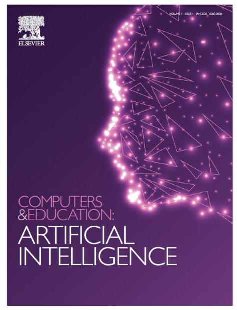
Journal Pre-proof




# Implementing Generative AI (GenAI) in Higher Education: A Systematic Review of Case Studies

Marina Belkina, Scott Daniel, Sasha Nikolic, Rezwanul Haque, Sarah Lyden, Peter Neal, Sarah Grundy, Ghulam M. Hassan

PII: S2666-920X(25)00047-5

DOI: https://doi.org/10.1016/j.caeai.2025.100407

Reference: CAEAI 100407

To appear in: *Computers and Education: Artificial Intelligence*

Received Date: 4 July 2024

Revised Date: 23 March 2025

Accepted Date: 7 April 2025

Please cite this article as: Belkina M., Daniel S., Nikolic S., Haque R., Lyden S., Neal P., Grundy S. & Hassan G.M., Implementing Generative AI (GenAI) in Higher Education: A Systematic Review of Case Studies, *Computers and Education: Artificial Intelligence*, https://doi.org/10.1016/j.caeai.2025.100407.

# Title page

**Article title:**

## Implementing Generative AI (GenAI) in Higher Education: A Systematic Review of Case Studies using LCF, SAMR and TPACK Frameworks

**Author names**

Marina Belkina[a], Scott Daniel[b], Sasha Nikolic[c], Rezwanul Haque[d], Sarah Lyden[e] , Peter Neal[a], Sarah Grundy[a], and Ghulam M. Hassan[f]

Affiliations
[a]University of New South Wales, [b]University of Technology Sydney, [c]University of Wollongong, [d]University of the Sunshine Coast, [e]University of Tasmania, [f]University of Western Australia,

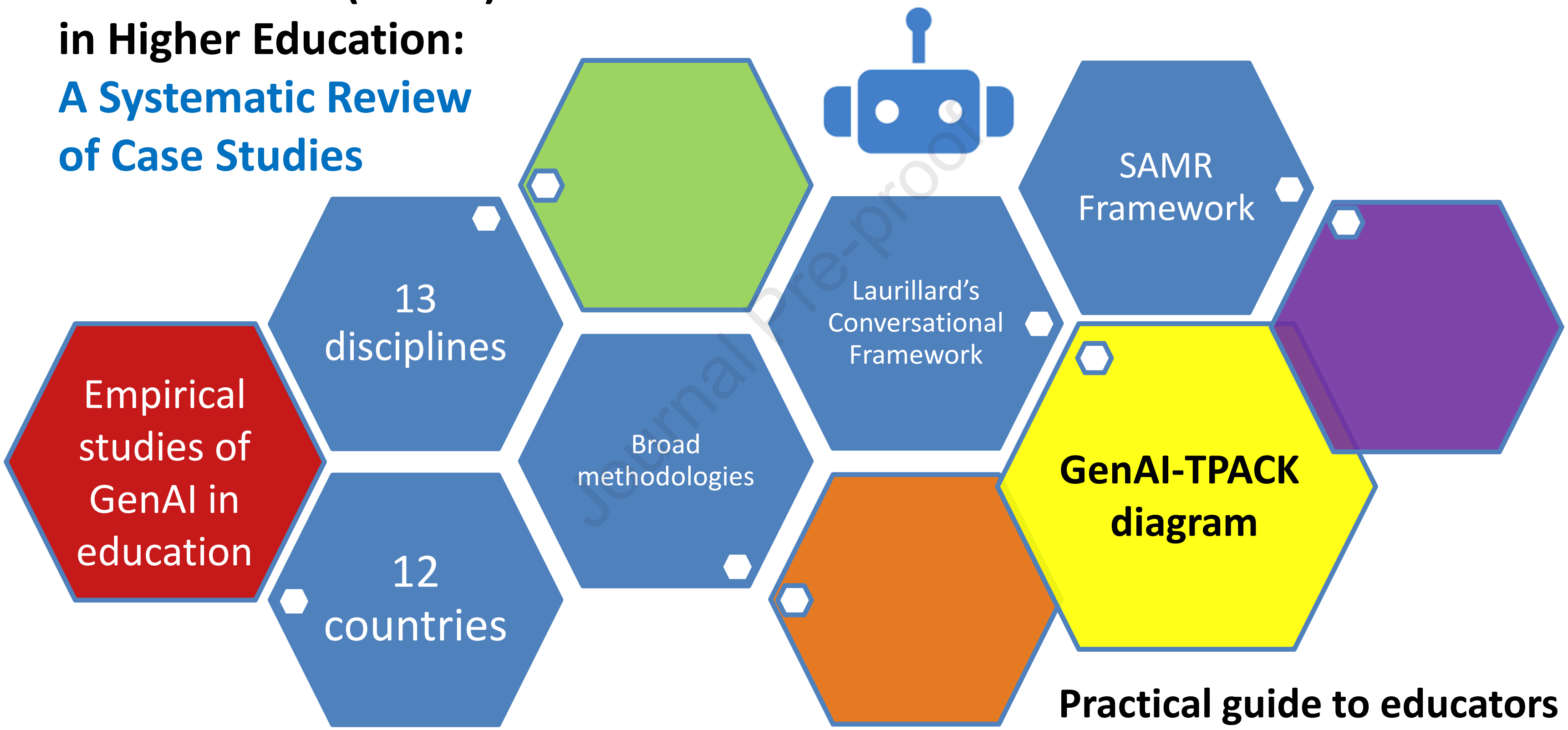
**Corresponding author – Marina Belkina**

m.belkina@unsw.edu.au

**Address**

High St Kensington NSW 2052 Australia

# Implementing Generative AI (GenAI) in Higher Education: A Systematic Review of Case Studies




**Practical guide to educators in effective incorporation GenAI tools into teaching**

# Implementing Generative AI (GenAI) in Higher Education: A Systematic Review of Case Studies

**Keywords:** Generative AI, education, GenAI, case study, university, implementation

**Abstract**

The introduction of Generative Artificial Intelligence (GenAI) tools, like ChatGPT, into higher education heralds a transformative era, reshaping instructional methods, enhancing student support systems, and redefining the educational landscape. Recent literature reviews on GenAI highlight a lack of focus on how these tools are being practically implemented in educational settings. Addressing this gap, the present study systematically examines empirical case studies that demonstrate the integration of GenAI into teaching and learning in higher education, offering actionable insights and guidance for academic practice.

We conducted a search of relevant databases and identified 21 empirical studies that met our inclusion criteria. The selected studies cover a diverse range of disciplines, locations, types of participants (from first-year students to postgraduates and academics), and a variety of methodologies. We classified the selected publications based on the pedagogic theory of Laurillard's Conversational Framework (LCF) and the Substitution, Augmentation, Modification, and Redefinition (SAMR) framework. We also synthesized definitions from selected empirical studies and recent research exploring Technological Pedagogical Content Knowledge (TPACK) in the age of GenAI, providing a comprehensive understanding of GenAI-TPACK factors. Limitations and future research opportunities are also discussed. The paper concludes by providing a GenAI-TPACK diagram to guide educators in effectively incorporating GenAI tools into their teaching practices, ensuring responsible and impactful use in higher education.



## 1. Introduction

Generative Artificial Intelligence (GenAI) is transforming higher education, by challenging traditional teaching approaches, improving student support systems, and reshaping the educational ecosystem. GenAI can be defined as a technology that leverages deep learning models to generate human-like content (e.g., images, words) in response to complex and varied prompts (e.g., languages, instructions, questions) Lim et al. (2023). The most popular type of GenAI model, ChatGPT, captured widespread attention across the global academic community. The claim of the developers that the ChatGPT-4 can pass any exam with a score around the top 10% of test takers (OpenAI et al., 2023) led to widespread discussions on academic integrity. As the implementation of ChatGPT and similar technologies in classrooms becomes more prevalent, a comprehensive examination of their effectiveness and integration is imperative.

A growing body of research across disciplines such as engineering (Nikolic et al., 2023; Nikolic, Sandison, et al., 2024), medical education (Currie (2023), Gilson et al. (2023), microbiology (Das et al., 2023), and economics (Geerling et al., 2023) has begun to evaluate ChatGPT's performance against university assessments. Findings generally indicate that with minimal input modifications, ChatGPT can produce acceptable responses, suggesting a need to recalibrate educational practices in anticipation of more advanced AI iterations. This evolving scenario underscores the importance of reassessing our pedagogical approaches as these tools become increasingly capable.

Recent investigations have highlighted the advantages of using GenAI in educational settings, with both teachers and students. With teachers, it can help in creating assessments (Baidoo-Anu et al., 2023; Kasneci et al., 2023; Zhai, 2023), enhancing flipped learning methodologies (Rudolph, 2023), developing curriculum (Simms, 2024), and identifying and developing superior learning resources (Muddam et al., 2023). With students, it can provide personal tutoring (Mhlanga, 2023), facilitate enhanced learning by providing answers to theoretical questions, stimulate creative thinking, summarise complex essays into understandable formats, and serve as a copyediting tool to aid students weak in language skills (Michel-Villarreal et al., 2023). Learning to use new GenAI tools is also important because students can learn to work more efficiently, accurately and make further advancements (Nikolic, Suesse, et al., 2024).

However, the use of GenAI is not devoid of challenges. Investigations have uncovered issues such as cheating, plagiarism, misleading information, and outdated content (Tlili et al., 2023). Overreliance on GenAI, and the ethical and pedagogical issues outlined above, emphasise the need for proper guidelines and policies to ensure responsible use of GenAI in education (Chan et al., 2023).

With these considerations in mind, understanding the functionality of GenAI within higher education classrooms becomes crucial. Due to the massive number of papers published in the field, educators often face challenges in finding practical examples of GenAI implementation. This makes it difficult to determine if the wheel is constantly being reinvented, or if diversity and originality in GenAI application is taking place. At the same time teachers, universities and regulators are seeking to engage and adapt educational offerings to the evolving requirements of the future workplace and the rapidly changing capabilities of GenAI technology. This paper aims to systematically review case studies on the implementation of GenAI in higher education, assessing the extent to which the current research addresses previously identified gaps. It also provides guidelines to academics on the workings and integration of these AI applications in higher education settings, emphasizing the need for a comprehensive understanding to enhance productivity and professional practice readiness.

## 2. Related literature

### 2.1 Theoretical Frameworks

Effectively leveraging GenAI capabilities requires more than just an understanding of the technology itself. Theoretical frameworks can be applied as a structured approach to assess, implement, and evaluate their use when integrated into educational contexts. They provide theoretical and practical structures to navigate the complexities of adoption, ensuring that educational technologies are used effectively, ethically, and inclusively (Nikolic, Sandison, et al., 2024). Some of the relevant frameworks that can be applied to GenAI follow.

The first theoretical framework utilised in this paper is Laurillard's Conversational Framework (LCF) and the concept of learning types (Laurillard, 2012). This framework was chosen because of its robust theoretical foundation in analysing learning processes across diverse educational settings. Studies have shown that it enables educators to align instructional strategies with desired learning outcomes, particularly in contexts involving emerging technologies (Laurillard (2012); Heinze et al. (2007). Laurillard's work helps emphasize the importance of communication and interaction between teachers and learners, helping teachers optimize technology-enhanced learning by providing theory-informed tools and scaffolding for adopting, adapting, and innovating effective pedagogical practices. Laurillard identifies six types of learning activities, each representing a fundamental way learners engage with material. They are acquisition, inquiry, discussion, practice, collaboration and production. By mapping these learning types onto the Conversational Framework, educators can evaluate whether all aspects of the learning process are being addressed.

The SMAR Model (Substitution, Augmentation, Modification, Redefinition) offers a valuable perspective on the transformative educational potential of GenAI by helping to describe and categorize the integration of digital technologies (Blundell et al., 2022). This framework categorizes the levels at which technology enhances or transforms educational practices, from the basic substitution of existing tools to the complete redefinition of learning experiences (Puentedura, 2009). By undertaking this analysis, insights can be gathered into whether the technology is being used for lower-level enhancement learning activities or higher-level transformation activities (Hamilton et al., 2016). When applied to GenAI, the SMAR model can help researchers and practitioners distinguish between superficial uses and more profound transformations, helping guide educators to maximize the value of GenAI technologies while avoiding shallow implementations that fail to address deeper educational goals.

