# Assessment and learning outcomes for generative AI in higher education: A scoping review on current research status and trends

**Xiaojing Weng**
Department of Curriculum and Instruction, The Education University of Hong Kong, Hong Kong, China

**Qi Xia**
College of Education, Zhejiang University, Hangzhou, China

**Mingyue Gu**
Department of English Language Education, The Education University of Hong Kong, Hong Kong, China

**Kumaran Rajaram**
Nanyang Business School, Nanyang Technological University, Singapore

**Thomas K. F. Chiu**
Department of Curriculum and Instruction, The Chinese University of Hong Kong, Hong Kong, China

Generative artificial intelligence (GenAI) impacts higher education assessment and learning outcomes, which are closely related and intertwined. Literature suggests that educators and researchers have many varied concerns regarding student assessment in the higher education GenAI context, such as how to assess students' learning and the new (refocused) learning outcomes that emerged in GenAI-facilitated learning environments. To provide evidence-based insights into and answers to these concerns, we conducted a scoping review by collating literature in relevant research areas. Following a five-stage scoping review framework, we collaboratively collected and coded 34 studies. The three assessment approaches identified in the review were traditional assessment, innovative and refocused assessment and GenAI-incorporated assessment. The new, refocused learning outcomes identified were career-driven competencies and lifelong learning skills. The review also revealed that most research designs were qualitatively oriented (e.g., with exploratory design, descriptive research, ethnographic research and phenomenological research). This study proposes a holistic diagram showing the current research status and trends. It suggests five future research directions: innovative assessment designs, collaborations among assessment approaches, new learning outcomes, relationships between assessment approaches and learning outcomes, and quantitative or mixed research studies.

*Implications for practice or policy:*

- Traditional assessment methods in higher education do not operate effectively in the GenAI era.
- Innovative and refocused assessment and GenAI-incorporated assessment are promising strategies to assess student learning.
- Career-driven competencies and lifelong learning skills are new focused learning outcomes evolved from the use of GenAI.
- More quantitative and mixed research studies should be conducted to provide additional empirical evidence on the impact of GenAI on student assessment and learning outcomes.

*Keywords:* generative AI, assessment, higher education, learning outcomes, scoping review

## Introduction

Generative artificial intelligence (GenAI) will revolutionise higher education (Baidoo-Anu & Ansah, 2023; Chen & Zhu, 2023; Chiu, 2024) by creating opportunities and challenges for teaching, learning and assessment (Michel-Villarreal et al., 2023; Ruiz-Rojas et al., 2023). Students could use GenAI technologies like ChatGPT or Sora to easily create essays, images and videos for their assignments and projects. They could complete assigned tasks that they initially thought they were unable to perform or explore different perspectives in their learning process (Kasneci et al., 2023; Ruiz-Rojas et al., 2023). For example, a research student could easily obtain an informed script from GenAI to complete a statistics project or get varying views to enhance their initial ideas. However, inappropriate use of or overreliance on this technology poses challenges for both teachers and students (Chiu, 2023; Kasneci et al., 2023). For instance, students might simply generate answers to complete their assignments instead of exploring and learning from different ideas or perspectives. Additionally, it is difficult for teachers to distinguish GenAI-generated work from student-generated work. These opportunities and challenges urge us to revisit learning outcomes, new literacies, pedagogies and assessment to address the evolving challenges emerging from GenAI usage (Dobrin, 2023). Among them, assessment emerges as a key topic to which higher education teachers and researchers must pay attention (Hodges & Kirschner, 2024).

Assessment plays a pivotal role in students' learning processes and learning outcomes (Bryan & Clegg, 2019; Gikandi et al., 2011; Ibarra-Sáiz et al., 2021; Lara et al., 2020; Rajaram, 2023), and its design should consider feedback approaches, student participation and self-regulation (Ibarra-Sáiz et al., 2021, Rajaram, 2021). A well-designed assessment not only measures students' understanding but also guides them in identifying their strengths and areas for improvement, thereby facilitating a more efficient and effective learning experience (Bryan & Clegg, 2019; Makransky et al., 2019; Rajaram, 2023). GenAI could provide personalised and real-time feedback experiences, adapting to each student's learning capacity and pace (Chiu, 2024; Su & Yang, 2023). Accordingly, the use of GenAI in assessment could potentially eliminate the one-size-fits-all approach, ensuring equitable access to an elevated level of quality education (Mintz et al., 2023). However, integrating GenAI in assessment is an innovative intervention that has transformed previous student learning approaches because assessment affects learning approaches. This has led to assessment challenges due to the mismatch between the innovative learning approach and its traditional assessment designs (Hodges & Kirschner, 2024; Luo, 2024).

There is an emerging discussion associated with assessment, including assessment methodologies (AI-based), and focus of current assessment (i.e., authentic), learning outcomes (e.g., AI literacy, critical thinking and creativity) and approaches to handling academic integrity (Alasadi & Baiz, 2023; Bannister et al., 2023b; Kasneci et al., 2023; Michel-Villarreal et al., 2023). However, there is insufficient research into assessment-related activities for higher education teachers and researchers, and few conclusions have been drawn based on relevant studies. Addressing the assessment-related challenges in higher education within the context of GenAI is an urgent research topic (Chiu, 2024) that requires more attention. To address this gap, this review study explores how to transform through (re)imagining and (re)designing assessment in higher education for GenAI-empowered learning environments.

## Literature review

### Assessing GenAI-facilitated learning

Student assessment is an integral part of the education system, aiding in the continuous improvement of teaching methods and learning outcomes. The goal of assessment is relatively stable – it enhances the learning and teaching process with intended learning outcomes, ensures accountability and prepares students for future challenges (Lau, 2016; Pereira et al., 2016; Rajaram, 2023; Rubin et al., 2004). Assessment practice needs to be matched with evolving education contexts (e.g., with GenAI integrated) to achieve high efficacy in facilitating educational innovation, attaining quality assessment and obtaining student assessment goals. This is an urgent issue that needs to be resolved and should be given much attention. Concerns about student assessment in GenAI contexts are particularly noticeable (Alasadi &

Baiz, 2023; Michel-Villarreal et al., 2023). These concerns arise from the mismatch between innovative GenAI-integrated education practices and traditional approaches to assessing students' learning, such as exams or written essays (Chaudhry et al., 2023; Dobslaw & Bergh, 2023; Nikolic et al., 2023; Yilmaz & Yilmaz, 2023).

Essentially, how to assess student learning when GenAI is used is a major concern in higher education (Chiu, 2024; Luo, 2024). GenAI enables students to access knowledge and information effortlessly; however, it also reduces the effort students must make to acquire and produce knowledge, which can lead to assessment concerns. Furthermore, GenAI's capacity to mimic human responses empowers students to generate content, which may increase the chances of students cheating on assigned coursework, tasks and exams (Michel-Villarreal et al., 2023). This raises concerns about academic integrity. As a result, the integration of GenAI in learning affects the assessment outcomes; essay-based assignments may not reflect student creativity when students directly submit GenAI's output as the assignment. Consequently, the design of assessment may lose its functions. This challenges teachers and educational institutions to guarantee equitable assessment and uphold academic integrity (Cotton et al., 2023). Moreover, as assessment drives learning outcomes, changes in assessment may lead to new or refocused learning outcomes.

## Learning outcomes in GenAI-facilitated environments

Current uses of AI in education, as elucidated by researchers, include adaptive learning platforms (Mavroudi et al., 2018), intelligent tutoring systems (Mousavinasab et al., 2021) and AI-driven administrative tasks (Ahmad et al., 2022). Researchers suggest that students can benefit from AI-assisted learning as these systems can provide automatic feedback, support intelligent tutoring and facilitate personalisation (Hwang et al., 2020; Weng & Chiu, 2023; Zhang & Aslan, 2021). As research in AI for education is ascendant, pivoting the focus to GenAI could potentially enhance these capabilities further. For example, GenAI enables students to engage in more creative, exploratory and self-guided learning processes (Ruiz-Rojas et al., 2023). Furthermore, some researchers recommend designing innovative and higher-order thinking tasks to enable students to develop essential thinking skills and competencies vital to their future development, including creativity and critical thinking (Alasadi & Baiz, 2023; Michel-Villarreal et al., 2023; Ruiz-Rojas et al., 2023). These efforts resemble the potential learning outcomes of GenAI-facilitated education.

