Harnessing the Generative AI Wave Towards Fair and Diverse Higher Education Assessments: A Comprehensive Analysis through an Innovative Lens of Students

Thashmee Karunaratne[1], Naghmeh (Nam)ghaee[2], and Mexhid Ferati[3]
[1]KTH Royal Institute of Technology, Stockholm, Sweden
[2]Lund University, Sweden
[3]Linnaeus University, Sweden

thashmee@kth.se
nam.aghaee@ics.lu.se
mexhid.ferati@lnu.se

Abstract: While Generative Artificial Intelligence (GAI), particularly tools powered by Large Language Models (LLMs), offer benefits in teaching and learning, they also raise critical concerns about academic integrity, fairness in examinations due to their potential for generating educational content. This evolving landscape requires higher education institutions to rethink their assessment models, ensuring they remain robust, inclusive, and aligned with the realities of AI-enhanced learning environments. In this backdrop, this study investigates the practical, GAI-resistant assessment frameworks in higher education. It explores how alternative, skill-focused methods such as oral exams (vivas) and AI-integrated tasks can be included in future assessment models. Central to the study is the understanding of how students perceive current assessments and envision future methods that fairly and effectively measure both knowledge and skills. The empirical investigation is based on a case study at a Swedish University. Research methodologies include a survey questionnaire administered to 30 students enrolled in a semi-theoretical course on innovation and technology, and a future workshop (FW) with 22 of them in five groups. The two research instruments corresponded to answering the two research questions, respectively. The survey results revealed students' clear concerns about the academic integrity challenges posed by essay and report-based take-home assessments, as well as online quizzes. They also expressed apprehension about the potential impact of relying solely on proctored and supervised exams, highlighting the risk of reducing diversity in assessment methods, and thereby raising red flags for the need for a new and innovative approach to assessment methods that is hardly affected by unauthorised assistance from GAI. Responses to open survey questions reflected their problem-solving mindset and deep thinking of how cheating can be minimised by increased peer collaboration and solving real problems, contextualised to specific and ongoing learning activities in class. The outcomes of the FW provided insights, such as active learning-based assessments, combined with real-world problem-solving or context-specific question-based assessments. These findings are intended to inform course design, policy-making, and broader discussions on educational reform in the digital age.

Keywords: AI-resilient Assessment, Academic Integrity, Fairness in assessment, Diversity, Higher Education

1. Introduction

Generative Artificial Intelligence (GAI) is rapidly transforming higher education, particularly in learning and assessment (Alam, 2021; Tlili et al., 2023). GAI, powered by Large Language Models (LLMs), enables new learning pathways by generating human-like text (Kasneci et al., 2023; Thorne, 2024). GAI variants, i.e., Generative Pre-trained Transformers (GPTs) such as ChatGPT and MS Copilot, have become handy in reshaping education (Ouaazki et al., 2023), but raising concerns about academic integrity (Gopane et al., 2024; Sullivan et al., 2023). The current research evidence pinpoints significant challenges to certain assessment methods in higher education due to GAI (Kasneci et al., 2023). Given that basic university courses predominantly aim to assess core conceptual understanding aligned with intended learning outcomes, the use of GAI during assessments raises concerns about academic integrity, particularly due to its capacity for reproducing content with minimal student input (Gamage et al., 2023). Non-proctored written exams, take-home assignments, and online quizzes, which have long been used to assess students' continuous progress in learning, are now increasingly vulnerable to AI-assisted academic misconduct (Xia et al., 2024). The situation is inarguably more troubling in online courses. However, dropping AI-vulnerable assessment types could negatively impact the quality and diversity of examinations since the potential of non-proctored examinations (especially with open-ended, application-based tasks) on deeper learning has been well justified in contemporary education research and practice (Gamage et al., 2023; Karunaratne et al., 2025). Such limitations will directly affect the diversity of the assessments since having scaffolded and different types of assessments is essential for fair measurement of skills and competencies (Karunaratne et al., 2025). Under these circumstances, teachers will now have to ensure they either use innovative complementary methods that are non-sensitive to AI or limit the assessment methods only to traditional proctored forms. Consequently, students may miss the opportunities to perform their assessment

tasks in a more explorative, relaxing and enjoyable environment. Limiting to a few assessment types will also negatively impact the overall education quality (Karunaratne et al., 2025). Therefore, an urgent need has emerged for a shift toward more authentic, dynamic, and practice-oriented assessment methods that also utilise human competencies, such as collaboration, creative thinking, critical analysis, and effective communication, which are soft skills becoming essential for students to successfully navigate and contribute to an AI-driven workforce (Xia et al., 2024).