The final Technological Pedagogical Content Knowledge (TPACK) framework provides a comprehensive lens through which to evaluate the interplay of technology, pedagogy, and content in integrating GenAI (Mishra et al., 2023). GenAI technologies often require educators to rethink their pedagogical strategies and the delivery of subject-specific content. TPACK facilitates this process by helping educators balance these domains effectively. For instance, in teaching with GenAI-driven writing assistants, educators must align the tool's capabilities with sound pedagogical approaches, such as fostering collaborative learning, and ensure it supports subject-specific goals like developing critical thinking skills in composition. TPACK ensures that technology integration occurs not in isolation but as part of a holistic educational strategy.

### 2.2 Recent systematic reviews on GenAI

The case for this systematic literature review is built upon insights from recent literature reviews that examine the role of Generative AI in education. The following paragraphs summarise key contributions from individual works, which collectively inform the rationale and scope of this review.

Bozkurt (2023) conducted an analysis of research trends and patterns in GenAI within the educational sector, utilising data mining and analytical techniques. The study identified seven key themes as promising areas for future research in educational praxis: interaction with GenAI-powered chatbots, the impact of large language models (LLMs) on teaching and learning, opportunities and challenges of conversational educational agents, the enhancement of social and cognitive learning processes through GenAI, the promotion of AI literacy to unlock future opportunities, the expansion of academic capabilities through AI, and the augmentation of educational experiences via human-AI interaction. These themes underscore the multifaceted potential of GenAI in transforming educational practices that need uncovering.

Building on the advancements of GenAI technologies, Park et al. (2024) anticipated an increase in research focusing on AI in blended learning environments. Their findings suggest that AI is primarily used in asynchronous online learning components but is less often applied to link these with face-to-face classroom activities. They recommend that future research should provide direction for educators on how to integrate GenAI to optimise blended learning implementations effectively.

Sohail et al. (2023) explored the practical applications of ChatGPT, emphasising its potential to solve real-world educational challenges. They highlighted critical issues such as biases and trustworthiness of the technology, advocating for further research and development to address these concerns. The study also identified potential future research directions, suggesting solutions to current challenges and forecasting advancements that could enhance the efficacy and reliability of ChatGPT in educational settings.

Castillo-Segura (2023) assessed the efficacy of AI tools and their respective LLMs during classification of 596 articles in the screening phase of a systematic literature review. The results highlight that while GPT-4 demonstrated the best overall performance, it is not yet entirely reliable, necessitating additional input from the research team for comprehensive classification. This finding points to the ongoing need for human oversight in AI-assisted research processes.

Kumar et al. (2024) performed a systematic literature review exploring how GenAI might drive innovation in higher education, and identified three main themes: Academic Integrity, Pedagogical Techno-Innovation, and Experiential Engagement. The literature review highlighted several benefits, including support for digital writing, automated writing evaluations, increased productivity, innovative assessment design, and enhancement of information literacy instruction. Concurrently, the study identified challenges such as managing indeterminate data, ensuring the ethical use of student data, distinguishing between human-written and AI-generated text, and addressing threats to academic integrity. These findings underscore the dual nature of GenAI's impact, and outlines many potential study directions, including exploring the potential of GenAI tools in enhancing student learning and engagement or enhancing readiness and adaptability of educational institutions to GenAI technologies and develop skills of faculty and students to handle these tools.

Hobensack et al. (2024) conducted a rapid literature review focusing on the current and potential uses of large language models in nursing. The review identified significant opportunities for applying these models but also highlighted several challenges, including ethical issues related to bias, misuse, and plagiarism. This review underscores the need for careful consideration of ethical implications in the deployment of AI technologies in specialised fields such as nursing.

Suryanto et al. (2023) reviewed the development, achievements, challenges, and emerging trends of chatbots, distinguishing between rule-based and GenAI approaches. The study concluded that while each approach has its advantages and disadvantages, there is a need for ongoing supervision in aspects such as language comprehension, bias, and ethical considerations. This highlights the importance of maintaining ethical standards in the development and deployment of chatbot technologies.

Alateyyat et al. (2024) analysed 295 articles from the Scopus database about GenAI in higher education, noting a predominant focus on providing general overviews with a shortage of research in specific topics, including integration with teaching practices, prediction models, AI in assessment, and support of administrative processes in higher education institutions. This recommendation points to a need for deeper exploration and practical application of AI technologies in educational settings.

In response to these identified research gaps, this study seeks to provide an overview of empirical research exploring the practical integration of GenAI into university teaching and learning. Unlike previous reviews, which often focus on broad trends (e.g., Alateyyat and Soltan (2024) or specific themes such as ethical considerations (Hobensack et al., 2024) or technological advancements (Bozkurt, 2023), this review uniquely examines the case studies of GenAI integration. By systematically analysing empirical studies, it aims to identify how GenAI can be effectively implemented to enhance student engagement, optimise teaching practices, and address the challenges of adoption in diverse higher education contexts. Furthermore, this review addresses critical gaps by mapping its findings against conceptual educational frameworks such as LCF, SMAR, and TPACK, providing a structured lens to evaluate GenAI's impact. To achieve this, the study seeks to answer the following research questions:

> **RQ1:** *What are the characteristics of the research conducted on GenAI implementation within higher education settings, including the geographical locations of the first authors, participant characteristics, and disciplinary focus?*

**RQ2:** *What study designs and research methods are used to evaluate the effect of GenAI integration in higher education?*

**RQ3:** *How can the effectiveness of the GenAI implementation in higher education can be analysed and categorised through the lens of the conceptual educational frameworks (LCF, SMAR and TPACK)?*

**RQ**4*: What recommendations to academics for the implementation of GenAI can be drawn from the analysis?*

**RQ5:** *What are the necessary methodological improvements and future research directions to better understand the impacts of GenAI in higher education?*

## 3. Methods

### 3.1 Research design

This review adopted PRISMA Statement (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) approach (Page et al., 2021) and proceeded in three steps: (i) article selection, (ii) article screening and inclusion, and (iii) data extraction and analysis.

First, we developed a research protocol which specified:

- The **population of interest**: Universities, irrespective of geographical location.
- The **phenomenon of interest**: GenAI in education.
- The **outcomes**: The use of GenAI in learning and teaching, associated examples and case studies, and the factors that helped or hindered it.

To the authors' knowledge, no similar review has been published or is in development. This was confirmed by searching academic databases.

The review is confined to the Scopus database exclusively, selected for its status as one of the largest curated abstract and citation databases that supports robust academic research in quantitative science studies with its comprehensive coverage of high-quality scholarly publications aligning with the review's objectives (Baas et al., 2020). Considering the novelty of the research question, book chapters or conference papers were not excluded from the search results. Table 1 summarises the inclusion criteria for article selection.

Table 1. Inclusion and exclusion criteria for article selection.

| Criterion | Inclusion |
|---|---|
| Topic | Focusing on the use GenAI in education or technologies with GenAI characteristics - technology that generates human-like content in response to complex and varied prompts |
| Study type | Empirical studies that demonstrate an authentic example of GenAI integration into the university teaching |

| | |
|---|---|
| Source | Journals, Conference papers, Book chapters |
| Period | January 1, 2023 to February 15, 2024 |
| Language | English |

### 3.2 Article selection

Given the massive increase in publications related to GenAI, the keyword structure was specifically designed to capture papers addressing the scope of the study. The following keywords were identified:

- **Keywords related to GenAI**: "artificial intelligence" OR "ChatGPT" OR "AI" OR "GenAI"
- **Keywords related to integration of GenAI**: "integration" OR "case study" OR "application" OR "implementation" OR "example"
- **Keywords related to education**: "teaching" OR "education" OR "classroom"
- **Keywords related to university**: "college" OR "faculty" OR "post-graduate" OR "postgraduate" OR "tertiary" OR "under-graduate" OR "undergraduate" OR "university" OR "HE"
- **Keyword related to the presence of students in the study**: "student"

To ensure the inclusion of the most recent and relevant insights, particularly following the significant advancements in GenAI marked by the launch of GPT-3 in late 2022, the article selection was limited to publications from 2023 onwards. Articles not published in English and review papers were excluded.

The Boolean search string we used was:

*TITLE-ABS-KEY("artificial intelligence" OR "ChatGPT" OR "AI" OR "GenAI")*

*AND*

*TITLE-ABS-KEY ("integration" OR "case study" OR "application" OR "implementation" OR "example")*

*AND*

*TITLE-ABS-KEY ("teaching" OR "education" OR "classroom")*

*AND*

*TITLE-ABS-KEY ("college" OR "faculty" OR "post-graduate" OR "postgraduate" OR "tertiary" OR "under-graduate" OR "undergraduate" OR "university" OR "HE")*

*AND*

*TITLE-ABS-KEY(student)*

*AND*

*PUBYEAR > 2022 AND PUBYEAR < 2026 AND ( LIMIT-TO ( LANGUAGE,"English" ) ) AND (EXCLUDE(DOCTYPE, “re”) OR EXCLUDE (DOCTYPE, “cr”)).*

### 3.2 Article screening and inclusion

The article search conducted in February 2024 identified 489 publications. No ineligible records were identified by the authors, and all articles were included for screening (Figure 1). Titles and abstracts of the publications were screened by eight authors to identify articles focused on findings from examples or case studies of the integration of GenAI in university teaching and learning. Each article was screened by two authors independently. Differences of opinion were resolved through group discussion. The exclusion criteria (EC) established for this review were as follows:

**EC1**: Publications not relevant to the scope of the study, for example, research focusing on school students instead of university students.

**EC2**: Publications lacking a case study with empirical data on GenAI implementation in tertiary education, such as for example, studies focusing solely on perceptions (rather than observations) of how GenAI can be used or misused in tertiary education.

**EC3:** Publication not accessible.

Following the initial title and abstract screening, 74 papers were included, 275 papers were excluded based on the exclusion criteria (239 by EC1 and 36 by EC2), and there was disagreement among the reviewers on 140 publications. These papers were then discussed further in a group, resulting in an additional 15 papers included and 125 papers excluded (73 by EC1 and 52 by EC2). The total number of papers excluded through the title and abstract screening process is 400.

After the initial exclusions, the full-text screening of the remaining 89 reports resulted in further exclusions: 28 reports were not relevant to the study's scope (EC1), 23 lacked evidence of GenAI implementation (EC2), and 1 report was inaccessible (EC3). Subsequently, there were differing opinions among reviewers on an additional 9 reports. These were discussed by the group of authors, resulting in all 9 being excluded (3 by EC1 and 6 by EC2). In total, out of the 89 papers reviewed at this stage, 28 were included for further analysis, and 52 were excluded for the reasons specified.

Ultimately, 7 studies were excluded post-extraction as they did not provide adequate evidence of GenAI implementation in higher education teaching (EC2), leaving a total of 21 studies for the final analysis.