However, there are concerns and ambiguity concerning what learning outcomes should be assessed in a GenAI-facilitated learning environment (Chiu, 2023; Hodges & Kirschner, 2024). The implementation of GenAI could redefine student learning performance and skills as knowledge creation becomes increasingly decentralised, collaborative and technology-driven (Luo, 2024). Meanwhile, learning processes could begin to emphasise new outcomes such as AI literacy and the necessary knowledge for GenAI learning, including critical reasoning and digital, media and information literacies (Chiu, 2023, 2024). Defining these new focal learning outcomes and determining the essential higher-level learning outcomes to be assessed in GenAI contexts are desirable research topics in this area. Moreover, a GenAI learning environment could potentially eliminate human interaction, unlike students' previous experiences (Michel-Villarreal et al., 2023). The skills and performance on which students should be assessed in such a dehumanised learning environment are unprecedented. However, the assessment criteria used to define these newly emerged learning outcomes in GenAI learning contexts can determine which skills require emphasis (Chiu, 2024). Though students in higher education are expected to acquire the skills needed to function effectively in their future workplaces (Abelha et al., 2020; Suleman, 2018), there is no definitive solution regarding how to assess GenAI-facilitated learning to foster a diverse range of future skills.

## Existing review studies

To date, review studies have focused on changes in or transformations of assessment in higher education influenced by GenAI tools. For example, Baidoo-Anu and Ansah's (2023) review highlighted that the GenAI tool ChatGPT could generate formative assessment prompts for improved teaching and learning. They suggested that educators need to revise current assessment methods and adopt more innovative

approaches. Similarly, Bahroun et al. (2023) examined the transformative impact of GenAI across disciplines. Their bibliometric study highlighted the need for revised assessment guidelines, emphasising the importance of refining AI models for precision, addressing biases, ensuring accuracy and setting firm ethical guidelines for AI's role in educational assessments. Besides, Bannister et al.'s (2023b) systematic review of GenAI in higher education highlighted its role in learning assessment and feedback and suggested that future research focus on GenAI-boosted evaluation and AI-aided self-assessment. In addition, Farrelly and Baker's (2023) review emphasised the need for fairness and equity in GenAI-based assessments. They encouraged students to become developers of AI tools instead of mere consumers and to collaborate with educators to develop resources regarding the effective use of GenAI tools in assessment tasks. Overall, review studies have defined assessment as a versatile application for GenAI and proposed possible remedies for its relevant and contextualised issues. However, these studies did not discuss how the changes in assessment could impact students' learning outcomes.

To enrich the relevant literature, we conducted a review study of assessment practices in the context of higher education GenAI (Bearman et al., 2024). Since GenAI in higher education is in its infancy phase, we used a scoping review in our study. Among the different types of review studies, scoping review is a method of accumulating and synthesising knowledge that employs a systematic and repetitive process to identify and summarise existing or emerging literature on a specific subject (Pham et al., 2014). Different from systematic reviews that respond to specific questions with established methodologies to evaluate the quality of articles, a scoping review is applicable in areas where there is a lack of solid evidence and may include literature from a wide variety of study designs (O'Flaherty & Phillips, 2015).

## Methodology

We adopted the five-stage framework proposed by Arksey and O'Malley (2005) to conduct the scoping review (as shown in Table 1). This framework features a rigorous procedure for ensuring openness and clarity to increase the search process's transparency. The search strategy guided by the framework can be replicated, which enhances the reliability of the research findings.

Table 1
*Five-stage scoping review framework, adopted from Arksey and O'Malley (2005)*

| Stage | Instruction |
|---|---|
| Identify research questions | This stage sets the path for the next steps. It is crucial to clearly specify relevant issues around the question, as they influence the search strategies. Research questions are typically wide-ranging as they aim to offer comprehensive coverage. |
| Identify relevant research | At this phase, the relevant research is identified, and a plan is created on where to look, what terms to use, which resources to explore, the time frame and the language. It is essential to have a wide and comprehensive search for the studies. |
| Select study | Selecting studies involves the implementation of criteria for inclusion and exclusion. These criteria are dependent on the specific details of the research question and the newfound understanding of the topic gained from reviewing the studies. |
| Chart the data | A form for charting data is created and utilised to gather information from each study. A narrative review or descriptive analytical approach is employed to obtain context-based or process-focused data from each study. |
| Collate, summarise and report results | An analytical framework or theme-building is applied to give a broad view of the literature. A quantitative analysis of the volume and characteristics of studies is displayed using tables and charts. A thematic analysis follows. It is crucial to maintain clarity and consistency when presenting the findings. |

## Identification of research questions

Our review centred on investigating how the use of GenAI impacts assessment and learning experiences in higher education. We formulated the following three research questions to direct our exploration, ensuring a comprehensive collection of literature related to the research focus:

- How do higher education teachers assess student learning when GenAI is used?
- What new focused learning outcomes have evolved from the use of GenAI?
- What research methods are used to investigate assessments when GenAI is used?

## Identification of relevant research

A broad range of keywords should be utilised as search terms to achieve extensive coverage of the available literature (Arksey & O'Malley, 2005). We formulated the main search strings to gather literature related to GenAI in higher education, evaluating from a global perspective lens. At this stage, we sought suggestions from a university librarian, who helped refine the search terms and pinpoint databases that were most likely to yield the desired results. We used the Boolean operators to refine, broaden and amalgamate the search areas. The key search terms designed for the search were ((“generative AI” OR “GenAI” OR “ChatGPT”) AND (“undergraduate*” OR “universit*” OR “higher education” OR “tertiary”)).

We designed the criteria for research inclusion and exclusion (as shown in Table 2), aiming to be as thorough as possible in identifying the most relevant and primary evidence while considering the practical constraints of time and financial resources. However, to include as many target studies as possible, we did not set any limitations on the publication’s date or cultural background.

Table 2
*Criteria for research inclusion and exclusion*

| | |
|---|---|
| Inclusion criteria | |
| (1) | Peer-reviewed journal papers |
| (2) | The written language is English, and the full text is accessible |
| (3) | The research investigates the use of GenAI in higher education settings |
| (4) | The overwhelming themes of the study relate to student assessment and GenAI |
| (5) | The study focuses on the assessment of student learning, student learning outcomes or the research methods used to investigate student development in GenAI-facilitated learning environments |
| Exclusion criteria | |
| (1) | The paper is not peer-reviewed |
| (2) | The body text of the article is written in a language other than English, or the full text is not available |
| (3) | The study uses secondhand resources to conduct data analysis, such as literature review or meta-analysis |
| (4) | The research is not conducted in higher education settings |
| (5) | The study is not relevant to the assessment of student learning, student learning outcomes, nor the research methods used to investigate student development in GenAI-facilitated learning environments |
| (6) | The study has no sufficient description of student learning in GenAI-facilitated learning environments or corresponding assessment strategies |

To collect high-quality peer-reviewed literature, we applied the search strings to six representative electronic databases in educational technology: Web of Science, Scopus, ERIC, ProQuest education database, ACM and IEEE Xplore. Furthermore, the reference lists of the collected articles were manually searched to locate any other original sources within the grey literature space. The research-searching stage ended in April 2024.

## Selection of study

As shown in Figure 1, we followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement when collecting studies for the scoping review. We collected 2,347 articles from the databases using the described search strings. The first two of us authors (XW & QX) collaborated on study selection, reading the abstracts of the identified reports independently before voting on whether to keep the paper. When disagreements arose, they sought advice from the moderators (i.e., the third and the last of us authors) until a consensus was reached.

The abstract reviewing process showed that many articles were not relevant, especially those focusing on the use of GenAI for primary and secondary education. Those articles mainly dealt with the ethical issues of adopting GenAI in education, assessed the accuracy rates and capabilities of GenAI in producing the desired content for different subjects, focused on instructors' or students' perceptions of or attitudes towards the application of GenAI to facilitate their study or systematically reviewed papers and other non-peer-reviewed publications; as such, they were ruled out. We excluded some articles from the search because they were identified repeatedly across the different databases. After applying the inclusion and exclusion standards to all identified studies, we deemed 32 papers as relevant to the research subject. The first two of us authors worked collaboratively to double-check the obtained studies and to review and confirm the full-text versions. This procedure also allowed for the identification of any additional relevant literature by reviewing the reference lists of the collected studies. We included an additional two articles through this reviewing procedure as they met the inclusion criteria but were not published in our target databases.

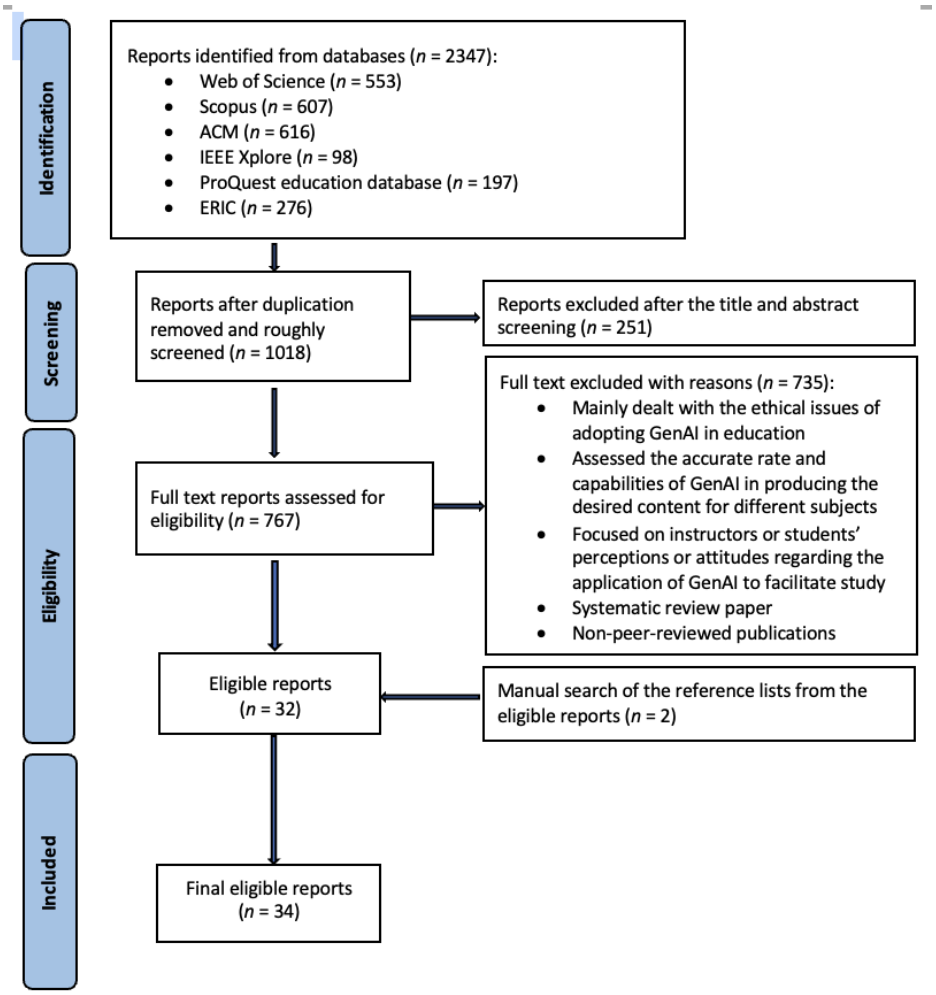


*Figure 1.* PRISMA flow chart for study collection

### Charting of the data

We utilised inductive methods when processing the data. Appendix 1 shows the specific summaries of the studies included (https://shorturl.at/mIcSZ). Each article was summarised from diverse perspectives, including its author, year, study location, intervention and resources, participants and sample size, assessment, student learning outcomes, research method, limitations and suggestions.

### Collation, summary and reporting of results

Compiling and presenting results is the fifth and final step of Arksey and O'Malley's (2005) scoping review process; see the next section.

## Results and implications

This scoping review resulted in 34 articles authored in various regions. Five of the studies were carried out in United Kingdom, four in Australia and two each in the United States of America, China, Canada, Sweden, Spain, Turkey and Saudi Arabia. Some other countries each contributed one article (e.g., Finland, the Netherlands, Singapore, Taiwan, Malaysia). Notably, insufficient literature originated from Asia or developing countries regarding how the application of GenAI impacts assessment in higher education. The articles cover a range of topics related to GenAI in higher education, such as teacher assessment approaches, emerging learning outcomes and research methods for assessing student learning. We reviewed the articles and concluded that they could answer the three main research questions.

### Research question 1: How do higher education teachers assess student learning when GenAI is used?

Though some studies (5/34) did not provide sufficient information regarding this research question, we gathered that higher education teachers generally used diverse ways to assess student learning in GenAI-facilitated settings. We categorised these assessment strategies into three approaches: traditional assessment, innovative and refocused assessment and GenAI-incorporated assessment.

*Traditional assessment approach*

In 10/34 of the reviewed articles, university teachers adopted long-existing assessment methods traditionally adopted in higher education, including quizzes, exams, assignments, case studies, projects and presentations (Chaudhry et al., 2023; Farazouli et al., 2023; Yilmaz & Yilmaz, 2023); in-class assignments that are difficult for AI to solve (Nikolic et al., 2023; Rajabi et al., 2023); open-book exams and home exams (Dobslaw & Bergh, 2023); in-person closed-book written quizzes (Rajabi et al., 2023); and observation (Yan, 2023). We found that teachers also used the same methods to assess student learning in the GenAI-facilitated learning environments; in other words, they did not change how they assessed students' learning.

*Innovative and refocused approach*

Some university educators (8/34) utilised strategies that incorporated interactive and creative elements to assess students' learning in GenAI-facilitated learning environments. Among them, some educators redesigned traditional assessment methods to encourage students' higher-order thinking skills (e.g., critical thinking and creativity) (Binh Nguyen et al., 2023; Kirwan, 2023; Salinas-Navarro et al., 2024). The assessment methods included personalised assessments in which students were challenged by customised exams or quizzes generated by GPT (Banihashem et al., 2024; Cotton et al., 2023; Günay et al., 2024); self-assessments in which students decided their next actions after getting varying perspectives from ChatGPT (Ali et al., 2024; Almasre, 2024; Chang et al., 2023); generic skill assessments in which students demonstrated critical thinking, problem-solving and communication skills (Almasre, 2024; Banihashem et al., 2024; Bearman et al., 2024; Blanke, 2024; Cotton et al., 2023; Fleckenstein et al., 2024); and open-ended assessments in which students were encouraged in aspects of originality and creativity (Eager & Brunton, 2023; Vicente-Yagüe-Jara et al., 2023). Overall, in this approach, the teachers addressed

challenges by redesigning their assessment methods, refocusing on the learning outcomes needed by students and emphasising and facilitating self-regulated learning.

*GenAI-incorporated assessment approach*

Some studies (11/34) reported that students participated in diverse GenAI-incorporated learning assessment activities. This assessment approach requires students to work with GenAI by interacting with the feedback and information provided by the platform. For instance, Exintaris et al. (2023) and Clark (2023) found that students needed to critique the ideas and solutions given by ChatGPT for their greater improvement. In French et al.'s (2023) study, students worked with GenAI tools to complete game development tasks, wherein they needed to demonstrate working solutions with GenAI and report their findings. In Mohamed's (2023) study, instead of traditional written summative assessments, students used ChatGPT to complete group-based projects that showed their communication and collaboration skills. Additionally, in Vartiainen and Tedre (2023), students used text-to-image GenAI tools to demonstrate their creative thinking during a hands-on workshop. These studies demonstrate the incorporation of GenAI into learning assessment tasks, representing a future direction that researchers could consider for GenAI's further development and more in-depth utilisation.

*Implications*

These three approaches indicate that assessment practices in higher education are diversified, but the development of changes in these assessments is a slow process. GenAI transforms how assessments are performed and redefines the intended learning outcomes; teachers need to redesign and embrace changes in their assessments (Chiu, 2024; Kasneci et al., 2023; Rajaram, 2023). Some teachers have already designed or redesigned new assessments to embrace the opportunities and address the challenges of GenAI-facilitated learning environments. The findings showed that many assessments implemented were still traditional – that is, teachers did not change their assessment methods, continuing to rely heavily on approaches favouring rote memorisation (e.g., quizzes and exams) and avoiding using GenAI. These traditional assessment approaches disregarded the benefits GenAI brings to student's long-term growth (e.g., self-regulated learning skills and higher-order thinking) and can have detrimental consequences (e.g., diminishing academic integrity, aggravating educational inequity, weakening critical thinking and increasing dehumanisation) when students have access to GenAI (Cotton et al., 2023; Michel-Villarreal et al., 2023). Overall, we concluded from our analysis that traditional assessment methods do not operate effectively in the GenAI era (Kasneci et al., 2023).

