education sciences

MDPI

Article

Redesigning Assessments for AI-Enhanced Learning: A Framework for Educators in the Generative AI Era

Zuheir N. Khlaif [1,*], Wejdan Awadallah Alkouk [2], Nisreen Salama [3] and Belal Abu Eideh [1]

[1] Faculty of Humanities and Educational Sciences, An Najah National University, Nablus P4110257, Palestine; beideh@najah.edu
[2] Department of Languages and Translation, Faculty of Humanities, Birzeit University, Bir Zeit P627, Palestine; wkook@birzeit.edu
[3] Faculty of Nursing, Arab American University, Jenin P.O. Box 840009, Palestine; nisreen.salama@aaup.edu
* Correspondence: zkhlaif@najah.edu

Abstract: The emergence of generative artificial intelligence (Gen AI) in education offers both opportunities and challenges, particularly in the context of student assessment. This study examines faculty members' motivations to redesign assessments for their courses in the Gen AI era and introduces a framework for this purpose. A qualitative methodology was employed, gathering data through semi-structured interviews and focus groups, along with examples of redesigned assessments. Sixty-one faculty members participated in the study, and the data were analyzed using both deductive and inductive thematic approaches. Key motivations for redesigning assessments included maintaining academic integrity, preparing learners for future careers, adapting to technological advancements, and aligning with institutional policies. However, the study also highlighted significant challenges, such as the need for professional development and addressing equity and accessibility concerns. The findings identified various innovative assessment approaches tailored to the requirements of the Gen AI era. Based on these insights, the study developed a conceptual framework titled "Against, Avoid, Adopt, and Explore". Future research is needed to validate this framework and further refine its application in educational contexts.

Keywords: assessment; STEM; Gen AI; human–AI collaboration; AI-resisted assessment; redesign assessment; AI-resisted assessment

Academic Editor: Jerome D'Agostino

Received: 27 November 2024
Revised: 20 January 2025
Accepted: 27 January 2025
Published: 2 February 2025

Citation: Khlaif, Z. N., Alkouk, W. A., Salama, N., & Abu Eideh, B. (2025). Redesigning Assessments for AI-Enhanced Learning: A Framework for Educators in the Generative AI Era. *Education Sciences*, *15*(2), 174. https://doi.org/10.3390/educsci15020174

1. Introduction

Generative artificial intelligence (Gen AI) has undergone significant advancements, transforming from simple text-generation tools into highly sophisticated systems capable of producing human-like content across a broad spectrum of domains (Feuerriegel et al., 2024). With the emergence of advanced models like the GPT-4o and GPT-01 preview model, AI is now able to perform a range of complex tasks, including conducting text analysis, understanding natural language, and even demonstrating creativity in writing and problem solving (Shahriar et al., 2024). The development in AI capabilities has unlocked thrilling opportunities for its integration into education, especially in delivering feedback and improving assessment practices (Dimitriadou & Lanitis, 2023).

Gen AI tools have empowered educators to enhance student engagement, revolutionize assessment methods, develop innovative teaching strategies, and redesign student tasks to align with the demands of the AI era (George & Wooden, 2023; Khlaif et al., 2024). Consequently, there is increasing interest among decision-makers, educators, practitioners, and researchers to investigate how Gen AI can be effectively integrated into higher and public

education systems, paving the way for substantial advancements in diverse educational settings (Noroozi et al., 2024). Furthermore, Gen AI has become a competitive focus for higher education institutions aiming to address global goals such as the United Nations' Sustainable Development Goals (SDGs), particularly SDG 4 (Quality Education) and SDG 5 (Gender Equality) (George & Wooden, 2023; Pisica et al., 2023). However, integrating AI into higher education poses considerable challenges. Beyond financial constraints, a lack of clear vision or strategic planning within higher education institutions often compounds key issues arising from the attitudes and perspectives of the academic community (Bozkurt et al., 2023; Jin et al., 2024).

1.1. Research Gap

Gen AI integration into education has been widely studied in areas like personalized learning, instructional design, and administrative automation (Sajja et al., 2024; Ruiz-Rojas et al., 2023). However, a gap remains in understanding Gen AI's effective incorporation into assessment practices, especially in higher education. While existing research predominantly focuses on AI applications for automating grading or providing instant feedback, limited attention has been given to how AI can fundamentally transform the nature and methodology of the assessment itself.

Moreover, the discourse surrounding Gen AI in education has been mainly concentrated in Global North contexts, overlooking its potential adaptation and implementation in unique cases with economic and political limitations, like Palestine and some other countries in the Middle East. The challenges and resource constraints faced by educational systems in Palestine necessitate tailored approaches to AI integration. This study seeks to address these gaps by exploring how Gen AI can be leveraged to rethink assessment methods. It focuses on early adopters in Palestine in humanities science, STEM fields, and medical sciences.

1.2. Purpose and Contribution of the Study

This study explores faculty members' motivations for redesigning their course assessments and presents a comprehensive framework for rethinking assessment practices in the era of generative AI (Gen AI). By examining the perspectives and experiences of higher education faculty who participated in a series of workshops on Gen AI conducted by the authors, this research captures the collective wisdom of practitioners regarding the adaptation of assessments to the Gen AI era. These workshops facilitated collaboration and the exchange of ideas, fostering innovative approaches to align course assessments with the evolving educational landscape.

The study contributes to the ongoing discourse on integrating Gen AI into higher education by addressing gaps in understanding AI's transformative role in assessment practices. While existing literature focuses primarily on AI's applications in personalized learning and administrative tasks, this research examines Gen AI's potential for reshaping assessments at a fundamental level. By engaging early adopters from diverse academic fields, the study highlights the motivations, challenges, and benefits of redesigning assessments to meet the demands of the Gen AI era.

A key outcome of this study is a comprehensive framework, grounded in theoretical insights and practical experiences from faculty training workshops, that provides actionable solutions for designing AI-enhanced assessments. This framework addresses critical issues such as academic integrity, student engagement, alignment with 21st-century skills, and the Sustainable Development Goals (SDGs). Additionally, by contextualizing its findings within the challenges faced by resource-constrained regions, the research promotes equitable and

effective AI integration in education, fostering global innovation in teaching, learning, and assessment.

1.3. Research Questions

What motivations are driving Palestinian higher education faculties to redesign assessments in response to Gen AI integration?

What challenges do faculty members encounter when redesigning AI-resistant assessments in higher education?

What types of assessment redesigns do faculty members perceive as most suitable for their teaching practices?

2. Literature Review

AI-Resistant Assessment

Defined by Khlaif et al. (2024) as tasks to evaluate students that minimize reliance on generative artificial intelligence (Gen AI) tools and emphasize skills that AI cannot replicate, like critical thinking, creativity, and ethical- decision-making, instructors assign AI-resistant assessments to reduce the likelihood that Gen AI can complete the whole assignment task (Moorhouse et al., 2024). Rudolph et al. (2023) provided examples on AI-resistant assessments such as analyzing images or videos, creating infographic for specific context and analyzing in-class discussions. These assessments prioritize authentic, process-driven tasks where students engage deeply with the material, including iterative projects, reflective writing, and oral evaluations Donaghy (2023). AI-resistant assessments align with educational principles by fostering academic integrity (Kostanek & Li, 2025), inclusivity (Ahmed et al., 2024), and real-world applicability, ensuring that students' learning outcomes demonstrate human capabilities (Abubakar et al., 2024; Awadallah Alkouk & Khlaif, 2024). By integrating these principles, AI-resistant assessments maintain fairness and originality while adapting to the evolving educational landscape (Abubakar et al., 2024; Miller, 2024).

Implementing AI-resistant assessments strengthens critical thinking, encourages independent problem-solving, and enhances key skills including adaptability and ethical reasoning (Kostanek & Li, 2025). This strategy mitigates plagiarism risks and excessive dependence on AI, promoting an informed approach to technology use (Eze, 2024; Shivshankar & Acharya, 2025). In the long run, AI-resistant assessments can foster critical thinking, enhance human–AI collaboration, and prepare individuals for future use of AI as a supportive tool. Thus, learners can evaluate and integrate AI technologies while experiencing authentic learning and skill development.

Social constructivism, grounded in the work of Vygotsky's work, provides a theoretical foundation for AI-resistant assessments by emphasizing learning as a collaborative and context-driven process. The theory views knowledge as constructed through social interaction, cultural engagement, and authentic experiences (Umar et al., 2024). AI-resistant assessments align with this perspective by encouraging tasks that foster active learning and meaningful interaction, stressing human cognition and collaboration AI reliance (Zhou et al., 2024). Researchers view Gen AI tools as mediating technologies supporting students' knowledge construction and providing scaffolding for learners to explore, analyze, and apply concepts. AI-resistant assessments promote students' ownership of their learning journey through the documentation of their processes and critical reflection regarding their decisions. This focus on the learning process over the final product ensures students build critical, adaptable skills while engaging deeply with the material (Abubakar et al., 2024; Khlaif et al., 2024).

The present researchers define AI-resistant assessment as strategically designed evaluation tasks that reduce reliance on Gen AI tools, focus on human-centered critical thinking,

creativity, and ethical decision-making skills, and enhance the learning process to prepare students for future jobs.

Gen AI in Higher Education

Gen AI technologies offer countless contributions to various elements of education, most notably in assessment (Zhang & Aslan, 2021; Chiu et al., 2023). ChatGPT, for instance, has already revealed its impact in the field of graduate studies (Al-Zahrani, 2024), which is tangibly experienced by instructors in their work. Research indicates a growing reliance among graduate students on Gen AI platforms, including ChatGPT, fostering negative attitudes toward these platforms among many educators.