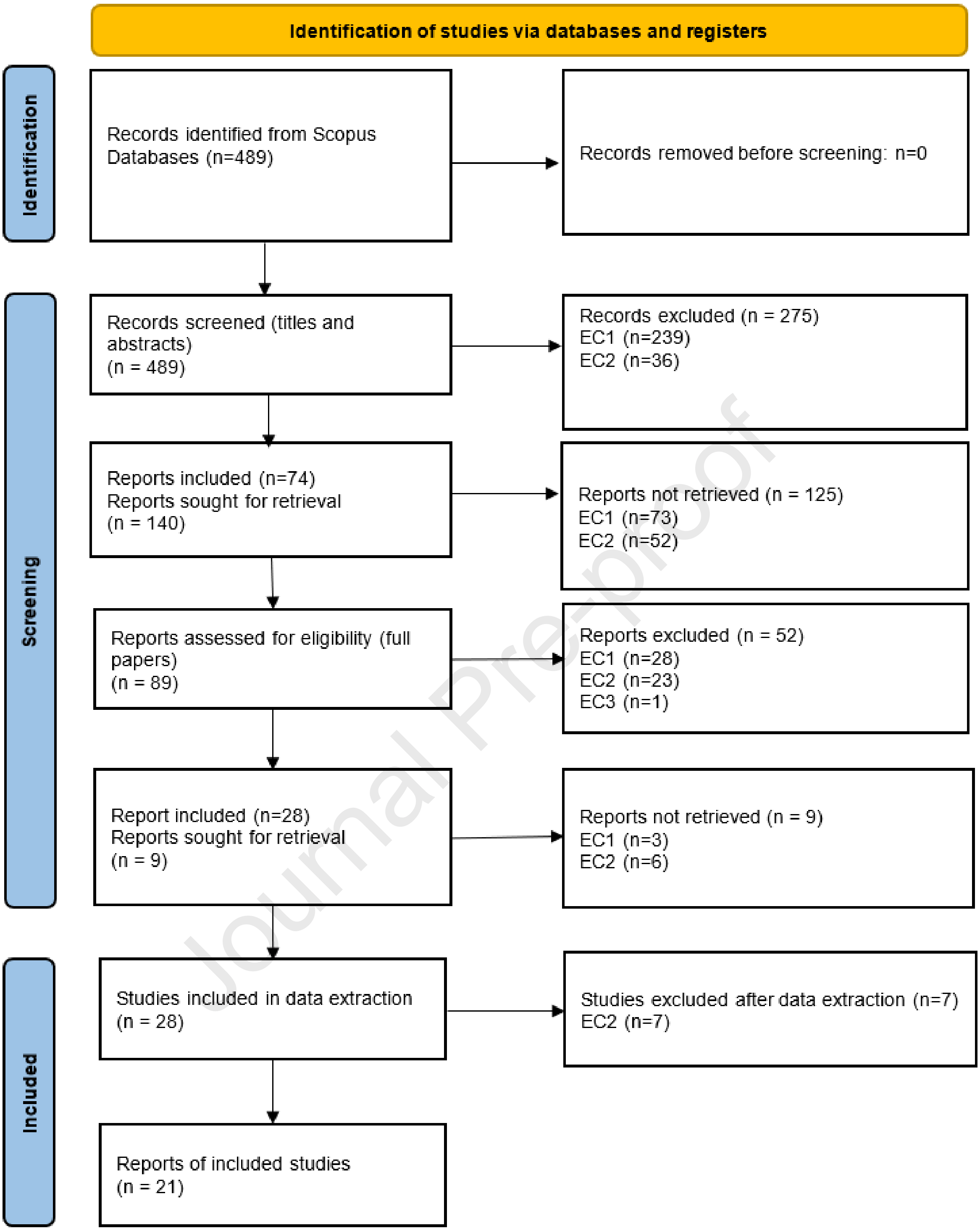



Figure 1. PRISMA flow diagram of article selection.

### 3.3 Data extraction

The final data extraction process was conducted on the remaining 21 studies. Data from included studies were extracted onto a pre-defined Excel data collection form. To address the research questions, data from each paper was collated as follows:

**RQ1** – title, author, year, location, number and type of participants, the aim of the study; the disciplines represented; reference to Gen AI used;

**RQ2** – study methods, educational theories, study outcomes and key findings;

**RQ3** – type of learning activities, study outcomes, summaries of GenAI implementation;

**RQ4** – author-identified limitations and future research opportunities;

**RQ 5** - established after the comprehensive study by the authors.

The summary of the included publications classified by the university educational level is shown in the appendix (Table A).

## 4. Results and discussion

### 4.1 Study characteristics to address RQ1

The first research question, aimed to determine *the characteristics of the research conducted on GenAI implementation in higher education (including the geographical locations of the first authors, participants' characteristics, and discipline distribution).* It is quite central of being aware of the characteristics of a research field in order to overcome possible methodological deficiencies, for example (Buchner et al., 2021). This is also important to know to understand the breadth of implementation (and any under-studied areas). As shown in appendix Table A, first-year students are represented in 4 studies, second and third-year students are also featured in 4 studies, postgraduate students are involved in 2 studies, mixed undergraduate and postgraduate groups appear in 1 study, unspecified undergraduate students are included in 7 studies, unspecified students are represented in 2 studies, and academics are the focus of 1 study.

The geographic distribution of studies summarised in Figure 2 is relatively diverse, with a significant representation from Asia and Europe (29% each; n=6), closely followed by Oceania (24%; n=5), North America (9%; n = 2), Africa and Middle East (9%; n = 2). The articles reviewed include 18 published in 2023 and three from early 2024, consistent with the literature search being conducted in February 2024. These data are derived from the included studies listed in Appendix Table A.

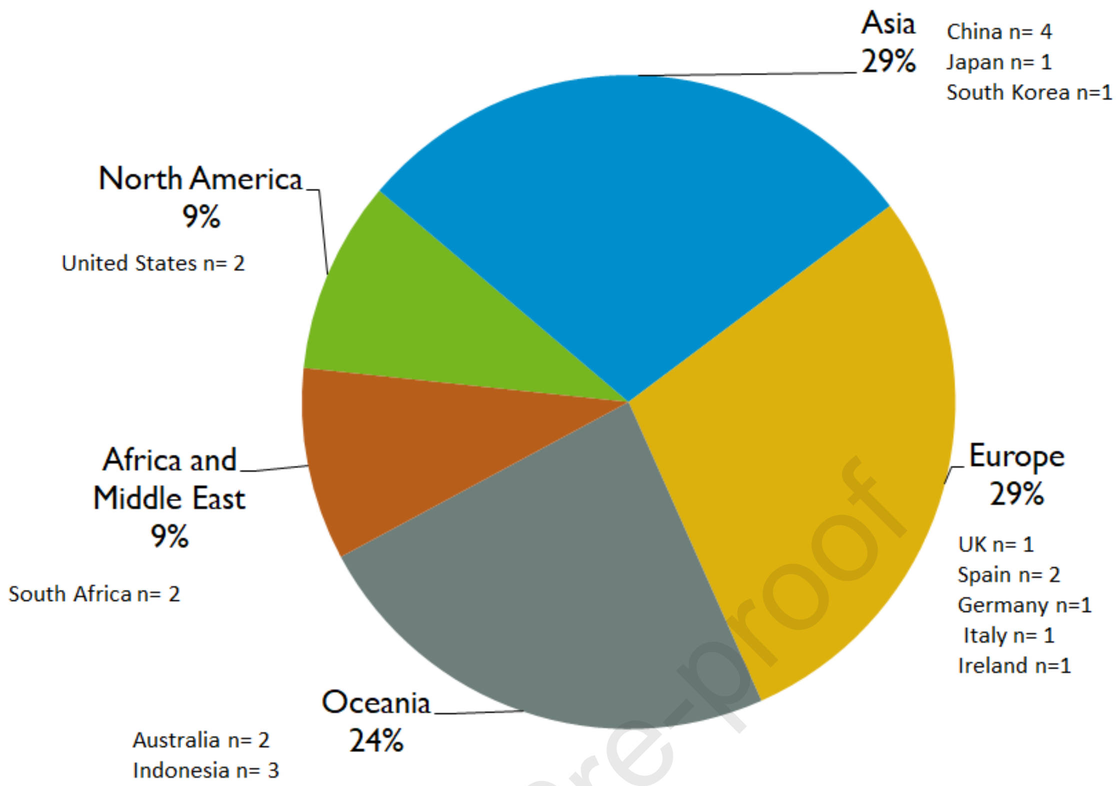



Figure 2. Geographical distribution of selected papers. (n = 21)

Discipline distribution (Figure 3), derived from the date listed in Appendix Table A, shows that language courses, of all disciplines, have the highest number of related articles at 33% (n=7). This focus is largely due to the significant impact GenAI tools like ChatGPT have on language-related academic tasks, particularly in writing. For example, Nikolic et al. (2023) noted that ChatGPT could produce work that satisfies marking rubrics when provided with the right prompts. This increased interest in the impact of GenAI on student writing is driven by academic concerns regarding breaches of academic integrity, as educators seek to understand and mitigate potential issues of academic dishonesty while still leveraging the benefits of these advanced tools for educational enhancement.

Following Languages, the fields of Information & Communication Technologies (29% ; n=6) and Engineering & Science (19% ; n=4), demonstrate the applicability of GenAI in more technical academic tasks, especially coding. One remarkable example of what GenAI can achieve in the field of coding is its ability to automate the creation of complex software applications. Given the many advantages of GenAI, academics are exploring alternative methods to traditional STEM education.

Education disciplines (9%, n=2), complemented by contributions from the Humanities and Social Sciences (5%, n=1) and Multidisciplinary case studies (5%, n=1), reflect the widespread impact of GenAI across the entire education system. This variety suggests that as more tailored applications of GenAI are developed, its influence could extend further, enhancing various fields of study.

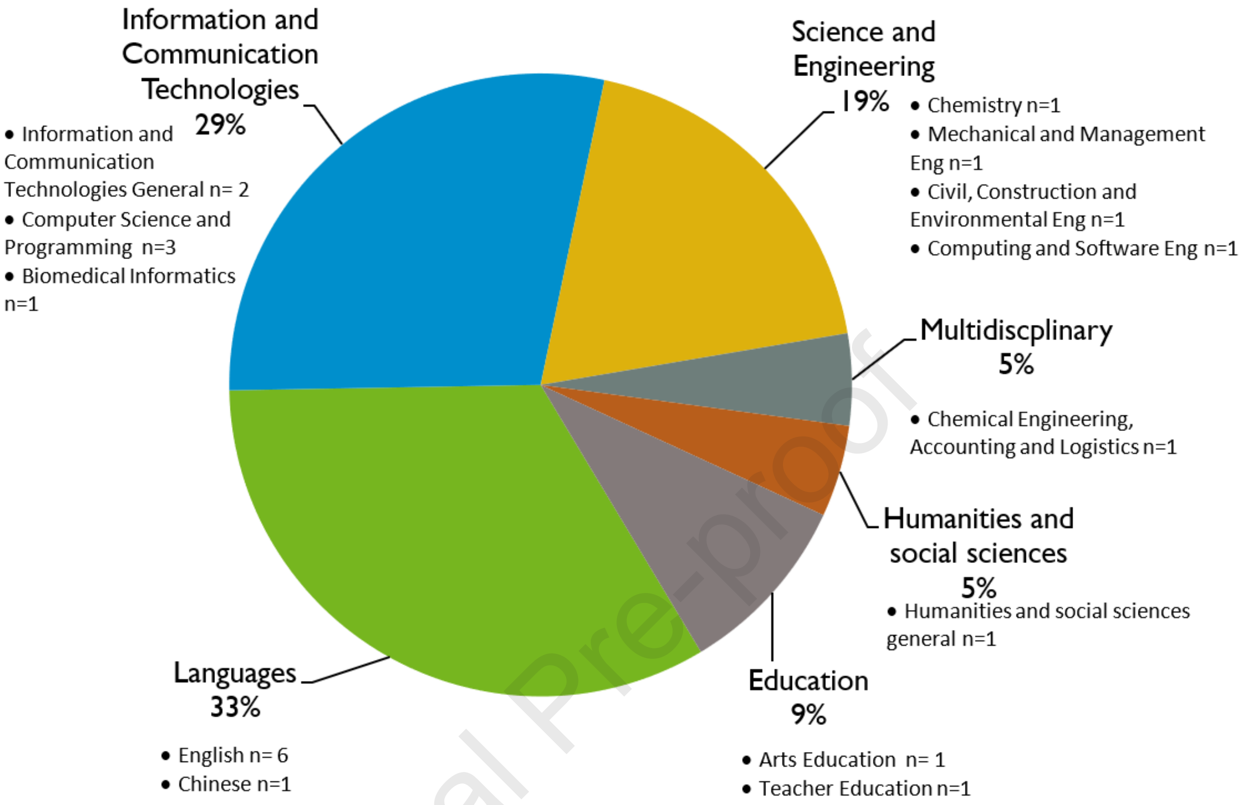



Figure 3. Discipline distribution of selected papers. (n = 21)

The aims of the selected research papers (Table A) are diverse, reflecting the wide-ranging potential of GenAI in transforming higher education practice across various disciplines. These studies collectively focus on leveraging GenAI to enhance learning experiences, develop specific skills, and assess the impact of this technology on both students and educators.