The last two approaches revealed that university teachers could revolutionise assessment methods for students, addressing assessment challenges in a GenAI-facilitated learning environment (Chiu, 2023; Kasneci et al., 2023, Salinas-Navarro et al., 2024). The findings suggested that teachers can refocus assessment designs by adopting more innovative practices that incorporate interactive and creative elements. Such strategies facilitate a more creative, exploratory and self-guided learning process that is beneficial for students (Ruiz-Rojas et al., 2023). The findings also highlight GenAI tools' ability to offer feedback and information on students' assessments, enabling them to identify areas for improvement. This echoes Chiu's (2024) findings that GenAI can facilitate real-time assessments, providing immediate feedback and tailored interventions. Essentially, these two assessment approaches are possible strategies for eliminating mismatches between innovative GenAI-integrated education practices and traditional approaches in assessing students' learning (Chiu, 2023; Hodges & Kirschner, 2024; Luo, 2024). More relevant studies are needed to benefit the field.

### Research question 2: What new focused learning outcomes have evolved from the use of GenAI?

The usage of GenAI has resulted in a wide range of novel, focused learning outcomes, including career-driven competencies and lifelong learning skills.

*Career-driven competencies*

Career-driven competencies are often domain-specific and directly related to the study programmes or students' desired profession route, such as subject-specific knowledge, technical skills or ethics and

values. Many reviewed articles (10/34) revealed that developing students' domain-specific competencies is a major learning outcome. Particularly, some articles reported that their outcomes were to develop students' programme coding competencies, including improving their game development programming skills (French et al., 2023), reducing their writing blocks in programming (Rajabi et al., 2023) and improving their programming self-efficacy and motivation (Yilmaz & Yilmaz, 2023). Some studies were concerned with enhancing students' writing skills using GenAI. For instance, Vicente-Yagüe-Jara et al. (2023) revealed that a GenAI tool improved student creative writing in language teaching, while Yan (2023) suggested that adopting ChatGPT improved the efficiency of student writing in an English as a Second Language course. It could be because students were inspired by GenAI's outputs and added the output to their original idea. The students were able to obtain a different perspective from GenAI in their creative writing. Some other articles proposed developing specific competencies such as work quality assurance capability (Bearman et al., 2024), real-world engineering thinking (Kong et al., 2023) and crafting with text-to-image generative models (Vartiainen & Tedre, 2023), as major intended learning outcomes of GenAI-incorporated learning. Additionally, one study reported that developing ethics and social responsibility values was critical for professional and career development (Chaudhry et al., 2023). Specifically, how students valued ethics in using GenAI tools (e.g., ChatGPT) should be an intended and crucial learning outcome. This focus on ethical considerations stems from the dual nature of GenAI as both a powerful educational resource and a potential source of misuse. By embedding these practices into the curriculum, teachers can help students appreciate the ethical dimensions of using GenAI tools, ultimately preparing them to harness these technologies in a responsible and informed manner.

*Lifelong learning skills*

Most articles (25/34) reported that students developed lifelong learning skills in the context of GenAI. Lifelong learning skills are relatively broader and more general than career-driven competencies, usually involving enhancing one's understanding of the world, fostering personal development and improving interpersonal relationships. These skills, such as problem-solving and emotional related competence, are foundational and can be applied across various contexts and throughout an individual's life. They can benefit not only students' professional lives but also their personal growth and social interactions.

GenAI has the capacity to transform lifelong learning through the provision of personalised, adaptable and easily accessible educational experiences. GenAI can assist individuals in continuously enhancing their relevant competencies throughout their careers as the labour market changes and the need for new skills grows. GenAI not only improves the learning experience but also enables individuals to embrace lifelong learning as a crucial part of personal and professional growth.

Higher-order thinking skills are predominantly reported as lifelong learning outcomes in most collected studies (22/34). The reviewed articles reported various skills that students have been exposed to and have developed in GenAI-facilitated learning environments. For example, some studies reported student creativity (Eager & Brunton, 2023; Nikolic et al., 2023; Smolansky et al., 2023; Vicente-Yagüe-Jara et al., 2023), reflective and feedback literacy – student ability to reflect their learning and give feedback to others (Lu et al., 2024; Salinas-Navarro et al., 2024) and critical thinking (Banihashem et al., 2024; Bannister et al., 2023a; Blanke, 2024; Chaudhry et al., 2023; Exintaris et al., 2023; Salinas-Navarro et al., 2024; Sánchez-Ruiz et al., 2023; Smolansky et al., 2023) as more focused intended learning outcomes (i.e., become more often and important). Noticeably, creativity was framed in diverse ways, such as including the dimensions of divergent and convergent thinking (Eager & Brunton, 2023) or from the aspects of fluency, flexibility and originality (Vicente-Yagüe-Jara et al., 2023). Meanwhile, some studies reported students' communication (Chaudhry et al., 2023; Michel-Villarreal et al., 2023; Mohamed, 2023) and collaboration skills (Chaudhry et al., 2023; Cotton et al., 2023; Mohamed, 2023; Sánchez-Ruiz et al., 2023) as new learning outcomes facilitated by the GenAI learning environment. Collaboration skills were cultivated in social learning settings such as teamwork (Chaudhry et al., 2023; Mohamed, 2023) and group work (Sánchez-Ruiz et al., 2023). Furthermore, several studies reported students' development of problem-solving skills (Chaudhry et al., 2023; Clark, 2023; French et al., 2023; Sánchez-Ruiz et al., 2023; Yilmaz & Yilmaz, 2023), in the form of problem conceptualisation, problem strategy and problem solution (Clark, 2023).

Further, a sizable minority of studies (6/34) included emotional related competence as a new focused intended learning outcome. For instance, some studies suggested that learning outcomes should include engagement and a sense of community (Cotton et al., 2023; Mohamed, 2023), while four studies suggested students' self-awareness was related to learning outcomes such as metacognition (Exintaris et al., 2023), self-regulation (Chang et al., 2023) and self-reflection (French et al., 2023; Nikolic et al., 2023).

*Implications*

Programming and writing emerged as two leading domains highly connected with students' career-driven competencies. Our findings evidenced the GenAI tools' outstanding ability to facilitate understanding of computer programming and natural language processing. Specifically, in programming, GenAI can assist with code generation, bug detection and the simplification of complex algorithms (French et al., 2023; Rajabi et al., 2023; Yilmaz & Yilmaz, 2023). In writing, it can aid with grammar checking, sentence structure and even content generation, improving writing efficiency and productivity (Barrett & Pack, 2023; Vicente-Yagüe-Jara et al., 2023; Yan, 2023). From a broader perspective, both programming and writing domains require high levels of precision, creativity and problem-solving skills – all of which can be enhanced by AI platforms (Luo, 2024). As GenAI tools become more prevalent in society, domain-specific skills in programming and writing are increasingly in demand, making them vital competencies for career advancement. This suggests the programming and writing domains' infinite development possibilities in the age of GenAI. Educators in these areas are encouraged to embrace the innovation's advantages and investigate its possible uses (Chiu, 2024; Xia et al., 2024).

According to the reviewed studies, developing students' lifelong skills is a more mainstream trend than developing their career-driven competencies. These studies, with their varied application contexts, highlight diverse lifelong-skills-development scenarios in GenAI-facilitated learning environments (Clark, 2023; Exintaris et al., 2023; French et al., 2023; Kolade et al., 2024; Nikolic et al., 2023; Sánchez-Ruiz et al., 2023; Yilmaz & Yilmaz, 2023). This aligns with the across-domain feature of lifelong skills (Kivunja, 2015). Additionally, among the two main categories of lifelong skills, students' higher-order thinking skills have received more attention than their emotional related competence.

Furthermore, creativity and critical thinking were the most frequently reported subcategories of higher-order thinking skills. This emphasis could potentially allow students to adapt better to evolving educational landscapes and equip them with skills that are applicable to the future world (Weng, Chiu, & Tsang, 2022; Weng, Cui et al., 2022). As illustrated, researchers have suggested implementing quality assessments to enable students to cultivate key cognitive abilities and proficiencies necessary for their growth and future-oriented prospects, such as creativity and critical thinking skills (Alasadi & Baiz, 2023; Binh Nguyen et al., 2023; Blanke, 2024; Michel-Villarreal et al., 2023; Ruiz-Rojas et al., 2023; Salinas-Navarro et al., 2024). We also re-iterate and advocate that teachers should design and improvise more assessment activities to enhance students' higher-order thinking skills in GenAI contexts.

### Research question 3: What research methods are used to investigate assessments when GenAI is used?

In the reviewed studies, researchers adopted diverse research methods to investigate student assessment in GenAI-facilitated learning environments. Some studies employed quantitative methods (7/34), including surveys, exams and tests, while other studies utilised a mixed method approach (12/34). However, the majority of the studies adopted qualitative methods (15/34). The research methods adopted can be further classified into three categories: exploratory design, descriptive research and ethnographic and phenomenological research.

*Exploratory design*

Most researchers (18/34) implemented exploratory design to enhance understanding of students' assessment in GenAI-facilitated learning environments (Fleckenstein et al., 2024; Günay et al., 2024; Kolade et al., 2024; Lu et al., 2024; Nikolic et al., 2023; Rajabi et al., 2023; Vartiainen & Tedre, 2023; Yan, 2023). Among them, the most frequently explored research issue was how educators and students perceived GenAI's impact on higher education students' assessment (Bannister et al., 2023a; Kolade et

al., 2024; Rajabi et al., 2023; Yan, 2023), which was addressed using qualitative focus group discussions. The researchers' exploration included the use of a strategic management framework (e.g., SWOT analysis) to investigate how ChatGPT responded to existing assessment prompts across domains and among subjects (Kolade et al., 2024; Nikolic et al., 2023). Moreover, researchers conducted exploratory case study using the research-creation approach to scaffold crafting with text-to-image generative models; their goal was to understand participants' perspectives on the possible advantages and hurdles of AI (Vartiainen & Tedre, 2023). These exploratory studies delved into previously unexplored aspects of student assessment in the context of GenAI. They provided fundamental insights, uncovered new findings and shaped preliminary theories or hypotheses. They can serve as foundational research for more definitive and detailed investigations, assisting in deciding the most suitable research design, data collection method and subject selection for future research.

*Descriptive research*

Some researchers (13/34) conducted descriptive studies to reveal students' assessment and learning experiences with GenAI (Bearman et al., 2024; Blanke, 2024; Chang et al., 2023; French et al., 2023; Smolansky et al., 2023; Vicente-Yagüe-Jara et al., 2023). For example, Exintaris et al. (2023) described a classroom exercise integrating metacognitive scaffolding and problem-solving practice, along with students' critique of solutions generated by ChatGPT. They also performed reflexive thematic analyses on students' perceptions of the activity involved. Kong et al. (2023) demonstrated using ChatGPT as a supplementary instrument to construct interactive learning settings and emulate students' authentic engineering thought processes. Additionally, using a multi-methods qualitative approach, Yan (2023) presented students' behaviours during and reflections on their exposure to ChatGPT in an English as a Second Language writing practicum. These descriptive studies investigated different aspects of student learning with GenAI, providing more precise and comprehensive understandings of the situation by examining experiences of students and educators. They also help identify patterns, relationships and trends in the emerging field of student assessment in a GenAI context.

*Ethnographic and phenomenological research*

Some researchers (3/34) adopted research methods especially suited to GenAI tools to facilitate their studies, such as thing ethnography and post-phenomenological approaches (Farazouli et al., 2023; Michel-Villarreal et al., 2023; Salinas-Navarro et al., 2024). For instance, Michel-Villarreal et al. used thing ethnography to indicate human-machine interactions with GenAI tools – through chatting with ChatGPT and reporting mock examples from AI chatbots to illustrate the tool's application. Similarly, Salinas-Navarro et al. employed an ethnographic approach to examine the intricacies involved in merging GenAI with experiential learning for authentic assessment. Further, Farazouli et al. employed a post-phenomenological approach to understand how ChatGPT, in a mediating role, influenced university teachers' assessment practices, particularly for digital technologies. These studies with ethnographic and phenomenological research methodologies aid in obtaining profound understandings of human behaviour and the motives behind it. Ethnographic research enriches understanding from the viewpoint of GenAI users, treating them as insiders, while phenomenological research investigates the experiences of individuals for insights regarding student assessment in GenAI context. The adoption of these approaches can result in detailed, contextualised and interpretive portrayals of people's experiences and interactions with GenAI.

*Implications*

Based on the review results, the research methods adopted to study student assessment in GenAI-facilitated learning environments appear to be unbalanced, with more focus on qualitative methodologies (Almasre, 2024; Bearman et al., 2024; Exintaris et al., 2023; Rajabi et al., 2023; Smolansky et al., 2023; Vartiainen & Tedre, 2023). This is consistent with previous research findings that our research topic – student assessment in GenAI-facilitated learning environments – is in its infancy (Bannister et al., 2023b; Chiu, 2023), attracting some preliminary research explorations. This finding may suggest that, as our target topic develops, more quantitative research (such as correlational and experimental designs) may be needed to examine the effects of different assessment approaches on Gen AI-facilitated learning.

Significantly, researchers can design methods that directly measure the effectiveness of assessment on student learning process and outcomes in GenAI contexts (Alasadi & Baiz, 2023; Blanke, 2024; Michel-Villarreal et al., 2023; Ruiz-Rojas et al., 2023; Weng, Chiu, & Tsang, 2022). The deficiency of relevant methods reduces the possibility of integrating GenAI into existing assessment designs and carrying out corresponding curriculum reform (Chiu & Chai, 2020), thereby hindering the development of this research area. From this perspective, educational researchers who are reflective and who actively participate in GenAI instruction and educational assessment are crucial to appropriately designing research methods to address interdisciplinary needs.

## A diagram for current research status and trends

The scoping review is summarised in Table 3, as shown below. We also suggest future research directions where the literature is insufficient and potential avenues for further study to advance our understanding of this emerging field.

Table 3
*Review summary*

| Research questions | Findings | Implications |
|---|---|---|
| (1) How do higher education teachers assess student learning when GenAI is used? | • Traditional assessment<br>• Innovative and refocused assessment<br>• GenAI-incorporated assessment | Traditional assessment methods do not operate effectively in the GenAI era.<br>The last two assessment approaches are possible strategies to eliminate mismatches between the innovative GenAI integrated education practice and the traditional approaches of assessing student learning. |
| (2) What new focused learning outcomes have evolved from the use of GenAI? | • Career-driven competencies<br>• Lifelong learning skills | Educators in programming and writing areas are encouraged to embrace the advantages and investigate possible uses of the innovation.<br>The development of lifelong skills in students is a mainstream trend compared with career-driven competencies.<br>Future researchers should design more assessment activities targeting students' higher-order thinking skills in GenAI contexts. |
| (3) What research methods are used to investigate assessments when GenAI is used? | • Exploratory research design<br>• Descriptive research<br>• Ethnographic and phenomenological research | The researchers in the area of research methods who examined students' assessment mainly focused on qualitative methodologies.<br>Researchers should design methods that directly measure the effectiveness of assessment on student learning process and outcomes in GenAI contexts. |

In this scoping review, we visually represented the connections between the results and findings in a diagram, as demonstrated in Figure 2. In the diagram, the dotted and solid lines show weak and strong support relationships, respectively. Three assessment approaches are identified on the left, while the two new focused learning outcomes are identified on the right.

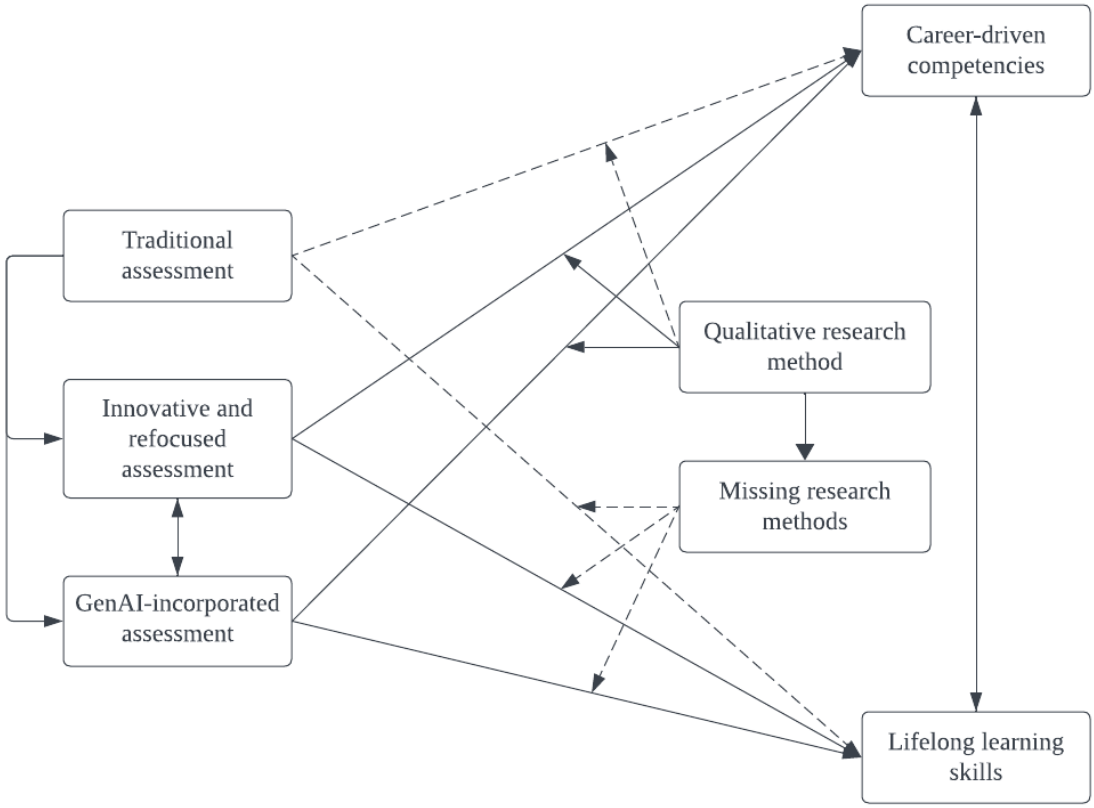


*Figure 2.* Concept mapping of the scoping review

Our findings suggest that greater attention should be given to two learning outcomes – career-driven competencies and lifelong learning skills – as they are becoming increasingly essential in GenAI-facilitated learning environments. Higher education educators and researchers must formally assess them, more often at the university level than at the program and course levels. We also emphasise that traditional assessment approaches may not be able to assess these two new focused learning outcomes with a high level of efficacy. These approaches should be redesigned and revised into appropriate innovative strategies, as they no longer meet the demands of GenAI-facilitated learning environments (Hodges & Kirschner, 2024; Luo, 2024). Traditional approaches failed to foster competencies and skills that are significant for students' future career development (Chaudhry et al., 2023) and lifelong flourishing (Alasadi & Baiz, 2023; Michel-Villarreal et al., 2023; Ruiz-Rojas et al., 2023). In this study, we suggest that innovative strategies can be categorised and viewed through two main approaches: innovative and refocused assessment and GenAI-incorporated assessment. We also suggest that these two approaches should complement each other. In other words, higher education educators and researchers should investigate how they work together effectively to foster student learning outcomes.

As shown in the diagram, the innovative and refocused assessment and GenAI-incorporated assessment approaches have a two-way interaction with the new focused intended learning outcomes. The new intended learning outcomes that emerged from educational practice require innovation and changes in assessment approaches to ensure appropriateness (Chiu, 2024; Kasneci et al., 2023), through redesigning the existing assessments or developing GenAI-incorporated assessments. The goal is to benefit students' assessment and intended learning outcomes in a GenAI-facilitated higher education learning environment. Future research should reveal and highlight these interactions. The diagram also shows that most reviewed studies used qualitative research methods to give researchers and designers insights into how GenAI impacts assessment in higher education. A plausible explanation is that other methods, such as correlational and experimental studies, need larger sample sizes and longer research periods. However, a qualitative design cannot confirm the assessment approaches' effectiveness or the relationships shown in the diagram. We thus expect more quantitative design research experiments to be conducted in the future.

Effective assessments of students' learning in a GenAI setting confirm their learning achievements, promoting the desired learning outcomes targeted by innovative intervention approaches. The full potential of GenAI teaching may not be achieved without appropriate and effective assessments. There should be a synergy between advancements in innovative technologies and existing educational practices (Hodges & Kirschner, 2024; Luo, 2024; Rajaram, 2023). This diagram provides insights into the current research status and trends in the existing literature. Based on our implications and the discussion in the diagram, we summarised the following future research directions:

- Innovative assessment designs: Traditional assessment may not appropriately assess the new focused learning outcomes. More studies should be conducted to explore new ideas of how to assess student learning in GenAI-facilitated learning environments.
- Collaborations among assessment approaches: We suggest that assessment approaches should have mutual relationships and complement each other. Therefore, studies should explore ways to enhance collaboration in assessing student learning outcomes and promoting their learning process.
- New learning outcomes: GenAI impacts intended learning outcomes (e.g., what and how our students will perform in the future in their workplaces). The diagram highlights only some of them. We suggest that more studies should identify learning outcomes that align with students' needs.
- Relationships between assessment approaches and learning outcomes: Although the diagram suggests the relationships are mutual, which is supported by studies using a small sample size, we suggest large sample size studies to validate the relationships by offering greater statistical power, increasing precision and reducing the impact of outliers. Accordingly, we suggest that studies examine the relationships in the diagram and investigate how they mutually affect each other.
- Quantitative or mixed research studies: Most of the extant research on how GenAI impacts assessment in higher education is qualitative. More quantitative or mixed research designs should be used to examine the effectiveness of different assessment approaches in the future.

## Conclusion

Although the significance of assessing student learning in GenAI contexts has been well recognised (e.g., by Bahroun et al., 2023; Baidoo-Anu & Ansah, 2023; Bannister et al., 2023b; Farrelly & Baker, 2023; Fleckenstein et al., 2024; Kolade et al., 2024; Netto, 2023), there is no exploration in the literature on educators' assessment approaches for university students, the more focused learning outcomes of universities that integrate GenAI in learning and assessment or the research methods adopted by relevant researchers. This scoping review contributes to the three issues mentioned above related to the adoption of GenAI in higher education and proposes implications for future research. However, there are some limitations to the current study. For instance, it includes literature from a wide variety of study designs, not all of which might respond to all three research questions. Additionally, as GenAI emerged as a research field only in 2023, the included articles covered a short period only; future publications may contribute more information for the review study. Further, we adopted although six representative databases to collect literature to ensure the quality of the collected studies, articles from other databases may have been missed. Future researchers could design their studies based on our experience and provide more insights into the research area.

## Author contributions

**Xiaojing Weng:** Conceptualisation, Investigation, Data curation, Formal analysis, Writing – original draft; **Qi Xia:** Investigation, Data curation, Formal analysis, Writing – review and editing; **Mingyue Gu:** Writing – review and editing; **Kumaran Rajaram:** Writing – review and editing; **Thomas K. F. Chiu:** Conceptualisation, Writing – review and editing.

## References

References marked with an asterisk * indicate studies included in the scoping review.


Abelha, M., Fernandes, S., Mesquita, D., Seabra, F., & Ferreira-Oliveira, A. T. (2020). Graduate employability and competence development in higher education—A systematic literature review using PRISMA. *Sustainability*, *12*(15), Article 5900. https://doi.org/10.3390/su12155900

Ahmad, S. F., Alam, M. M., Rahmat, M. K., Mubarik, M. S., & Hyder, S. I. (2022). Academic and administrative role of artificial intelligence in education. *Sustainability*, *14*(3), Article 1101. https://doi.org/10.3390/su14031101

Alasadi, E. A., & Baiz, C. R. (2023). Generative AI in education and research: Opportunities, concerns, and solutions. *Journal of Chemical Education*, *100*(8), 2965–2971. https://doi.org/10.1021/acs.jchemed.3c00323

*Ali, K., Barhom, N., Tamimi, F., & Duggal, M. (2024). ChatGPT—A double-edged sword for healthcare education? Implications for assessments of dental students. *European Journal of Dental Education, 28*(1), 206–211. https://doi.org/10.1111/eje.12937

*Almasre, M. (2024). Development and evaluation of a custom GPT for the assessment of students' designs in a typography course. *Education Sciences, 14*(2), Article 148. https://doi.org/10.3390/educsci14020148

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology, 8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Bahroun, Z., Anane, C., Ahmed, V., & Zacca, A. (2023). Transforming education: A comprehensive review of generative artificial intelligence in educational settings through bibliometric and content analysis. *Sustainability*, *15*(17), Article 12983. https://doi.org/10.3390/su151712983

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, *7*(1), 52–62. https://doi.org/10.61969/jai.1337500

*Banihashem, S. K., Kerman, N. T., Noroozi, O., Moon, J., & Drachsler, H. (2024). Feedback sources in essay writing: Peer-generated or AI-generated feedback? *International Journal of Educational Technology in Higher Education, 2*, Article 23. https://doi.org/10.1186/s41239-024-00455-4

*Bannister, P., Santamaría-Urbieta, A., & Alcalde-Peñalver, E. (2023a). A Delphi study on generative artificial intelligence and English Medium instruction assessment: Implications for social justice. *Iranian Journal of Language Teaching Research*, *11*(3), 53–80. https://doi.org/10.30466/IJLTR.2023.121406

Bannister, P., Santamaría-Urbieta, A., & Alcalde-Peñalver, E. (2023b). Una revisión sistemática de la IA generativa y la educación superior (en inglés como medio de instrucción) [A systematic review of generative AI and (English medium instruction) higher education]. *Aula Abierta*, *52*(4), 401–409. https://doi.org/10.17811/rifie.52.4.2023.401-409

Barrett, A., & Pack, A. (2023). Not quite eye to AI: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education*, *20*, Article 59. https://doi.org/10.1186/s41239-023-00427-0

*Bearman, M., Tai, J., Dawson, P., Boud, D., & Ajjawi, R. (2024). Developing evaluative judgement for a time of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, *49*(6), 893–905. https://doi.org/10.1080/02602938.2024.2335321

*Binh Nguyen, T., Diem Thi Hong, V., Minh Nguyen, N., Thi Thu Tra, P., Hieu Thai, T., & Son Ha, X. (2023). Race with the machines: Assessing the capability of generative AI in solving authentic assessments. *Australasian Journal of Educational Technology*, *39*(5), 59–81. https://doi.org/10.14742/ajet.8902

*Blanke, J. M. (2024). ChatGPT: The sky is not falling. *Journal of Legal Studies Education*, *41*(1), 39–47. https://doi.org/10.1111/jlse.12145

Bryan, C., & Clegg, K. (Eds.). (2019). *Innovative assessment in higher education: A handbook for academic practitioners*. Routledge. https://doi.org/10.4324/9780429506857

*Chang, D. H., Lin, M. P. C., Hajian, S., & Wang, Q. Q. (2023). Educational design principles of using AI chatbot that supports self-regulated learning in education: Goal setting, feedback, and personalization. *Sustainability*, *15*(17), Article 12921. https://doi.org/10.3390/su151712921

*Chaudhry, I. S., Sarwary, S. A. M., El Refae, G. A., & Chabchoub, H. (2023). Time to revisit existing student's performance evaluation approach in higher education sector in a new era of ChatGPT—A case study. *Cogent Education*, *10*(1), Article 2210461. https://doi.org/10.1080/2331186X.2023.2210461

Chen, B., & Zhu, X. (2023). Integrating generative AI in knowledge building. *Computers and Education: Artificial Intelligence*, *5*, Article 100184. https://doi.org/10.1016/j.caeai.2023.100184

Chiu, T. K. F. (2023). The impact of Generative AI (GenAI) on practices, policies and research direction in education: A case of ChatGPT and Midjourney. *Interactive Learning Environments*, 1–17. https://doi.org/10.1080/10494820.2023.2253861

Chiu, T. K. F. (2024). Future research recommendations for transforming higher education with generative AI. *Computers and Education: Artificial Intelligence, 6*, Article 100197. https://doi.org/10.1016/j.caeai.2023.100197

Chiu, T. K. F., & Chai, C. S. (2020). Sustainable curriculum planning for artificial intelligence education: A self-determination theory perspective. *Sustainability*, *12*(14), Article 5568. https://doi.org/10.3390/su12145568

*Clark, T. M. (2023). Investigating the use of an artificial intelligence chatbot with general chemistry exam questions. *Journal of Chemical Education*, *100*(5), 1905–1916. https://doi.org/10.1021/acs.jchemed.3c00027

*Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, *61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Dobrin, S. I. (2023). *Talking about Generative AI: A Guide for Educators*. Broadview Press.

*Dobslaw, F., & Bergh, P. (2023, June). Experiences with remote examination formats in light of GPT-4. In J. Mottock (Ed.), *Proceedings of the 5th European Conference on Software Engineering Education* (pp. 220–225). Association for Computing Machinery. https://doi.org/10.1145/3593663.3593695

*Eager, B., & Brunton, R. (2023). Prompting higher education towards AI-augmented teaching and learning practice. *Journal of University Teaching & Learning Practice*, *20*(5), Article 2. https://doi.org/10.53761/1.20.5.02

*Exintaris, B., Karunaratne, N., & Yuriev, E. (2023). Metacognition and critical thinking: Using ChatGPT-generated responses as prompts for critique in a problem-solving workshop (SMARTCHEMPer). *Journal of Chemical Education*, *100*(8), 2972–2980. https://doi.org/10.1021/acs.jchemed.3c00481

*Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K., & McGrath, C. (2023). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots' impact on university teachers' assessment practices. *Assessment & Evaluation in Higher Education*, *49*(3), 363–375. https://doi.org/10.1080/02602938.2023.2241676

Farrelly, T., & Baker, N. (2023). Generative artificial intelligence: Implications and considerations for higher education practice. *Education Sciences*, *13*(11), Article 1109. https://doi.org/10.3390/educsci13111109

*Fleckenstein, J., Meyer, J., Jansen, T., Keller, S. D., Köller, O., & Möller, J. (2024). Do teachers spot AI? Evaluating the detectability of AI-generated texts among student essays. *Computers and Education: Artificial Intelligence, 6*, Article 100209. https://doi.org/10.1016/j.caeai.2024.100209

*French, F., Levi, D., Maczo, C., Simonaityte, A., Triantafyllidis, S., & Varda, G. (2023). Creative use of OpenAI in education: Case studies from game development. *Multimodal Technologies and Interaction*, *7*(8), Article 81. https://doi.org/10.3390/mti7080081

Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & Education*, *57*(4), 2333–2351. https://doi.org/10.1016/j.compedu.2011.06.004

*Günay, S., Öztürk, A., Özerol, H., Yiğit, Y., & Erenler, A. K. (2024). Comparison of emergency medicine specialist, cardiologist, and chat-GPT in electrocardiography assessment. *American Journal of Emergency Medicine, 80*, 51–60. https://doi.org/10.1016/j.ajem.2024.03.017

Hodges, C. B., & Kirschner, P. A. (2024). Innovation of instructional design and assessment in the age of generative artificial intelligence. *TechTrends*, *68*(1), 195–199. https://doi.org/10.1007/s11528-023-00926-x

Hwang, G. J., Xie, H., Wah, B. W., & Gaˇsevi´c, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 1*, Article 100001. https://doi.org/10.1016/j.caeai.2020.100001

Ibarra-Sáiz, M. S., Rodríguez-Gómez, G., & Boud, D. (2021). The quality of assessment tasks as a determinant of learning. *Assessment & Evaluation in Higher Education*, *46*(6), 943–955. https://doi.org/10.1080/02602938.2020.1828268

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*, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

*Kirwan, A. (2023). ChatGPT and university teaching, learning and assessment: Some initial reflections on teaching academic integrity in the age of Large Language Models. *Irish Educational Studies*, 1–18. https://doi.org/10.1080/03323315.2023.2284901

Kivunja, C. (2015). Teaching students to learn and to work well with 21st-century skills: Unpacking the career and life skills domain of the new learning paradigm. *International Journal of Higher Education, 4*(1), 1–11. https://doi.org/10.5430/ijhe.v4n1p1

*Kolade, O., Owoseni, A., & Egbetokun, A. (2024). Is AI changing learning and assessment as we know it? Evidence from a ChatGPT experiment and a conceptual framework. *Heliyon, 10*(4), Article e25953. https://doi.org/10.1016/j.heliyon.2024.e25953

*Kong, Z. Y., Adi, V. S. K., Segovia-Hernández, J. G., & Sunarso, J. (2023). Complementary role of large language models in educating undergraduate design of distillation column: Methodology development. *Digital Chemical Engineering*, *9*, Article 100126. https://doi.org/10.1016/j.dche.2023.100126

Lara, J. A., Aljawarneh, S., & Pamplona, S. (2020). Special issue on the current trends in e-learning assessment. *Journal of Computing in Higher Education*, *32*, 1–8. https://doi.org/10.1007/s12528-019-09235-w

Lau, A. M. S. (2016). 'Formative good, summative bad?'–A review of the dichotomy in assessment literature. *Journal of Further and Higher Education*, *40*(4), 509–525. https://doi.org/10.1080/0309877X.2014.984600

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 Educationn, 49*(5), 651–664. https://doi.org/10.1080/02602938.2024.2309963

*Lu, Q., Yao, Y., Xiao, L. H., Yuan, M. Z., Wang, J., & Zhu, X. H. (2024). Can ChatGPT effectively complement teacher assessment of undergraduate students' academic writing? *Assessment & Evaluation in Higher Education, 49*(5), 616–633. https://doi.org/10.1080/02602938.2024.2301722

Makransky, G., Borre-Gude, S., & Mayer, R. E. (2019). Motivational and cognitive benefits of training in immersive virtual reality based on multiple assessments. *Journal of Computer Assisted Learning*, *35*(6), 691–707. https://doi.org/10.1111/jcal.12375

Mavroudi, A., Giannakos, M., & Krogstie, J. (2018). Supporting adaptive learning pathways through the use of learning analytics: Developments, challenges and future opportunities. *Interactive Learning Environments*, *26*(2), 206–220. https://doi.org/10.1080/10494820.2017.1292531

*Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Education Sciences*, *13*(9), 856. https://doi.org/10.3390/educsci13090856

Mintz, J., Holmes, W., Liu, L., & Perez-Ortiz, M. (2023). Artificial intelligence and K-12 education: Possibilities, pedagogies and risks. *Computers in the Schools*, *40*(4), 325–333. https://doi.org/10.1080/07380569.2023.2279870

*Mohamed, A. M. (2023). Exploring the potential of an AI-based chatbot (ChatGPT) in enhancing English as a Foreign Language (EFL) teaching: Perceptions of EFL faculty members. *Education and Information Technologies*, *29*(3), 3195–3217. https://doi.org/10.1007/s10639-023-11917-z

Mousavinasab, E., Zarifsanaiey, N., R. Niakan Kalhori, S., Rakhshan, M., Keikha, L., & Ghazi Saeedi, M. (2021). Intelligent tutoring systems: A systematic review of characteristics, applications, and evaluation methods. *Interactive Learning Environments*, *29*(1), 142–163. https://doi.org/10.1080/10494820.2018.1558257

*Netto, N. R. (2023). Use of case studies in social work assessments - ChatGPT's kryptonite? *Social Work Education*, 1–12. https://doi.org/10.1080/02615479.2023.2266461

*Nikolic, S., Daniel, S., Haque, R., Belkina, M., Hassan, G. M., Grundy, S., Lynden, S., Neal, P., & Sandison, C. (2023). ChatGPT versus engineering education assessment: A multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity. *European Journal of Engineering Education, 48*(4), 559–614. https://doi.org/10.1080/03043797.2023.2213169

O'Flaherty, J., & Phillips, C. (2015). The use of flipped classrooms in higher education: A scoping review. *The Internet and Higher Education*, *25*, 85–95. https://doi.org/10.1016/j.iheduc.2015.02.002

Pereira, D., Flores, M. A., & Niklasson, L. (2016). Assessment revisited: A review of research in Assessment and Evaluation in Higher Education. *Assessment & Evaluation in Higher Education*, *41*(7), 1008–1032. https://doi.org/10.1080/02602938.2015.1055233

Pham, M. T., Rajić, A., Greig, J. D., Sargeant, J. M., Papadopoulos, A., & McEwen, S. A. (2014). A scoping review of scoping reviews: Advancing the approach and enhancing the consistency. *Research Synthesis Methods*, *5*(4), 371–385. https://doi.org/10.1002/jrsm.1123

*Rajabi, P., Taghipour, P., Cukierman, D., & Doleck, T. (2023). Exploring ChatGPT's impact on post-secondary education: A qualitative study. In O. Ola & B. Harrington (Eds.), *Proceedings of the 25th Western Canadian Conference on Computing Education* (pp. 1–6). Association for Computing Machinery. https://doi.org/10.1145/3593342.3593360

Rajaram, K. (2021). *Evidence-based teaching for the 21st-century classroom and beyond: Innovation-driven learning strategies*, Springer.

Rajaram, K. (2023). *Learning intelligence: Innovative and digital transformative learning strategies – cultural and social engineering perspectives*, Springer.

Rubin, D. B., Stuart, E. A., & Zanutto, E. L. (2004). A potential outcomes view of value-added assessment in education. *Journal of Educational and Behavioral Statistics*, *29*(1), 103–116. http://www.jstor.org/stable/3701308

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), Article 11524. https://doi.org/10.3390/su151511524

*Sánchez-Ruiz, L. M., Moll-López, S., Nuñez-Pérez, A., Moraño-Fernández, J. A., & Vega-Fleitas, E. (2023). ChatGPT challenges blended learning methodologies in engineering education: A case study in mathematics. *Applied Sciences*, *13*(10), Article 6039. https://doi.org/10.3390/app13106039

*Salinas-Navarro, D. E., Vilalta-Perdomo, E., Michel-Villarreal, R., & Montesinos, L. (2024). Using generative artificial intelligence tools to explain and enhance experiential learning for authentic assessment. *Education Sciences, 14*(1), Article 83. https://doi.org/10.3390/educsci14010083

*Smolansky, A., Cram, A., Raduescu, C., Zeivots, S., Huber, E., & Kizilcec, R. F. (2023). Educator and student perspectives on the impact of generative AI on assessments in higher education. In D. Spikol (Chair), *Proceedings of the Tenth ACM Conference on Learning@ Scale* Conference (pp. 378–382). Association for Computing Machinery. https://doi.org/10.1145/3573051.3596191

Su, J., & Yang, W. (2023). Unlocking the power of ChatGPT: A framework for applying generative AI in education. *ECNU Review of Education*, *6*(3), Article 20965311231168423. https://doi.org/10.1177/20965311231168423

Suleman, F. (2018). The employability skills of higher education graduates: Insights into conceptual frameworks and methodological options. *Higher Education*, *76*, 263–278. https://doi.org/10.1007/s10734-017-0207-0

*Vartiainen, H., & Tedre, M. (2023). Using artificial intelligence in craft education: Crafting with text-to-image generative models. *Digital Creativity*, *34*(1), 1–21. https://doi.org/10.1080/14626268.2023.2174557

*Vicente-Yagüe-Jara, M. I., López-Martínez, O., Navarro-Navarro, V., & Cuéllar-Santiago, F. (2023). Escritura, creatividad e inteligencia artificial. ChatGPT en el contexto universitario [Writing, creativity, and artificial intelligence: ChatGPT in the university context]. *Comunicar*, *31*(77), 47–57. https://doi.org/10.3916/C77-2023-04

Weng, X., & Chiu, T. K. F. (2023). Instructional design and learning outcomes of intelligent computer assisted language learning: Systematic review in the field. *Computers and Education: Artificial Intelligence*, *4*, Article 100117. http://doi.org/10.1016/j.caeai.2022.100117

Weng, X., Chiu, T. K. F., & Tsang, C. C. (2022). Promoting student creativity and entrepreneurship through real-world problem-based maker education. *Thinking Skills and Creativity*, *45*, Article 101046. https://doi.org/10.1016/j.tsc.2022.101046

Weng, X., Cui, Z., Ng, O. L., Jong, M. S., & Chiu, T. K. F. (2022). Characterising students' 4C skills development during problem-based digital making. *Journal of Science Education and Technology*, *31*(3), 372–385. https://doi.org/10.1007/s10956-022-09961-4

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

*Yan, D. (2023). Impact of ChatGPT on learners in an L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, *28*(11), 13943–13967. https://doi.org/10.1007/s10639-023-11742-4

*Yilmaz, R., & Yilmaz, F. G. K. (2023). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Computers and Education: Artificial Intelligence*, *4*, Article 100147. https://doi.org/10.1016/j.caeai.2023.100147

Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers and Education: Artificial Intelligence*, *2*, Article 100025. https://doi.org/10.1016/j.caeai.2021.100025

---

**Corresponding author:** Qi Xia, ixia@zju.edu.cn



**Please cite as**: Weng, X., Xia, Q., Gu, M., Rajaram, K., & Chiu, T. K. F. (2024). Assessment and learning outcomes for generative AI in higher education: A scoping review on current research status and trends. *Australasian Journal of Educational Technology*, *40*(6), 37–55. https://doi.org/10.14742/ajet.9540