Several studies, including Cabellos et al. (2024) and Cabero-Almenara et al. (2024), have examined higher education instructor's beliefs regarding Gen AI in teaching. These studies have found that higher education program instructors have not yet formed clear stances on using these platforms in their teaching; some view them as harmful, others see them as beneficial, and many remain undecided.

Despite skepticism, some studies have highlighted the role of motivation among instructors to use Gen AI platforms in teaching, particularly in assessment. Farrelly and Baker (2023) identified several factors enhancing graduate program instructors' motivation to adopt AI platforms in education. These include contextual factors (e.g., autonomy-supportive leadership, autonomy-thwarting leadership), occupational experiences (e.g., professional growth striving, change-related stress), and background factors (e.g., gender, age, teaching experience, contract length, class size, school level).

Nevertheless, Fakhar et al. (2024) emphasized that adopting Gen AI tools in higher education programs is not solely dependent on instructors' motivation. It also requires professional development, robust technological infrastructure, and addressing concerns related to data privacy. Meanwhile, Sembey et al. (2024) examined the use of AI, learning analytics, and XR in higher education, focusing on assessment and feedback practices. The systematic literature review analyzed 3038 studies, wherein journal articles made up 71% of the studies, with 50% related to learning analytics. The research findings revealed that the primary motivation for integrating emerging technologies (ETs) was to analyze student learning performance, provide personalized support, and reduce teacher workloads. In the future, the use of XR technologies, formative assessment practices, and ethical perspectives should be researched in greater depth.

While Gen AI might demonstrate the potential to reform education by enabling the context-based evaluation of student learning outcomes, current Gen AI tools have limitations, such as social biases in data sets. However, the Gen AI revolution is moving away from memorization-based systems to focus on fostering knowledge application. To ensure the validity of Gen AI-generated assessments, it is vital to align these assessments with the vision and performance expectations of learning environments. Future research should establish guidelines and methodologies for assessing AI-based assessments (Kaldaras et al., 2024).

Student learning assessment, a cornerstone of the educational process, is often based on a standardized assessment procedure (SAP) utilizing various types of questions, including constructed-response and selected-response questions (Stăncescu, 2017). However, this type of assessment is fraught with issues, with the most important being the effort teachers must invest in preparing these questions, creating answer keys, and deriving insights from regarding student performance. These results are often prone to errors, leading to difficulty in understanding student performance (Cazan & Indreica, 2014).

The emergence of Gen AI applications prompts the urgent need and opportunity to rethink the philosophy of assessment, its forms, and the tools employed (Nadeem et al.,

2024). AI-based technologies can fully or partially automate parts of traditional assessment practices, and AI can create assessment tasks, select appropriate peers to grade work, and automatically evaluate student work. These technologies help transition tasks from humans to AI, making assessment practices more executable and maintainable (Swiecki et al., 2022). Thus, in the AI age, teachers can design assessment tasks and evaluate student work automatically, potentially making the assessment process more accurate and objective.

Indeed, technology advancements necessitate a flexible, trial-and-error methodology. While a multi-modal approach using various technologies enhances learning, challenges like device incompatibility and accessibility can arise; teachers can save time by becoming familiar with new technologies and their related studies (Deeley, 2018). Research on Gen AI and assessment since ChatGPT's release has focused on the ability to answer multiple-choice or open-ended questions and differentiate AI-generated outputs from human and AI detection tools. Despite variations, Gen AI can pass professional exams and generate coherent texts, offering new opportunities in higher education assessment (Luo, 2024).

Despite the advantages of using AI applications to assess student performance, numerous downsides and limitations have been indicated (Martínez-Comesaña et al., 2023). According to Swiecki et al. (2022), these challenges include excluding human elements and professional expertise by automatically relying on results derived from AI-based applications without human verification. Other significant aspects entail the use of alternative assessments, peer assessments, and human judgements based on teachers' understanding of their students' performances through direct observation (Swiecki et al., 2022).

Notably, AI evaluations frequently encounter ethical issues related to bias, privacy, and the ownership of data. When taught improperly, AI systems might negatively impact disadvantaged learners by enforcing biases in assessment criteria and threatening data privacy, necessitating high prioritization from technology providers and educational institutions (Khlaif et al., 2024). AI in higher education also faces risks relating to data quality, safety, misinformation, dependency, accountability, legality, and integrity (Donnell et al., 2024). The work of Zawacki-Richter et al. (2019) highlighted bias and unfairness as among the most significant ethical challenges related to AI use in education; as the algorithms rely on historical data to build their models, existing prejudices in the data may introduce biases. Likewise, AI-supported education systems may amplify racial, economic, and social gaps by potentially relying on data with systematic biases against certain groups.

Previous studies have raised concerns about AI resulting in potential dependency, laziness, and lower learning quality, as well as the lack of clear university guidance on ethical AI use (Khlaif et al., 2024). They have also differentiated between AI's use in a supportive context versus direct assessment completion (Donnell et al., 2024). Similarly, Huber et al. (2024) highlighted six key online assessment considerations in business education—academic integrity, quality feedback, positive learning experience, student information integrity, equal opportunity, and authenticity—while addressing challenges educators face in balancing academic integrity with a positive student experience. Such challenges included resource limitations, institutional policies, and accreditation requirements. Meanwhile, Boud and Bearman (2022) suggested normalizing collaborative experiences within course structures, providing assessable tasks from diverse experiences, avoiding inequity, and allowing for risk and failure, emphasizing trust in collaborative approaches and the need for holistic, course-wide attention to communicate learning beyond individuality. Many of its insights can be extrapolated to the use of AI in higher education.

Azevedo and Gašević (2019) claimed that AI applications cannot analyze higher-order thinking skills, leading to a narrow perception of the aspects of learning a student has achieved. According to Rudolph et al. (2023), one of the most significant challenges of Gen AI tools in assessment is their inability to detect plagiarism and cheating. Their study

found that many of the tools could not identify text generated by ChatGPT, complicating teachers' work and threatening the educational process.

While Williamson and Eynon (2020) indicated that Gen AI applications can be used by teachers to assess student learning, this capability is not comprehensive across all learning aspects. Indeed, it is limited to measuring aspects like recall and did not extend to evaluating implicit aspects of personality, including critical and creative thinking. Similarly, Couldry (2020) highlighted the loss of strong social relationships between teachers and students, which enhances teachers' understanding of the personal and social context of students. Couldry (2020) described the hidden nature of AI tools when used in student assessments as a "black box", referring to the lack of transparency in the evaluation criteria, which makes it difficult for students, teachers, and administrators to understand how assessment results are reached. Thus, the academic community must engage in broader discussions about AI assessment limitations and recognize forms of learning that may remain invisible to technology.

Therefore, this study builds on prior research by addressing gaps in the integration of Gen AI in assessment practices, focusing on practical solutions grounded in faculty workshops. While existing literature highlights challenges like ethical concerns and plagiarism, our research offers a comprehensive framework promoting academic integrity, student engagement, and alignment with 21st-century skills and SDGs. Additionally, it explores AI-resistant assessments to foster critical thinking and adaptability, contextualizing findings within resource-constrained settings to support equitable AI integration in higher education. This contribution bridges theoretical and practical insights, advancing the discourse on Gen AI's transformative role in education.

3. Methodology

This study employed a qualitative research approach to develop a framework for educators to redesign their assessments in the Gen AI era. The approach included three primary components: semi-structured interviews, focus group sessions, and document analyses (content analyses of redesigned assessments). Each method was chosen to gather insights into faculty members' experiences, perspectives, and how they applied workshop principles to redesign their assessments (Akyıldız & Ahmed, 2021; Berndtsson, 2017).

3.1. Context of the Study

Over a year, the first author conducted training workshops and public lectures to rethink educational assessment in the AI age, engaging 155 faculty members from various Palestinian universities both online and in-person. Table 1 presents the demographic information of the participants who attended the training workshops. These workshops combined theoretical frameworks with practical applications to prepare educators for AI's impact on teaching and assessment. The purpose of the workshops was to educate and support faculty members to redesign assessments tasks to improve the learning outcomes in the Gen AI era.

The theoretical component introduced the AI Assessment Scale (AIAS), a framework informed by prior research (Grassini, 2023; Lye & Lim, 2024; Petihakis et al., 2024). This tool guided educators in systematically evaluating student work that incorporated AI tools. Participants were also trained in prompt engineering to craft effective prompts for eliciting high-quality AI responses. Supporting this, a GPT-4-powered chatbot was developed to assist educators in aligning AI-resistant assessments with course requirements, fostering critical thinking, creativity, and ethical AI use.

Table 1. Demographic information of the participants who attended the training workshops.

Variable		Frequency	%
Gender	Male	89	57.4
	Female	66	42.6
Age (Years)	25–35	43	27.7
	36–45	68	43.9
	46–55	25	16.1
	56+	19	12.3
Frequency of Gen AI Use	Daily	58	37.4
	Weekly	63	40.6
	Monthly	19	12.3
	Occasionally	15	9.7
Discipline	Medical Sciences	15	9.7
	Humanities and Educational Sciences	30	19.4
	Engineering Sciences	24	15.5
	Social Sciences	35	22.6
	Natural Sciences (Physics, Math, etc.)	28	18
	Business and Communication	23	14.8
University	An Najah National University	25	16.1
	Birzeit University	22	14.2
	Hebron University	16	10.3
	Al-Quds University	17	11
	Ministry of Higher Education	15	9.7
	Palestine Technical University	20	12.9
	Arab American University	20	12.9
	Palestine Ahliya University	20	12.9

The practical component focused on hands-on activities and educators applied the AIAS to design assessments promoting human–AI collaboration while maintaining academic integrity. Emphasizing Gen AI's role in brainstorming, the workshops encouraged students' accountability for AI-generated content by assessing their cognitive processes in creating effective prompts and critically engaging with AI outputs.

