect of professors' experience, this study provides nuanced insights into how familiarity with technology influences its adoption.

Moreover, this research addresses both the benefits and challenges of integrating Gen AI, offering practical implications for educators, institutions, and policymakers to

enhance the effective and ethical use of AI in educational assessment. These findings lay the groundwork for future studies to explore long-term impacts, diverse contexts, and the development of robust frameworks for the ethical use of Gen AI in education.

3. Literature Review

Creating exams and assessments for students is an integral part of a teacher's job. It is often a dull, time-consuming, and tiresome process and requires follow-up work that involves correcting these tests and monitoring and checking students' grades. With the development of AI techniques, it has become possible to rely on generative multiple-choice question tools to complete this tedious work. Recent studies show that AI is transforming how students are assessed, making the process more personalized, efficient, and effective [14–19].

3.1. Assessment in Tertiary Education

In tertiary education, assessment and feedback serve as pivotal elements of education, influencing a diverse array of stakeholders, including students, faculty, and administrative personnel. Assessment and feedback not only augment students' grading proficiency, motivational engagement, and academic performance but also facilitate enhanced learning trajectories [20–23]. Many studies highlight the transformative impact of assessment on the pedagogical outcomes of students within tertiary educational settings. Assessment and feedback have moved beyond a student-centered focus to encompass elements such as curricular frameworks, pedagogical methodologies, and administrative mechanisms. These assessments are critical to elevating students' academic outcomes and are integral to their subsequent academic and professional achievements [23].

Diagnostic, formative, summative, and electronic assessments are among the assessment types used in higher education, with self-evaluation and peer assessment being common as well. The appropriate assessment type is determined by the course's specific educational objectives and expected learning outcomes. Formative assessments, or "assessments for learning", are implemented throughout the educational process, whereas summative assessments, or "assessments of learning", are typically administered at the culmination of the educational activities [24,25].

Traditionally, evaluations in higher education have generally focused on the retention and application of knowledge within narrowly defined contexts, typically measured through conventional examinations and academic tasks like term paper composition. These assessment methodologies are fundamental to instructional strategies that furnish feedback, thereby allowing both students and educators to refine ongoing teaching and learning processes to better achieve the intended instructional outcomes. Moreover, educators are increasingly adopting cognitive analytics as a foundation for robust e-learning evaluations, thereby enabling learners to substantially improve their academic performance [26].

Assessment involves collecting and analyzing students' exams and work according to a certain criterion and rubric so as to form judgments. Assessment or evaluation is an essential component of education, closely linked to learning, teaching, and the curriculum, and plays a critical role in generating successful learning outcomes and increasing student satisfaction. In many Arab nations, where students care about passing their tests, which are centered on memorization and rote learning, assessment is the most essential aspect of their study [27]. Because assessment is so important in education, many studies have looked into the characteristics of assessments, including both summative and formative assessments, and how they affect students' learning performance, motivation, and learning quality [28,29]. Moreover, research on peer and self-assessment has highlighted that these methods enhance students' skills, their participation, and eventually their academic excellence.

Summative evaluation analyzes students at the end of a course by summarizing both their learning and teaching experiences. This assessment method provides critical feedback on students' achievements during the course, using grades, projects, term papers, and standardized tests [30]. The advantages of summative assessment include (1) assisting

instructors in minimizing errors; (2) enhancing their ability to correct mistakes; (3) offering dependable data such as grades and mid-term marks for accountability purposes among various stakeholders, including learners, teachers, and administrators, in higher education; and (4) aiding in the development of educational strategies such as curricula or funding departments [31]. Traditionally, higher education institutions have relied on summative assessments to gauge students' learning outcomes [32].

The integration of generative AI tools into higher education assessment practices has sparked discussions on the adaptation of traditional assessment strategies. Educators and students have different attitudes towards the use of AI in assessments, with educators favoring tailored assessments that promote critical thinking and questioning. In contrast, students have expressed concerns about the potential lack of creativity [33]. Studies have shown that AI-generated content, such as that from GPT-4, poses challenging situations in detecting academic misconduct, highlighting the need for improved AI detection tools and enhanced awareness amongst faculty and students [34]. To address these challenges, there is a call for adjustments in assessment strategies to enhance the resistance against AI tools, the incorporation of AI-inclusive assessments, and the implementation of comprehensive training programs for both faculty and students [35].

Assessment is a vital method in education, providing valuable statistical information for students, parents, educators, and policymakers to track progress, regulate efforts, and make knowledgeable decisions [36–38]. It is crucial in making choices based on data, demonstrating worth, and highlighting the impact on student achievement [39]. Nevertheless, challenges such as inadequate data accuracy and contested data might impede the assessment process. Assessment provides information to students, parents, educators, and schools on the progress of learning, which helps to guide study efforts and course choices [38]. Assessment in college and university settings promotes learning and generates formal records of achievement by utilizing grading frameworks such as norm-referenced and criterion-referenced frameworks [40,41]. The development of assessment systems, legislation, and policies in education, as well as the incorporation of international assessment systems, highlights the significance of assessment in enhancing the quality and outcomes of education. Effective assessment design is essential in impacting students' future lives and professional careers [40].

AI tools offer significant advantages in educational assessment. They help students to monitor their learning progress and prepare for tests, while also enabling teachers to efficiently create, correct, and give feedback on assessments. There is a concern that AI may undermine traditional forms of evaluation, like essay writing, because AI tools can generate high-quality written content that performs well in standardized settings [4]. To mitigate these concerns, students could be tasked with correcting or improving texts generated by AI, which not only tests their understanding but also engages them in critical thinking. These methods guarantee that AI instruments are employed to supplement, and not to replace, the teaching process. It is also critical to understand the limitations of AI, which include the potential for "hallucinations" and erroneous information to be produced. Instructors must create exams that not only make good use of AI but also push students to think critically and use their knowledge in novel ways. This is especially important in fields like language and literature courses, where the content is subjective and creative.

3.2. Evolution of AI in Education

Artificial intelligence (AI) has significantly changed education since it was first developed in 1956 [42]. It offers tailored solutions to enhance student learning, while improper usage might have detrimental effects [43]. While many contend that artificial intelligence (AI) has the potential to replace human educators, others maintain that human educators are irreplaceable due to their unique qualities, such as creativity and critical thinking [44]. Therefore, it is essential to integrate AI effectively, without viewing it as a substitute for the provision of a comprehensive learning experience [45]. Moturu and Nethi [46] argued that AI applications are increasingly essential in education, aiming to personalize learning expe-

riences, make them effective, focus on output, promote integration, and foster long-term retention. The integration of AI in education is directly linked to advancements in educational methodologies, underscoring the importance of AI-based e-learning applications in revolutionizing the educational environment.

Significant progress has been made in student-facing systems, personalized adaptive learning, automatic grading, and teacher feedback thanks to AI in education over the last 30 years. The integration of AI with cutting-edge technologies like the Internet of Things and immersive technology is essential to the future of AI in education [47]. Undergraduate nursing education currently incorporates artificial intelligence simulations [48]. By using this technology, students have the opportunity to engage with a virtual patient that is constantly changing as they advance through the curriculum, beginning with essential skills.

3.3. Using AI in Student's Assessment

Many academics have noted that the COVID-19 pandemic has greatly increased the acceptance of asynchronous and remote learning approaches [49]. Such approaches are at the forefront of the ongoing digital revolution in education [50]. Artificial intelligence has resulted in the creation of innovative tools with the capability to revolutionize various aspects. A notable example is GPT-3, which has demonstrated its ability to create educational materials, debug software, and compose detailed texts that rival the quality of human writing [51].

The definition of AI varies depending on the context, but it generally refers to machines' ability to execute tasks that would normally need human intelligence. Some scholars define it as the intelligence demonstrated by computers and programs to mirror human mental capacities and work habits, such as the ability to learn, infer, and react to unprogrammed conditions. Furthermore, it is the name of an academic discipline that studies how to develop programs and computers that are capable of exhibiting intelligent behavior.

Education is not an exception to the growing presence of AI in many spheres of society. The use of AI in education started in 1960, when computers were brought into classrooms to handle administrative duties including scheduling, record-keeping, and evaluation [52]. It has been improved over many years to help teachers to assess students and determine their strengths and weaknesses more effectively. With the tremendous development brought about by the Internet, educational chatbots that support artificial intelligence have recently emerged. They can simulate learners and answer their queries. This resulted in the development of individualized learning algorithms that are tailored to the needs of each individual learner.