For example, in science education, Exintaris et al. (2023) aim to enhance problem-solving and critical thinking in chemistry via ChatGPT-generated prompts. In engineering, Zhao et al. (2023) assess GenAI's potential and limitations, focusing on student experiences with ChatGPT. Kirwan (2023) in humanities addresses concerns about Large Language Models like ChatGPT by exploring their applications and related discussions, and in Information Technology, Wang and Feng (2024) study the impact of prompt engineering on information retrieval skills.

Lu et al. (2024) compare teacher and ChatGPT feedback on Chinese writing to understand its effectiveness and student perceptions, while Kuramitsu et al. (2023) evaluate the impact of AI-based assistance in programming education, focusing on how GenAI can supplement traditional teaching methods by helping students address unresolved errors and clarify unknown terms.

These studies collectively reveal a significant interest in understanding how GenAI can transform teaching and learning practices across various disciplines, highlighting both opportunities and challenges in its integration into university curricula.

### 4.2 Study designs and research methods to address RQ2

The second research question aimed to determine the *study designs and research methods used to evaluate the effect of GenAI integration in higher education.* This is important to know to gauge the reliability and validity of findings that can guide effective implementation and policy decisions. The selected publications were classified into qualitative, quantitative, and mixed-methods designs (Table 2) according to Creswell's framework (Creswell, 2014). The classification utilized extracted data that covered study methods, educational theories, study outcomes and key findings.

The majority of publications (15 of the 21 studies) implemented mixed-method research designs. For example, Kuramitsu et al. (2023) conducted qualitative and quantitative analysis of student interactions with coding assistant GenAI tool, and Uddin et al. (2023) combined case study quantitative evaluations with qualitative student surveys in a construction program.

Solely quantitative research designs were used in two studies (Wang, Wang, et al., 2024),(Qureshi, 2023) to measure performance metrics objectively. Wang, Wang, et al. (2024) assessed the impact of prompt engineering on students' ability to find information using ChatGPT by conducting statistical analysis of the task completion qualities, and Qureshi (2023) scored performance of student teams in programming challenges.

Qualitative observations were conducted in four studies (Kirwan, 2023; Pitso, 2023; Widiati et al., 2023; Zhao et al., 2023) often involving qualitative student surveys and classroom performance observations. For example, Pitso (2023) studied how multidisciplinary students tested out ChatGPT on their assignments and conducted post-study interviews to check lessons learnt and whether students were ready to embrace learning based on critical thinking, empowerment theory and GenAI systems.

The most common research methods across the studies were surveys and interviews, utilised in fifteen papers, of which fourteen researchers reported data from student's survey (Belda-Medina et al., 2023; Bernabei et al., 2023; Exintaris et al., 2023; French et al., 2023; Guo et al., 2023; Lu et al., 2024; Murillo-Ligorred et al., 2023; Pitso, 2023; Silitonga et al., 2023; Speth et al., 2023; Uddin et al., 2023; Wang & Feng, 2024; Zhao et al., 2023) and Widiati et al. (2023) interviewed teachers.

Eight studies evaluated the correlation between subject knowledge and GenAI integration into classroom activities. Five of these studies assessed student performance through pre- and post-GenAI intervention subject knowledge tests and assignments (Elkhodr et al., 2023; Guo et al., 2023; Qureshi, 2023; Uddin et al., 2023). Assignments completed with the assistance of

GenAI were evaluated in two studies: Bernabei et al. (2023) and Lu et al. (2024). Exintaris et al. (2023) studied students' ability to critically evaluate solutions generated by ChatGPT and to identify errors. Finally, Wang, Wang, et al. (2024) investigated whether subject-related prior knowledge affects the quality of the answers obtained from ChatGPT by the students.

Six publications employed a quasi-experimental design to compare groups of students with and without GenAI interventions. Of these, the studies by Wang, Wang, et al. (2024) and Qureshi (2023) focused on quantitative assessments, while the studies by Silitonga et al. (2023), Elkhodr et al. (2023), Khang et al. (2023) and Guo et al. (2023) used mixed methods to provide a more comprehensive analysis of how GenAI enhances student performance and engagement.

Table 2 Classification of final papers by methodology (n=21)

| Study design | Citation | Research methods |
|---|---|---|
| **Quantitative** | Wang, Wang, et al. (2024) | Quasi-experimental research (experimental vs control group), statistical analysis of student performance |
| | Qureshi (2023) | Quasi-experimental research, Quantitative evaluation of student performance |
| **Qualitative** | Zhao et al. (2023) | Qualitative student survey |
| | Kirwan (2023) | Student performance observation |
| | Widiati et al. (2023) | Qualitative teacher interviews |
| | Pitso (2023) | Qualitative Empowering Education method called Evaluation Design, semi-structured student interviews |
| **Mixed** | Silitonga et al. (2023) | Quasi-experimental research,qualitative student performance and survey motivation evaluations |
| | Bernabei et al. (2023) | Student survey, quantitative and qualitative student assignment evaluation |
| | Exintaris et al. (2023) | Evaluation of student performance, qualitative student interviews, metacognitive tools |
| | Lu et al. (2024) | Quantitative evaluation of student performance, qualitative student survey followed by the quantitative analysis of the responses |
| | Kuramitsu et al. (2023) | Qualitative and quantitative evaluation of student interaction and engagement with GenAI |
| | Speth et al. (2023) | Student survey, observation of the case study |
| | Elkhodr et al. (2023) | Quasi-experimental research, qualitative reflective exercise analysis and instructor observations, qualitative rubric scores |
| | Murillo-Ligorred et al. (2023) | Semi-structured student interviews, data analysis using the constant comparative method |
| | Belda-Medina and Kokošková (2023) | Qualitative and quantitative student surveys and interviews |
| | French et al. (2023) | Qualitative student survey and semi-structured interviews, quantitative data analysis |
| | Khang et al. (2023) | Quasi-experimental design, student observations, inferential statistical analysis |

| Study design | Citation | Research methods |
|---|---|---|
| | Guo et al. (2023) | Quasi-experimental design, student performance evaluation through pre-test and post-test, student's survey |
| | Uddin et al. (2023) | Case study observations, results evaluation, qualitative student survey |
| | Han et al. (2023) | Quantitative student survey. Student performance observations, focus group interviews |
| | Wang and Feng (2024) | Student survey and semi-structured group interviews, qualitative skills assessment |



### 4.3 Laurillard's Conversational Framework (LCF) classification of selected studies to address RQ3

The third research question was to determine *how can the effectiveness of the GenAI implementation in higher education can be analysed through the lens of the conceptual educational frameworks (LCF, SMAR and TPACK)?* This is important to check the alignment of GenAI integration with established pedagogical theories in seeking to enhance educational outcomes.

First we classified the final 21 papers based on the pedagogic theory of Laurillard's Conversational Framework (LCF) and the concept of learning types (Laurillard, 2012). This framework has proven effective in aiding educators to describe and discuss the student learning process comprehensively. Laurillard (2012) identifies six distinct learning types that facilitate different aspects of the educational experience, each fostering unique skills and competencies in learners:

**Acquisition**: This type of learning involves students exploring ideas provided by their teachers. It is characterized by the transmission of knowledge from educator to student, typically through lectures, readings, or multimedia content.

**Investigation**: Students engage in learning through investigation where they explore, compare, and critique texts, documents, and resources. This process encourages learners to delve deeper into the subject matter, reflecting on the concepts and ideas being taught and developing critical thinking skills.

**Practice:** Learning through practice requires learners to engage in activities where they must apply what they have learned to complete specific tasks. This learning type emphasizes the importance of iterative feedback and adapting one's approach based on this feedback to meet the learning objectives effectively.

**Discussion**: This learning type necessitates learners articulating their ideas and questions and responding to those posed by their teachers or peers. Discussion fosters a deeper understanding through dialogue and can often lead to new insights and perspectives.

**Collaboration**: In collaborative learning, students work together to solve problems or complete projects. This type of learning is about the process of engaging with others, sharing ideas, and developing solutions as a group, which mirrors many real-world work environments.

**Production:** Production involves students creating a tangible or digital product that demonstrates their understanding of the subject matter. This type of learning is driven not just by the feedback from teachers but also by the motivation to create a public output that has real-world application or academic value.

Table 3 demonstrates the varied applications and outcomes of the final set of papers integrating GenAI tools into different learning environments across the educational spectrum.

Table 3 Classification of final papers by Laurillard's Conversational Framework (n=21)

| LCF Learning type | Reference | Learning Activities Implemented | Outcome |
|---|---|---|---|
| **Acquisition** | Wang and Feng (2024) | ChatGPT used for reading assistance and analysis in an English reading class | Explored whether ChatGPT could facilitate the exploration of English original masterpieces in a comparative study |
| **Inquiry** | Kirwan (2023) | Introductory class to demonstrate how to operate ChatGPT and identify the affordances and limitations | Provided an overview of ChatGPT in academia and moderated fears about its potential |
| | Elkhodr et al. (2023) | Allowing students to use ChatGPT in tutorials as a tool, compared to the traditional approach of web searches and lecture material | ChatGPT proved to be a valuable tool in assisting students in generating user flows and ideas |
| | Han et al. (2023) | Using ChatGPT as instructor assisting with essay revision | Most students reported positive experience with ChatGPT, however students with lower technology skills faced challenges |
| **Discussion** | Exintaris et al. (2023) | Workshop to critique provided ChatGPT-generated solutions for problems | Enhanced metacognitive and critical thinking skills |
| | Murillo-Ligorred et al. (2023) | Classroom discussions and analysis of deepfake images | Develop critical thinking about deepfakes |
| | Wang, Wang, et al. (2024) | Prompt engineering applied in flipped classrooms for information retrieval | Observed positive influence of mastering prompt engineering on the effectiveness of information retrieval from ChatGPT |
| **Practice** | Kuramitsu et al. (2023) | AI-based assistance for programming education | Reduced number of basic questions to teachers |
| | Widiati et al. (2023) | Students used AI tools to assist them in creative and other writing tasks in English | Teachers reported positive impact on student writing, such as idea generation, vocabulary and |

| | | | |
|---|---|---|---|
| | | | organisation |
| | Bernabei et al. (2023) | Students engaged with ChatGPT to write an essay and then ChatGPT used to detect if the essay AI generated | Students enhanced understanding of GenAI benefits and limitations and emphasized the crucial role of teachers in the educational process |
| | Silitonga et al. (2023) | AI chatbot-based learning in English writing classroom | Increased students' motivation after using ChatGPT |
| | Speth et al. (2023) | AI-generated exercises implemented in programming courses | Evaluated the use of AI teaching materials in coding education |
| | Qureshi (2023) | Students were encouraged to engage with ChatGPT for help with the programming problems | Students using ChatGPT improved their scores, but teams using ChatGPT-generated code faced difficulties with test cases in the Programming Contest Control environment |
| | Khang et al. (2023) | Students learning English using different chatbots | Students may lack experience with GenAI |
| | Guo et al. (2023) | Chatbot-assisted debates to enhance students' argumentation skills and motivation | Chatbot improves the argumentative skills and motivation in students |
| | Uddin et al. (2023) | ChatGPT integrated into construction hazard recognition curriculum | ChatGPT significantly improved students' construction hazard recognition ability |
| | Lu et al. (2024) | Combination of teacher and ChatGPT feedback | Evaluated the effectiveness of ChatGPT and teacher feedback in assessing student writing. Improved student's writing |
| **Collaboration** | Belda-Medina and Kokošková (2023) | Students completed assessment report with the assistance of Chatbots | Found moderate level of student satisfaction with linguistic chatbots |
| **Production** | French et al. (2023) | Students were given a research and development assignment that explicitly required them to engage with OpenAI tools | The integration of OpenAI tools into the curriculum was both productive and popular |
| | Pitso (2023) | Students using ChatGPT to complete assignment and resolve social problem | ChatGPT significantly lessens assignment completion time and improves problem-solving abilities |
| | Zhao et al. (2023) | Using ChatGPT to design a course; students creating materials using ChatGPT | Explored the potentials and limitations of AI in the classroom, learning students' perceptions and experiences |

Specific implementations of AI technologies were extracted from the selected publications and categorised by the LCF learning types, and collated along with the reported outcomes, which commonly range from improved student performance to enhanced engagement with course content.