Despite growing recognition of these challenges in higher education, established and scalable models for AI-resilient assessment remain scarce, especially in semi-technical courses that contain intended learning goals to balance the evaluation of theoretical knowledge with the development of practical skills. (Aghaee & Karunaratne, 2023). The pragmatic challenge lies in developing innovative, diverse assessment frameworks that ensure fairness, sustain engagement, and accurately measure students' abilities while preventing academic misconduct related to AI. What is significant here is the dual and interconnected view on innovative and authentic assessments, particularly emphasising the soft skills students should develop and demonstrate during both formative and summative activities within an academic course. Numerous studies in the literature, including the study by Aghaee & Karunaratne, (2023), highlight the significance of excelling in skills such as communication, collaboration, teamwork, leadership, creativity, participation, academic writing, and so forth to succeed in assessments. For example, in an oral exam (examination vivias or viva-voce), students must not only deliver the content verbally but also express it more concisely, clearly, and progressively. Such soft skills are increasingly being explicitly assessed in future education scenarios to enhance employability (SkillsMatch, 2018). Therefore, in this study, we argue that innovative assessment methods cannot be effective without considering the soft skills embedded within them.

In this backdrop, this research essentially emphasises the lack of insights into pragmatic frameworks for innovative assessment methodologies resistant to GAI. Therefore, it investigates how and what alternative, skill-focused assessment models, including oral exams, activity-based evaluations, and AI-integrated critical analysis tasks, can be suitably integrated for future university education. In doing so, we examine how students envision their learning can be better assessed in the age of GAI. However, for students to envision an ideal future assessment structure, they must first understand the current examination landscape within both formative and summative assessments and their view of the potential of the current assessments in measuring their skills and subject competencies. Thus, this study will first explore how students envision the roles of soft skills in their education and employability in the job market. Specifically, this study is driven by the following research questions:

[RQ1]: How do students perceive knowledge and skills assessment in the current higher education examination landscape?

[RQ2]: How do students envision the assessment methods that can reflect students' learning, knowledge, and skills while promoting fairness in grading in higher education?

By answering these RQs, we intend to derive insights to form guidelines that can be utilised in various courses and used to examine the effectiveness of these methods in real-world course settings. Furthermore, the outcomes will contribute to the future development of scalable, AI-resilient assessment strategies that enhance learning integrity and align with the evolving demands of higher education and the workforce. As universities, policymakers, and employers seek solutions, the study outcomes are also expected to provide insights for timely alternatives that will inform education policies, improve assessment practices, and contribute to wider discussions on digitalisation and learning integrity, making it highly relevant to academia, industry, and public debate.

2. Methodological approach

As described above, this research was formulated to explore how students envision the innovative composition of an AI-resistant assessment framework. Considering that understanding how students perceive the current assessment landscape is a mandatory component, the first step is to investigate how students envision their typical assessment frameworks in their courses, as a basis for problematisation of the assessment frameworks for the future. Thus, the methodological design combined preparatory background data collection with a short survey, followed by a deep dive into AI-resilient assessments through the Future Workshop (FW) that developed creative collaboration, critical reflection, and the co-design of potential solutions. This study employed FW methodology, following the structure outlined by Vidal, (2006) and based on the original model developed by Jungk & Müllert, (1987). The approach was selected for its alignment with democratic problem-solving and

participatory design principles, particularly its capacity to create collaborative ideation and critical engagement among participants with shared learning contexts.

2.1 The case study

In this empirical research, a case study strategy is applied for several reasons, including that case studies enable a deep and nuanced exploration of the problem being investigated. The selected case study is a course in informatics that integrates design thinking into digital transformation at a higher education institution in southern Sweden. The specific reason for choosing this course is that it is semi-technical, requiring the evaluation of both theoretical knowledge and practical skills to meet the intended learning goals. As mentioned above, such courses demand diverse assessment methods to capture students' knowledge and skills. Additionally, the course also addresses the transformative aspects of technology and includes learning objectives for developing innovation competence. Since the students are the subjects of this study, selecting such a course helps to minimise bias caused by students' design methodological knowledge. In other words, the students involved already possess knowledge of the technical and social aspects of digitalisation and innovation methodologies, which reduces bias related to skills and competence in design thinking and innovation design.