Artificial intelligence contributed to minimizing the administrative burdens on the educational system. It helped in creating tables and records and therefore improved the process of distributing tasks. This can develop the curriculum and provide academic content for learners in accordance with their needs and capabilities. It also prepares tests and corrects questions, as in most universities that adopt distance learning, where the learner answers random questions from a bank of questions covering all aspects of the subject, whether the questions are multiple-choice or essay questions. In addition, learners can also have access to recorded lectures of their courses at any time and with the least effort. This also comes with instant translations that the learner can use in any language to avoid difficulties in receiving and understanding the information. All of these advantages have contributed to creating a more advanced educational system where the learner does not feel overwhelmed and tired, leading to higher levels of productivity. In general, AI could transform education into a more effective and interesting experience in the future.

The transition from traditional, in-class education to an online mode of learning has sparked concerns regarding student evaluation, which remains a core aspect of e-learning. Previously, in-class evaluations were conducted using conventional methods such as paper-based exams, which required extensive effort and time from educators to prepare, administer, correct, and archive. These traditional methods often fell short in achieving educational objectives, particularly those involving technology. In contrast,

the modern digital era, with its advancements in Gen AI technologies, has enhanced the educational process for both teachers and students. While AI tools have significantly aided in meeting diverse educational needs, it is crucial to view them as supplements rather than replacements.

3.4. The Potential of Gen AI in Assessment

Assessing knowledge stands as a crucial component of the educational journey, offering quantifiable insights into the knowledge acquired by learners. It highlights areas needing improvement, thereby facilitating positive reinforcement and motivating students to excel further. Technological advancements have revolutionized not just dynamic teaching and learning methodologies but also the conventional approach to examinations. Traditional pen-and-paper tests are increasingly giving way to automated online assessments, which are more inclusive, accessible, and precise. Furthermore, these automated examinations encapsulate the refined best practices in evaluation and testing that have evolved over time.

3.5. The Factors Influencing the Usage of Gen AI in Students' Assessment

Scholars have pinpointed assessment as a key domain where AI and related technological advancements hold substantial promise for educational innovation [3,53]. Nonetheless, the broad adoption of AI is not without its challenges, introducing both practical and ethical issues that must be addressed [54].

Students' perceptions considerably influence the usage of Gen AI in assessment in higher education. Many students have a positive opinion regarding Gen AI, recognizing its potential for individualized learning support, as a writing aid, and its research capabilities [33]. Furthermore, the intention to use Gen AI correlates positively with its perceived value and negatively with its perceived cost. Concerns regarding accuracy, privacy, ethical issues, personal growth, and societal values also influence students' desire to use Gen AI tools [55,56]. Educators and policymakers must evaluate these elements in order to appropriately customize Gen AI technologies, address students' concerns, and promote responsible assessment integration, thereby improving teaching and learning experiences in tertiary education [57].

3.6. The Proposed Model of the Study

Based on the discussion in the previous sections and the framework of this study, namely the UTAUT2, the researchers developed the proposed model (Figure 1). The figure below represents a conceptual research model examining the factors influencing behavioral intention (BI) and its effect on use behavior (UB). The model includes constructs such as hedonic motivation (HM), habit (HT), performance expectancy (PE), effort expectancy (EE), and social influence (SI), each hypothesized to affect BI through paths H1a to H1e. Additionally, experience (EXP) is posited to have direct influences on BI and moderate other paths, as indicated by the blue arrows. The relationships between BI and UB are defined by paths H2a, H2b, and H2c. SI also has a direct relationship with EE through H3. This model seeks to investigate how different motivational, behavioral, and experiential factors drive behavioral intentions and the actual usage of a system or technology.

3.7. The Context of the Study

The current study was carried out in a number of Middle Eastern higher education institutions. We recruited participants who were early adopters of Gen AI in teaching. The snowball technique was used to recruit the potential participants for this study based on predefined conditions, including the use of different Gen AI applications for teaching, the use of these tools frequently in his/her courses, and the use of these tools in assessing students' work/assignments. Therefore, the present investigation focuses on the following research questions.

RQ1: How do teachers in higher education institutions use Gen AI to assess their students' performance?

RQ2: What are the factors that drive instructors to use Gen AI in assessing students' performance in higher education institutions from teachers' perspectives?

RQ3: What is the relationship among these drivers of the use of Gen AI in assessing students' performance?

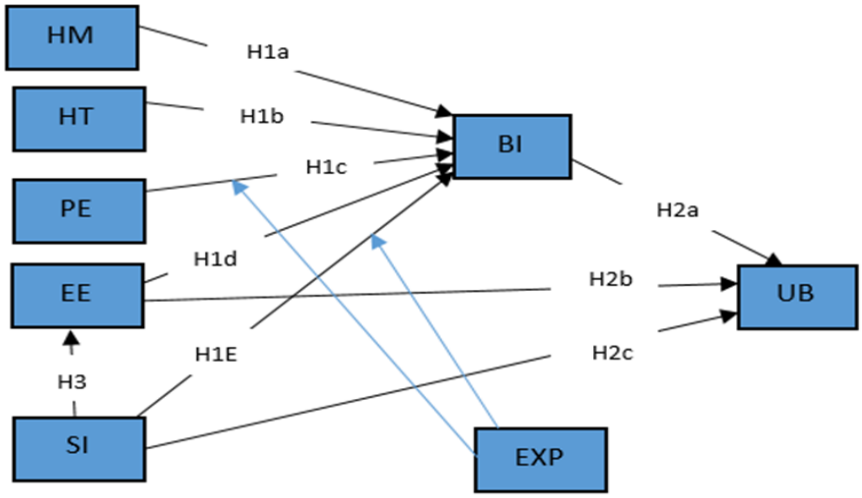

Figure 1. Proposed model where the arrows present the relationship between the constructs of the model.

4. Methodology

A mixed-methods design was used in this study, using a survey composed of closed questions and open-ended questions. Open-ended questions enable the exploration of participants' perspectives, experiences, and reasoning in greater depth, providing context and insights into the quantitative findings. The combination of closed and open-ended questions facilitates a more comprehensive understanding of the research problem, allowing for triangulation, enhancing the validity of the results, and offering a more nuanced perspective on the phenomena being studied [58].

4.1. Participants

The participants of this study were faculty members who were teaching in higher education institutions in in the Middle East. We devised a set of criteria for the recruitment of the participants to ensure that our study included a broad and relevant group of individuals from various universities, which would enrich the insights and increase the application of the study's findings. The participants in the current study had to be instructors in a higher education institution, have experience and familiarity with Gen AI tools, have used AI tools in teaching their courses, be early adopters of Gen AI technology, come from various fields and universities, and have attended professional development workshops on the integration of Gen AI applications in teaching settings. Based on the proposed criteria, the participants in the current study consisted of 358 instructors (47.2% female and 52.8% male). Participants from the social sciences amounted to 34% of the total participants, those from the natural and engineering sciences amounted to 41%, and those from the medical sciences amounted to 25%.

4.2. Open-Ended Questions: Data Analysis

We considered the participants' responses to the open-ended questions as qualitative data, which we utilized to address the first and second research questions. To address the first question, we conducted an inductive thematic analysis, following the six-step methodology proposed by [58]. Prior to the analysis, we cleaned and organized the

responses. The qualitative data analysis during this phase was guided by the six phases of thematic analysis outlined by [58].

4.3. Translation Process

Because the survey items and the participants' responses to the open-ended questions were in Arabic, the researchers employed a conceptual equivalence translation method combined with a back-translation approach to ensure the accuracy and reliability of the translated themes and survey items. Initially, the survey items, quotations, and themes were translated from Arabic to English by an initial translator. The researchers then meticulously reviewed these translations to identify and correct any errors or technical inaccuracies. This review process was crucial in addressing potential misunderstandings related to technical concepts, ensuring that the translated content faithfully represented the original material.

Following this initial translation, the researchers implemented the back-translation technique. The English versions of the quotations, themes, and survey items were given to a second translator who was not privy to the original Arabic content. This second translator then translated the English text back into Arabic. The purpose of this step was to evaluate the consistency and conceptual equivalence between the original Arabic text and the back-translated version.

Upon receiving the back-translated Arabic version, the researchers conducted a detailed comparison with the original data and survey items. This comparison focused on ensuring that the meaning and concepts conveyed in the translations aligned with the original content. The results showed a high degree of agreement, with a 93% match between the translated and original versions. This high level of similarity indicated that the translations were conceptually equivalent and accurately reflected the intended meaning of the survey items.

Due to time constraints and the substantial agreement observed between the translations, the researchers decided not to convene a panel of translation experts for further validation. The rigorous review and back-translation processes were deemed sufficient to ensure the quality and accuracy of the translations, allowing the researchers to proceed with confidence in the reliability of the translated survey items.

4.4. Research Instrument

The research instrument was composed of closed-ended questions based on a Likert scale that included five points on assessments and open-ended questions to allow the participants to write about their perspectives and experiences in more detail. The researchers developed the survey (see Appendix A) based on the findings of previous studies and the UTAUT model, with slight changes to reflect the context of this study (see Table 1).

Table 1. Study measurement scale items.