For instance, under ***Acquisition***, Wang and Feng (2024) used ChatGPT to assist in reading and analysing English literature, highlighting the tool's capacity to facilitate a deeper understanding of complex texts. In the ***Inquiry*** learning type, multiple studies such as those by Kirwan (2023) and Kong et al. (2023) illustrate how ChatGPT serves as a powerful tool for enhancing student investigation skills, aiding in everything from introductory tutorials to the development of methodologies in chemical engineering design. This not only improves understanding but also engagement with course material.

***Discussion*** activities, represented by Exintaris et al. (2023) and Murillo-Ligorred et al. (2023), focus on utilizing ChatGPT for critical thinking and discussing contemporary issues like deepfakes, which plays a crucial role in developing students' analytical skills. In the ***Practice*** learning type, studies such as those by Wang, Wang, et al. (2024) and Kuramitsu et al. (2023) reveal how ChatGPT aids in practical applications like prompt engineering and programming assistance, significantly reducing basic inquiries to instructors and enhancing information retrieval effectiveness.

***Collaboration*** is exemplified by Lu et al. (2024), where the combination of teacher and ChatGPT feedback was found to improve student writing, showing the tool's effectiveness in collaborative settings. And finally, studies corresponding to the ***Production*** type of learning, such as those by Belda-Medina and Kokošková (2023)) and Pitso (2023), explore how ChatGPT can assist in creating more practical outputs such as research reports and assignments, demonstrating its utility in producing tangible academic products.

By classifying the selected papers through the LCF framework, we can identify how GenAI tools facilitate various learning types and their potential to support holistic educational experiences. For example, acquisition learning types benefit from GenAI's ability to deliver content interactively, while collaborative learning activities leverage GenAI for group problem-solving. For educators, this classification is practical as it highlights specific learning types where GenAI has demonstrated success, providing actionable insights to incorporate these tools effectively into their teaching practices.

### 4.4 Substitution, Augmentation, Modification and Redefinition (SAMR) analysis to address RQ3

Our next approach to address RQ3 is to analyse the included papers against the ***Substitution, Augmentation, Modification and Redefinition*** (SAMR) framework.

The SAMR framework was developed and first promoted by Dr Ruben Puentedura in 2009 (Puentedura, 2009). The four levels are grouped in to two main layers representing different degrees of technology integration into learning experiences: ***enhancement*** and ***transformation*** (for more information see Blundell et al. (2022)).

At the ***enhancement*** layer, technology enhances the existing educational process without fundamentally changing it. The focus is on improving efficiency and adding features. Its two levels are:

- **Substitution:** Technology acts as a direct substitute for traditional tools, with no functional change. For example, a word processor can be used to type an essay instead of handwriting it.
- **Augmentation:** Technology acts as a direct substitute, but with functional improvements. For example, using a word processor with spell check and grammar suggestions to type an essay.

At the ***transformation*** layer, technology transforms the learning experience by creating new opportunities and ways of learning that were not possible before. This can lead to deeper understanding and engagement. Its levels are:

- **Modification:** Technology allows for significant task redesign. For example, students collaborate on a shared Google Doc to write and edit an essay in real-time, providing immediate feedback to each other.
- **Redefinition:** Technology allows for the creation of new tasks that were previously inconceivable. For example, students create a multimedia presentation or a video essay incorporating various digital tools, conduct research online, and collaborate with peers from around the world.

To conduct this analysis, we analysed extracted data, such as types of learning activities, study outcomes, summaries of GenAI implementation, and classified each learning activity mentioned in the selected articles through the SAMR Framework. Examples of the technology integration extracted from the GenAI implementation case studies are shown in Table 4, along with the justification for the allocation of each example to the different SAMR levels.

Most of the selected studies reported multiple teaching activities where GenAI was implemented, and, therefore, each study was allocated to all possible SAMR levels of integration that are reflected in the publication. For example, Exintaris et al. (2023) uses ChatGPT to generate solutions to chemical problems, classified as ***Substitution*** because this task substituted the teacher's work without causing any significant changes to the teaching and learning process. However, the solutions created were used by students to analyse and determine errors that enhance students' critical thinking. As this level of integration describes a teaching activity where technology is used to redesign learning activities with significant functional improvement compared to human assistance, it is classified as ***Modification***.

At the ***Substitution*** level, GenAI is used to directly replace traditional methods without significantly enhancing or changing the educational outcome. As shown in Table 4, examples include using GenAI for basic skills delivery in subjects like math and language, and simulating conversations for practice in language learning or customer service.

The ***Augmentation*** level, frequently referenced across the studies, enhances traditional educational tools by integrating advanced functionalities such as grammar checks and assistance with writing tasks. This focus is logical, considering the third of the selected studies are from language disciplines.

Moving to the higher tiers of the SAMR model, ***Modification*** and ***Redefinition*** also feature prominently. Modification, as evidenced in practices like real-time feedback and dynamic brainstorming sessions, significantly alters traditional educational activities. Redefinition goes a step further by creating completely new tasks and methods of engagement, such as integration of ChatGPT into the course material to enhance student learning (Zhao et al., 2023) or GenAI integration into Computing and Digital Media research and development assignment (French et al., 2023). The examples of Modification and Redefinition demonstrate the potential for GenAI tools to profoundly shape and enhance the future of education.

For educators, SAMR can be used to assess the impact of GenAI on pedagogy and learning outcomes. For example, the framework allows researchers and educators to distinguish between basic uses of GenAI, such as grammar checking (augmentation), and transformative uses, like creating entirely new learning tasks (redefinition). This distinction is particularly valuable for educators who are just beginning to explore the integration of GenAI into their teaching practices. At the substitution and augmentation levels, educators can experiment with simple implementations, such as using GenAI for automated feedback on writing assignments or generating sample problems for students. These initial steps provide a low-risk entry point, enabling educators to become familiar with the capabilities of GenAI without significantly altering their existing teaching methods.

As educators gain confidence, the framework guides them toward more advanced applications in the modification and redefinition stages. For instance, in modification, educators might redesign a traditional brainstorming activity by incorporating GenAI to provide real-time, GenAI-generated prompts that inspire deeper critical thinking or collaborative problem-solving. At the redefinition level, Ge19, nAI could be used to create entirely new tasks, such as virtual role-playing scenarios or interactive simulations that were previously impossible to achieve with traditional tools.

Table 4. An implementation analysis against the SAMR framework

| **SAMR Classification** | **Example of technology integration** | **Justification** | **Reference** |
|---|---|---|---|
| **Substitution** | Teach fundamental skills | Using GenAI to deliver fundamental skills, such as basic math or language rules, can directly replace traditional teaching methods without altering the educational function. | Wang et al. (2023)<br>Kuramitsu et al. (2023)<br>Silitonga et al. (2023)<br>Elkhodr et al. (2023)<br>Belda-Medina and Kokošková (2023)<br>Han et al. (2023) |
| | Simulate conversations | GenAI can simulate conversations for language learning or customer service training, substituting for human interaction without adding significant functionality. | Widiati et al. (2023)<br>Belda-Medina and Kokošková (2023)<br>Khang et al. (2023) |
| | Answer student enquiries/ generate | This use of GenAI acts as a direct substitute for teacher or tutor responses, providing information without changing the learning | Kirwan (2023)<br>Wang and Feng (2024)<br>Kuramitsu et al. (2023) |

| | | | |
|---|---|---|---|
| | solutions | dynamics. | Bernabei et al. (2023)<br>Qureshi (2023) |
| **Augmentation** | Performs grammar checks | GenAI enhances traditional spell-checkers by understanding context and suggesting more accurate grammatical corrections | Widiati et al. (2023)<br>Kuramitsu et al. (2023)<br>Belda-Medina and Kokošková (2023)<br>Khang et al. (2023) |
| | Assist with writing tasks | GenAI provides enhancements like suggestions, improvements, or structural help that goes beyond simple text processing | Widiati et al. (2023)<br>Kirwan (2023)<br>Bernabei et al. (2023)<br>Belda-Medina and Kokošková (2023)<br>Han et al. (2023) |
| | Assist with problem solving | GenAI offers tools or methods that improve the problem-solving process, such as step-by-step guides or interactive aids. | Qureshi (2023) |
| **Modification** | Provide immediate feedback and suggestions | GenAI can analyse student responses in real-time and offer tailored feedback | Wang et al. (2023)<br>Lu et al. (2024)<br>Silitonga et al. (2023)<br>Elkhodr et al. (2023)<br>Belda-Medina and Kokošková (2023)<br>Khang et al. (2023)<br>Guo et al. (2023)<br>Han et al. (2023) |
| | Facilitates brainstorming | GenAI provides unique ways to enhance or manage the brainstorming process that wouldn't be possible with traditional tools. | Elkhodr et al. (2023);<br>Zhao et al. (2023)<br>Pitso (2023)<br>Guo et al. (2023) |
| | Assist with the analysis of errors | GenAI allows for deeper analysis or a more comprehensive review than traditional methods. | Lu et al. (2024)<br>Kuramitsu et al. (2023)<br>Uddin et al. (2023) |
| | Enhance critical thinking | GenAI challenges students or provides scenarios that require higher-order thinking that goes beyond traditional tasks. | Zhao et al. (2023)<br>Kirwan (2023)<br>Murillo-Ligorred et al. (2023) |
| **Redefinition** | Generate new ideas | GenAI uniquely generates ideas that would not have been possible without it, potentially through AI-driven insights. | Zhao et al. (2023)<br>Pitso (2023) |
| | Generate new work samples (text, images etc.) | Uses GenAI to create unique content or samples that extend beyond simple templates or modifications. | Speth et al. (2023)<br>Zhao et al. (2023)<br>Murillo-Ligorred et al. (2023)<br>French et al. (2023) |
| | Generates new method of learning | GenAI introduces completely new ways of learning or interacting with the material that fundamentally transforms the educational experience. | Belda-Medina and Kokošková (2023)<br>Uddin et al. (2023) |
| | Generative problem-solving environments | GenAI creates complex, real-world problems that students must solve, offering a platform for innovative thinking and solution development | French et al. (2023)<br>Pitso (2023) |

| |
|---|
| that wouldn't be possible without AI. |

### 4.5 Technological Pedagogical Content Knowledge (TPACK) framework analysis (RQ3) and recommendations to academics (RQ4)

In our final approach to address RQ3 and to answer RQ4, we utilised the TPACK (Technological Pedagogical Content Knowledge) framework, a model that describes the knowledge and skills needed by teachers to effectively integrate technology into their teaching practices (Niess, 2002). The TPACK framework can provide a scaffold for considering what teachers need to know to use any technology effectively (Mishra et al., 2023).

In the TPACK framework, there are four principal areas of knowledge related to technology: ***Technological Knowledge (TK)*** as well as in the overlapping spaces that constitute ***Technological Content Knowledge (TCK)***, ***Technological Pedagogical Knowledge (TPK)***, and ***Technological Pedagogical Content Knowledge (TPACK).*** Additionally, a recent and significant shift in the research on TPACK has been an emphasis on understanding ***Contextual Knowledge (XK)***—a recognition of the fact that context matters and impacts what educators can and cannot do (Mishra, 2019).

To provide the description of the TPACK knowledge domains in the context of GenAI integration, we synthesized definitions from the Celik (2023) empirical study on the ethical integration of artificial intelligence (AI) and Mishra et al. (2023) research exploring TPACK in the age of ChatGPT and GenAI. The descriptions of GenAI-TPACK factors are as follows:

- **GenAI-TK**: Focuses on understanding and proficiency with GenAI tools themselves. It involves knowing how to interact with these tools, execute tasks, and utilize them effectively in specific subject area or classroom.
- **GenAI-TPK**: Addresses the integration of GenAI tools in teaching practices. It includes understanding how these tools can enhance teaching methods, such as providing adaptive and personalised feedback, real-time assessments, and monitoring student learning. Teachers should know how to leverage GenAI to support diverse teaching strategies and scaffold students' learning experiences.
- **GenAI-TCK**: Combines knowledge of GenAI tools with subject-specific content. It encompasses the ability to use GenAI to search for, curate, and create educational material relevant to the teacher's field. Teachers should be familiar with best practices for using GenAI to enhance their understanding and explanation of subject content and know how to integrate these tools into their teaching.
- **GenAI-TPACK:** Represents the comprehensive understanding of how to effectively combine GenAI tools with pedagogical strategies and subject content. It involves creating lessons that integrate GenAI tools to enhance student engagement and learning outcomes.