3.2. Semi-Structured Interviews

Semi-structured interviews were utilized to explore redesigning assessment methods to integrate Gen AI in education, a relatively new phenomenon (Yin, 2013). From a pool of 155 faculty members who attended training sessions, 25 participants were selected for interviews using purposive sampling to ensure diversity and representation. The selection process considered several key factors. Participants were chosen to reflect a range of experiences with Gen AI tools, from novice to those with substantial AI expertise, granting comprehensive understanding of different approaches to integrating AI in assessments. Additionally, the study emphasized disciplinary diversity, involving faculties from various academic fields from humanities to engineering. This diversity enabled an exploration of how AI-driven assessment redesigns vary across disciplines. Institutional representation was another important consideration, with participants drawn from multiple universities to capture a variety of institutional policies and technology-in-education resources. These criteria ensured the study could provide a detailed perspective on faculty members' experiences with incorporating AI into their assessment practices.

3.3. Focus Group Sessions

Focus groups are particularly useful for exploring how social dynamics and group interactions shape participants' perspectives and experiences (Lambert & Loiselle, 2008; Geampana & Perrotta, 2024), encouraging participants to engage in discussions, share their experiences, and build on one another's ideas. This fosters an environment where opinions can be debated, knowledge co-constructed, and complex phenomena explored in depth. In this study, focus groups provided a platform to investigate how faculty members are utilizing Gen AI in teaching and rethinking assessment practices in the AI era. Unlike individual interviews, focus groups enabled the collection of rich, multifaceted data; participants responded to both the facilitator's prompts and the ideas and insights creating a dynamic and reflective dialogue (Poliandri et al., 2023).

Four focus group sessions were conducted, each comprising nine participants, selected from a pool of 130 faculty members, excluding those who participated in semi-structured interviews. This resulted in a total of 36 participants. The selection criteria for the focus groups were as follows:

- Experience with AI-Resistant Assessment Design: Priority was given to faculty members with substantial experience adapting assessments to AI considerations. This included tangible modifications to assessment practices or insightful reflections on integrating Gen AI into teaching and evaluation.
- Openness to Collaboration: Participants were chosen based on their willingness to engage collaboratively, ensuring meaningful contributions to group discussions. This criterion emphasized participants who were proactive shared best practices and actively engaged in problem-solving challenges.

The objective of the focus group was to create a collaborative learning environment where participants could reflect collectively, share experiences, discuss challenges, and exchange innovative strategies for leveraging Gen AI in education (Gundumogula & Gundumogula, 2020). This interactive approach enriched the study's findings by capturing the dynamic interplay of ideas and shared learning among faculty members. Therefore, the total number of participants in this study was 61. Table 2 provides the demographic information for the participants in this study.

Table 2. Demographic information about the participants in the semi-structured interviews and focus group sessions.

Variable		Frequency	%
Gender	Male	38	62.3
	Female	23	37.7
Age (Years)	25–35	11	18
	36–45	19	29.5
	46–55	21	36.1
	56+	10	16.4
Frequency of Gen AI Use	Daily	18	29.5
	Weekly	16	26.2
	Monthly	12	19.7
	Occasionally	15	24.6
Discipline	Medical Sciences	7	11.5
	Humanities and Educational Sciences	12	19.7
	Engineering Sciences	10	16.4
	Social Sciences	8	13.1
	Natural Sciences (Physics, Math, etc.)	13	21.3
	Business and Communication	11	18

3.4. Document Analysis: Redesigned Assessment Analysis

Redesigned assessments were analyzed based on submissions from 61 participants, providing a comprehensive perspective on how faculty members integrated the training principles. Each assessment was systematically collected and evaluated to ensure alignment with two key aspects, as follows.

- Incorporation of AI-Resistant Design: Elements that directly reflected the training focus, including task complexity, prompts designed to promote critical thinking, and strategies for assessing student originality and engagement.
- Alignment with Workshop Principles: How closely each assessment adhered to the workshop's objectives, particularly redesigning assessments to integrate AI considerations.

The approach offered valuable insights into the practical applications of the workshop's concepts across diverse educational contexts, demonstrating the impact and adaptability of the training principles.

Therefore, the total number of participants across the semi-structured interviews and the four focus group sessions was 61.

3.5. Data Collection

Individual semi-structured interviews were conducted with 25 participants. Each interview lasted 30 to 45 min and was audio-recorded for accuracy. Participants were first asked to share their experiences with Gen AI technology, specifically their efforts to design new assessments addressing the challenges posed by AI integration. This allowed participants to narrate their stories and journeys in adapting to Gen AI. Follow-up questions delved deeper into their experiences, particularly regarding how students' interact with newly designed assessments and projects.

Participants were also asked to reflect on traditional assessment methods they deemed outdated and describe new assessment strategies they had implemented. The interviews aimed to collect comprehensive insights from participants' experiences, including their practices and the challenges of these new assessments.

Additionally, participants shared a variety of artifacts from their work, such as assessment designs and student engagement examples. These served as valuable secondary data sources that complemented the qualitative data gathered during the interviews. This combination of narrative and tangible materials enriched the study's understanding of how educators adapt assessment practices in the age of Gen AI.

Four focus group sessions, each lasting approximately one hour, were conducted to explore faculty members' perspectives on adapting their assessment methods. Two sessions were in person and two took place online via video conferencing. Two researchers facilitated and guided the conversations using derived from the interview data (refer to File S1 for examples of these prompts). The focus groups provided a deeper understanding of how faculty members perceived the need to evolve their assessment practices in response to their experiences with Gen AI. With participants' consent, all sessions were audio-recorded. The data collected from these discussions offered additional insights that enriched and complemented the findings from the semi-structured interviews, providing a more comprehensive view of the challenges and opportunities in redesigning assessments for the AI era.

Samples of redesigned assessments were collected and analyzed to classify the types and components of these assessments. These samples provided tangible evidence of the shifts in assessment design, allowing for a deeper understanding of the practical applications and implications of incorporating AI into educational contexts.

These samples not only served as secondary data sources but also enriched the primary data by offering examples that complemented the qualitative findings. These artifacts in the study were instrumental to triangulating data, ensuring a comprehensive exploration of the ways educators transformed their assessment practices to remain relevant in the Gen AI era. The secondary data provided a unique lens through which to assess the innovation, creativity, and challenges associated with redesigning assessments, reinforcing the significance of these contributions to the overall research findings.

3.6. Data Analysis Procedures

Inductive thematic analysis was conducted to examine the data collected from the semi-structured interviews and focus group discussions. This analysis adhered to the six-step procedure outlined by Braun and Clarke (2006), encompassing 11.65 h of recorded interview data and 3.75 h of recorded focus group discussions.

The researchers transcribed the audio recordings, which were subsequently validated by the participants to ensure accuracy and reliability. The data were then analyzed using NVivo software, enabling the systematic coding of ideas and concepts. During this process, the researchers identified and categorized subthemes, which were further organized into overarching main themes, drawing on insights from relevant existing research.

The analysis closely aligned with the study's research questions, providing a structured and methodical framework for uncovering key findings. This approach ensured themes and insights emerged organically while addressing the study's objectives. An example of the coding framework is provided in (Table A1).

The study also incorporated a detailed examination of the collected samples of redesigned assessments. These artifacts were analyzed to identify patterns and classified based on their structures, objectives, types of assessment, and integration of AI tools. This artifact analysis provided a complementary perspective, allowing the researchers to contextualize and validate the themes derived from the qualitative data. Thus, the study gained deeper insights into how educators rethought assessment practices in response to the challenges and opportunities of the Gen AI era.

3.6.1. Triangulation Process

Coding data from individual interviews provided rich insights into participants' experiences and perspectives. These initial codes were then cross-referenced with data from focus group discussions to confirm or refine themes, ensuring consistency and validity. For example, themes such as "equity concerns" and "time and resource demands" identified in interviews were further supported by patterns emerging from focus group data. Additionally, documented examples of AI-resistant assessments were analyzed as a third source, validating and contextualizing the codes generated from the verbal data. This triangulation process allowed us to compare findings across sources, ensuring that the final themes were robustly reflective of diverse perspectives. By intentionally cross-verifying data and examining convergence and divergence among sources, we upheld methodological rigor and strengthened the credibility of the study's findings.

3.6.2. Reporting the Themes Emerged in This Study

Participant responses were coded into themes using a systematic approach. Quantitative terms such as "most", "majority", "many", "several" and "few" were defined as follows: "most" and "majority" represented over 75% of participants, "many" represented 50–75%, "some" indicated less than 30–49%, "several" represented 20–29% and "few" indicated less than 20% of the participants. Where participants expressed multiple perspectives (e.g., concerns and exploratory attitudes toward AI), their responses were coded under each relevant theme.

3.7. Trustworthiness

The researchers prioritized confirmability, credibility, dependability, and transferability throughout the research process to ensure trustworthiness. Two distinct data collection methods ensured the triangulation of the findings, strengthening credibility. Confirmability was achieved by documenting the research process and data analysis, ensuring participants' perspectives were faithfully represented without undue influence from researchers' interpretations. The transcribed files were shared with participants for member checking, allowing them to review and amend their statements as needed, thereby enhancing accuracy.

The interview protocol was based on the study's research questions, a pilot interview, and input from experts in computer science and educational technology. The researchers adopted a code-recode strategy in which they coded the data independently three times and compared the results to maintain consistency, ensuring dependability. As the interviews and focus group discussions were conducted in Arabic, the researchers employed a backward translation process paired with conceptual equivalence to ensure linguistic and contextual accuracy. The calculated interrater reliability was 89%, demonstrating robust agreement among coders. For transferability, participants were selected through purposive sampling to ensure the relevance and applicability of findings across similar contexts.

3.8. Ethical Considerations

The study was conducted following approval from the Institutional Review Board (IRB) at An Najah University Committee, with approval number Intr. April. 2024/18. All participants provided informed consent, which was facilitated by a consent form included in the interview and focus group protocols. The protocol began with a clear explanation of the study's purpose, emphasizing that participation was voluntary and responses would remain confidential and anonymous. Participants were informed that their agreement to proceed would serve as consent, and they were reminded of their right to withdraw from the study at any time. This approach ensured ethical compliance and participant autonomy throughout the research.

4. Results

In this section, the findings are presented based on the research questions (Table A2).

4.1. Research Question 1

What motivations are driving Palestinian higher education faculties to redesign assessments in response to Gen AI integration?

The participants reported various motivations for rethinking assessments while integrating Gen AI into their teaching. The researchers categorized these motivations into five main themes: preparing students for future work, technological adaptation, academic integrity, institutional policy, and ethical considerations (Table A3). Each theme has subthemes, and these themes and their associated subthemes are detailed below.

4.1.1. Preparing Students for Future Work

Most participants highlighted the need to emphasize creativity, critical thinking, and analytical skills, areas in which Gen AI falls short. Participants acknowledged that, while AI can support students' initial drafts, the assessments' focus should be on skills that require human originality and depth: "In my course, AI can answer factual questions, but it can't evaluate or synthesize ideas. That is where we want students to shine." (F17).

Many participants identified enhancing creativity and innovation in student work since students can collaborate with AI in innovative ways, encouraging unique problem-

solving approaches: "I ask them to use AI creatively, like brainstorming alternative solutions, to show their innovative thinking." (F23).

Many participants also reported their intention to equip learners with practical AI skills and competencies for their future careers: "I have to prepare my students in my courses to be familiar with AI tools; if I do not do so, they will be behind AI, and you know AI is everywhere now." (F2).

4.1.2. Technological Adaptation

Several participants mentioned promoting active and authentic learning by engaging students in real-world tasks and project-based assessments. Examples were shared in which students collaborated with peers on projects while using tools like ChatGPT for ideation. Additionally, many assessments highlighted the deliberate use of Gen AI tools, requiring students to utilize these tools and critically evaluate the outputs: "By making projects more interactive and real-world, I find my students stay engaged, and AI becomes a tool, not a shortcut. They also critique the content they generate using tools like Bard." (F7).

Incorporating AI tools for enhanced learning was also reported by several participants in the focus group sessions, with AI being a source to improve learners' engagement and learning and understanding, as reported by many of the participants in the interviews: "Using AI for refining ideas helps learners learn more effectively and enhance critical thinking." (F2).

4.1.3. Academic Integrity

Many faculty members expressed a desire to guarantee that student work reflects their understanding and the skills taught in the course, even when they incorporate AI into their learning: "My biggest concern is that students might just copy and paste AI outputs without really engaging with the material." (FG4).

Additionally, several participants emphasized that redesigning assessments is crucial to distinguishing between AI-assisted work and authentically student-generated work: "It's about leveling the playing field, ensuring assessments remain fair regardless of students' access to AI tools." (FG3).

4.1.4. Institution Policy

The majority of the participants noted they are forced to adapt to new policies and rethink improving the assessment process: "It is required by my institution to design new assessments which are different from previous ones we are familiar with. . .so I need to reflect that in my courses assessments." (F21).

Some participants were motivated to align Palestinian higher education with global trends. They emphasized that rethinking assessments in the era of AI is part of a broader international movement, and Palestinian institutions must actively participate: "To stay relevant on a global stage, we must show we're embracing technology responsibly and riding the AI wave." (F1).

4.1.5. Ethical Considerations

Participants highlighted the importance of addressing ethical issues and promoting responsible AI using among students, ensuring they understand AI tools' capabilities and limitations. Many participants expressed concern that, without proper guidance, students might misuse or rely on AI inappropriately, compromising their learning: "Ethical use means knowing when to rely on AI and when to step back and rely on your own judgment." (F2).

A critical ethical aspect involves balancing the use of AI technology with the cultivation of independent learning skills. Faculty members emphasized that while AI provides

valuable support in tasks like brainstorming and drafting, it cannot supplant essential cognitive processes such as critical thinking and self-directed learning. As part of their ethical responsibility, educators highlighted the need to redesign assessments to ensure that students engage meaningfully with the material and develop original ideas, even when leveraging AI as a supplementary tool: "AI can't replace the learning process. Assessments need to encourage students to think independently." (FG3).

4.2. Research Question 2

What challenges do faculty members encounter when redesigning AI-resistant assessments in higher education?

Most participants indicated various challenges while redesigning assessments. The researchers categorized these challenges into five main themes, including maintaining academic integrity, time and resource demands, equity concerns, resistance to change, and lack of training (Table A4).

4.2.1. Maintaining Academic Integrity

Many participants reported difficulties ensuring that redesigned assessments effectively distinguish between AI-assisted work and authentic student contributions. Gen AI tools can mimic high-quality academic writing, making it challenging to verify originality: "It's becoming increasingly difficult to identify where the student's work ends and the AI's begins." (F5).

4.2.2. Time and Resource Demands

Another challenge mentioned by participants was the time and effort required to redesign AI-resistant assessments, including creating innovative tasks, training educators, and adopting new evaluation methods. These additional demands strain already limited resources, as supported by the discussion in the focus group sessions: "We're expected to innovate while balancing heavy workloads—there's just not enough time to experiment and redesign." (FG4).

4.2.3. Equity and Accessibility Concerns

Equity and accessibility were important concerns for most of the participants. Since students have diverse backgrounds, not all have equal access to AI tools or the digital infrastructure. This disparity creates challenges in designing fair assessments: "Not all students are familiar with AI, and some don't even have reliable Internet. It's hard to make assessments equitable under these circumstances." (F13).

4.2.4. Resistance to Change

Some participants were disinclined to change their assessment modes due to familiarity with the traditional assessments. Other participants mentioned that learners accustomed to traditional assessments may resist transitioning to AI-resistant formats. However, faculties often need to overcome institutional inertia and skepticism from colleagues and learners, as mentioned by some participants: "Some colleagues don't see the urgency of redesigning assessments, and students often push back on tasks that demand more effort." (F25).

4.2.5. Lack of Clear Guidelines and Training

Many institutions lack clear policies or guidelines on incorporating or resisting in assessments, leaving the faculty to navigate this challenge. This was considered an important challenge by the majority of the participants Other participants reported that professional

development opportunities on this topic are often insufficient: "We're left to figure this out on our own; there's no institutional framework to guide us." (FG1).

4.3. Research Question 3

What types of assessment redesigns do faculty members perceive as most suitable for their teaching practices?

Here, the researchers utilized three data sources, prioritizing the analysis of redesigned assessments developed by participants for their courses. They identified four key principles, namely against, avoid, adapt, and explore, that participants reported during discussion sessions and interviews for redesigning assessments. These principles are detailed in Table 3, which includes examples of assessment types and the associated levels of Gen AI integration.

Table 3. Types of assessments and examples from participants' practices.

Assessment Type	Example	Purpose of Assessment	Gen AI Integration Level
Against	Exams, oral exams, discussion	Evaluate students lower skills	Not allowed to use Gen AI
Avoid	Performance based assessment, Personal reflection, Individual portfolio,..	Evaluate of higher order levels	Lower level of integrating Gen AI
Adapt	Brainstorming. Ideation	Higher order thinking	Higher level of Gen AI integration (part, full)
Explore	Solving-problem, co-design, learner-AI partners, creativity in learning.	Gen AI partner in assignments	Higher level of Gen AI integration (part, full)

Below is a description of each theme as reported by the participants in the study and as analyzed from the samples from participants' practices: Example of assessments redesigned for various courses during and after the training sessions). Table 3 presents the categories and examples for each type based on the samples provided by the participants in the study.

Additionally, the researchers classified the redesigned assessments into these four categories. Building upon these principles and the degrees of Gen AI integration in the assessments, they developed a conceptual framework illustrated in Figure 1. The Against, Avoid, Adopt, and Explore (AAAE) framework is accompanied by a comprehensive description of each principle and the types of assessments suitable for each category.

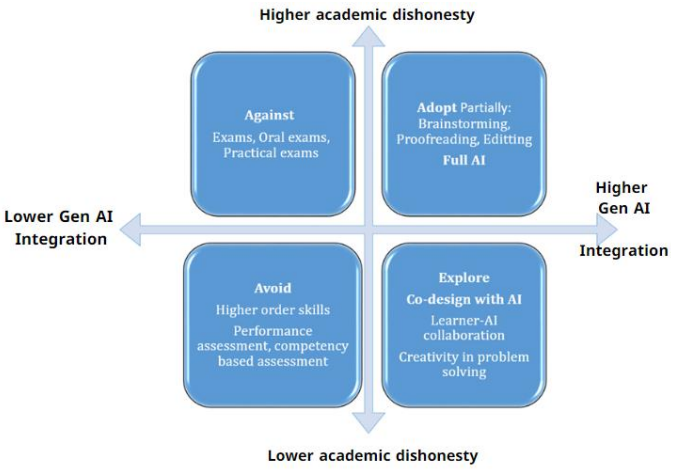

Figure 1. The Against, Avoid, Adopt, and Explore (AAAE) framework.

4.3.1. Against: No Use of Gen AI in Assessments

Most participants reported that they still used traditional assessment methods such as traditional exams, oral exams, and presentations to measure lower-level thinking, including remembering and other related skills to reduce academic dishonesty. In this type of assessment, students were not allowed to use any Gen AI tool: "I am still using traditional assessment methods such as oral, presentation, and paper-based exams to focus on lower-order thinking skills such as the basics and recalling information." (F15).

4.3.2. Avoid: Assessments with Which Gen AI Currently Struggles

Some participants emphasized the importance of developing and implementing innovative assessment methods that minimize the use of Gen AI. Examples included performance-based assessments, individual portfolios, and reflection papers. These approaches aim to evaluate students' higher-order thinking skills, which participants claim Gen AI tools struggle to address effectively. One faculty member shared an example of a performance-based assessment, as shown in Appendix B: "This assessment is tied to personal experiences and simulates a real-life scenario. Generative AI cannot replicate self-awareness as students can." (F22).

4.3.3. Adopt: Incorporating Gen AI in Assessments

Many participants developed various types of assessments that integrated Gen AI. These assessments often leveraged Gen AI to refine the final version of the work or generate ideas, particularly in project-based assessments. Interestingly, even participants who predominantly used traditional assessment methods incorporated Gen AI into their tasks. The extent of integration varied between partial and full depending on the assessment type, as reflected in the interviews, discussions, and assessment instructions.

Participants highlighted several justifications for integrating Gen AI into assessments, including teaching students how to use AI responsibly, fostering an understanding of AI ethics, enhancing students' engagement in the generated content, and encouraging its use in specific parts of the assessment process: "AI is real, and students are using it. I include it to encourage them to declare their use of AI rather than hide it. I noticed they were happy to share their experiences with it and how it benefited their assessments." (F4). Another faculty member added: "I encourage them to use it partially, like generating initial ideas, and then continue with their own thoughts or expand on ideas provided by Gen AI." (F9).

Furthermore, it was observed that faculty members designed rubrics to accompany the assessments (Appendix B). These rubrics served as guidelines for students on how to use Gen AI in their work, and they outlined the criteria faculty would use to evaluate the final product.

4.3.4. Explore: Gen AI Partner in Assessments

Participants focused on exploratory assessments that encourage learners to uncover the potential of Gen AI in the learning process. These assessments involved activities such as problem solving, co-design, learner–AI collaboration, and fostering creativity. This approach aimed to cultivate higher-order thinking, critical reflection, and the ability to adapt AI capabilities for innovative problem solving while also emphasizing the responsible use of technology: "My philosophy in redesigning assessments is to encourage my students to explore the additional value Gen AI can offer and how they can develop their ideas based on it." (F25).

4.4. Types of the New Assessment Approaches

After analyzing and categorizing the assessments, as outlined in Figure 1, a clear consensus emerged regarding AI-resistant assessments. The majority of participants indicated that these assessments must address the challenges of the Gen AI era, mitigating the AI misuse risk, while leveraging its potential to support productive learning.

During the focus group discussions, participants elaborated on the core concept of AI-resistant assessments, drawing from their experiences in redesigning evaluation methods. Participants also highlighted that these assessments entail designing tasks that AI cannot easily solve, thereby encouraging deeper cognitive engagement from students. This approach was consistently reported as a key strategy to ensure the integrity of student work.

The researchers presented the types of assessments redesigned during the training workshops, as detailed below. Figure 2 presents various types of assessments collected from the participants in workshops.

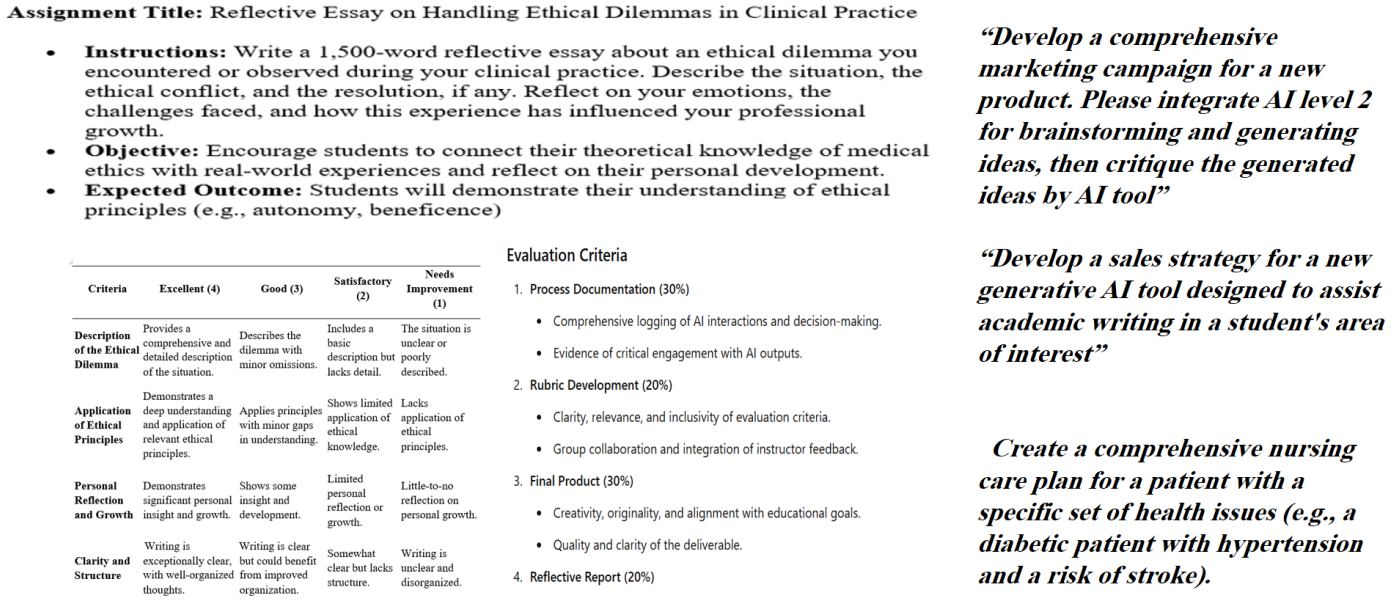

Figure 2. Various types of assessments developed by the participants in the workshops.

4.4.1. Product–Process Assessment

The focus was placed on emphasizing critical engagement with Gen AI, encompassing the final product and the process involved in completing a task, as reported by a faculty member in educational sciences. Many participants placed importance on this approach, focusing on learners working with AI tools as a form of human AI collaboration. Vital to integrating Gen AI in teaching, this type of assessment encourages students to engage in developing the prompts to enhance the outcome within the adopted and explored principles: "I need to evaluate the process the student used while generating the content by using Gemini." (F17).

4.4.2. Competency-Based Assessment

Some of the participants presented a new competency-based assessment approach to assess learners' application of theoretical knowledge in practical situations. One professor reported that AI enhanced virtual partners' ability to assess clinical decision-making and problem-solving skills in high-pressure scenarios. The examples in Figure 2 illustrate the potential of AI to transform assessments in ways that prepare students for real-world challenges, offering a dynamic and interactive learning experience.

4.4.3. Authentic Assessment: Real-World Applications in Education

One of the key insights that emerged was the need for authentic assessments that reflect real-world challenges, requiring students to apply knowledge in complex, context-rich situations: "Instead of a traditional essay, I might ask my students to create podcasts, develop multimedia presentations, or solve real-world problems related to the course using ChatGPT or Gemini." (F23).

The participants expressed the importance of authentic assessment for enhancing creativity, collaboration, and critical thinking, but they noted that the faculty should design accompanying rubrics.

5. Discussion

This study highlights the opportunities and challenges of redesigning assessment practices in the generative artificial intelligence (Gen AI) era, drawing connections and highlighting the dynamic, interrelated nature of such challenges. The proposed Against, Avoid, Adopt, and Explore (AAAE) framework builds on previous work (e.g., Lye & Lim, 2024; Grassini, 2023; Waladi & Lamarti, 2024) to assist educators in navigating these issues and granting equitable attention to ethical considerations, institutional demands, and student needs. Higher-order thinking, creativity, and active learning are key areas that can be influenced when transforming assessment approaches via Gen AI. In line with previous studies (Aziz et al., 2020; Khlaif et al., 2024; Waladi & Lamarti, 2024) on technological advancements to improve assessment authenticity and creativity, many participants focused on addressing Gen AI's weaknesses. This entails redesigning assessments to account for elements like critical thinking and originality, promoting improvements to student engagement and learning experiences via AI's use as a brainstorming or co-creation tool. This resonates with prior studies emphasizing the importance of AI-enhanced learning environments (Boud & Bearman, 2022; Feuerriegel et al., 2024).

Regardless, Gen AI requires faculties to reevaluate traditional assessment paradigms; the importance of human–AI collaboration in education is emphasized by examining how students interact with AI to refine outputs that align with the "Product–Process Assessment Model" (Luo, 2024).

In line with previous studies, including Swiecki et al. (2022) and Rudolph et al. (2023), participants identified noteworthy obstacles concerning the misuse of AI: maintaining academic integrity, addressing equity issues, overcoming resistance to change, and preventing the obfuscation of student contributions. These challenges highlight the critical need for clarity, fairness, and accessibility in those guidelines and institutional policies that support ethical AI use.

Echoing the findings of Sembey et al. (2024), participants listed disparities in access to AI tools and digital infrastructure as particularly salient concerns. Furthermore, as suggested by Deeley (2018) and Spivakovsky et al. (2023), systemic efforts are required if policymakers are to ensure that professional development opportunities that bridge the digital divide are provided to educators.

Many participants identified institutional policies as the driving force behind their decision to redesign assessments. This finding aligns with global educational practices, reflecting broader trends in aligning educational practices with global movements like the United Nations' Sustainable Development Goals (George & Wooden, 2023). In addition, the vitality of ethical considerations also featured prominently, with importance placed on responsible AI use. Such a finding emphasizes the need to embed ethical reasoning into AI-enhanced assessments if educators are to sufficiently prepare students for the real-world challenges that await us (Waladi & Lamarti, 2024).

5.1. Implications for Future Research and Practice

Future studies ought to explore the impacts of varying levels of AI integration on learning outcomes, for which the AAAE framework offers a valuable starting point. Another area in which the AAAE is applicable was elucidated by the work of Couldry (2020) and Azevedo and Gašević (2019), which addressed critical questions about the balance between leveraging AI's capabilities while preserving human-centric educational values. Future research could expand on these themes by investigating the long-term effects of AI-enhanced assessments on student learning and employability.

5.2. Practical Implications

Ongoing faculty development and collaboration are essential to sufficiently address challenges and share best practices as a collective, forming the first—and arguably most important—practical implication. This study also underscores the importance of co-designing policies and guidelines alongside educators, students, and technology providers alike to create a sustainable and inclusive framework for AI integration in education.

6. Conclusions

This study emphasizes the dynamic opportunities and challenges associated with redesigning assessment practices in the era of generative artificial intelligence (Gen AI). The proposed Against, Avoid, Adopt, and Explore (AAAE) framework provides educators with a structured approach to navigating these complexities, offering equitable consideration of ethical considerations, institutional demands, and student needs.

One of the key opportunities highlighted is the ability of Gen AI to enhance higher-order thinking, creativity, and active learning. By integrating AI into assessments, educators can design tasks that promote critical thinking and originality while improving student engagement. Gen AI can serve as a tool for brainstorming and co-creation, enabling learners to refine and expand their ideas collaboratively. However, this transformation requires a fundamental reevaluation of traditional assessment paradigms, ensuring that human–AI collaboration becomes a cornerstone of future educational practices.

At the same time, the study identifies significant challenges related to the integration of Gen AI. These include maintaining academic integrity, addressing equity issues, and preventing the obfuscation of student contributions. Resistance to change and disparities in access to AI tools and digital infrastructure further complicate the process. These findings underscore the importance of clear, fair, and accessible institutional policies that guide the ethical use of AI in education. Systemic efforts, including professional development opportunities, are necessary to bridge the digital divide and support educators in adapting to these changes.

Institutional policies and ethical considerations play a central role in driving the redesign of assessments. Embedding ethical reasoning into AI-enhanced assessments is essential to prepare students for real-world challenges and align educational practices with broader global movements toward sustainability.

Looking ahead, future research should explore the long-term impacts of AI integration on learning outcomes and employability. Ongoing faculty collaboration and the co-design of inclusive policies with stakeholders will be critical to creating sustainable and effective AI-driven educational frameworks. Through thoughtful implementation, the AAAE framework provides a valuable foundation for addressing the complexities of the Gen AI era.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/educsci15020174/s1: File S1: Interview protocol; File S2: Focus Group Protocol for Discussing Redesign Assessment in the Gen AI Era.

Author Contributions: Conceptualization, Z.N.K.; Methodology, Z.N.K. and W.A.A.; Formal analysis, Z.N.K., W.A.A., N.S. and B.A.E.; Investigation, W.A.A., N.S. and B.A.E.; Resources, Z.N.K., W.A.A., N.S. and B.A.E.; Data curation, W.A.A.; Writing—original draft, Z.N.K., N.S. and B.A.E.; Writing—review & editing, Z.N.K. and W.A.A. 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 study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of An Najah University Committee (protocol code Intr. April. 2024/18 and 15 April 2024).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data used in this study will be available upon request from the corresponding author.

Acknowledgments: During the preparation of this work, the authors used ChatGPT in order to organize and proofread the final version of the manuscript. After using this tool, the author(s) reviewed and edited the content as needed and take full responsibility for the content of the publication.

Conflicts of Interest: The authors declare there are no conflicts of interest among them.

Appendix A

Table A1. Coding example.

Script	Theme
F 5: Well, it's been a mix of trial and error, honestly. For example, in one course, I introduced an assignment where students used AI to generate a draft of their research question and then refine it through peer feedback and their own critical thinking. Another time, I asked them to analyze the output of an AI-generated summary of an article and compare it with the original content to identify gaps or biases. These activities helped them engage with the material on a deeper level while also evaluating the tool itself. I've also done some group projects where students collaborate to create content using generative AI and then reflect on the process and outcomes.	Attitude Adopt Develop students Motivation Explore Explore Type of new assignments Product–process: Avoid Explore: co-author Human–AI collaboration
F 11: Absolutely. One example that comes to mind is from a creative writing class I teach. I asked students to use ChatGPT to generate a story prompt and then write their own short story based on that prompt. Afterward, they had to reflect on how AI's ideas influenced their creative process what they kept, what they changed, and why. It was fascinating because, um, some students said AI pushed them to think outside the box, while others said they found its suggestions too generic and wanted to make them more personal. That reflective component was really insightful and showed me how students are engaging with these tools on multiple levels.	Product–process Motivation Explore

Table A2. Examples of the Three Research Questions (Coding).

Theme	Subtheme	Quotation	Research Question
Preparing Students for Future Work	Fostering Higher-Order Thinking Skills	In my course, AI can answer factual questions, but it can't evaluate or synthesize ideas. That is where we want students to shine. (F17)	Motivations
Preparing Students for Future Work	Enhancing Innovation and Creativity	I ask them to use AI creatively, like brainstorming alternative solutions, to show their innovative thinking. (F23)	Motivations
Technological Adaptation	Preparing Learners for AI-Enhanced Workplaces	I have to prepare my students in my courses to be familiar with AI tools; if I do not do so, they will be behind AI, and you know AI is everywhere now. (F2)	Motivations
Ethical Considerations	Balancing Innovation with Learning Integrity	Students might rely too heavily on AI, risking the undermining of their learning processes.	Challenges
Ethical Considerations	Promoting Responsible AI Use	Encouraging learners to use AI ethically ensures they develop the judgment needed to navigate these tools effectively.	Challenges
Assessment Redesign	Creating AI-Resistant Assessments	Redesigning assessments to focus on the process and not just the product helps in mitigating over-reliance on AI.	Redesign Practices
Assessment Redesign	Product–Process Collaboration	For creative writing, students used AI for initial prompts and reflected on how they revised it, emphasizing personal input over generic AI suggestions. (F11)	Redesign Practices
Assessment Redesign	Co-Design with Students	Including students in redesigning assessments that integrate AI encourages transparency and ownership.	Redesign Practices
Engagement	Exploring AI as a Creative Partner	Group projects where students collaborate to create content using AI helped them reflect on the strengths and limitations of the tools. (F5)	Motivations
Institutional Policy	Faculty and Student Training	Workshops on ethical AI integration help clarify the boundaries and expectations for using these tools.	Challenges

Table A3. Coding book for the first question: Motivation.

Theme	Subtheme	Quotation
Preparing Students for Future Work	Fostering Higher-Order Thinking Skills	In my course, AI can answer factual questions, but it can't evaluate or synthesize ideas. That is where we want students to shine. (F17)
	Enhancing Innovation and Creativity	I ask them to use AI creatively, like brainstorming alternative solutions, to show their innovative thinking. (F23)
	Active Learning Engagement	Students need to collaborate with AI in ways that emphasize their original contributions and unique problem-solving skills.
Technological Adaptation	Preparing Learners for AI-Enhanced Workplaces	I have to prepare my students in my courses to be familiar with AI tools; if I do not do so, they will be behind AI, and you know AI is everywhere now. (F2)
	Integrating AI Tools for Enhanced Learning	Using tools like ChatGPT in projects allowed students to explore real-world tasks and learn collaboratively.
Academic Integrity	Addressing Plagiarism Challenges	We are exploring ways to assess originality and ensure students engage deeply with the material rather than outsourcing their learning to AI.
	Authentic Assessments to Mitigate Misuse	We design tasks requiring critical thinking and creativity, skills that AI tools currently cannot authentically replicate.
Institutional Policy	Developing AI Usage Guidelines	Policies need to guide ethical AI use while promoting innovation and maintaining academic standards.
	Faculty and Student Training	Workshops on ethical AI integration help clarify the boundaries and expectations for using these tools.
Ethical Considerations	Balancing Innovation with Learning Integrity	Students might rely too heavily on AI, risking the undermining of their learning processes.
	Promoting Responsible AI Use	Encouraging learners to use AI ethically ensures they develop the judgment needed to navigate these tools effectively.
Assessment Redesign	Creating AI-Resistant Assessments	Redesigning assessments to focus on the process and not just the product helps in mitigating over-reliance on AI.
	Co-Design with Students	Including students in redesigning assessments that integrate AI encourages transparency and ownership.

Table A4. Coding book for the second research question (Challenges).

Theme	Subtheme	Quotation
Maintaining Academic Integrity	Distinguishing AI from Student Work	It's becoming increasingly difficult to identify where the student's work ends and the AI's begins. (F5)
Time and Resource Demands	Strain on Faculty Time and Efforts	We're expected to innovate while balancing heavy workloads—there's just not enough time to experiment and redesign. (FG4)
Time and Resource Demands	Limited Institutional Resources	Developing innovative assessments requires resources that are simply not available in most cases. (Observation from focus groups)
Equity and Accessibility Concerns	Inequality in AI Access	Not all students are familiar with AI, and some don't even have reliable Internet. It's hard to make assessments equitable under these circumstances. (F13)
Equity and Accessibility Concerns	Digital Infrastructure Limitations	Students from rural areas face issues even accessing online resources, let alone experimenting with AI. (Paraphrased from focus group discussions)
Resistance to Change	Faculty Reluctance to Innovate	Some colleagues don't see the urgency of redesigning assessments, and students often push back on tasks that demand more effort. (F25)
Resistance to Change	Learner Resistance to New Formats	Traditional exams are what they're used to—when we introduce AI-resistant tasks, there's often backlash. (Observation from interviews)
Lack of Clear Guidelines and Training	Absence of Institutional Frameworks	We're left to figure this out on our own; there's no institutional framework to guide us. (FG1)
Lack of Clear Guidelines and Training	Insufficient Professional Development	There's minimal training provided for understanding how to integrate or avoid AI in assessments. (General comment from faculty)

Appendix B. Sample of Redesigned Assignments

اسم المهمة: تحليل معمق لأحداث الثورة الفلسطينية

وصف المهمة:

قم بدراسة الثورة الفلسطينية منذ الانطلاقة ، حلل الأسباب، النتائج من خلال استخدام المصادر الأولية والثانوية. يمكنك الاستفادة من أدوات الذكاء الاصطناعي لدعم البحث، مثل توليد الأفكار وتحليل النصوص الكبيرة، مع تقديم استنتاجاتهم حول الجوانب المتنوعة للثورة، مع التركيز على وجهات النظر المختلفة.

المعيار	الوصف	ممتاز(4)	جيد جدًا(3)	جيد(2)	يحتاج تحسين(1)	الوزن
جمع وتحليل المصادر	القدرة على جمع وتحليل مصادر أولية وثانوية متعددة وموثوقة تتناول الثورة الفلسطينية من زوايا متعددة.	جمع مصادر متنوعة وموثوقة وتحليلها بعمق ووضوح.	جمع مصادر متنوعة وتحليلها بوضوح نسبي.	جمع مصادر محدودة أو تفتقر للتنوع والتحليل البسيط.	جمع مصادر ضعيفة أو غير موثوقة وغياب التحليل الواضح.	25%
استخدام الذكاء الاصطناعي	الاستفادة من أدوات الذكاء الاصطناعي في البحث، مثل تحليل النصوص وتوليد الأفكار لتعزيز فهم الجوانب المختلفة للثورة الفلسطينية.	استخدام فعال للذكاء الاصطناعي، مما يعزز الفهم ويوفر رؤى جديدة.	استخدام الذكاء الاصطناعي بشكل جيد، مع تقديم نتائج مفيدة.	استخدام محدود أو غير فعال للذكاء الاصطناعي.	لم يتم استخدام الذكاء الاصطناعي بشكل مؤثر.	20%
تحليل الأسباب والنتائج	تحليل شامل للأسباب والنتائج المترتبة على الثورة الفلسطينية، مع ربط منطقي بين	تحليل عميق ومتناسق للأسباب والنتائج، وربط منطقي ممتاز.	تحليل جيد للأسباب والنتائج مع بعض الربط بين الأحداث.	تحليل جزئي أو غير مكتمل للأسباب أو النتائج.	غياب التحليل الواضح أو عدم الترابط بين الأحداث.	30%

	الأحداث والتطورات.					
التنظيم والعرض	تنظيم الأفكار وعرضها بشكل منظم، مع تدرج منطقي للأفكار والتحليلات وسهولة القراءة.	تنظيم متميز وسرد واضح وسلس للأفكار والتحليلات.	تنظيم جيد للأفكار، مع بعض الصعوبات في التتابع أو الوضوح.	التنظيم غير محكم، مما يؤثر على تدفق الأفكار وفهم التحليل.	عرض غير منظم، وصعوبة في متابعة الأفكار والتحليلات.	15%
وجهات النظر المختلفة	إدراج وتحليل وجهات نظر متعددة حول الثورة، مع تقييم نقدي لتلك الرؤى واستخلاص النتائج.	عرض شامل ومنطقي لوجهات نظر متعددة مع تقييم نقدي متميز.	عرض وجهات نظر متعددة مع بعض التقييم النقدي.	عرض وجهة نظر واحدة أو بعض التحليل النقدي فقط.	عدم عرض وجهات نظر مختلفة، وغياب التقييم النقدي.	10%

Assignment Title: Reflective Essay on Handling Ethical Dilemmas in Clinical Practice

- **Instructions:** Write a 1,500-word reflective essay about an ethical dilemma you encountered or observed during your clinical practice. Describe the situation, the ethical conflict, and the resolution, if any. Reflect on your emotions, the challenges faced, and how this experience has influenced your professional growth.
- **Objective:** Encourage students to connect their theoretical knowledge of medical ethics with real-world experiences and reflect on their personal development.
- **Expected Outcome:** Students will demonstrate their understanding of ethical principles (e.g., autonomy, beneficence) and their application in clinical settings.

Criteria	Excellent (4)	Good (3)	Satisfactory (2)	Needs Improvement (1)
Description of the Ethical Dilemma	Provides a comprehensive and detailed description of the situation.	Describes the dilemma with minor omissions.	Includes a basic description but lacks detail.	The situation is unclear or poorly described.
Application of Ethical Principles	Demonstrates a deep understanding and application of relevant ethical principles.	Applies principles with minor gaps in understanding.	Shows limited application of ethical knowledge.	Lacks application of ethical principles.
Personal Reflection and Growth	Demonstrates significant personal insight and growth.	Shows some insight and development.	Limited personal reflection or growth.	Little-to-no reflection on personal growth.
Clarity and Structure	Writing is exceptionally clear, with well-organized thoughts.	Writing is clear but could benefit from improved organization.	Somewhat clear but lacks structure.	Writing is unclear and disorganized.

"Develop a comprehensive marketing campaign for a new product. Please integrate AI level 2 for brainstorming and generating ideas, then critique the generated ideas by AI tool"

"Develop a sales strategy for a new generative AI tool designed to assist academic writing in a student's area of interest"

Create a comprehensive nursing care plan for a patient with a specific set of health issues (e.g., a diabetic patient with hypertension and a risk of stroke).

Evaluation Criteria

1. **Process Documentation (30%)**
   - Comprehensive logging of AI interactions and decision-making.
   - Evidence of critical engagement with AI outputs.
2. **Rubric Development (20%)**
   - Clarity, relevance, and inclusivity of evaluation criteria.
   - Group collaboration and integration of instructor feedback.
3. **Final Product (30%)**
   - Creativity, originality, and alignment with educational goals.
   - Quality and clarity of the deliverable.
4. **Reflective Report (20%)**

References

Abubakar, U., Falade, A. A., & Ibrahim, H. A. (2024). Redefining student assessment in Nigerian tertiary institutions: The impact of AI technologies on academic performance and developing countermeasures. *Advances in Mobile Learning Educational Research*, *4*(2), 1149–1159. [CrossRef]

Ahmed, S., Zaki, A., & Bentley, Y. (2024). AI and personalised grading criteria. In *Utilizing AI for assessment, grading, and feedback in higher education* (pp. 85–113). IGI Global. [CrossRef]

Akyıldız, S. T., & Ahmed, K. H. (2021). An overview of qualitative research and focus group discussion. *International Journal of Academic Research in Education*, *7*(1), 1–15. [CrossRef]

Al-Zahrani, A. M. (2024). The impact of generative AI tools on researchers and research: Implications for academia in higher education. *Innovations in Education and Teaching International*, *61*(5), 1029–1043. [CrossRef]

Awadallah Alkouk, W., & Khlaif, Z. N. (2024). AI-resistant assessments in higher education: Practical insights from faculty training workshops. In *Frontiers in education* (Vol. 9, p. 1499495). Frontiers Media SA. [CrossRef]

Azevedo, R., & Gašević, D. (2019). The role of self-regulation in learning with technology. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Reimann (Eds.), *International handbook of the learning sciences* (pp. 449–460). Routledge.

Aziz, M. N. A., Yusoff, N. M., & Yaakob, M. F. M. (2020). Challenges in using authentic assessment in 21st century ESL classrooms. *International Journal of Evaluation and Research in Education*, *9*(3), 759–768. [CrossRef]

Berndtsson, J. (2017). Combining semi-structured interviews and document analysis in a study of private security expertise. In *Researching non-state actors in international security* (pp. 81–95). Routledge.

Boud, D., & Bearman, M. (2022). The assessment challenge of social and collaborative learning in higher education. *Educational philosophy and theory*, *56*, 1–10. [CrossRef]

Bozkurt, A., Junhong, X., Lambert, S., Pazurek, A., Crompton, H., Koseoglu, S., Farrow, R., Bond, M., Nerantzi, C., Honeychurch, S., Bali, M., Dron, J., Mir, K., Stewart, B., Costello, E., Mason, J., Stracke, C. M., Romero-Hall, E., Koutropoulos, A., . . . Jandrić, P. (2023). Speculative futures on ChatGPT and generative artificial intelligence (AI): A collective reflection from the educational landscape. *Asian Journal of Distance Education*, *18*(1), 53–130.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. [CrossRef]

Cabellos, B., De Aldama, C., & Pozo, J. (2024). University teachers' beliefs about the use of generative artificial intelligence for teaching and learning. *Frontiers in Psychology*, *15*, 1468900. [CrossRef]

Cabero-Almenara, J., Palacios-Rodríguez, A., Loaiza-Aguirre, M. I., & Del Rosario De Rivas-Manzano, M. (2024). Acceptance of educational artificial intelligence by teachers and its relationship with some variables and pedagogical beliefs. *Education Sciences*, *14*(7), 740. [CrossRef]

Cazan, A., & Indreica, E. (2014). Traditional assessment of learning versus online assessment. *eLearning and Software for Education*, *3*, 96–101. [CrossRef]

Couldry, N. (2020). The (in)visibility of AI in educational assessments: Examining the implications of datafication for educational practices. In K. R. L. Baumberger, & T. M. George (Eds.), *Digital transformations in education: The role of AI in educational innovation* (p. 1139). Springer.

Deeley, S. J. (2018). Using technology to facilitate effective assessment for learning and feedback in higher education. *Assessment & Evaluation in Higher Education*, *43*(3), 439–448.

Chiu, T. K., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *4*, 100118. [CrossRef]

Dimitriadou, E., & Lanitis, A. (2023). A critical evaluation, challenges, and future perspectives of using artificial intelligence and emerging technologies in smart classrooms. *Smart Learning Environments*, *10*(1), 12. [CrossRef]

Donaghy, K. (2023, November 16). *Multimodality and multimodal literacy: What are they and why are they important in ELT?* Oxford University Press. Available online: https://teachingenglishwithoxford.oup.com/2023/11/16/multimodality-and-multimodal-literacy-elt/#:~:text=Multimodal%20literacy%20refers%20to%20the,curricula%20to%20include%20mult (accessed on 7 October 2024).

Donnell, F. O., Porter, M., & Fitzgerald, S. (2024). The role of artificial intelligence in higher education: Higher education students use of ai in academic assignments. *Irish Journal of Technology Enhanced Learning*, *8*(1). [CrossRef]

Eze, C. A. (2024). The role of educators in upholding academic integrity in an ai-driven era. In *AI and ethics, academic integrity and the future of quality assurance in higher education*.

Fakhar, H., Lamrabet, M., Echantoufi, N., Ouadrhiri, K., El Khattabi, K., & Ajana, L. (2024). ChatGPT as an intelligent self-continuous professional development tool for teachers. *Statistics, Optimization & Information Computing*, *13*(2), 488–507.

Farrelly, T., & Baker, N. (2023). Generative artificial intelligence: Implications and considerations for higher education practice. *Education Sciences*, *13*(11), 1109. [CrossRef]

Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative AI. *Business & Information Systems Engineering*, *66*(1), 111–126. [CrossRef]

Geampana, A., & Perrotta, M. (2024). Using interview excerpts to facilitate focus group discussion. *Qualitative Research*, 14687941241234283. [CrossRef]

George, B., & Wooden, O. (2023). Managing the strategic transformation of higher education through artificial intelligence. *Administrative Sciences*, *13*(9), 196. [CrossRef]

Grassini, S. (2023). Development and validation of the AI attitude scale (AIAS-4): A brief measure of general attitude toward artificial intelligence. *Frontiers in Psychology*, *14*, 1191628. [CrossRef]

Gundumogula, M., & Gundumogula, M. (2020). Importance of focus groups in qualitative research. *International Journal of Humanities and Social Science (IJHSS)*, *8*(11), 299–302. [CrossRef]

Huber, E., Harris, L., Wright, S., White, A., Raduescu, C., Zeivots, S., Cram, A., & Brodzeli, A. (2024). Towards a framework for designing and evaluating online assessments in business education. *Assessment & Evaluation in Higher Education*, *49*(1), 102–116.

Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2024). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *arXiv*, arXiv:2405.11800. [CrossRef]

Kaldaras, L., Akaeze, H. O., & Reckase, M. D. (2024). Developing valid assessments in the era of generative artificial intelligence [Conceptual Analysis]. In *Frontiers in Education* (Vol. 9). Frontiers Media SA. [CrossRef]

Khlaif, Z. N., Ayyoub, A., Hamamra, B., Bensalem, E., Mitwally, M. A. A., Ayyoub, A., Hattab, M. K., & Shadid, F. (2024). University teachers' views on the adoption and integration of generative AI tools for student assessment in higher education. *Education Sciences*, *14*(10), 1090. [CrossRef]

Kostanek, E., & Li, F. (2025). The false sense of achievement: Navigating academic integrity and assessment challenges of GenAI. In *Educational Assessments in the Age of Generative AI* (pp. 33–58). IGI Global Scientific Publishing.

Lambert, S. D., & Loiselle, C. G. (2008). Combining individual interviews and focus groups to enhance data richness. *Journal of advanced nursing*, *62*(2), 228–237. [CrossRef] [PubMed]

Luo, J. (2024). A critical review of GenAI policies in higher education assessment: A call to reconsider the "originality" of students' work. *Assessment & Evaluation in Higher Education*, *49*(5), 651–664.

Lye, C. Y., & Lim, L. (2024). Generative artificial intelligence in tertiary education: Assessment redesign principles and considerations. *Education Sciences*, *14*(6), 569. [CrossRef]

Martínez-Comesaña, M., Rigueira-Díaz, X., Larrañaga-Janeiro, A., Martínez-Torres, J., Ocarranza-Prado, I., & Kreibel, D. (2023). Impact of artificial intelligence on assessment methods in primary and secondary education: Systematic literature review. *Revista De Psicodidáctica (English Ed)*, *28*(2), 93–103. [CrossRef]

Miller, W. (2024). Adapting to AI: Reimagining the role of assessment professionals. *Intersection: A Journal at the Intersection of Assessment and Learning*, *5*(4), 99–113. [CrossRef]

Moorhouse, B. L., Wan, Y., Wu, C., Kohnke, L., Ho, T. Y., & Kwong, T. (2024). Developing language teachers' professional generative AI competence: An intervention study in an initial language teacher education course. *System*, *125*, 103399. [CrossRef]

Nadeem, M., Farag, W., & Helal, M. (2024, May 2–4). *Rethinking assessment methodologies in the era of artificial intelligence: Expanding beyond ChatGPT's scope*. 2024 Mediterranean Smart Cities Conference (MSCC), Tetuan, Morocc. [CrossRef]

Noroozi, O., Soleimani, S., Farrokhnia, M., & Banihashem, S. K. (2024). Generative AI in education: Pedagogical, theoretical, and methodological perspectives. *International Journal of Technology in Education*, *7*(3), 373–385. [CrossRef]

Petihakis, G., Farao, A., Bountakas, P., Sabazioti, A., Polley, J., & Xenakis, C. (2024, July 30–August 2). *AIAS: AI-ASsisted cybersecurity platform to defend against adversarial AI attacks*. 19th International Conference on Availability, Reliability and Security (pp. 1–7), Vienna, Austria.

Pisica, A. I., Edu, T., Zaharia, R. M., & Zaharia, R. (2023). Implementing artificial intelligence in higher education: Pros and cons from the perspectives of academics. *Societies*, *13*(5), 118. [CrossRef]

Poliandri, D., Perazzolo, M., Pillera, G. C., & Giampietro, L. (2023). Dematerialized participation challenges: Methods and practices for online focus groups. *Frontiers in Sociology*, *8*, 1145264. [CrossRef]

Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, *6*(1), 342–363. [CrossRef]

Ruiz-Rojas, L. I., Acosta-Vargas, P., De-Moreta-Llovet, J., & Gonzalez-Rodriguez, M. (2023). Empowering education with generative artificial intelligence tools: Approach with an instructional design matrix. *Sustainability*, *15*(15), 11524. [CrossRef]

Sajja, R., Sermet, Y., Cikmaz, M., Cwiertny, D., & Demir, I. (2024). Artificial intelligence-enabled intelligent assistant for personalized and adaptive learning in higher education. *Information*, *15*(10), 596. [CrossRef]

Sembey, R., Hoda, R., & Grundy, J. (2024). Emerging technologies in higher education assessment and feedback practices: A systematic literature review. *Journal of Systems and Software*, *211*, 111988. [CrossRef]

Shahriar, S., Lund, B. D., Mannuru, N. R., Arshad, M. A., Hayawi, K., Bevara, R. V. K., Mannuru, A., & Batool, L. (2024). Putting GPT-4o to the sword: A comprehensive evaluation of language, vision, speech, and multimodal proficiency. *Applied Sciences*, *14*(17), 7782. [CrossRef]

Shivshankar, S., & Acharya, N. (2025). AI in assessment and feedback. In *Next-generation AI methodologies in education* (pp. 119–146). IGI Global Scientific Publishing.

Spivakovsky, O. V., Omelchuk, S. A., Kobets, V. V., Valko, N. V., & Malchykova, D. S. (2023). Institutional policies on artificial intelligence in university learning, teaching and research. *Information Technologies and Learning Tools*, *97*(5), 181. [CrossRef]

Stăncescu, I. (2017). The importance of assessment in the educational process—Science teachers' perspective. *The European Proceedings of Social & Behavioural Sciences*, *27*, 753–759. [CrossRef]

Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., Selwyn, N., & Gašević, D. (2022). Assessment in the age of artificial intelligence. *Computers and Education Artificial Intelligence*, *3*, 100075. [CrossRef]

Umar, M., Congman, R., & Akram, H. (2024). Perceived entrepreneurial orientation and perceived academic entrepreneurial intention: A mediating role of knowledge creation. *Policy Research Journal*, *2*(4), 953–967.

Waladi, C., & Lamarti, M. S. (2024). Adaptive AI-Driven Assessment for Competency-Based Learning Scenarios. In *Innovative Instructional Design Methods and Tools for Improved Teaching* (pp. 215–226). IGI Global.

Williamson, B., & Eynon, R. (2020). Historical Threads, Digital Transitions, and the Implications of AI for Learning, Education, and the Field of the Learning Sciences. *British Journal of Educational Technology*, *51*(4), 1117–1130. [CrossRef]

Yin, R. K. (2013). Validity and generalization in future case study evaluations. *Evaluation*, *19*(3), 321–332. [CrossRef]

Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers and Education: Artificial Intelligence*, *2*, 100025.

Zhou, X., Schofield, L., Zhang, J., Abuelmaatti, A., & Howell, L. (2024). Building bridges in AI: Enhancing AI literacy for students and staff across disciplines. In *PGR Student Partners–Empowering doctoral students through partnership and co-creation in institutional equality, diversity and inclusion change projects* (Vol. 20). Educational Developments.

Zawacki-Richter, O., Bäcker, E. M., & Vogt, S. (2019). Systematic review of research on artificial intelligence in higher education. *International Journal of Educational Technology in Higher Education*, *16*(1), 39. [CrossRef]