2.2 Study participants

The participant group comprised 30 bachelor's students, organised into five groups for the workshop activity. Of these, 22 students provided informed consent for their contributions to be included in this publication. One group participated in the full workshop activity but later indicated that they did not wish their input to be included in the publication. Consequently, the analysis for this study is based on the inputs from the four remaining groups. Although the gender proportion was fair, no demographic information, such as age or gender, was directly collected to preserve anonymity. All participants had prior familiarity with key concepts such as Generative AI, co-design, and assessment diversity, which enabled them to engage meaningfully with the workshop content. Furthermore, no personal data, nor any reflections derived from students' course results or assessments, were collected or utilised in the FW activity, or neither in the preparation of this publication.

2.3 Pre-workshop survey

The pre-workshop survey consisted of a brief questionnaire administered to 30 students via the interactive digital platform Mentimeter. Participation was voluntary, and informed consent was presented and obtained prior to the submission of responses. This survey included three closed-ended and three open-ended questions. The questions assessed students' awareness of how different assessment types reflect the skills and subject competencies, the challenges students face generally in higher education regarding demonstrating their learning, the assignment types they perceived to be best showcasing their knowledge acquisition, and whether supervised exams or vivas would make the assessments fairer. The open questions allowed students to state other assessment types, challenging them to demonstrate their knowledge, and how they envision re-designing exams to allow them to showcase their competencies. The data collected served to contextualise the upcoming workshop activities by revealing the participants' conceptual entry points and general orientation toward the topic., While the questionnaire serves as a basis for understanding the perception of assessment types and their consequences, it also sensitises facilitators about students' abilities of design thinking.

2.4 Future workshop for participatory co-design

The Future Workshop followed the classical three-stage structure as outlined by Vidal (2006). Preparatory activities were an official initial phase to start the workshop, serving as a warm-up and providing information to the participants. The short survey was also conducted as part of the preparation phase. During this stage, students were also asked to problematise the topic, having an awareness of the task to build. In the second phase, the critique phase, participants in groups collectively identified the challenges and shortcomings in current assessment types in higher education, reflecting their own experiences and perceptions triangulated with the existing accessible knowledge available. These reflections, grounded in students' academic experiences, enabled them to critically examine prevailing challenges in the practices of higher education. The moderator additionally introduced and facilitated brief discussions on the role of generative AI and the growing importance of soft skills in the context of ongoing educational transformation.

In the third phase, the fantasy phase, participants generated speculative ideas for innovative assessments that could address the previously identified problems and develop solutions connected to AI-resilient exams. In line with the FW model, this phase encouraged participants to suspend constraints of realism and explore aspirational, creative design directions. The emphasis was placed on open-ended thinking rather than immediate feasibility, creating space for imaginative alternatives to conventional learning strategies. The fourth and final phase in this workshop, referred to as the implementation phase, required groups to refine and consolidate

their concepts into more concrete forms. This was done using the MoSCoW prioritisation framework (Waters, 2009), which helped participants distinguish between essential and non-essential features of their proposed solutions. MoSCoW is a requirements prioritisation method that categorises features into Must have, Should have, Could have, and Won't have (at this time). Each group created low-fidelity prototypes (due to the complexity of the problem, limited time for design activity, and the outcomes are not complete) to illustrate their concepts, using digital tools such as Padlet, Canva, and Miro, or traditional tools like pen and paper, to visualise interaction flows, motivational elements, and intended learning outcomes.

The workshop phases marked a transition from problem and abstract ideation to grounded design work, allowing students to articulate how their ideas might be realised in practice. As there was no direct prototype or product resulting from this workshop, the follow-up phase was not included in the workshop, and participants were informed accordingly. Throughout all phases of the workshop, the process was designed to encourage democratic participation, mutual learning, and the negotiation of meaning among group members (Spinuzzi, 2005). These elements are central to both participatory design and the pedagogical goals of developing soft skills.

3. Results and Discussion

The results were organised in two sections, i.e., the survey study outcomes intended to answer RQ1: the students' understanding of the current assessment landscape and the viewpoint for the future, and the FW activity for innovation of AI resilient assessment, answering RQ2.

3.1 Pre-workshop survey on the current assessment landscape

The pre-survey is guided by a preamble of the problem, namely, that "*the rise of GAI tools is disrupting the traditional assessment methods in higher education, making take-home assignments and online quizzes increasingly vulnerable to AI-assisted cheating. We are studying how to design fair and engaging university assessments in the age of AI. The survey asks for your opinions on current exams and ideas for new assessment models. The survey is to gather students' input on designing AI-resilient exams for the next round of this course and similar courses in higher education*".

The questions were constituted with an overview of the exam landscape. Firstly, capturing the view of common types of (continuous) assessments, such as project work, project reports, project oral presentations, oral examinations (Viva-voce), online quizzes, supervised exams (in exam halls) and other types that students feel provide the opportunity to demonstrate their knowledge and skills acquisition. As Figure 1 shows, students think project works enable their knowledge reflection the most. The interesting outcome here is that online quizzes are not among the most attractive assessment types.

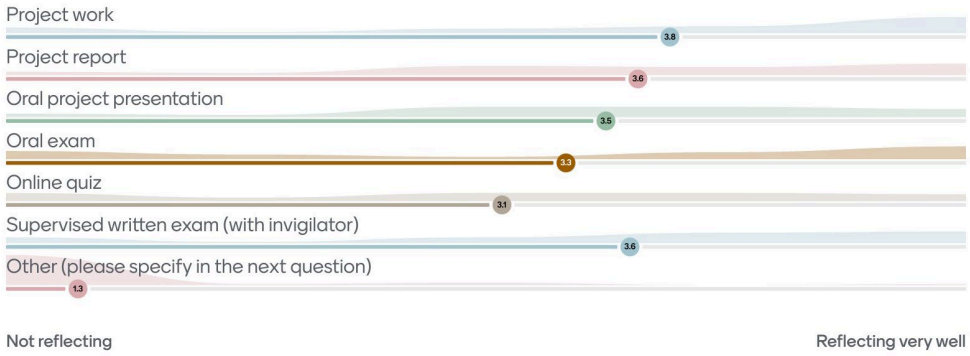

Figure 1: Student perception of the assessment types reflecting the knowledge and skills acquired

Students were allowed to give their views on other types of assessments in the following questions. Students also believed that combined assessments, such as an individual project plus presentation, group projects (not essays but actual projects), more verbal exams, home exams with context-specific questions which are harder to answer with GAI support, and Industry/real-life related problems/projects, could be useful to showcase their competencies. Students were also promoting code-checking questions in verbal forms for specific programming courses.

Regarding the challenges when demonstrating students' learning (cf. Figure 2), students think that the risk of AI-assisted results affecting assessment fairness is the biggest among the provided factors. They also foresee the lack of practical/real-world problems/applications within the assessment frameworks.

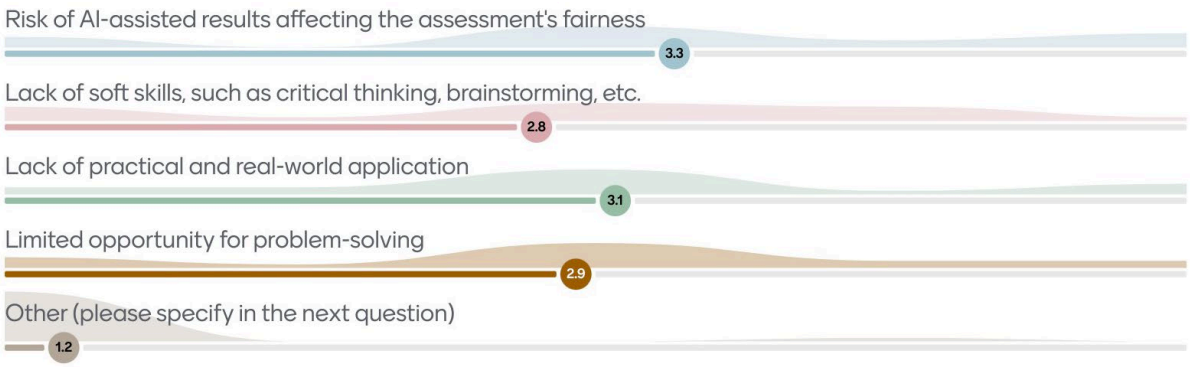

Figure 2: How challenging the given options are in demonstrating students' learning on a scale of 1-5.

Students also believe that the project works, and AI-assisted assignments are the most beneficial among the other types of oral examinations and real-time discussions, project oral presentations, and pitching ideas/products. In the question of whether students perceive that oral exams and supervised (proctored) exams would make assessments fairer, it was surprising to see that only 6 students thought that it is. 10 students have not even answered that question. This result reflects that there could be significant resistance among students to increasing oral and supervised assessment types in response to AI and cheating. Students also recognise the academic burden of conducting oral exams. The question about what types they think are harder to reflect their knowledge acquisition, written examinations (proctored), was one of them. Thus, a clear caution emerges here; simply increasing or restricting assessment types to only supervised, written exams is not a sufficient or effective solution.

In the open question about how other types of assessment approaches students view beyond that already mentioned, they focused on more of methodologies and processes for assessments: i.e., "*All types of exams give us some confidence and push to prepare ourselves better to deal with the exams. So in my opinion, AI is helping to learn things easily and in up-to-date way*". Another similar open response was "*Develop an AI-based examination system where AI agents loop into learning materials, exam policies, grading and generate dynamic questions based on the answers or explanations provided by the students*", so the students envision an interconnected facility of learning and assessment in contrast to typical views for standalone decisions for simply "using AI" based on cases or exam types. Another interesting response was "*In coding courses the best examination in my opinion is when the code is reviewed at an oral examination and the student has to explain the code and expected to answer questions about it*". What is interesting in these responses is that they are really in the "problem solving" mindset in their responses, deeply reasoning on the possible solution space and reflecting on learning opportunities, instead of a pure focus on the "problem space".

These outcomes justify the potential of students to provide impactful solutions when they immerse themselves in problematisation, problem-solving, and design thinking situations.

3.2 Deep dive outcomes

The deep dive activity is conducted in a future workshop (FW) with 22 students following the methodology described above. Of the 5 FW work groups, four have delivered their outcomes. Figure 3 shows two of those deliverables.

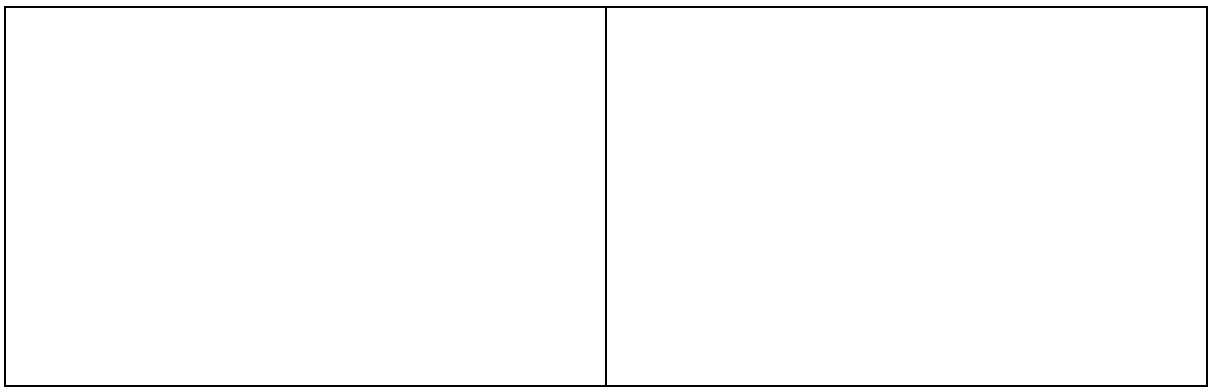

Figure 3. FW deliverables

According to the FW framework, problematisation, design, and implementation steps of a low-fidelity prototype are carried out. The summary of the delivered design processes is presented in Table 1.

Table 1: The prototypes delivered

G#	Critique	Fantasy	Implementation (MSCW)
1	1. Too easy to cheat on take -home exams/quizzes 2. Traditional Methods focus on memorisation 3. Memorisation won't get you far in work environments 4. Critical Thinking is not tested efficiently using traditional methods 5. Subjective marking for open-ended questions	1. Workshops to be able to discuss (Practical critical thinking) 2. Oral assessment (expanding how and why) 3. Use cases (Link concepts into real life contexts /situations 4. Pen & paper (focus on memorisation)	M : Use cases (#3 of Fantasy) S: Oral Assessment (#2) C: Workshops (#1) W: Pen & paper
2	1. Hard to evaluate all written assignment that are not fully supervised 2. Hard to measure the soft skills of writing/ being creative and analytical thinking 3. Same assessments feel uninspiring and that students feel that they just go through the motions. No problem solving. Feels like predetermined path	1. Students being assessed by solving a larger more difficult challenge/task while being allowed to use any means or tools available with not necessarily one correct answer/solution 2. More discussion /opposition-based assessment, not presentations. Students have to verbally defend their work individually 3. Teachers need to hold students accountable for their learning 4. There is a clear and good purpose with every assignment /assessment	M: 1. No definitive answers 2. Hold students accountable for their work and learning 3. Clear purpose with all assignments S: Larger tougher, projects allowing AI usage C: Examinations in MA so no AI is available W: Access to AI in exams
3	1. The assessment could be considered unfair (in cases of AI-use) 2. No tools for identifying AI-generated text. 3. Grading inflation - the grades get better because student use AI and therefore also the requirements for higher grades increase 4. If students use AI to complete school work, there is a risk that their soft skills, such as collaboration, communication and problem solving, won't develop	1. Real-time assessments 2. Points for attendance in workshops / seminars 3. Encourage participation in case exercises and workshops 4. Create an AI tool similar to Urkund that can analyze whether the text is AI-generated without mistakes 5. More oral examinations as they make it impossible for AI to influence the outcome.	1. Develop a standardized format for real-time assessments (e.g., oral exams, live problem-solving,) 2. Encourage active participation through structured discussions, workshops, quizzes, or reflections to develop soft skills 3. Some form of opposition from the grading teacher regarding the submitted work (exam, written assignment) to allow the student to prove that they can stand by what they have written and that it represents their own thoughts and knowledge - Negative aspect: very time consuming
4	1. Students may be less informed of the course material because of the functions of AI 2. Harder for teachers to assess the level students are during examinations 3. Students risk losing/stagnate in their critical thinking 4. Almost all forms of examinations taken at home have the possibility to be manipulated by AI 5. LLMs can be used to help with programming writing text and proposing solutions to exam questions 6. Students that use AI tools during their studies face the risk of not being able to conduct future work tasks in a way that reflect their own ideas given by ChatGPT etc. 7. Lack of integration of AI tools in classroom or assignments	Supervised examinations where the students are able to explain their way of thinking when presenting projects, they may have written at home while still using AI	Could use AI tools to create a collaboration assignments where they work together. But importantly with predefined goals for the human and AI

In answering the second research question of how students envision the assessment methods that can reflect students' learning, knowledge, and skills while promoting fairness in grading in higher education, firstly, the identified critiques are transformed into design baseline requirements, which serve as the guiding principles of the solution space as follows.

1. Authenticity of the assessments: Move beyond mere memorisation to real-world, problem-based scenarios.
2. Accountability: Ensure students can defend and demonstrate their work, including the thinking process.
3. Include both skills and competencies in the assessment: Assess both hard skills (knowledge, technical ability) and soft skills (critical thinking, communication, collaboration).
4. Fairness & Integrity: Ensure to minimise the AI-related cheating risks while acknowledging legitimate AI use.
5. Improve Engagement: Create meaningful assessments with clear purposes to prevent "going through the motions."

3.2.1 Design Elements (Fantasy mapped to Implementation prioritised via MoSCoW)

From the design concepts identified in the prototypes, the following strategies are derived according to the MoSCoW priorities the students envisioned (cf. Table 2).

Table 2: The strategies for the assessments

MoSCoW Priority	Assessment Strategies	Purpose
Must Have	1. Real-life use case tasks linked to professional practice. 2. Oral/real-time assessments (presentations, live problem-solving, viva voce). 3. Larger, open-ended projects with no single "correct" answer. 4. Explicit accountability checks (student defends their work). 5. Clear objectives for every assignment.	Ensure authenticity, prevent over-reliance on AI, promote deep learning.
Should Have	1. AI-permitted assignments with guidelines and transparency. 2. Structured opposition/discussion instead of static presentations. 3. Active participation grading via workshops, case exercises. 4. Collaboration tasks integrating AI under controlled conditions.	Encourage productive AI use while building soft skills.
Could Have	1. Periodic supervised exams without AI tools. 2. AI-detection systems to flag suspicious work (similar to plagiarism checkers). 3. Teacher-led "challenge" questions post-submission to validate authorship.	Strengthen integrity and safeguard fairness.
Would Have	1. Purely memorisation-based pen-and-paper tests without real-world context. 2. Fully unsupervised AI-free exams without alternative verification methods.	Avoid outdated, low-impact assessment methods.

3.2.1 Implementation Pathway

Lastly, from summaries of the implementation priorities in Table 2, triangulated with the design principles derived from the critiques, the following pathways have emerged.

- **Integrate** diverse assessments: combine real-time oral, collaborative, and AI-enabled tasks to create a balanced evaluation.
- **Embed** accountability moments: post-submission viva, peer opposition, or structured debates.
- **Design** for transparency: make AI use permissible in certain contexts, but require process documentation.
- **Prioritise** soft skill development: Wherever possible, add communication, collaboration, and creativity measurements as explicit assessment criteria.
- **Iterate & evaluate:** pilot in small courses, gather feedback, adjust weighting of AI-permitted vs AI-free components.

3.3 Discussion: outcomes impact on formative and summative assessment practices

Generative AI applications are shown to be risky to adopt, especially when it is used for assessment-related activities. Academic dishonesty and over-reliance on AI have been highlighted as negative impacts of AI in several studies (Dempere et al., 2023; Gopane et al., 2024; Kasneci et al., 2023). This study instrumented a deep dive study into students' perceptions on the challenges and opportunities of AI in examinations, and echoes the fact that universities are in need to adapt to innovative assessments, as resisting them may lead to disruption (Lucas, 2016). There are also abundant examples of prior research showing how students can enhance their learning with the use of GAI-based tools (Dempere et al., 2023; Kasneci et al., 2023), which is recognised by the students

of this study. Essentially, the generative tools enable students to generate content effortlessly, which means they challenge the integrity of conventional assessments such as take-home and online examinations and assignments (Akintande, 2024; Gopane et al., 2024), but the rapid adoption of GAI in educational settings is raising concerns about academic integrity and the effectiveness of traditional assessment practices (Dempere et al., 2023; Kasneci et al., 2023).

The Low-fidelity prototypes from FW showcased potential design solutions in response to these challenges, specifically, AI-resistant assessment strategies, such as connecting examinations with real-world problems and progressive tasks monitored and assessed using multiple or combined methods, i.e., project work, peer review, oral presentations, etc. A process-product assessment approach, as such, demonstrates the transformative influence of GAI on assessment practices according to the FW outcomes. The innovative potential of such a process framework is also explored in related research (Awadallah Alkouk & Khlaif, 2024; Nartgün & Kennedy, 2024). It is also a fact that the integration of AI in educational measurement has altered assessment methods, and enabling automated scoring, rapid content analysis, and personalised feedback through machine learning and natural language processing also raises significant ethical concerns regarding validity, reliability, transparency, fairness, and equity (Bulut et al., 2024).

As a response to the challenges of academic integrity, assessment fairness, limited diversity in assessment formats, and the lack of robust methods for evaluating soft skills, higher education institutions must critically rethink assessment models. They must be resilient against AI-assisted misconduct while effectively assessing both disciplinary knowledge and essential skills such as critical thinking, problem-solving, communication, and collaboration (Aghaee & Karunaratne, 2023). The outcomes of the FW provided insights for priorities: **must-have approaches** include embedding real-world use cases, oral and real-time assessments, and clear accountability for students' learning; **should-have approaches** emphasize challenging, open-ended projects where AI can be used as a tool under guided conditions; **could-have approaches** involve workshops and structured discussions that foster participation and develop soft skills; while **would-like-to-have approaches** include minimising traditional pen-and-paper elements with specific memorisation purposes. By integrating these priorities into assessment design, educators can not only safeguard academic integrity but also create richer, more diverse, and student-centred assessment environments that foster deeper learning and transferable skills.

4. Concluding remarks

This research focused on a deep exploration into AI-resilient assessment through the perspectives of higher education students, addressing two core questions: the significance of authentic, skill-focused methods such as oral examinations and AI-integrated tasks, and how students envision diverse assessments in future. The outcomes regarding the current assessment practices, students perceived authentic and compound assessments as a part of a continuous process, would mitigate the challenges of the AI-influenced academic misconduct. The low-fidelity prototypes designed by students in FW revealed potential implementation pathways to future AI-resilient assessments. These strategies can be piloted, iteratively refined, and scaled to build assessment systems that encourage integrity while developing deeper, transferable learning. Moving forward, higher education institutions should view such designs not only as a safeguard against misconduct but as an opportunity to cultivate more authentic, equitable, and future-ready learning environments as a step for transformation in education.

Ethics declaration

This study followed the ethical standards and research integrity guidelines of the host university, with confirmation from the head of unit. As the research was conducted within an educational context and did not involve the collection of sensitive personal data, no formal ethical approval was required. Participants were informed about the purpose of the study, its potential for a research publication, and how their data would be handled. Participation was voluntary, anonymity was ensured by excluding personal identifiers, and informed consent was obtained through their submissions, with the option to withhold material from inclusion in the study.

AI declaration

No AI tools were directly used in conducting the research. Language correction tools were employed during the development of the workshop and the writing of the manuscript to proofread and correct language errors.

References

Aghaee, N., & Karunaratne, T. (2023). Soft Skills Demand and Supply Through the Lens of Higher Education Students. European Conference on E-Learning, 22(1), 1–10.

Akintande, O. J. (2024). Artificial versus natural intelligence: Overcoming students' cheating likelihood with artificial intelligence tools during virtual assessment. Future in Educational Research, 2(2), 147–165. https://doi.org/10.1002/fer3.33

Alam, A. (2021). Possibilities and Apprehensions in the Landscape of Artificial Intelligence in Education. 2021 International Conference on Computational Intelligence and Computing Applications (ICCICA), 1–8. https://doi.org/10.1109/ICCICA52458.2021.9697272

Awadallah Alkouk, W., & Khlaif, Z. N. (2024). AI-resistant assessments in higher education: Practical insights from faculty training workshops. Frontiers in Education, 9, 1499495.

Bulut, O., Beiting-Parrish, M., Casabianca, J., Slater, S., Jiao, H., Song, D., Ormerod, C., Fabiyi, D., Ivan, R., Walsh, C., Rios, O., Wilson, J., Yildirim-Erbasli, S., Wongvorachan, T., Liu, J., Tan, B., & Morilova, P. (2024). The Rise of Artificial Intelligence in Educational Measurement: Opportunities and Ethical Challenges. Chinese/English Journal of Educational Measurement and Evaluation | 教育测量与评估双语期刊, 5(3). https://doi.org/10.59863/MIQL7785

Dempere, J., Modugu, K., Hesham, A., & Ramasamy, L. K. (2023). The impact of ChatGPT on higher education. Frontiers in Education, 8. https://doi.org/10.3389/feduc.2023.1206936

Gamage, K. A. A., Dehideniya, S. C. P., Xu, Z., & Tang, X. (2023). ChatGPT and higher education assessments: More opportunities than concerns? Journal of Applied Learning and Teaching, 6(2), 358–369. Scopus. https://doi.org/10.37074/jalt.2023.6.2.32

Gopane, S., Kotecha, R., Obhan, J., & Pandey, R. K. (2024). CHEAT DETECTION IN ONLINE EXAMINATIONS USING ARTIFICIAL INTELLIGENCE. ASEAN Engineering Journal, 14(1), Article 1. https://doi.org/10.11113/aej.v14.20188

Jungk, R., & Müllert, N. (1987). Future workshops: How to create desirable futures. Inst. for Social Inventions.

Karunaratne, T., Axelsson, P., Lindblad, E., & Enoksson, F. (2025). THE DEVIL IS IN THE DETAILS: DIGITAL ASSESSMENT OPPORTUNITIES AND CHALLENGES IN HIGHER EDUCATION - A CASE STUDY FROM SWEDISH ENGINEERING EDUCATION. EDULEARN25 Proceedings, 5060–5070. 17th International Conference on Education and New Learning Technologies. https://doi.org/10.21125/edulearn.2025.1279

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Lucas, B. (2016). A Five-Dimensional Model of Creativity and its Assessment in Schools. Applied Measurement in Education, 29(4), 278–290. https://doi.org/10.1080/08957347.2016.1209206

Nartgün, Z., & Kennedy, E. (2024). Artificial Intelligence (AI) and Education: Contributions Opportunities and Challenges. Political Economy and Management of Education, 5(2), Article 2.

Ouaazki, A., Bergram, K., & Holzer, A. (2023). Leveraging ChatGPT to Enhance Computational Thinking Learning Experiences. 2023 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE), 1–7. https://doi.org/10.1109/TALE56641.2023.10398358

SkillsMatch. (2018). SkillsMatch | Developing a prototype that helps users adapt their non-cognitive skills to the demands of the market. http://skillsmatch.eu/

Spinuzzi, C. (2005). The Methodology of Participatory Design. Technical Communication, 52(2), 163–174.

Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. Journal of Applied Learning and Teaching, 6(1), Article 1. https://doi.org/10.37074/jalt.2023.6.1.17

Thorne, S. L. (2024). Generative artificial intelligence, co-evolution, and language education. The Modern Language Journal, 108(2), 567–572. https://doi.org/10.1111/modl.12932

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learning Environments, 10(1), 15. https://doi.org/10.1186/s40561-023-00237-x

Vidal, R. V. V. (2006). The future workshop: Democratic problem solving. Economic Analysis Working Papers, 5(4), 21.

Waters, K. (2009). Prioritization using moscow. Agile Planning, 12, 31.

Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. International Journal of Educational Technology in Higher Education, 21(1), 40. https://doi.org/10.1186/s41239-024-00468-z