- **GenAI-XK:** Focuses on the broader context within which GenAI tools are used, including policies, ethical considerations, and practical constraints. This domain also involves assessing the fairness and inclusivity of GenAI tools and advocating for their responsible use in education.

From selected publications, the examples of **GenAI-TK** include delivering introductory classes demonstrating how to operate ChatGPT (Kirwan, 2023) or studying the advantages and limitations of ChatGPT (Exintaris et al., 2023; Wang & Feng, 2024). For **GenAI-TPK**, the studies emphasise enhancing teaching practices, like providing personalised feedback on student work (Belda-Medina & Kokošková, 2023; Elkhodr et al., 2023; Guo et al., 2023; Han et al., 2023; Khang et al., 2023; Lu et al., 2024; Silitonga et al., 2023; Wang et al., 2023; Widiati et al., 2023), supporting flipped-learning classrooms (Wang & Feng, 2024), or assisting with writing tasks by offering structural suggestions and improvements, thereby augmenting traditional teaching methods and enabling more dynamic and interactive learning experiences (Belda-Medina & Kokošková, 2023; Bernabei et al., 2023; Han et al., 2023; Kirwan, 2023; Widiati et al., 2023).

In the context of **GenAI-TCK**, the majority of selected studies implemented GenAI with subject-specific content, as this is the main scope of the systematic literature review. The examples range across disciplines, including ChatGPT integration into a construction hazard recognition curriculum (Uddin et al., 2023), using GenAI to assist in learning English (Khang et al., 2023; Widiati et al., 2023), or applying GenAI in programming studies (Speth et al., 2023).

The examples of how teachers integrate pedagogical strategies, content knowledge, and technology to create transformative learning experiences (**GenAI-TPACK**) include the integration of ChatGPT into a course, where the instructor used the tool to facilitate student learning and develop course materials, demonstrating the comprehensive application of TPACK (Zhao et al., 2023), or when students used ChatGPT to complete assignments and resolve social problems (Pitso, 2023).

Finally, for **GenAI-XK**, we provide examples where studies emphasise the importance of understanding broader personal, cultural, political, and ethical implications of GenAI. For instance, one study investigates the moral and legal aspects of GenAI-generated deepfakes (Murillo-Ligorred et al., 2023), and another explores the level of trust towards GenAI, the social influence that content generates on students, and whether students judge ChatGPT fairly and with moral standards (Bernabei et al., 2023).

Summarising the TPACK analysis and to answer RQ4, we developed the GenAI-TPACK diagram (Figure 4) to assist academics in GenAI implementation in teaching practices. The diagram is based on the canonical TPACK diagram (Mishra, 2019) and expanded with the key teacher skills and knowledge recommended for GenAI implementation in the different TPACK technical domains. The lists of recommended knowledge is adapted from the Intelligent-TPACK scale developed by Celik (2023) and amended to address the characteristics of GenAI. Each overlapping domain in the diagram—such as GenAI-TPK, GenAI-TCK, and GenAI-TPACK—is further explained in the accompanying text to clarify how GenAI tools contribute to specific pedagogical, content, and technological teaching practices. This model provides a comprehensive framework for educators, outlining the necessary technological, pedagogical, and content knowledge, as well as contextual understanding, required to effectively integrate GenAI tools into educational settings. By following this model, educators can enhance their teaching strategies, personalise learning experiences, and ensure ethical and effective use of GenAI technologies in the classroom.

This study does not expand on the TPACK areas of Content Knowledge (CK), Pedagogical Knowledge (PK), and their overlap in Pedagogical Content Knowledge (PCK), as these are beyond the scope of this study and therefore are not detailed with required teacher skills.

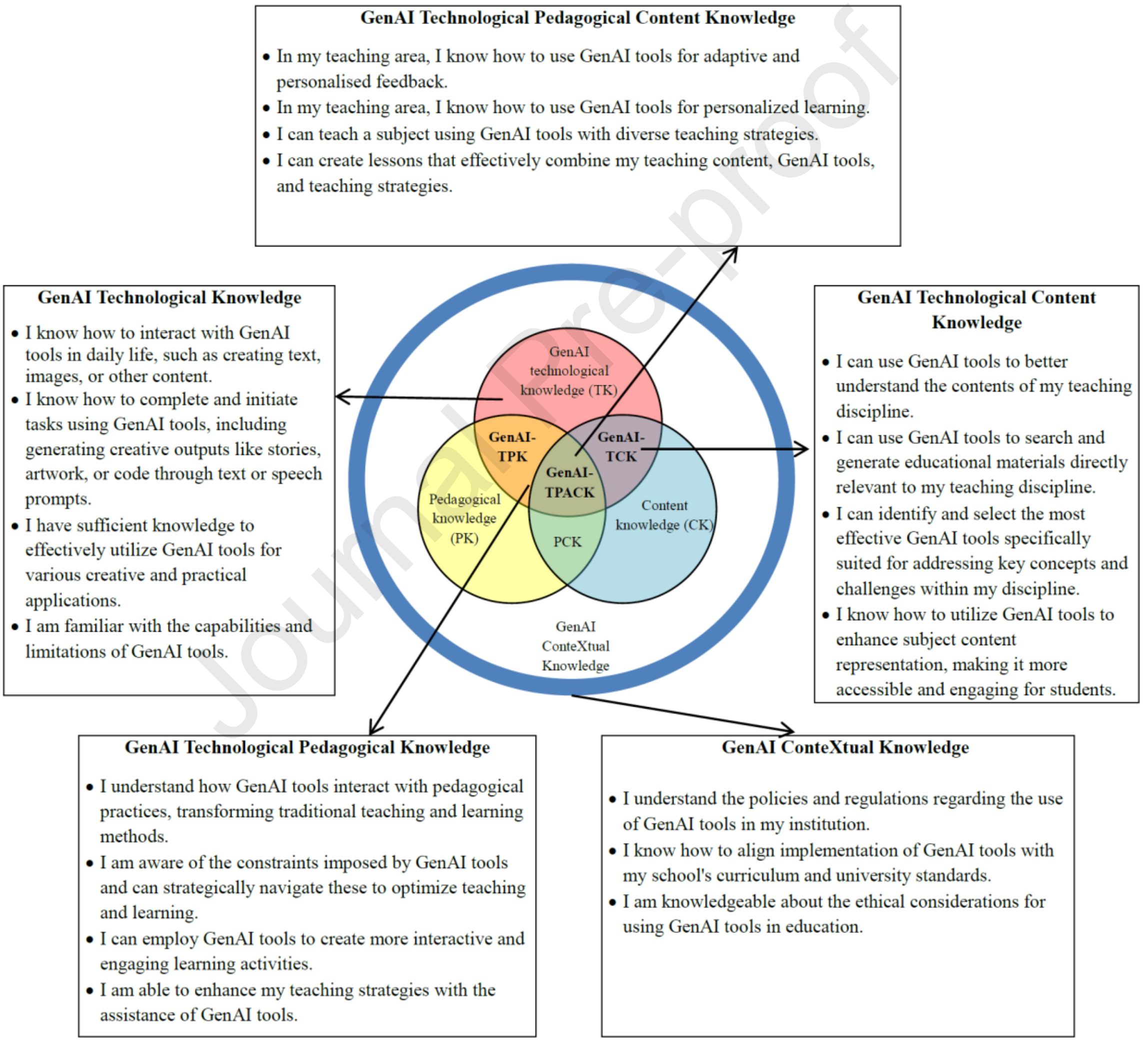



Figure 4. GenAI-TPACK diagram with the key teacher skills and knowledge recommended for GenAI implementation.

The developed GenAI-TPACK framework offers a structured approach for educators to integrate generative AI tools into their teaching practices effectively and ethically. It divides the required knowledge into distinct domains - each addressing specific aspects of GenAI integration. Educators start by developing foundational technological skills, such as understanding how to interact with GenAI tools and creating content like text or images. These basic capabilities allow educators to explore GenAI's potential in routine tasks, such as automating feedback or generating lesson materials, while building their confidence. As educators progress, they can leverage TPK to incorporate GenAI into teaching strategies, such as creating interactive activities or providing personalized learning experiences, which transform traditional approaches and enhance student engagement. TCK, on the other hand, focuses on applying GenAI within specific subject areas, enabling educators to generate discipline-relevant content, address unique challenges, and enrich their instructional methods.

At the heart of the framework is Technological Pedagogical Content Knowledge, which represents the integration of all domains. Educators operating at this level design lessons that seamlessly combine technology, pedagogy, and content to maximize learning outcomes. The inclusion of Contextual Knowledge (XK) further strengthens the framework by addressing the broader environment in which GenAI is used. Educators are guided to consider ethical issues, institutional policies, and regulatory requirements, ensuring that their use of GenAI is both responsible and aligned with institutional goals.

### 4.5 Reported limitations and future research opportunities to address RQ5

The fifth and final research question was to determine *the necessary methodological improvements and future research directions needed to better understand the impacts of GenAI in higher education.* This is important for gaining a deeper and more accurate understanding of the impacts of GenAI in higher education, thereby enhancing its effectiveness and integration. The key reported limitations include small sample sizes and the inconsistency in students' ability with prompt engineering. In addition, several studies have identified inherent technological limitations within GenAI systems used in educational research, including inconsistent responses, general unreliability, and outright errors in the AI's output (Elkhodr et al., 2023; Kong et al., 2023; Qureshi, 2023; Wang & Feng, 2024).

Numerous reports highlight scope limitations in research, particularly citing small sample sizes (Elkhodr et al., 2023; Lu et al., 2024; Murillo-Ligorred et al., 2023; Silitonga et al., 2023; Widiati et al., 2023; Zhao et al., 2023) and narrow subject focus (Kuramitsu et al., 2023; Uddin et al., 2023). Additionally, the short duration of many studies further impedes a comprehensive understanding of the long-term implications of GenAI technologies on higher education and student learning.

Wang et al. (2023) reported a limitation in assuming that students are already adept at using GenAI technologies such as prompt engineering. This could affect research results, as not all students may be at the same level of proficiency. Further compounding this issue is a lack of control over other significant variables like  motivation levels and prior knowledge (Elkhodr et al., 2023), which can greatly influence the outcome of educational interventions.

The generalizability and precision of research findings are noted in two studies (Belda-Medina & Kokošková, 2023; Pitso, 2023) to be limited by the specific conditions under which research is conducted, which may not accurately represent broader educational environments. This is certainly a limitation of all the papers, even if not explicitly acknowledged.

To overcome the limitations of narrow scopes and small sample sizes, future research should aim to involve larger and more diverse groups of participants as well as an evaluation based on students' proficiency. Extending studies across various educational areas (Belda-Medina & Kokošková, 2023; Kuramitsu et al., 2023; Zhao et al., 2023) could help validate the effectiveness of GenAI tools in a range of contexts. There is also a significant opportunity to investigate newer and broader versions of GenAI (Belda-Medina & Kokošková, 2023; Guo et al., 2023; Kuramitsu et al., 2023), which could help uncover additional capabilities of GenAI that may be beneficial in educational settings. This will remain an ongoing challenge while the powers of GenAI are evolving so rapidly and dynamically diversifying and strengthening.

Understanding the long-term impacts of GenAI on student learning is crucial (Belda-Medina & Kokošková, 2023; Widiati et al., 2023). Longitudinal studies could provide insights into how sustained use of GenAI affects learning outcomes, student engagement, and educational equity. The responsible use of GenAI in unsupervised settings (Elkhodr et al., 2023) and the ethical implications of GenAI in education (French et al., 2023; Kong et al., 2023) are highlighted as critical areas for future investigation.

Research focusing on the quantification of time savings for academics when employing GenAI in teaching could demonstrate some practical benefits (Speth et al., 2023). Likewise, exploring the automation of routine teaching practices, such as grading (Han et al., 2023), could offer insights into how GenAI can streamline educational processes and free up educator time for additional student-focused activities.

Specifically in terms of our study, the main limitation is due to the dynamic nature of the field of GenAI, this systematic literature review captures papers published from January 2023 to February 2024, potentially overlooking more recent developments. Additionally, it includes only papers written in English, which may exclude relevant research published in other languages. Future work should include a revised study based on the stated limitations which would gain further insights on the dynamic nature of GenAI through comparison studies between timeframes.