Construct	Item	Items	Source
Performance expectancy (PE)	PE1	"I believe that Gen AI is useful in assessing students' assignments"	[59]
	PE2	"Using Gen AI increases my chances to evaluate my students' assignments professionally"	
	PE3	"Using Gen AI helps students get tasks and projects done faster"	
	PE4	"Using Gen AI increases students' productivity in their assignments"	
Effort expectancy (EE)	EE1	"Learning how to use Gen AI in assessment is easy for me"	[60]
	EE2	"My interaction with Gen AI is clear and understandable"	
	EE3	"I find Gen AI easy to design rubrics for assessing my students' projects"	
	EE4	"It is easy for me to become skillful at using Gen AI in assessment"	
Social influence (SI)	SI1	"People who are important to me think I should Gen AI in assessment"	[61]
	SI2	"People who influence my behavior believe that I should use Gen AI"	
	SI3	"People whose opinions I value prefer me to use Gen AI for students' assessment"	

Table 1. *Cont.*

Construct	Item	Items	Source
Facilitating conditions (FC)	FC1	"I have the resources necessary to use Gen AI in assessing my students"	[59]
	FC2	"I have the knowledge necessary to use Gen AI for students' assessment"	
	FC3	"Gen AI is compatible with the technologies I use in teaching"	
	FC4	"I can get help from others when I have difficulties using Gen AI"	
Hedonic motivation (HM)	HM1	"Using Gen AI for assessment is fun"	[59]
	HM2	"Using Gen AI for assessment is enjoyable"	
	HM3	"Using Gen AI for assessment is very entertaining"	
Price value (PV)	PV1	"Gen AI is reasonably priced"	[62]
	PV2	"Gen AI is good value for the money"	
	PV3	"At the current price, Gen AI provides good value"	
Habit	HT1	"The use of Gen AI for students' assessment has become a habit for me"	[59]
	HT2	"I am addicted to using Gen AI in teaching and assessment"	
	HT3	"I must use Gen AI for students' assessment"	
	HT4	"Using Gen AI for assessment has become natural for me"	
Behavioral intention (BI)	BI1	"I intend to continue using Gen AI for assessment in the future"	[59]
	BI2	"I will always try to use Gen AI in my teaching and assessment"	
	BI3	"I plan to continue to use Gen AI for assessment frequently"	
Personal innovativeness (PI)	PI1	"I like experimenting with new information technologies"	[63,64]
	PI2	"If I heard about a new information technology, I would look for ways to experiment with it"	
	PI3	"Among my family/friends, I am usually the first to try out new information technologies"	
	PI4	"In general, I do not hesitate to try out new information technologies"	
Use behavior (UB)	UB1	"Please choose your usage frequency for Gen AI: 1. Never; 2. Once a month; 3. Several times a month; 4. Once a week; 5. Several times a week; 6. Once a day; 7. Several times a day"	[59]

5. Results

5.1. RQ1: How Do Faculty Members in Higher Education Institutions Use Gen AI to Assess Students' Performance?

In their responses, the participants focused on two main themes. The first one reflected the procedures used to obtain high-quality assignments and assessments through their dialogue with Gen AI, and the second theme focused on strategies for the use of Gen AI in assessing students' performance. Further descriptions of these findings are given below.

5.1.1. Procedures for Use of Gen AI to Generate Assignments and Ways to Assess Students

Based on the analysis of the participants' responses to the open-ended questions, various procedures were identified to obtain high-quality outputs from generative AI (Gen AI) applications in generating assignments and assessing students' performance. The participants described a range of approaches, which were not uniform across the board. Each participant utilized a different set of steps to achieve their goals, indicating the diversity in the strategies and techniques employed.

The researchers organized and synthesized these varied approaches into a cohesive flowchart (see Figure 2). This flowchart serves as a guide, outlining the essential steps that educators can follow to optimize the use of Gen AI applications in their teaching practices. The process begins with a deep understanding of the context in which the AI will be used,

followed by the careful crafting of prompts to generate relevant and high-quality responses. The participants emphasized the importance of revising and iterating these prompts to continually improve the quality of the generated outputs.

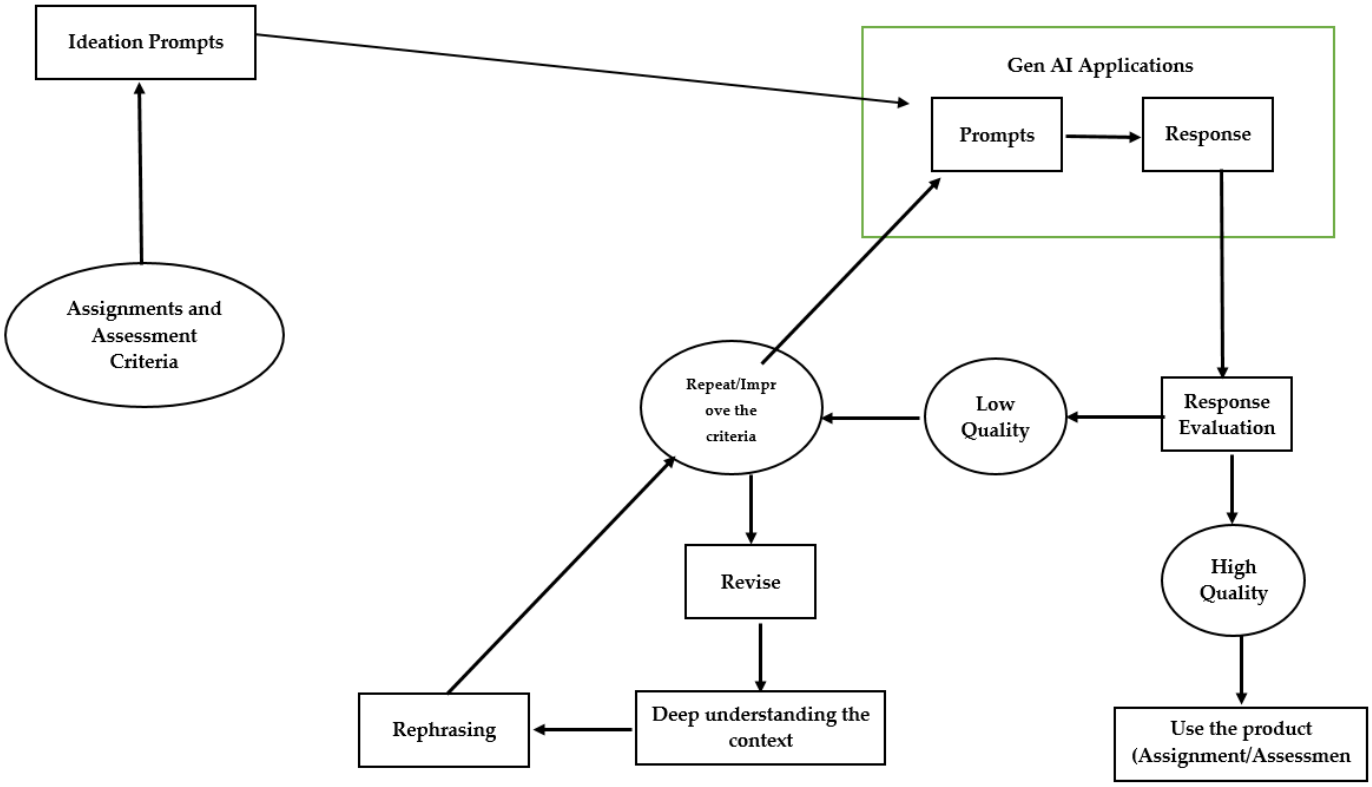

Figure 2. Flowchart of procedures to develop assignments using Gen AI as reported by the participants. The arrows represent the steps to be taken while developing the assignments using Gen AI.

Moreover, the evaluation of responses was a critical step, involving assessing the quality and relevance of the AI-generated content. This process includes determining the effectiveness of the AI in meeting the educational objectives and adjusting the prompts or approaches as necessary. The flowchart highlights the iterative nature of this process, encouraging educators to continually refine their methods.

Finally, the participants' approaches culminated in the practical application of the generated content, whether for assignments or assessments. The use of Gen AI applications in this context requires not only technical proficiency but also a nuanced understanding of how to integrate these tools into pedagogical practices effectively. The resulting flowchart provides a valuable resource for educators and researchers, offering a structured framework to harness the potential of Gen AI in educational settings.

The participants in this study identified various strategies for the integration of Gen AI into their assessment practices. Their responses varied based on their teaching fields, experience with technology, and institutional policies on the use of Gen AI. The strategies that they shared reflected their understanding of Gen AI's capabilities and how to incorporate them into their assessments to promote fair, authentic, and inclusive practices.

5.1.2. Categorizing Assignments in the Gen AI Era

The researchers categorized the responses of the participants in terms of the type of assignment into three themes, including assignments without Gen AI, Gen AI-assisted assignments, and Gen AI-empowered assignments. These types of assignments emerged based on the development of Gen AI in teaching and learning, as reported by most of the participants.

5.1.3. Assignments without Gen AI

In this type of assessment, as described by the responses of the participants, students were not allowed to use any type of Gen AI application. Some of the assessment tools are in-person exams, oral exams, practical applications, and final presentations. A professor in nursing stated that, "In the course description, I informed my students where they can use Gen AI and where they cannot use it. So already they know in exams and oral presentations they are not allowed to use it".

5.1.4. Gen AI-Assisted Assignments

The second type of assignment involves the use of Gen AI as a form of assistance for students. Students can use it to generate ideas for their assignments and teamwork projects and can use it for brainstorming but must write the final text themselves. This type of assignment is intended to develop their skills and critical thinking, especially while critiquing the ideas generated by the Gen AI application. A professor in computer engineering wrote that, "allowing students to use ChatGPT or any generated application enhances the learning ethics of AI and its literacy. My students were not allowed to copy and paste the ideas; I requested that they critique and develop the materials, and also include the generated text as an appendix in the final project".

In addition, some faculty members reported using these tools to generate assignments for their students. An assistant professor in social sciences emphasized that, "I used generative AI to create project-based assignments, saving time by providing the tool with criteria and the final outcome".

5.1.5. Gen AI-Empowered Assignments

In this type of assignment, as reported by the participants in their responses to the open-ended questions, students were allowed to create all aspects in their assignments, focusing on critical thinking and problem solving. All students had to write a reflection paper about their project and their experience with Gen AI in completing the assignment. An IT professor wrote, "students are allowed to fully integrate Gen AI into their assignments, but they are also required to write a reflection paper on their experience using it". Another professor in the humanities wrote, "using Gen AI to empower students in the assignment will make them lazy and depend on it in the future. I do not like to use it in this way; my idea is that educators use it only for assessment, teaching, and scientific research".

5.1.6. Utilizing Gen AI in Assessing Students' Performance

The majority of the participants in this study confirmed in their responses to the open-ended questions that they used different Gen AI applications to generate various types of exam questions, such as false/true questions, multiple-choice questions, and short essay questions. For example, a professor in biology stated, "I used ChatGPT-4 to generate multiple-choice questions for my courses and shared the presentations online, making slight changes to the generated questions. Students were not allowed to use any Gen AI during the exam".

Some participants mentioned challenges such as accuracy and redundancy in AI-generated questions. "When I use Gemini to generate multiple-choice and true/false questions, I notice high similarity in questions. These tools are not professional enough to generate exams", wrote an assistant professor in IT.

5.1.7. Utilizing Gen AI for Assignment and Rubric Design

Many participants used Gen AI tools to design assignments and create grading rubrics, aiming to ensure fairness and accurately assess students' knowledge and skills. A professor in computer engineering noted, "I used these tools to generate ideas for my assignments to ensure fairness and make sure my students complete the assignment without Gen AI assistance".

5.1.8. Grading Handwritten Assignments

Some participants reported using ChatGPT-4 to grade handwritten assignments. They defined criteria for the assessment of students' work and reviewed the AI-generated feedback to ensure its suitability. "I used ChatGPT-4 to grade students' work and provide feedback, usually reviewing and modifying the feedback", noted a faculty member in educational sciences. Another mentioned that tools like "GradeScope combined with ChatGPT can save time and effort, especially during the busy end-of-semester grading period".

5.1.9. Generating Varied Assignments

The participants highlighted the power of Gen AI to create diverse assignments, including images, text, and videos. "I used it to generate assignments with audio, video, and images. You just need to know how to use it to generate assignments based on your teaching style", explained an assistant professor in electronics.

5.2. RQ2: What Are the Factors That Drive Instructors to Use Gen AI in Assessing Students' Performance in Higher Education Institutions from Teachers' Perspectives?

The researchers utilized both qualitative and quantitative data to address the second question in this study, which aimed to explore the factors influencing educators' adoption of Gen AI to assess students' academic performance. The analysis of the open-ended responses revealed several factors affecting educators' use of Gen AI in assessment, including institutional policies, Gen AI application features, pricing considerations, and integration with learning management systems (LMS). Furthermore, the researchers identified the main survey themes as additional factors that could influence the adoption of Gen AI in student assessment.

5.2.1. PLS-SEM Analysis

Partial least squares structural equation modeling was used to investigate the complex relationships among the variables via the Smart PLS software.

5.2.2. Model Estimation

We assessed the specified constructs by analyzing the indicator loadings (Table 2). An indicator loading greater than 0.7 indicates an acceptable level of item reliability.

Validity and construct reliability are established as part of the measurement model of assessment. Composite reliability (CR) and Cronbach's alpha were utilized to establish construct reliability. Table 3 displays the sample's construct reliability and convergent validity. For every construct, the Cronbach alpha values and CR exceeded the suggested threshold of 0.7. Convergent validity is supported by the constructs' average variance extracted (AVE), which was more than 0.5 [65]. The goodness of fit (GoF) for the overall hypothesized model was 0.69, confirming that the model met the global criterion of 0.3 as proposed by [65].

To assess the discriminant validity of PLS-SEM, we used the criterion and HTMT proposed by [66], which states that the square root of the average variance that is extracted by a construct must be larger than its correlation with any other construct. The heterotrait–monotrait (HTMT) ratio of correlations technique from [65] was used. It is recommended to use the HTMT threshold of 0.90 when the constructs are conceptually comparable. A lower requirement of 0.85 is suggested by [65] for more distinct entities. Because every value in Table 4 is below the 0.85 threshold, the discriminant validity is maintained.

The diagonal and italicized elements are the square roots of the average variance extracted (AVE). Above the diagonal, the elements represent the correlations between the constructs. Below the diagonal are the HTMT values.

The examination of the proposed relationships was the next stage of our investigation. Direct relationships were examined first. Table 5 presents the analysis's results in detail.

Table 2. Item loadings.

Construct	Item	Loading
Behavioral intention	BI1	0.96
	BI2	0.96
	BI3	0.97
Effort expectancy	EE1	0.96
	EE2	0.95
	EE3	0.96
	EE4	0.96
Hedonic motivation	HM1	0.97
	HM2	0.97
	HM3	0.98
Habit	HT1	0.93
	HT2	0.9
	HT3	0.92
	HT4	0.95
Performance expectancy	PE1	0.96
	PE2	0.97
	PE3	0.97
	PE4	0.95
Social influence	SI1	0.95
	SI2	0.93
	SI3	0.94
Use behavior	UB	

Table 3. Reliability and convergent validity.

Construct	Cronbach's α	CR	AVE
Behavioral intention (BI)	0.96	0.96	0.92
Effort expectancy (EE)	0.97	0.97	0.92
Hedonic motivation (HM)	0.97	0.97	0.95
Habit (HT)	0.95	0.96	0.86
Performance expectancy (PE)	0.97	0.97	0.93
Social influence (SI)	0.94	0.94	0.88

Table 4. Discriminant validity using the criterion of [66] and the heterotrait–monotrait method (HTMT).

Construct	BI	EE	HM	HT	PE	SI
BI	***0.96***	0.89	0.89	0.82	0.9	0.85
EE	0.85	***0.96***	0.83	0.82	0.84	0.83
HM	0.86	0.83	***0.97***	0.78	0.88	0.81
HT	0.85	0.85	0.81	***0.93***	0.74	0.79
PE	0.85	0.84	0.84	0.77	***0.96***	0.81
SI	0.84	0.85	0.85	0.83	0.85	***0.94***

Table 5. Direct relationships.

Hypothesis	Path	B	*t*	*p*	Result
H2a	BI -> UB	0.99	30.83	0.00	supported
H1d	EE -> BI	0.19	3.92	0.00	supported
H2b	EE -> UB	−0.14	3.72	0.00	supported
H1a	HM -> BI	0.14	2.49	0.01	supported
H1b	HT -> BI	0.07	2.02	0.04	supported
H1c	PE -> BI	0.3	6.83	0.00	supported
H1e	SI -> BI	0.26	4.73	0.00	supported
H2c	SI -> UB	0.11	3.51	0.00	supported
H3	SI -> EE	0.83	40.7	0.00	supported

Table 6 compares the initial result of the bootstrap-based test for the exact overall model fit (i.e., d_ULS and d_G) with the confidence interval derived from the sampling distribution. The initial value should be included in the confidence interval. Therefore, to show that the model has a "good fit", the upper bound of the confidence interval needs to be greater than the initial value of the d_ULS and d_G fit requirements. The confidence interval is selected so that the 95% or 99% point represents the upper bound. Moreover, satisfactory fitting is indicated by an SRMR value of 0.05, which is less than 0.08.

Table 6. Goodness of fit.

		Original Sample (O)	Sample Mean (M)	95%	99%
SRMR	Saturated model	0.05			
d_ULS	Saturated model	0.3	0.15	0.21	0.44
	Estimated model	0.71	0.43	0.69	0.87
d_G	Saturated model	0.6	0.45	0.57	0.64
	Estimated model	0.72	0.46	0.59	0.76

The results revealed that all hypotheses were supported, as shown in Figure 3, with positive coefficients, with the exception of H2b: EE -> UB effort expectancy → use behaviour ($\beta = -0.14$, t = 3.72, $p = 0.00$) was found to be significant, with a negative coefficient. Next, the model's explanatory power was assessed. The study's endogenous variables have R2 values of 0.52 for effort expectancy, 0.87 for behavioral intention, and 0.89 for use behavior. The R^2 value ranges from substantial to very good. We evaluated the predictive relevance by utilizing the Q-square (Q^2) value. The endogenous constructs' Q^2 values varied between 0.69 and 0.89. According to [67], this study's Q^2 values can be characterized as high.

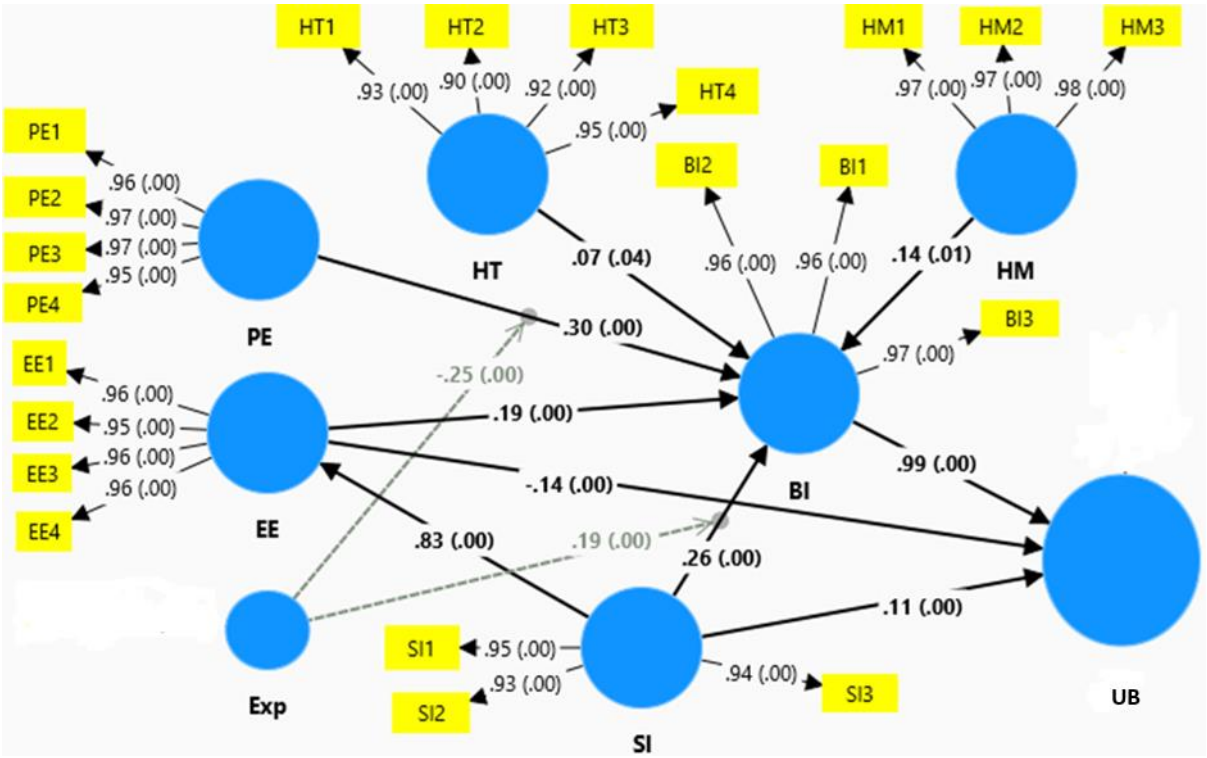

Figure 3. Path result with factor loadings (p) for outer model and (β, p) for inner model where arrows in the figure represents the relationships between the constructs.

RQ3: What is the relationship among these drivers of the use of Gen AI in assessing students' performance?

The researchers answered the third question, which considered the relationships among the factors that guided the adoption and acceptance of Gen AI in assessing students'

performance, using Smart PLS. The findings of the mediated relationships are presented in Table 5. They revealed the significant mediating role of behavioral intention between use behavior and effort expectancy (β = 0.19, t = 3.79, p = 0.00), hedonic motivation (β = 0.14, t = 2.49, p = 0.01), habit (β = 0.07, t = 2.04, p < 0.04), performance expectancy (β = 0.29, t = 6.64, p = 0.00), and social influence (β = 0.25, t = 4.71, p = 0.00). Effort expectancy also significantly mediated the relationship between use behavior and social influence (β = −0.12, t = 3.73, p = 0.00) and between social influence and behavioral intention (β = 0.16, t = 3.89, p = 0.00). Effort expectancy and behavioral intention together significantly mediated the relationship between social influence and use behavior (β = 0.15, t = 3.78, p = 0.00).

There was a significant and moderate relationship between performance expectancy and behavioral intention, with a positive coefficient (β = 0.19, t = 4.30, p = 0.00), indicating that professors' experience strengthens the relation PE -> BI. The R^2 change in behavioral intention was 0.02, and f^2 = 0.153. Additionally, there was a negative and moderate relationship between social influence and behavioral intention (β = -.25, t = 5.17, p = 0.00), suggesting that professors' experience weakens the relation SI -> BI. The R^2 change in behavioral intention was 0.03, and f^2 = 0.31. Professors' experience had a medium effect size (Cohen, 1988), as shown in Tables 7 and 8.

Table 7. Mediation analysis.

Path	B	*T*	*P*
EE -> BI -> UB	0.19	3.79	0.00
HM -> BI -> UB	0.14	2.49	0.01
SI -> EE -> BI	0.16	3.89	0.00
HT -> BI -> UB	0.07	2.04	0.04
SI -> EE -> UB	−0.12	3.73	0.00
PE -> BI -> UB	0.29	6.64	0.00
SI -> BI -> UB	0.25	4.71	0.00
SI -> EE -> BI -> UB	0.15	3.78	0.00

Table 8. Moderation analysis.

Path	β	*t*	*p*	R^2	f^2
Exp ×PE -> BI	0.19	4.30	0.00	0.02	0.153
Exp ×SI -> BI	−0.25	5.17	0.00	0.04	0.304

Effect size (f^2) (≥0.02 is small; ≥0.15 is medium; ≥0.35 is large), where β is the coefficient of the relationship among the constructs; t is the value of the coefficient; p is the probability; and f^2 is the effect size.

6. Discussion

The findings of this study highlight the complex landscape regarding the use of Gen AI in the assessment of students in higher education. The results underscore both the potential benefits and challenges associated with integrating this new technology, providing valuable insights for educators, policymakers, and academic institutions. Moreover, the participants reported different procedures used to improve the quality of assignments and assessment methods. Implementing new assessment methods with varied assignments in the era of AI can enhance the empowerment of learning and assessment procedures through Gen AI.

6.1. Potential Benefits

One of the key advantages of Gen AI in education is its ability to provide immediate and detailed feedback to students, which can enhance learning outcomes and reduce the workload for educators [11]. The high composite reliability (CR) and average variance extracted (AVE) values in the measurement model indicate that the constructs were well measured and reliable, supporting the robustness of the findings. In addition, the study revealed significant positive relationships between several constructs and the use of Gen AI in assessments. Performance expectancy, effort expectancy, hedonic motivation, habits, and social influences all positively influenced the behavioral intention to use Gen AI,

which in turn positively influenced the actual use behavior. These findings align with previous research emphasizing the usefulness and efficiency of Gen AI tools in educational settings [10,68].

6.2. Challenges and Concerns

Despite these benefits, this study also identified significant challenges. Effort expectancy negatively impacted use behavior, suggesting that the ease of use of Gen AI tools might not always translate into actual usage. This outcome emphasizes the importance of thorough training and assistance for educators in effectively integrating these tools into their teaching practices [3]. Additionally, concerns about academic integrity and the potential for AI-generated content to undermine traditional assessment methods were prevalent among participants. This result is consistent with previous research that emphasizes the need for improved AI detection tools and awareness programs to reduce the dangers of plagiarism and academic dishonesty [34,35].

6.3. Moderating Role of Experience

This study also found that professors' experience mediated the relationship between performance expectancy, social influences, and behavioral intention. Experienced educators were more likely to see the benefits of Gen AI tools, suggesting that familiarity and comfort with technology play crucial roles in its adoption. This finding suggests that targeted professional development and continuous learning opportunities are essential to foster positive attitudes towards Gen AI in education [44].

6.4. Implications

This study's findings have many implications for practice and policy. First, robust frameworks for the use of AI in evaluations must be developed in order to address concerns about academic integrity. This entails the formulation of clear guidelines and the education of teachers and students on the ethical application of AI tools [36]. Secondly, higher education institutions should allocate resources to professional development programs aimed at improving educators' digital skills, especially in utilizing Gen AI tools. By fostering a supportive environment, institutions can encourage the effective integration of these technologies, thereby enhancing teaching and learning outcomes [39].

Finally, policymakers should take into account the various factors that impact the adoption of Gen AI in education, such as effort expectancy, performance expectancy, and social influences. By understanding these factors, strategies can be developed to promote the adoption of AI tools, ensuring that they are used to enhance, rather than replace, traditional educational practices [56].

6.4.1. Theoretical Implications

This study adds to the expanding body of literature on the integration of Gen AI in higher education by giving empirical evidence on the factors that influence its use for student evaluations. Using the UTAUT model, this study validates the role of performance expectancy, effort expectancy, social influences, and hedonic motivation in altering instructors' behavioral intentions and actual usage of Gen AI tools. The constructs utilized in this study were very reliable and valid, reaffirming the UTAUT model's robustness in the context of educational technology adoption.

Furthermore, this study contributes to a better understanding of how instructors' experience moderates the relationship between performance expectancy, social influences, and behavioral intention. This finding highlights the importance of considering user experience when investigating technology adoption, suggesting that experienced educators are more likely to perceive the benefits and integrate Gen AI tools into their teaching practices effectively.

6.4.2. Practical Implications for Educators and Institutions

Professional Development and Training: The negative impact of effort expectancy on use behavior underscores the need for comprehensive training programs. Educators should be given continuous professional development opportunities to improve their digital competencies and facilitate the integration of Gen AI tools into their assessment methods. This training should include both the technical aspects of using these technologies and their educational applications.

Addressing Academic Integrity: Concerns about academic dishonesty and plagiarism require the creation of rigorous standards and norms for the ethical application of Gen AI in evaluations. Institutions should implement policies that promote academic integrity and provide tools to detect and mitigate the misuse of AI-generated content. Educators should be trained to recognize AI-generated work and incorporate strategies that encourage original student contributions.

Enhanced Support Systems: Facilitating conditions, such as access to resources and support, are critical to the adoption of Gen AI solutions. Institutions should provide educators with the required technological infrastructure and support services. Peer support networks and collaborative platforms can also help educators to share best practices and troubleshoot common issues.

6.4.3. Practical Implications for Policymakers

Policy Development: Clear guidelines and standards for the use of Gen AI in educational assessments should be devised by policymakers. These regulations should address ethical concerns, data protection, and the equitable application of AI tools in various educational settings. By establishing a regulatory framework, policymakers can help to mitigate the risks associated with Gen AI and promote its responsible use.

Incentives for Adoption: To encourage the adoption of Gen AI tools, policymakers can provide incentives such as grants, funding for technology integration projects, and recognition programs for innovative teaching practices. These incentives can encourage instructors to try out and accept new technology, thereby improving the educational quality.

Research and Development: Continued research into the effects of Gen AI on educational outcomes is critical. Policymakers should encourage research projects that look into the long-term effects of AI integration in education, identify best practices, and create new tools and methodologies. Collaborative efforts between academia, industry, and the government can drive innovation and ensure that AI tools are aligned with educational goals.

6.5. Limitations

Generalizability: While the study included a large number of participants from Middle Eastern higher education institutions, the results may not be applicable to other educational contexts. The cultural, technological, and educational differences across regions may influence participants' adoption and successful integration of Gen AI tools.

Self-Reported Data: This study relied on self-reported data from the participants, which are susceptible to biases such as social desirability and recollection bias. The participants may have inflated or understated their use and perceptions of Gen AI tools, potentially impacting the results' accuracy.

Cross-Sectional Design: This study used a cross-sectional approach, collecting data at a single point in time. This technique limits the ability to infer causality or track changes in perceptions and behaviors over time. Longitudinal studies would be more effective in understanding the evolution of educators' attitudes and the long-term impacts of Gen AI integration.

Focus on Early Adopters: This study focused on early adopters of Gen AI tools, who may have more positive perceptions and experiences compared to the general population of educators. This could lead to the overestimation of the benefits and the underestimation of the challenges associated with Gen AI adoption.

Technological Variability: This study did not account for the variability in the types and functionalities of Gen AI tools used by the participants. Different tools may have varying levels of effectiveness and user-friendliness, influencing educators' perceptions and experiences.

6.6. Future Research

Longitudinal Studies: Future studies should use longitudinal designs to track changes in educators' perceptions, attitudes, and use of Gen AI tools over time. This would provide further insights into the long-term effects of Gen AI on educational assessment and learning outcomes.

Comparative Studies: Conducting comparative studies across different regions, educational levels, and cultural contexts could help to identify universal and context-specific factors influencing Gen AI adoption. This would enhance the generalizability of the findings and inform tailored implementation strategies.

Exploring Diverse Populations: Future studies should include a broader range of individuals, including those who are skeptical about or neutral towards Gen AI techniques. Understanding the barriers and facilitators for different groups can provide a more complete picture of the problems and potential of integrating Gen AI into education.

Impact on Student Learning: While the current study focused on instructors' perceptions, future research should investigate the influence of Gen AI tools on students' learning results, engagement, and satisfaction. Investigating students' opinions and experiences in using AI-driven evaluations can provide useful insights about the instruments' effectiveness.

Technological Advancements: As Gen AI technologies continue to evolve, future research should examine the latest advancements and their implications for educational assessment. Studies should investigate the integration of emerging technologies, such as immersive VR and adaptive learning systems, with Gen AI tools to enhance educational experiences.

Ethical and Policy Considerations: Further research is needed to investigate the ethical implications of utilizing Gen AI in education, including concerns about data privacy and academic integrity and the possibility of AI bias. Developing comprehensive ethical rules and policies will be critical to the responsible use of Gen AI capabilities.

Training and Support: Investigating the effectiveness of different training and support programs for educators using Gen AI tools can provide insights into best practices for professional development. Future research should explore the most effective ways to build educators' digital competencies and confidence in using AI-driven technologies.

By addressing these limitations and exploring the recommended future research topics, scholars can contribute to a more comprehensive understanding of Gen AI integration in educational evaluations, as well as its potential to improve higher education teaching and learning.

7. Conclusions

This study offers a thorough understanding of the factors influencing the use of generative artificial intelligence (Gen AI) in student assessments in higher education. Although there are evident benefits, such as increased efficiency and personalized feedback, challenges related to ease of use and academic integrity need to be addressed. By leveraging these insights, educational institutions and policymakers can develop strategies to effectively integrate Gen AI, thereby enhancing both teaching practices and student learning experiences.

Author Contributions: Conceptualization, Z.N.K. and M.K.H.; methodology, Z.N.K., M.A.A.M.; validation, Z.N.K. and A.A. (Abdelkarem Ayyoub); formal analysis, Z.N.K.; investigation; Z.N.K., B.H.; resources, M.A.A.M., M.K.H., A.A. (Abdelkarem Ayyoub)., Z.N.K., F.S. and E.B.; data curation, A.A. (Ahmed Ayyoub)., A.A. (Abdelkarem Ayyoub), F.S., E.B., Z.N.K. and M.K.H.; writing—original draft preparation, Z.N.K., B.H., M.K.H., F.S., A.A. (Abdelkarem Ayyoub), A.A. (Ahmed Ayyoub), F.S., M.A.A.M. and E.B.; writing—review and editing, E.B., M.K.H. and B.H.; visualization, Z.N.K. and

A.A. (Abdelkarem Ayyoub).; supervision, Z.N.K. project administration, Z.N.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The authors of this study obtained approval from the Institutional Review Board (IRB) committee atAn Najah National University. The IRB approval reference is Intr. Jan.2024/43 (approval date: 14 January 2024).

Informed Consent Statement: Informed consent was obtained from all individuals who participated in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. (The data are not publicly available due to privacy or ethical restrictions. And we informed the participants that their data will not be shared with third party. Only the researchers access the data.

Conflicts of Interest: The authors declare that there is no conflict of interest.

Appendix A

Gender: Male Female
Teaching experience:
Teaching experience with ICT:

Items	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
"I believe that Gen AI is useful in assessing students' assignments"					
"Using Gen AI increases my chances to evaluate my students' assignments professionally"					
"Using Gen AI helps students get tasks and projects done faster"					
"Using Gen AI increases students' productivity in their assignments"					
"Learning how to use Gen AI in assessment is easy for me"					
"My interaction with Gen AI is clear and understandable"					
"I find Gen AI easy to design rubrics for assessing my students' projects"					
"It is easy for me to become skillful at using Gen AI in assessment"					
"People who are important to me think I should Gen AI in assessment"					
"People who influence my behavior believe that I should use Gen AI"					
"People whose opinions I value prefer me to use Gen AI for students' assessment"					
"I have the resources necessary to use Gen AI in assessing my students"					
"I have the knowledge necessary to use Gen AI for students' assessment"					
"Gen AI is compatible with technologies I use in teaching"					
"I can get help from others when I have difficulties using Gen AI"					

Items	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
"Using Gen AI for assessment is fun"					
"Using Gen AI for assessment is enjoyable"					
"Using Gen AI for assessment is very entertaining"					
"Gen AI is reasonably priced"					
"Gen AI is good value for the money"					
"At the current price, Gen AI provides good value"					
"The use of Gen AI for students' assessment has become a habit for me"					
"I am addicted to using Gen AI in teaching and assessment"					
"I must use Gen AI for students' assessment"					
"Using Gen AI for assessment has become natural for me"					
"I intend to continue using Gen AI for assessment in the future"					
"I will always try to use Gen AI in my teaching and assessment"					
"I plan to continue to use Gen AI for assessment frequently"					
"I like experimenting with new information technologies"					
"If I heard about a new information technology, I would look for ways to experiment with it"					
"Among my family/friends, I am usually the first to try out new information technologies"					
"In general, I do not hesitate to try out new information technologies"					
"Please choose your usage frequency for Gen AI: 1. Never; 2. Once a month; 3. Several times a month; 4. Once a week; 5. Several times a week; 6. Once a day; 7. Several times a day"					

Based on your experience, please answer the following open questions. Can you describe your experience with using generative AI tools for students assessment in your courses? How do you use Gen AI tools in your teaching and assessing your students (please write the Gen AI tools you use).

References

1. Stokel-Walker, C. AI bot ChatGPT writes smart essays—Should professors worry? *Nature* **2022**. [CrossRef] [PubMed]
2. Casal, J.E.; Kessler, M. Can linguists distinguish between ChatGPT/AI and human writing?: A study of research ethics and academic publishing. *Res. Methods Appl. Linguist.* **2023**, *2*, 100068. [CrossRef]
3. Swiecki, Z.; Khosravi, H.; Chen, G.; Martinez-Maldonado, R.; Lodge, J.M.; Milligan, S.; Selwyn, N.; Gašević, D. Assessment in the age of artificial intelligence. *Comput. Educ. Artif. Intell.* **2022**, *3*, 100075. [CrossRef]
4. Hamamra, B.; Mayaleh, A.; Khlaif, Z.N. Between tech and text: The use of generative AI in Palestinian universities—A ChatGPT case study. *Cogent Educ.* **2024**, *11*, 2380622. [CrossRef]
5. Chan, C. A comprehensive AI policy education framework for university teaching and learning. *Int. J. Educ. Technol. High. Educ.* **2023**, *20*, 38. [CrossRef]
6. Warschauer, M.; Tseng, W.; Yim, S.; Webster, T.; Jacob, S.; Du, Q.; Tate, T. The Affordances and Contradictions of AI-Generated Text for Second Language Writers. *SSRN Electron. J.* **2023**, *62*. [CrossRef]
7. Chan, C.; Lee, K. The AI generation gap: Are gen Z students more interested in adopting generative AI such as ChatGPT in teaching and learning than their Gen X and millennial generation teachers? *Smart Learn. Environ.* **2023**, *10*, 60. [CrossRef]

8. Chiu, T. The Impact of Generative AI (GenAI) on practices, Policies and Research Direction in education: A Case of ChatGPT and Midjourney. *Interact. Learn. Environ.* **2023**, 1–17. [CrossRef]
9. Kohnke, L.; Moorhouse, B.L.; Zou, D. ChatGPT for Language Teaching and Learning. *RELC J.* **2023**, *54*, 003368822311628. [CrossRef]
10. Mate, K.; Weidenhofer, J. Considerations and strategies for effective online assessment with a focus on the biomedical sciences. *FASEB BioAdv.* **2021**, *4*, 9–21. [CrossRef]
11. Celik, I.; Dindar, M.; Muukkonen, H.; Järvelä, S. The Promises and Challenges of Artificial Intelligence for Teachers: A Systematic Review of Research. *TechTrends* **2022**, *66*, 616–630. [CrossRef]
12. Gerard, L.F.; Linn, M.C. Using Automated Scores of Student Essays to Support Teacher Guidance in Classroom Inquiry. *J. Sci. Teach. Educ.* **2016**, *27*, 111–129. [CrossRef]
13. Lee, H.-S.; Gweon, G.-H.; Lord, T.; Paessel, N.; Pallant, A.; Pryputniewicz, S. Machine Learning-Enabled Automated Feedback: Supporting Students' Revision of Scientific Arguments Based on Data Drawn from Simulation. *J. Sci. Educ. Technol.* **2021**, *30*, 168–192. [CrossRef]
14. Braiki, B.A.; Harous, S.; Zaki, N.; Alnajjar, F. Artificial intelligence in education and assessment methods. *Bull. Electr. Eng. Inform.* **2020**, *9*, 1998–2007. [CrossRef]
15. Chen, L.; Chen, P.; Lin, Z. Artificial Intelligence in Education: A Review. *IEEE Access* **2020**, *8*, 75264–75278. [CrossRef]
16. Gardner, J.; O'Leary, M.; Yuan, L. Artificial intelligence in educational assessment: "Breakthrough? or buncombe and ballyhoo?" *J. Comput. Assist. Learn.* **2021**, *37*, 1207–1216. [CrossRef]
17. González-Calatayud, V.; Prendes-Espinosa, P.; Roig-Vila, R. Artificial Intelligence for Student Assessment: A Systematic Review. *Appl. Sci.* **2021**, *11*, 5467. [CrossRef]
18. Hooda, M.; Rana, C.; Dahiya, O.; Rizwan, A.; Hossain, M.S. Artificial Intelligence for Assessment and Feedback to Enhance Student Success in Higher Education. *Math. Probl. Eng.* **2022**, *2022*, 7690103. [CrossRef]
19. Talan, T.; Kalinkara, Y. The Role of Artificial Intelligence in Higher Education: ChatGPT Assessment for Anatomy Course. *Uluslararası Yönetim Bilişim Sist. Ve Bilgi. Bilim. Derg.* **2023**, *7*, 33–40. [CrossRef]
20. Gamage KA, A.; Dehideniya SC, P.; Xu, Z.; Tang, X. ChatGPT and higher education assessments: More opportunities than concerns? *J. Appl. Learn. Teach.* **2023**, *6*, 358–369. [CrossRef]
21. Heywood, J. *Assessment in Higher Education*; Jessica Kingsley Publishers: London, UK, 2000; Volume 56.
22. Pereira, D.; Flores, M.A.; Niklasson, L. Assessment revisited: A review of research in Assessment and Evaluation in Higher Education. *Assess. Eval. High. Educ.* **2016**, *41*, 1008–1032. [CrossRef]
23. Umar, A.M.A.-T. The Impact of Assessment for Learning on Students' Achievement in English for Specific Purposes A Case Study of Pre-Medical Students at Khartoum University: Sudan. *Engl. Lang. Teach.* **2018**, *11*, 15. [CrossRef]
24. Jacoby, J.C.; Heugh, S.; Bax, C.; Branford-White, C. Enhancing learning through formative assessment. *Innov. Educ. Teach. Int.* **2013**, *51*, 72–83. [CrossRef]
25. Svensäter, G.; Rohlin, M. Assessment model blending formative and summative assessments using the SOLO taxonomy. *Eur. J. Dent. Educ.* **2022**, *27*, 149–157. [CrossRef] [PubMed]
26. Chen, D.W.; Jeng, A.; Sun, S.; Kaptur, B. Use of technology-based assessments: A systematic review covering over 30 countries. *Assess. Educ. Princ. Policy Pract.* **2023**, *30*, 396–428. [CrossRef]
27. Hamamra, B.; Alawi, N.; Daragmeh, A.K. COVID-19 and the decolonisation of education in Palestinian universities. *Educ. Philos. Theory* **2021**, *53*, 1477–1490. [CrossRef]
28. Holmes, N. Engaging with assessment: Increasing student engagement through continuous assessment. *Act. Learn. High. Educ.* **2017**, *19*, 23–34. [CrossRef]
29. Sharma, R.; Jain, A.; Gupta, N.; Garg, S.; Batta, M.; Dhir, S. Impact of self-assessment by students on their learning. *Int. J. Appl. Basic Med. Res.* **2016**, *6*, 226. [CrossRef]
30. Mahshanian, A.; Shoghi, R.; Bahrami, M. Investigating the Differential Effects of Formative and Summative Assessment on EFL Learners' End-of-term Achievement. *J. Lang. Teach. Res.* **2019**, *10*, 1055. [CrossRef]
31. Kıncal, R.Y.; Ozan, C. Effects of Formative Assessment on Prospective Teachers' Achievement, Attitude and Self-Regulation Skills. *Int. J. Progress. Educ.* **2018**, *14*, 77–92. [CrossRef]
32. Fischer, J.; Bearman, M.; Boud, D.; Tai, J. How does assessment drive learning? A focus on students' development of evaluative judgement. *Assess. Eval. High. Educ.* **2023**, *49*, 233–245. [CrossRef]
33. Smolansky, A.; Cram, A.; Raduescu, C.; Sandris Zeivots Huber, E.; Kizilcec, R.F. Educator and Student Perspectives on the Impact of Generative AI on Assessments in Higher Education. *Educ. Stud. Perspect. Impact Gener. AI Assess. High. Educ.* **2023**, *27*, 149–157. [CrossRef]
34. Mills, A.; Bali, M.; Eaton, L. How do we respond to generative AI in education? Open educational practices give us a framework for an ongoing process. *J. Appl. Learn. Teach.* **2023**, *6*, 16–30. [CrossRef]
35. Tenakwah, E.S.; Boadu, G.; Tenakwah, E.J.; Parzakonis, M.; Brady, M.; Kansiime, P.; Said, S.; Ayilu, R.K.; Radavoi, C.N.; Berman, A.L. Generative AI and Higher Education Assessments: A Competency-Based Analysis. *Res. Sq. (Res.Sq.)* **2023**. [CrossRef]
36. Perkins, M.; Roe, J.; Postma, D.; McGaughran, J.; Hickerson, D. Detection of GPT-4 Generated Text in Higher Education: Combining Academic Judgement and Software to Identify Generative AI Tool Misuse. *J. Acad. Ethics* **2023**, *22*, 89–113. [CrossRef]

37. Magdalena, I.; Nurchayati, A.; Mustikawati, R. Kompetensi Pengetahuan dan Teknik Penilaian dalam Evaluasi Pembelajaran di Sekolah Dasar. *Tsaqofah* **2023**, *3*, 794–801. [CrossRef]
38. Sievertsen, H.H. Assessments in Education. *arXiv* **2022**, arXiv:2208.05826.
39. Lundquist, A.E.; Kelly, C. Assessment: Using Data to Support Graduate Student Success and Program Effectiveness. In *A Practitioner's Guide to Supporting Graduate and Professional Students*; Shepard, V.A., Perry, A.L., Eds.; Routledge: New York, NY, USA, 2022; pp. 185–208. [CrossRef]
40. Allen, N. Assessment in Higher Education. *Ref. Libr.* **1992**, *17*, 57–68. [CrossRef]
41. Sarkar, T.K. Assessment in Education in India. *SA-eDUC J.* **2012**, *9*, 1–37.
42. Sanabria-Navarro, J.-R.; Silveira-Pérez, Y.; Pérez-Bravo, D.-D.; de-Jesús-Cortina-Núñez, M. Incidences of artificial intelligence in contemporary education. *Comun. Media Educ. Res. J.* **2023**, *31*, 93–103. [CrossRef]
43. You, Y.; Chen, Y.; You, Y.; Zhang, Q.; Cao, Q. Evolutionary Game Analysis of Artificial Intelligence Such as the Generative Pre-Trained Transformer in Future Education. *Sustainability* **2023**, *15*, 9355. [CrossRef]
44. Chan, C.K.Y.; Tsi, L.H.Y. The AI Revolution in Education: Will AI Replace or Assist Teachers in Higher Education? *arXiv* **2023**, arXiv:2305.01185.
45. McCalla, G. The history of artificial intelligence in education–the first quarter century. In *Handbook of Artificial Intelligence in Education*; du Boulay, B., Mitrovic, A., Yacef, K., Eds.; Edward Elgar Publishing: Cheltenham, UK, 2023; pp. 10–29. [CrossRef]
46. Moturu, V.R.; Nethi, S.D. Artificial Intelligence in Education. In *Emerging IT/ICT and AI Technologies Affecting Society*; Chaurasia, M.A., Juang, C.F., Eds.; Springer: Berlin/Heidelberg, Germany, 2023; pp. 233–244. [CrossRef]
47. Tan, S. Harnessing Artificial Intelligence for Innovation in Education. In *Learning Intelligence: Innovative and Digital Transformative Learning Strategies*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 335–363. [CrossRef]
48. Lebo, C.; Brown, N. Integrating Artificial Intelligence (AI) Simulations Into Undergraduate Nursing Education. *Nurs. Educ. Perspect.* **2024**, *45*, 55–56. [CrossRef] [PubMed]
49. Lockee, B.B. Shifting digital, shifting context: (re)considering teacher professional development for online and blended learning in the COVID-19 era. *Educ. Technol. Res. Dev.* **2020**, *69*, 17–20. [CrossRef] [PubMed]
50. Khlaif, Z.N.; Mousa, A.; Hattab, M.K.; Itmazi, J.; Hassan, A.A.; Sanmugam, M.; Ayyoub, A. The Potential and Concerns of Using AI in Scientific Research: ChatGPT Performance Evaluation. *JMIR Med. Educ.* **2023**, *9*, e47049. [CrossRef]
51. Sharples, M. Automated Essay Writing: An AIED Opinion. *Int. J. Artif. Intell. Educ.* **2022**, *32*, 1119–1126. [CrossRef]
52. Doroudi, S. The Intertwined Histories of Artificial Intelligence and Education. *Int. J. Artif. Intell. Educ.* **2022**, *33*, 885–928. [CrossRef]
53. Zhai, X.; Nehm, R.H. AI and formative assessment: The train has left the station. *J. Res. Sci. Teach.* **2023**, *60*, 1390–1398. [CrossRef]
54. Khlaif, Z.N.; Sanmugam, M.; Hattab, M.K.; Bensalem, E.; Ayyoub, A.A.; Sharma, R.C.; Joma, A.; Itmazi, J.; Najmi, A.H.; Ahmed, M.A.; et al. Mobile technology features and technostress in mandatory online teaching during the COVID-19 crisis. *Heliyon* **2023**, *9*, e19069. [CrossRef]
55. Chan, C.; Hu, W. Students' Voices on Generative AI: Perceptions, Benefits, and Challenges in Higher Education. *Int. J. Educ. Technol. High. Educ.* **2023**, *20*, 43. [CrossRef]
56. Chan CK, Y.; Zhou, W. Deconstructing Student Perceptions of Generative AI (GenAI) through an Expectancy Value Theory (EVT)-based Instrument. *arXiv* **2023**, arXiv:2305.01186. [CrossRef]
57. Bulut, O.; Wongvorachan, T. Feedback Generation through Artificial Intelligence. *Open/Technol. Educ. Soc. Scholarsh. Assoc. Conf.* **2022**, *2*, 1–9. [CrossRef]
58. Braun, V.; Clarke, V. Using thematic analysis in psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101. [CrossRef]
59. Venkatesh, V.; Thong, J.Y.L.; Xu, X. Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Q.* **2012**, *36*, 157–178. [CrossRef]
60. Venkatesh, V.; Morris, M.G.; Davis, G.B.; Davis, F.D. User Acceptance of Information Technology: Toward a Unified View. *MIS Q.* **2003**, *27*, 425. [CrossRef]
61. Kandoth, S.; Shekhar, S.K. Social influence and intention to use AI: The role of personal innovativeness and perceived trust using the parallel mediation model. In *Forum Scientiae Oeconomia*; Wydawnictwo Naukowe Akademii WSB: Dąbrowa Górnicza, Poland, 2022; Volume 10, pp. 131–150.
62. Gansser, O.A.; Reich, C.S. A new acceptance model for artificial intelligence with extensions to UTAUT2: An empirical study in three segments of application. *Technol. Soc.* **2021**, *65*, 101535. [CrossRef]
63. Agarwal, R.; Prasad, J. The antecedents and consequences of user perceptions in information technology adoption. *Decis. Support Syst.* **1998**, *22*, 15–29. [CrossRef]
64. Floruss, J.; Vahlpahl, N. *Artificial Intelligence in Healthcare: Acceptance of AI-Based Support Systems by Healthcare Professionals*; Jonkoping University: Jönköping, Sweden, 2020.
65. Henseler, J.; Sarstedt, M. Goodness-of-fit indices for partial least squares path modeling. *Comput. Stat.* **2013**, *28*, 565–580. [CrossRef]
66. Fornell, C.; Larcker, D.F. Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* **1981**, *18*, 39–50. [CrossRef]

67. Hair, J.F.; Ringle, C.M.; Sarstedt, M. Partial least squares structural equation modeling: Rigorous applications, better results and higher acceptance. *Long Range Plan.* **2013**, *46*, 1–12. [CrossRef]
68. Moorhouse, B.L.; Yeo, M.A.; Wan, Y. Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Comput. Educ. Open* **2023**, *5*, 100151. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.