## Conclusions

This systematic review examines the actual impacts, challenges, and opportunities of integrating GenAI into higher education. By analysing selected papers through the lenses of the LCF, SAMR, and TPACK frameworks, this study demonstrates how these analytical tools enhance various learning environments. Our investigation, driven by five research questions, uncovered a variety of GenAI applications, strategies for implementation, and outcomes. These findings reveal GenAI's potential to elevate educational pedagogy, boost student motivation, and enrich the overall learning experience.

The findings from RQ1 highlight the diversity of GenAI applications across regions, participant demographics, and disciplines. However, gaps remain, particularly in underexplored fields such as the social sciences and interdisciplinary studies, as well as in geographic areas with limited representation. For educators, this emphasizes the opportunity to extend GenAI research and implementation into underrepresented areas while leveraging existing insights from more well-studied contexts.

The study designs and methods reviewed in RQ2 reveal the value of mixed-method approaches in capturing both quantitative and qualitative outcomes of GenAI integration. Educators can use these methods to evaluate the impacts of GenAI tools on student learning, engagement, and performance in their own teaching contexts.

Through RQ3, the use of conceptual frameworks such as LCF, SAMR, and TPACK provides structured guidance for educators seeking to align GenAI integration with pedagogical theories. These frameworks allow educators to start with basic uses, such as grammar checks, and progress toward transformative applications like collaborative simulations or interdisciplinary problem-solving. By adopting these frameworks, educators can strategically implement GenAI in ways that enhance teaching practices and improve learning outcomes.

Addressing RQ4, GenAI-TPACK framework was developed, which offers specific recommendations for the knowledge and skills required at different levels of GenAI integration. The framework emphasizes foundational skills, such as understanding GenAI tools and their capabilities, progressing toward more advanced knowledge of integrating these tools into pedagogy and subject-specific content. By following the framework, educators can develop expertise in areas such as adaptive learning design, ethical considerations, and personalized student feedback, ensuring a phased and effective implementation of GenAI in higher education.

Finally, RQ5 highlights the critical need for future research to address current limitations, such as narrow scopes, small sample sizes, and short study durations. Longitudinal studies are particularly important to understand the sustained impacts of GenAI on student learning, engagement, and equity. Moreover, research exploring the time-saving benefits of GenAI for educators and its potential to automate routine tasks, such as grading, could provide practical evidence for its value. Addressing ethical concerns, such as the reliability of GenAI in unsupervised settings, remains a key priority as these tools become more widely adopted.

Despite encountering issues such as bias and reliability, this research argues that critical thinking and structured pedagogical approaches can harness GenAI to tailor and enhance each student's learning journey. By aligning GenAI integration with pedagogical frameworks and addressing gaps in research, this study provides valuable insights for educators, academics, and policymakers. This review contributes to bridging gaps in current knowledge and lays the groundwork for leveraging GenAI's transformative potential in higher education.

## Acknowledgements

This work is an initiative of the Australasian Artificial Intelligence in Engineering Education Centre (AAIEEC), a special interest group of the Australasian Association of Engineering

Education (AAEE). This work was supported by the 2023 Australian Council of Engineering Deans (ACED) and the Australasian Association for Engineering Education (AAEE) Grant.

## Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGPT in order to optimize the language style in the original draft. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

## Declaration of Competing Interest

There are no conflicts of interest to disclose.

## Appendix

Table A. Characteristics of Included Studies Classified by Educational Level (n=21)

| Nº | Reference | Disciplines | GenAI used | Country | № of participants | Aim(s) |
|---|---|---|---|---|---|---|
| | | | **First-year students** | | | |
| 1 | Exintaris et al. (2023) | Science (chemistry) | ChatGPT 3.5 | Australia | 213 | To develop problem-solving skills and critical thinking by using ChatGPT-generated responses as prompts for critiques in a problem-solving workshop |
| 2 | Kirwan (2023) | Humanities and social sciences | ChatGPT 3.5 | Ireland | Not reported | To provide understanding of how GenAI technologies functions, their strengths and weaknesses, and to provide insights on how to best encourage students to think critically about their own potential use of GenAI. |
| 3 | Wang, Wang, et al. (2024) | Information Technology | ChatGPT 3.5 | China | 26 | To evaluate the impact of prompt engineering on college students' information retrieval skills using ChatGPT. |
| 4 | Khang et al. (2023) | English | My virtual Dream Friend and John English Bot | Indonesia | 36 | To determine how AI chatbots can help to learn English as a foreign language. |
| | | | **Second and Third-Year Students** | | | |
| 5 | Wang and Feng (2024) | English | ChatGPT | China | 83 | To investigate the effectiveness of ChatGPT in assisting with reading comprehension, analysis of narrative structure and language style, word explanation, and translation of sentences and paragraphs. |

| Nº | Reference | Disciplines | GenAI used | Country | № of participants | Aim(s) |
|---|---|---|---|---|---|---|
| 6 | Kuramitsu et al. (2023) | Computer science | ChatGPT 3.5 | Japan | 127 | To evaluate the impact of providing AI-based assistance to students for addressing unresolved errors, clarifying unknown terms, and explaining or modifying code, as an alternative to traditional support from teaching staff. |
| 7 | Belda-Medina and Kokošková (2023) | Teacher Education (English as a foreign language) | Chatbots: Mondly, Andy, John Bot and Buddy.ai | Spain, Czech Republic | 237 | To compare various linguistic and technological aspects of four App-Integrated Chatbots (AICs) and to investigate the perceptions of English as a Foreign Language teacher candidates regarding these chatbots. |
| 8 | Uddin et al. (2023) | Civil, Construction and Environmental Engineering | ChatGPT | USA | 42 | To explore if ChatGPT can aid hazard recognition when integrated into the curriculum of students pursuing a career in the construction industry. |
| | | | **Postgraduate students** | | | |
| 9 | Murillo-Ligorred et al. (2023) | Education (Arts) | Technology generating 'deepfake images' | Spain | 100 | To assess university students' (training to be teachers) ability to recognise deepfakes and their level of knowledge about this technology. |
| 10 | Bernabei et al. (2023) | Mechanical and management engineering | ChatGPT 3.5 | Italy | 31 | To examine the effectiveness of using ChatGPT as a learning tool among engineering students. Specifically, it involves students using ChatGPT to generate essays, assessing the detectability of these essays as AI-generated, and exploring student perceptions of large language models (LLMs) in the learning process. |
| | | | **Mixed undergraduate and postgraduate** | | | |
| 11 | Elkhodr et al. (2023) | Information and Communication Technologies | ChatGPT 3.5 | Australia | 52 | To explore the outcomes of using GenAI as an assistive tool in tutorials. |
| | | | **Unspecified undergraduate students** | | | |
| 12 | Lu et al. (2024) | Chinese writing | ChatGPT 3.5 | China | 46 | To compare teacher and ChatGPT feedback on student writing, examining the nature of the feedback, student perceptions of it, and how students use this feedback to revise their work. |
| 13 | Speth et al. (2023) | Computer programming | ChatGPT 3.5 | Germany | 9 | To evaluate the effectiveness of using GenAI teaching materials in coding education. |
| 14 | French et al. (2023) | Computing and Software Engineering | ChatGPT 3.5 or Dall-E-2 | UK | Not reported | To describe and evaluate students' experiences using AI tools. |

| Nº | Reference | Disciplines | GenAI used | Country | № of participants | Aim(s) |
|---|---|---|---|---|---|---|
| 15 | Pitso (2023) | Chemical Engineering, Accounting and Logistics | ChatGPT 3.5 | South Africa | 15 | To qualitatively examine the use of ChatGPT in students' assignments and problem-solving processes, focusing on the emerging learning dynamics and benefits associated with its integration into learning. |
| 16 | Qureshi (2023) | Computer science | ChatGPT 3.5 | Saudi Arabia | 24 | To investigate the effectiveness of ChatGPT in improving students' learning outcomes in the initial programming courses of a Computing degree. |
| 17 | Guo et al. (2023) | English | Argumate chatbot | China | 44 | To examine the impact of chatbots on students' argumentation skills and motivation. |
| 18 | Han et al. (2023) | English | RECIPE that uses ChatGPT | South Korea | 231 | To introduce and evaluate a novel learning platform called RECIPE (**R**evising an **E**ssay with **C**hatGPT on an **I**nteractive **P**latform for **E**nglish as a Foreign Language learners). |
| | | | **Unspecified students** | | | |
| 19 | Zhao et al. (2023) | Biomedical informatics | ChatGPT 3.5 and available AI tools | United States | 6 | To explore the potential and limitations of AI in the classroom; to investigate students' perceptions of and experiences with the application of ChatGPT in teaching and learning. |
| 20 | Silitonga et al. (2023) | English | ChatGPT 3.5 | Indonesia | 109 | To investigate the impact of AI chatbots on students' motivation to learn English. |
| | | | **Academics** | | | |
| 21 | Widiati et al. (2023) | English | Jenni AI<br>Quillbot<br>WordTune<br>ChatGPT<br>Copy.ai<br>Paperpal<br>Essay Writer | Indonesia | 4 | To investigate the types of AI writing tools used by English as a Foreign Language (EFL) teachers to enhance student writing quality, specifically in content and organization, and to explore teachers' perceptions of the impact of these tools on students' writing. |



## References


Alateyyat, S., & Soltan, M. (2024, 28-29 Jan. 2024). *Utilizing Artificial Intelligence in Higher Education: A Systematic Review.* Paper presented at the 2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETSIS).

Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies, 1*(1), 377-386. doi:10.1162/qss_a_00019 %J Quantitative Science Studies

Baidoo-Anu, D., & Ansah, L. (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*. doi:10.61969/jai.1337500

Belda-Medina, J., & Kokošková, V. (2023). Integrating chatbots in education: insights from the Chatbot-Human Interaction Satisfaction Model (CHISM). *International Journal of Educational Technology in Higher Education, 20*(1), 62. doi:10.1186/s41239-023-00432-3

Bernabei, M., Colabianchi, S., Falegnami, A., & Costantino, F. (2023). Students' use of large language models in engineering education: A case study on technology acceptance, perceptions, efficacy, and detection chances. *Computers and Education: Artificial Intelligence, 5*, 100172. doi:https://doi.org/10.1016/j.caeai.2023.100172

Blundell, C. N., Mukherjee, M., & Nykvist, S. (2022). A scoping review of the application of the SAMR model in research. *Computers and Education Open, 3*, 100093. doi:https://doi.org/10.1016/j.caeo.2022.100093

Bozkurt, A. (2023). Unleashing the potential of generative AI, conversational agents and chatbots in educational praxis: A systematic review and bibliometric analysis of GenAI in education. *Open Praxis*. doi:10.55982/openpraxis.15.4.609

Buchner, J., Buntins, K., & Kerres, M. (2021). A systematic map of research characteristics in studies on augmented reality and cognitive load. *Computers and Education Open, 2*, 100036. doi:https://doi.org/10.1016/j.caeo.2021.100036

Castillo-Segura, P., Alario-Hoyos, C., Kloos, C.D., Fernández Panadero, C. (2023, 23-27 Oct. 2023). *Leveraging the Potential of Generative AI to Accelerate Systematic Literature Reviews: An Example in the Area of Educational Technology.* Paper presented at the 2023 World Engineering Education Forum - Global Engineering Deans Council (WEEF-GEDC).

Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior, 138*, 107468. doi:https://doi.org/10.1016/j.chb.2022.107468

Chan, C. K. Y., & Lee, K. K. W. (2023). The AI generation gap: Are Gen Z students more interested in adopting generative AI such as ChatGPT in teaching and learning than their Gen X and millennial generation teachers? *Smart Learning Environments, 10*(1). doi:10.1186/s40561-023-00269-3

Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*: SAGE Publications.

Currie, G. M. (2023). Academic integrity and artificial intelligence: is ChatGPT hype, hero or heresy? *Semin Nucl Med, 53*(5), 719-730. doi:10.1053/j.semnuclmed.2023.04.008

Das, D., Kumar, N., Longjam, L. A., Sinha, R., Deb Roy, A., Mondal, H., & Gupta, P. (2023). Assessing the Capability of ChatGPT in Answering First- and Second-Order Knowledge Questions on Microbiology as per Competency-Based Medical Education Curriculum. *Cureus, 15*(3), e36034. doi:10.7759/cureus.36034

Elkhodr, M., Gide, E., Wu, R., & Darwish, O. (2023). ICT students' perceptions towards ChatGPT: An experimental reflective lab analysis. *STEM Education, 3*(2), 70-88. doi:10.3934/steme.2023006

Exintaris, B., Karunaratne, N., & Yuriev, E. J. J. o. C. E. (2023). Metacognition and Critical Thinking: Using ChatGPT-Generated Responses as Prompts for Critique in a Problem-Solving Workshop (SMARTCHEMPer). *100*, 2972-2980. doi:10.1021/acs.jchemed.3c00481

French, F., Levi, D., Maczo, C., Simonaityte, A., Triantafyllidis, S., & Varda, G. (2023). Creative Use of OpenAI in Education: Case Studies from Game Development. *7*(8), 81.

Geerling, W., Mateer, G. D., Wooten, J., & Damodaran, N. (2023). ChatGPT has Aced the Test of Understanding in College Economics: Now What? *The American Economist, 68*(2), 233-245. doi:10.1177/05694345231169654

Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., & Chartash, D. (2023). How Does ChatGPT Perform on the United States Medical Licensing Examination? The Implications of Large Language Models for Medical Education and Knowledge Assessment. *JMIR Med Educ, 9*, e45312. doi:10.2196/45312

Guo, K., Zhong, Y., Li, D., & Chu, S. K. W. (2023). Effects of chatbot-assisted in-class debates on students' argumentation skills and task motivation. *Computers & Education, 203*, 104862. doi:https://doi.org/10.1016/j.compedu.2023.104862

Hamilton, E. R., Rosenberg, J. M., & Akcaoglu, M. (2016). The Substitution Augmentation Modification Redefinition (SAMR) Model: a Critical Review and Suggestions for its Use. *TechTrends, 60*(5), 433-441. doi:10.1007/s11528-016-0091-y

Han, J., Yoo, H., Kim, Y., Myung, J., Kim, M., Lim, H., . . . Oh, A. (2023). *RECIPE: How to Integrate ChatGPT into EFL Writing Education*. Paper presented at the Proceedings of the Tenth ACM Conference on Learning @ Scale, Copenhagen, Denmark. https://doi.org/10.1145/3573051.3596200

Heinze, A., Procter, C., & Scott, B. (2007). Use of conversation theory to underpin blended learning. *1*(1-2), 108-120. doi:10.1504/ijtcs.2007.014213

Hobensack, M., von Gerich, H., Vyas, P., Withall, J., Peltonen, L. M., Block, L. J., . . . Song, J. (2024). A rapid review on current and potential uses of large language models in nursing. *Int J Nurs Stud, 154*, 104753. doi:10.1016/j.ijnurstu.2024.104753

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., . . . Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. doi:https://doi.org/10.1016/j.lindif.2023.102274

Khang, A., Muthmainnah, M., Seraj, P. M. I., Al Yakin, A., & Obaid, A. J. (2023). AI-Aided Teaching Model in Education 5.0. In A. Khang, V. Shah, & S. Rani (Eds.), *Handbook of Research on AI-Based Technologies and Applications in the Era of the Metaverse* (pp. 83-104). Hershey, PA, USA: IGI Global.

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. doi:10.1080/03323315.2023.2284901

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*, 100126. doi:https://doi.org/10.1016/j.dche.2023.100126

Kumar, S., Rao, P., Singhania, S., Verma, S., & Kheterpal, M. (2024). Will artificial intelligence drive the advancements in higher education? A tri-phased exploration. *Technological Forecasting and Social Change, 201*, 123258. doi:10.1016/j.techfore.2024.123258

Kuramitsu, K., Obara, Y., Sato, M., & Obara, M. (2023). *KOGI: A Seamless Integration of ChatGPT into Jupyter Environments for Programming Education*. Paper presented at the Proceedings of the 2023 ACM SIGPLAN International Symposium on SPLASH-E, Cascais, Portugal. https://doi.org/10.1145/3622780.3623648

Laurillard, D. (2012). *Teaching as a design science : building pedagogical patterns for learning and technology*. New York, NY: Routledge.

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education, 21*(2), 100790. doi:https://doi.org/10.1016/j.ijme.2023.100790

Lu, Q., Yao, Y., Xiao, L., Yuan, M., Wang, J., & Zhu, X. (2024). Can ChatGPT effectively complement teacher assessment of undergraduate students' academic writing? *Assessment & Evaluation in Higher Education*, 1-18. doi:10.1080/02602938.2024.2301722

Mhlanga, D. (2023). Open AI in Education, the Responsible and Ethical Use of ChatGPT Towards Lifelong Learning. *SSRN Electronic Journal*. doi:10.2139/ssrn.4354422

Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, Ernesto, D., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and Opportunities of Generative AI for Higher Education as Explained by ChatGPT. *13*(9), 856.

Mishra, P. (2019). Considering Contextual Knowledge: The TPACK Diagram Gets an Upgrade. *Journal of Digital Learning in Teacher Education, 35*(2), 76-78. doi:10.1080/21532974.2019.1588611

Mishra, P., Warr, M., & Islam, R. (2023). TPACK in the age of ChatGPT and Generative AI. *Journal of Digital Learning in Teacher Education, 39*, 1-17. doi:10.1080/21532974.2023.2247480

Muddam, A., & Reddy, S. (2023). *A Guide to Conversational AI: ChatGPT for Higher Education and Professional Development*.

Murillo-Ligorred, V., Ramos-Vallecillo, N., Covaleda, I., & Fayos, L. (2023). Knowledge, Integration and Scope of Deepfakes in Arts Education: The Development of Critical Thinking in Postgraduate Students in Primary Education and Master's Degree in Secondary Education. *13*(11), 1073.

Niess, M. L. (2002). Preparing Teachers to Teach Science and Mathematics with Technology. In D. Watson & J. Andersen (Eds.), *Networking the Learner: Computers in Education* (pp. 689-697). Boston, MA: Springer US.

Nikolic, S., Daniel, S., Haque, R., Belkina, M., Hassan, G. M., Grundy, S., . . . 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. doi:10.1080/03043797.2023.2213169

Nikolic, S., Sandison, C., Haque, R., Daniel, S., Grundy, S., Belkina, M., . . . Neal, P. (2024). ChatGPT, Copilot, Gemini, SciSpace and Wolfram versus higher education assessments: an updated multi-institutional study of the academic integrity impacts of Generative Artificial Intelligence (GenAI) on assessment, teaching and learning in engineering. *Australasian Journal of Engineering Education*, 1-28. doi:10.1080/22054952.2024.2372154

Nikolic, S., Suesse, T. F., Grundy, S., Haque, R., Lyden, S., Hassan, G. M., . . . Lal, S. (2024). Laboratory learning objectives: ranking objectives across the cognitive, psychomotor and affective domains within engineering. *European Journal of Engineering Education, 49*(3), 454-473. doi:10.1080/03043797.2023.2248042

OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., . . . and, S. B. a. V. B. a. P. B. a. H. B. a. M. B. a. J. B. a. I. B. a. J. B. a. G. B.-S. a. C. B. (2023). GPT-4 Technical Report. doi:https://doi.org/10.48550/arXiv.2303.08774

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., . . . Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *372*, n71. doi:10.1136/bmj.n71 %J BMJ

Park, Y., & Doo, M. Y. (2024). Role of AI in Blended Learning: A Systematic Literature Review. *The International Review of Research in Open and Distributed Learning, 25*, 164-196. doi:10.19173/irrodl.v25i1.7566

Pitso, T. (2023). Post-COVID-19 Higher Learning: Towards Telagogy, A Web-Based Learning Experience. *IAFOR Journal of Education, 11*, 39-59. doi:10.22492/ije.11.2.02

Puentedura, R. R. (2009). As we may teach: educational technology, from theory into practice. In: Apple.

Qureshi, B. (2023). *ChatGPT in Computer Science Curriculum Assessment: An analysis of Its Successes and Shortcomings*. Paper presented at the Proceedings of the 2023 9th International Conference on e-Society, e-Learning and e-Technologies, <conf-loc>, <city>Portsmouth</city>, <country>United Kingdom</country>, </conf-loc>. https://doi.org/10.1145/3613944.3613946

Rudolph, J., Tan, Samson., Tan, Shannon (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of applied learning and teaching, 6*(1), 342-363.

Silitonga, L. M., Hawanti, S., Aziez, F., Furqon, M., Zain, D. S. M., Anjarani, S., & Wu, T.-T. (2023, 2023//). *The Impact of AI Chatbot-Based Learning on Students' Motivation in English Writing Classroom.* Paper presented at the Innovative Technologies and Learning, Cham.

Simms, R. C. (2024). Work With ChatGPT, Not Against: 3 Teaching Strategies That Harness the Power of Artificial Intelligence. *Nurse Educator, 49*(3), 158-161. doi:10.1097/NNE.0000000000001634

Sohail, S. S., Farhat, F., Himeur, Y., Nadeem, M., Madsen, D. Ø., Singh, Y., . . . Mansoor, W. (2023). Decoding ChatGPT: A taxonomy of existing research, current challenges, and possible future directions. *Journal of King Saud University - Computer and Information Sciences, 35*(8), 101675. doi:https://doi.org/10.1016/j.jksuci.2023.101675

Speth, S., Meißner, N., & Becker, S. (2023, 7-9 Aug. 2023). *Investigating the Use of AI-Generated Exercises for Beginner and Intermediate Programming Courses: A ChatGPT Case Study.* Paper presented at the 2023 IEEE 35th International Conference on Software Engineering Education and Training (CSEE&T).

Suryanto, T. L. M., Wibawa, A. P., Hariyono, H., & Nafalski, A. (2023). Evolving Conversations: A Review of Chatbots and Implications in Natural Language Processing for Cultural Heritage Ecosystems. *2023, 3*(4), 52 %J International Journal of Robotics and Control Systems. doi:10.31763/ijrcs.v3i4.1195

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. doi:10.1186/s40561-023-00237-x

Uddin, S. M. J., Albert, A., Ovid, A., & Alsharef, A. (2023). Leveraging ChatGPT to Aid Construction Hazard Recognition and Support Safety Education and Training. *15*(9), 7121.

Wang, & Feng, Y. (2024). *An Experimental Study of ChatGPT-Assisted Improvement of Chinese College Students' English Reading Skills: A Case Study of Dear Life*. Paper presented at the Proceedings of the 15th International Conference on Education Technology and Computers, <conf-loc>, <city>Barcelona</city>, <country>Spain</country>, </conf-loc>. https://doi.org/10.1145/3629296.3629300

Wang, Wang, M., Xu, X., Yang, L., Cai, D., & Yin, M. (2023). Unleashing ChatGPT's Power: A Case Study on Optimizing Information Retrieval in Flipped Classrooms via Prompt Engineering. *IEEE Transactions on Learning Technologies, PP*, 1-13. doi:10.1109/TLT.2023.3324714

Wang, Wang, M., Xu, X., Yang, L., Cai, D., & Yin, M. (2024). Unleashing ChatGPT's Power: A Case Study on Optimizing Information Retrieval in Flipped Classrooms via Prompt Engineering. *IEEE Transactions on Learning Technologies, 17*, 629-641. doi:10.1109/TLT.2023.3324714

Widiati, U., Rusdin, D., Indrawati, I., Marzuki, & Govender, N. J. C. E. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *10*.

Zhai, X. (2023). ChatGPT for Next Generation Science Learning. *29*(3 %J XRDS), 42–46. doi:10.1145/3589649

Zhao, X., Aydeniz, M., & Yuan, F. (2023). *Exploring Opportunities and Challenges of AI-incorporated Biomedical Informatics Education: A Qualitative Study *.*

**Highlights**

- Synthesizes empirical studies of Generative AI implementation in Higher Education.
- Highlights the existence of research shortages in GenAI applications and innovative uses of AI tools in education.
- Covers a broad spectrum of disciplines, participant demographics, geographical locations, and methodologies in selected case studies.
- Classifies publications based on Laurillard's Conversational Framework (LCF) and the Substitution, Augmentation, Modification, and Redefinition (SAMR) framework.
- Offers a GenAI-TPACK diagram as a practical tool for educators to effectively incorporate GenAI into their practices.

**Declaration of Competing Interests**

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: