How Large Language Models Can Support Learning

Baminihannadige Sunekshi Lasuni Bhareti [redacted]

University Canada West

MBAR 661: Academic Research Project

Research Supervisor: Mohsen Ghodrat

June 25, 2025

Table of Contents

Introduction

Background of the study

The educational systems have been transformed with the integration of AI and its subsets. Over the past decade, AI powered LLMs have been developed and its application has been expanded across various aspects of education. OpenAI models such as GPT series, Google Gemini and many more tools have taken over both the teaching and learning processes proving to be very effective in offering numerous benefits that bridge the gaps of the traditional education system.

LLMs are a significant improvement of advanced AI systems which are trained on massive data sets and consist of powerful Natural Language processes and generative skills (Imaran & Almusharraf, 2024; see also Shahzad et al., 2025). With the introduction of transformer architectures in 2017, LLMs such as BERT, & GPT series have proven greater standards of NLP offering text analysis, interpretation, and content generation (Yan et al., 2024; see also Shahzad et al., 2025; Imran & Almusharraf, 2024). With such improvements to the LLMs it has paved the way for their vast application in a range of learning and education activities going beyond content creation to grading, providing feedback and creating assessments (Yan et al., 2024; see also Imran & Almusharraf, 2024). The improved integration of LLMs in education has been able to address many limitations and challenges posed by traditional learning systems such as heavy workloads of educators, inefficiencies in assessments, lack of customized learning plans, and inequitable educational access (Bucea-Manea-Tonis et al., 2022).

The existing literature on the impact of LLMs on learning has highlighted the key advantages they offer to both educators and learners. Accordingly, LLMs have helped the quality of education to improve through real-time feedback, convenient and automated grading, Academic profiling, recommendation of resources, generation of curriculum activities and assistance provided in collaborative activities (Yan et al., 2024; see also Shahzad et al., 2025; Ruiz-Rojas et al., 2024). For example, GPT-3 and its succeeding LLMs are currently generating assessment questions, grade them, and provide real-time feedback. Further, models like Google Gemini go beyond just text generation but also create images, handle audios and videos and therefore offer more enhanced learning experiences, especially suiting the different needs, capabilities and backgrounds of students (Imran & Almusharraf, 2024).

One cannot number the advantages offered by LLMs in learning. However, one of the most discussed and highlighted benefits is personalized learning offered through the integration of LLMs. LLMs are able to curate the learning experience to match the pace, complexity, and capacity of students and provide scaffolding where necessary (Shahzad et al., 2025; see also Praveena & Anupama, 2025; Ruiz-Rojas et al., 2024). Further, LLMs are greatly appreciated by the existing literature for promoting collaborative learning and improving critical thinking skills. LLMs such as ChatGPT can be leveraged as virtual moderators fostering different perspectives and enhancing debate in different group work settings and therefore creating space for peer feedback (Ruiz-Rojas et al., 2024).

Despite the valuable advantages brought in by the LLMs, their application can give rise to many practical limitations and challenges. Most of the LLMs are still in early developmental stages and digital illiteracy of both educators and learners can prevent them from achieving the expected outcomes. Most of these models have not been tested in real-world classroom settings and therefore, lack of transparency in those models can raise doubts about replicability and accessibility (Yan et al., 2024). Additionally, overreliance and dependence on LLMs can hinder the critical thinking skills, creativity and individualism in the work produced by students (Baltaci et al., 2024). Furthermore, deployment of LLMs in the educational setting has given rise to significant ethical concerns as well. The existing literature has highlighted the biases resulting from the data on which LLMs are trained and therefore, reinforcing stereotypes and social inequalities, concerns relating to data protection, authorship of work generated, and intellectual property rights issues. Accordingly, they have emphasized the importance of human-integration and judgement to balance the drawback of the LLMs.

Looking at the current state of the existing research, we can see that many authors have focused on discussing the benefits offered by the LLMs in learning and how they support different aspects of education. However, most of the literature consists of small pilot studies or laboratory testing rather than actual deployment of LLMs in broader classroom settings. It can be seen that most studies have highlighted some general recommendations to improve the future of LLMs in learning but have not dived into the specifications or strategic approaches to address the limitations and challenges posed by LLMs. The availability of limited knowledge and resources can be one of the main reasons for this gap in the research and focus should be shifted on to discussing a human-centred approach for integrating LLMs in education. Such approach will include stakeholder involvement from development to implementation stages in order ensure the right standards, transparency, and policies to promote ethical use and equitable access to learning powered by LLMs (Yan et al., 2024; see also Imran & Almusharraf, 2024; Bucea-Manea-Tonis et al., 2022). The ultimate goal should be to reflect on LLMs as augmenting the experience of educators and learners but not replacing them, supplementing human thinking, judgement, creativity but not supplanting those skills.

Objective of the study

This research is focused on addressing the following objectives:

- To analyze the impact of LLMs on supporting learning processes, particularly through interactive feedback and tutoring.
- To explore methods for integrating LLM-based tools in educational settings to improve learning outcomes.
- To identify potential improvements in LLM feedback systems to foster deeper learning and critical thinking.

Further, the below questions are intended to be explored by this research:

- How can LLMs be effectively integrated into learning environments to enhance feedback and student engagement?
- What are the specific features and techniques of LLMs that could support the goals of education?
- How does the use of LLMs in educational contexts impact students' critical thinking and problem-solving skills?

Structure of the research report

The first section of the discussion section draws the attention of the reader to a brief introduction of AI and its broader application in learning and education. While this part mainly describes how AI in general have impacted the education sector, it will also briefly outline the relationship between AI and its subset of LLMs. The next section will narrow the focus of the research exclusively to LLMs, starting off with an overview of LLMs, their origin and development, main sources of knowledge and functionalities. Next, the focus will be shifted to the application of LLMs in education and discuss the existing literature on the role of LLMs in different types of learning methods. The subsequent section will outline the benefits and impact of using LLMs in learning, following with a section emphasizing the limitations and ethical concerns identified by the existing literature in relation to the use of

LLMs in educational settings. The final section of the report will be devoted to discussing the recommendations that must be implemented to mitigate the effects of challenges and limitations posed by LLMs and the future of LLMs in learning.

Literature Review

Use of LLMs in learning

The integration of AI and most specifically, LLMs has transformed the traditional approaches to education and fostering wider space for innovative solutions for prolonging challenges posed by traditional learning systems. The existing literature has highlighted the development of LLMs from their original state of probabilistic models to more refined transformer-based architectures such as GPT-4 and Google Gemini. These sophisticated models are currently being used in varying learning settings for generation of content, interactive tutoring, generation of real-time feedback, student engagement and many more advantages. Studies undertaken by Shahzad et al. (2025) and Yan et al. (2024) have emphasized how LLMs have supported learning and enhanced the experience of both educators and learners through personalized feedback, personalized assessment processes, and adaptive delivery of content. Bewersdorff et al. (2025) building on the same aspects have highlighted how multimodal LLMs such as Gemini go beyond the basic functionalities and process images, audios, and videos to offer a richer and more engaging learning experience by drawing examples from science education.

Benefits of LLMs in learning

Speaking of the benefits offered by LLMs in the learning environment, Research by Ruiz-Rojas et al. (2024) and Richer et al. (2025) have highlighted the ability of LLMs to deliver complex concepts in simpler manners to suit the learning capacities of different learners, offer real-time assistance and therefore providing scaffolding where necessary. The LLMs are therefore able to foster inclusive education to support multilingual learners and widen accessibility to learners with disabilities or coming from distinct backgrounds and learning proficiencies (Wu et al. 2025). Sarangi et al. (2024) has further focused on the flexibility offered by LLMs through 24/7 availability. Researches such as Belkina et al. (2025) and Lang et al. (2024) have highlighted the importance of LLMs in enhancing efficiency by automating most of the routine and administrative tasks that take up a considerable amount of time of educators that can be alternatively used for curriculum innovation and mentoring of students. Shen et al. (2025)'s work on LLMs have focused on how LLMs foster collaborative learning environments by playing the role of a 'third collaborator/moderator' in group projects. Studying the interactions of nursing students with LLMs, Shen et al. (2025) has observed how they keep on comparing the insights generated by LLMs against the perspectives of their peers to improve the nursing care plans.

Identification of drawbacks

The existing literature has however identified the drawbacks in the LLMs and the functionalities they perform. Yan et al. (2024), carrying out excellent research on this matter has confirmed that LLMs tend to create false information when they are not able to generate relevant and accurate information and that such hallucinations pose a risk at academic integrity. This is just one limitation of LLMs that standout along with the rest of the issues discussed by existing literature. Researchers have also shed light on the ethical concerns surrounding the usage of LLMs in learning and Nguyen Van Viet et al. (2024) and Claman & Sezgin. (2024) have highlighted the risks of giving rise to systemic biases and social inequalities through the content generated by LLMs due to the broader biases in their training data. Erkılıç & Ibrahim Cifci (2024) and Guizani et al. (2025) have further discussed authorship and attribution of intellectual property rights of work generated by LLMs.

Recommendations and future of LLMs

The existing literature has also identified the suitable solution for the discussed issues and Lee & Palmer (2025) have highlighted the importance of building institutional policies related to the use and standards of LLMs. While Wang et al. (2021) has emphasized the importance of enhancing prompt engineering skills of the LLM users, Yan et al. (2024) has looked for solutions that focus on creating empirical validity of the efficiency of LLMs through application in actual classroom settings.

Finally, the existing researchers have shed light on the potential role of LLMs in the future paving the way for more efficient advancements in education. While Al-Dokhny et al. (2024) has thoroughly discussed how multimodal LLMs can improve student engagement and the quality of learning experience through more interactive learning, Cain (2025) has envisioned the future role of LLMs as a cognitive partner rather than mere source of generating content.

In summary, the existing literature has underlined the transformative nature of LLMs through their numerous benefits offering to the learning environments. However, the researchers have also carefully navigated through the limitations of LLMs while identifying the suitable solutions. The common understanding of these studies is that LLMs must be used in learning as a tool to compliment the skills of both educators and learners but not to replace them.

Methodology

The methodology that was used to develop this report is a systemic literature review to analyze how LLMs are supporting the learning and drink experience related to both students and educators. The approach we have adopted has paved the way for identifying, selecting, evaluating and integrating the relevant academic literature covering the topic to build a robust and evidence-based understanding of different aspects of the topic.

In order to ensure that the search for the relevant academic journals was broad but focused, major academic databases were used. Appendix 1 presents the keywords, synonyms and the relevant search strings that were used in combination to search for the appropriate and relevant academic studies. Further, the search was limited to peer reviewed articles published in the past five years to ensure that the most recent discussions and developments were taken into consideration as the field of AI and LLMs is constantly evolving.

Inclusion and Exclusion criteria

The following inclusion criteria was used to select the academic articles:

- **Type of publication** – For enhanced validity, only peer reviewed academic articles were selected for the research.
- **Publication date** – in order to promote contemporary research and to capture the most recent developments, the academic articles published in last five years were selected.
- **Relevance** – Academic articles covering the application of AI and LLMs in different methods of learning, benefits of use of LLMs in learning, limitations rising from the LLMs in the context of education, recommendations to mitigate the effects of limitations and future of LLMs in the context of learning.
- **Language –** Academic articles published in English was used for the study.

Below lies the exclusion criteria that was used for the search of the papers:

- Focus was restricted to peer reviewed academic articles and therefore, book chapters, conference papers, reports, dissertations and theses and other opinion pieces were excluded from the search.
- Articles covering general aspects of LLMs or not solely focused on the application of LLMs in learning were excluded.
- Articles that were not peer reviewed or not published within the last five years were excluded from the search.

Explanation of the Databases

A comprehensive approach was adopted to access a wide range of scholarly articles through different academic databases. The Publication Finder for University Canada West was used as the main path of access to a number of subscribed resources to make sure that access was allowed to peer reviewed, valid, and relevant journal articles. Directory of Open Access Journals was a main database that was used to have access to peer-reviewed open access journals. Another data base that was used is the Google Scholar to search for many introductory and preliminary studies that covers many academic formats. Academic Search Ultimate was another multidisciplinary database that was heavily used in the search process as it allows access to almost every field and their studies. This database was proven to cover a wide range of journals and extremely helpful to ensure that the search was done extensively. For more general areas such as the introduction, origin and development of AI and LLMs, JSTOR proved to be helpful in providing access to primary resource articles. Sage Open was a similar platform but introduced more extensive studies done in different fields and peer reviewed.

Further, few specialized databases were also used to narrow the search and improve relevancy. ScienceDirect was one such specialized database that was used to gain access, particularly through Elsevier, to studies done in the science and medical field and the application of LLMs in those fields. Another key database that was used is PubMed to gain access to articles related to biomedical education, life sciences and other medical studies. The agreement to use a multi-database strategy ensured that a more robust, recent, and comprehensive set of journal articles were accessed to successfully complete the research and satisfy the research objectives.

Selection and Screening Process

The articles were carefully scrutinized to extract the relevant data and these data were categorized based on the relevant subtopics such as key themes and methodologies, application of LLMs in different types of learning, reported benefits, discussed limitations, suggested solutions. The extracted data from all the articles were then combined to build a comprehensive review of the research topic and relevant sub-topics. The apparent strengths, and weaknesses in the existing literature and the methodologies these studies have deployed were analyzed top identify the gaps in the current research to recommend practical solutions.

Screening steps and selection process

The screening steps utilized in the process of selecting the most relevant journal articles from the many databases are listed below:

1. **Initial searches through keywords**

   The selection process was started by conducting initial searches across the many databases identified in the preceding section. The search was based on the keywords, their synonyms and search strings created from these keywords as displayed in appendix 1. The keywords and search strings were formulated based on the key themes, concepts, research questions and objectives used for the research study.

2. **Screening of titles and abstracts**

   Upon gaining access to a vast number of journal articles, the titles and abstracts of the articles were screened with the help of Notion to see the potential relevance. Articles were taken into the next stage of the study if they discussed about the application of AI in educational contexts, particularly LLMs, their benefits, limitations, and future trends. Any articles that discussed about AI or LLMs in general without specific focus on educational application were excluded at this stage.

3. **Full-Text review**

   The next step was reviewing the full text of the selected articles to make sure that they were within the inclusion criteria. The Notion tool was of assistance at this stage to thoroughly study the contents of the selected articles and exclude any article that did not fit the inclusion criteria.

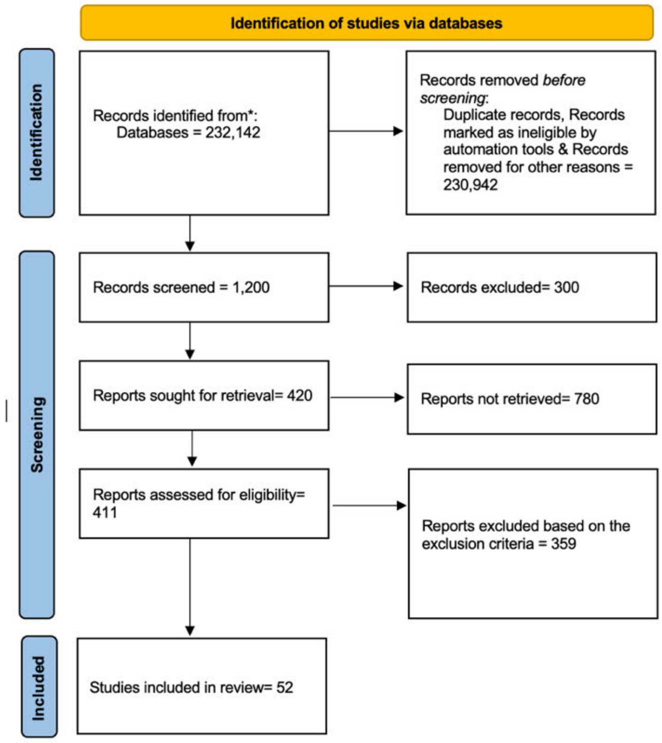
4. **Extraction of Data and Synthesis**

   Upon reviewing the full text of the articles, around 52 articles were short listed, a summary of which is included in Appendix 2. The data supporting the research study was systematically extracted from these articles and this included information covering the main theses of the study such as recurring themes, emerging trends, challenges, proposed solutions in the application of AI and LLMs in different educational contexts.

This whole process has ensured that the methodology of this study is transparent rigorous and well-aligned with the objectives of the research. Table 1 below includes the PRISMA flow diagram for ease of understanding of the numbers that were screened in the selection process.

Figure 1

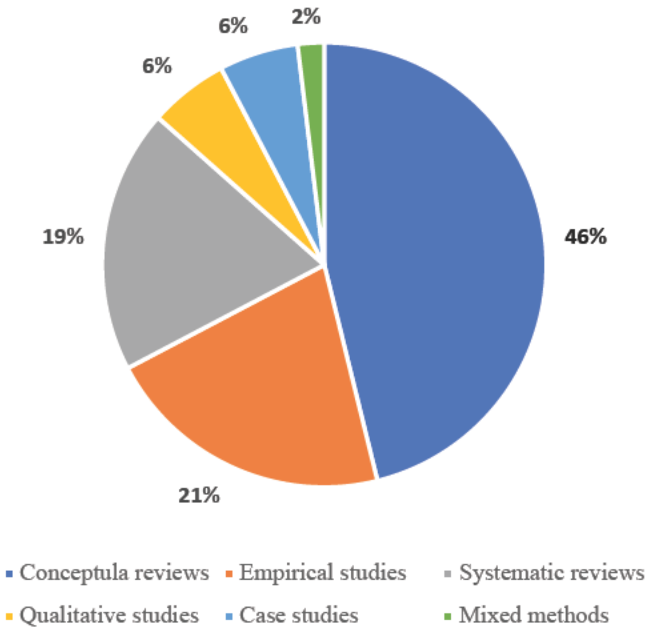
PRISMA Flow diagram of the selection process.



The characteristics of the articles included in the study can be categorized according to the research methodology that has been used in each of these articles. This will further provide a clear view as to the type of the studies contributing to this research paper. Figure 2 illustrates this categorization, and it can be seen that close to half of the studies are conceptual reviews and only 21.15% of the studies carry empirical evidence.

Figure 2

Distribution of research methodology of the reviewed articles



Results

Following the screening process, below are the key results or findings emanating from the selected articles.

Key characteristics of the reviewed literature

The core focus of the selected articles was on the application of LLMs in different educational settings and the resulting implications. The majority of the literature comprised of qualitative studies, systematic reviews, and theoretical discussions. Further, they contained more pilot or initial studies rather than any large-scale empirical experiments in actual classroom settings.

Key Applications of LLMs in different Learning Environments

The scrutinized literature identified different types of learning environments and how LLMs are applied in each of these settings. In-person or traditional classroom settings was one and several studies identified how LLMs are sued in such settings to create visualizations, assist multimodal science learnings, help educators to curate assessments in a way to enhance student engagement, critical thinking, and real-time learning.

Most of the literature has identified the application of LLMs in online learning experiences through many activities such as gamified quizzes, and LLM powered interactive activities. Further, these educators using these online environments are currently using LLM powered tools to obtain detailed feedback on student performances, to grade assignments and curate content according to the varying abilities of students.

Finally, self-directed learning is also being identified as a growing method of learning, especially in higher educational settings, and the existing literature has highlighted the important role played by LLMs in promoting autonomy, and independence through different tools used to generate ideas, find resources, structuring and evaluating assignments and obtain feedback.

Benefits of using LLMs in learning

Some key benefits have been identified by most of the reviewed literature and these individual benefits fall under few common categories. Nearly all the articles have discussed how LLMs enhance personalized learnings that best suit the different needs of students. These studies have also highlighted how the educators are also benefited by such personalization to identify the many different performance levels of students and cater accordingly.

A considerable number of articles have also highlighted how LLMs have eliminated the barriers of traditional learning systems through facilitating wide access for learners with disabilities and removing language barriers through multimodal and translating tools. The 24/7 availability has also been highlighted frequently as supportive in increasing flexibility.

Studies have also repeatedly highlighted how LLMs have increased efficiency through automation of routine tasks and promoted life-long learning through independent platforms without any locational constraints.

Almost all the studies have highlighted how student engagement is enhanced through the integration of LLMs in learning experiences, especially through personalized feedback which mitigates comparison within student groups and increases the identification of individual learning capacities.

Finally, a key aspect the has been discussed in several studies is how LLMs improve collaborative learning and plays the role of a moderator in learning settings, especially in group projects, by feeding into critical discussions among peers and refining their outputs.

Limitations and ethical concerns

The selected literature has identified recurring limitations, including ethical concerns surrounding the use of LLMs in education.

Complexity in the content generated has been highlighted by several studies as resulting in lack of clarity especially in learning settings of younger students. The issue of creating incorrect information or 'hallucinations' is also constantly discussed in majority of the literature.

Further, over-reliance on LLMs has also been discussed in several articles to show how it can reduce the critical thinking skills, and creativity among students and reduce authenticity and individualism in the work produced by them.

Another issue highlighted is the lack of digital literacy of both students and educators that leads to obstructions in obtaining highest levels of effective utilization of LLMs.

Turning to the ethical concerns, the existing literature has discussed the systematic biases resulting from the existing biases in the training data and how it leads to reinforcement of stereotypes. Further, studies have shown how potential leakage of private data of students can lead to significant privacy concerns. Finally, the issue of authorship of content generated by LLMs and the related issues surrounding intellectual property rights have also been identified by several studies.

Theoretical frameworks supporting the study

This report analyses the application of LLM in the context of learning and education. The objective is to identify the benefits of such application, what limitations and challenges they pose and what are the practical solutions to overcome these challenges. However, before the resources can be analyzed to discuss the findings, there are few important theoretical frameworks that we need to lay down. These theoretical frameworks assist the research in laying a foundation for the main concepts involved in the study and to ensure that the scope of the study is broadened yet focused and not strayed. They also provide validity to the concepts discussed in the report and a sense of origin or understructure to the relevant topics.

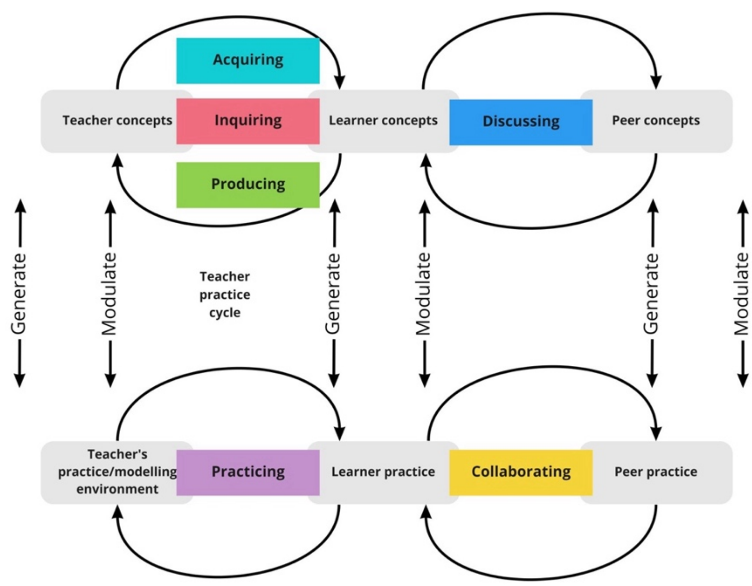
The focus of this study is narrowed to the application of LLMs in 'learning'. Therefore, in order to understand and restrict the scope of 'learning' we discuss here, we have used Diana Laurillard's conversational framework (LCF) to guide us. The LCF lays down a structure for including different learning experiences and emphasize the importance of collaborative learning and engagement. The focal idea of this concept is that the learner is central in the learning environment and the development occurs as a result of interactions between the educator and learner and amongst the learners (Yu et al., 2020). This framework identifies six different learning types that optimize these interactions (Yu et al., 2020):

1. **Acquiring** – Refers to the knowledge gaining activities such as reading from resources (books, websites) or listening to lectures or watching educational videos.

2. **Inquiring** – this step refers to the stage where learners formulate questions, explore, compare and contrast the concepts acquired to widen their understanding of the knowledge gained.
3. **Discussing** – this is the process where learners share their opinions, ideas, concerns, and engage in conversations with the educators and peers to express their thoughts.
4. **Practicing** – in this stage, learners utilize different tools such as feedback, self-reflection, external sources to evaluate their learning and refine their understanding or output.
5. **Collaborating** – This refers to act of collaborating with peers to create shared outcomes, such as reports, papers, designs, diagrams and illustrations. This involves the practice of negotiating and re-creating work until a solution that is agreed by everyone is achieved. This process involves challenging ach other's work to find the best outcome.
6. **Producing** – this is the process where learners utilize what they have learnt for the generation of a new output to that is evaluated by the educator.

Figure 3

Diana Laurillard's conversational framework (LCF)

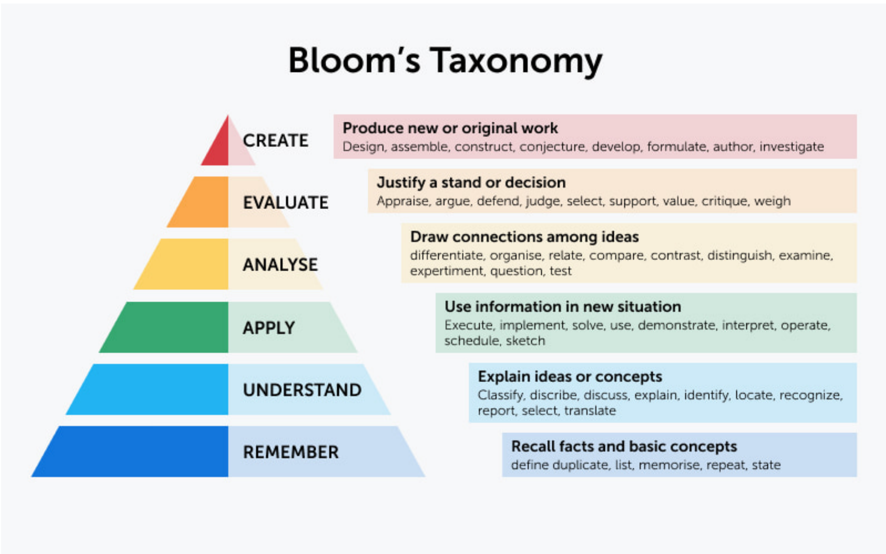


Further, in order to restrict our understanding on the objectives of learning, we have Bloom’s taxonomy that describes six levels of learning objectives. Accordingly, the first level is ‘Remember’ which refers to recalling facts, information an d definitions learnt. Second level is ‘Understand’ which refers to the skill of comprehending the meaning of concepts and information acquired (Eber & Parker, 2007). The learner should be able to paraphrase the information learnt to express the skill of understanding. The next objective is ‘Apply’ (Adams, 2015). This level requires the utilizing of knowledge for new situations to find the best solutions in a given scenario. Fourth level is ‘Analyze’ and this refers to dissecting concepts into sub components and understanding how each component relates to the other and compare and contrast between the concepts (Adams, 2015). The next is ‘evaluate’ that refers to making judgements related to the ideas learnt and evaluating facts, ideas and information learnt against given standards (Adams, 2015). The final level is ‘Create’ and this means the generation of new knowledge or outcomes with the knowledge acquired and putting different elements together to build a new creation (Eber & Parker, 2007).

While the application of LLMs in learning will be studies in the context of different types of learning, their benefits and limitations will be discussed in the context of their impact on the achievement of objectives of learning.

Figure 4

Bloom’s taxonomy



Discussion of the findings

Introduction of AI and its broader application in Learning and Education

Modern technology and its advancements have proven the incorporation of technological tools in many industries. One such important field where technology has offered numerous benefits and enhanced quality and experience is education and learning. Artificial Intelligence (AI) has taken the forefront of the progression of technology and it has proven a remarkable transformative effect in learning by opening unprecedented opportunities to bridge the gaps in traditional education systems (Reicher et al., 2025). AI is considered a critical component of the digital transformation process of education that facilitates agile solutions for educational market needs (Bucea-Manea-Tonis et al., 2022). AI has provided reliable and effective solutions to challenges posed by traditional learning systems such as gaps in pedagogical alignment, lack of self-paced learning and inadequacy of inclusivity of diverse learners (Reicher et al., 2025). It is said that broader Sustainable Development goals are intended to be achieved through the integration of AI in education and therefore promoting quality of education and life-long learning (Ruiz-Rojas et al., 2024). One may believe that AI supports learning only through automation. However, it has redefined learning systems through different approaches by offering personalized instructions for educators that suit distinctive needs of students (Ruiz-Rojas et al., 2024). Creating more adaptive systems with innovative learning and assessment tools, and empowering both educators and learners have been the central focus of AI advancements in education (Ajani et al., 2024). AI's position or purpose in education is best reflected by stating that it does not replace the educators but augment their performance (Reicher et al., 2025). AI is believed to have shaped the critical thinking of learners, promoted teamwork and collaboration through platforms facilitating virtual meeting spaces, fostered inclusivity and enhanced problem solving (Ruiz-Rojas et al., 2024).

AI utilizes its key domains such as natural language processing (NLP), machine learning, immersive technology, IoT, and conversational agents in education and for better creativity, socio-emotional intelligence and cultural competence in learning, these tools must be functioned with human-led instructions (Praveena & Anupama, 2025). Both the learning experience of students and teaching techniques of educators are advanced through different forms and uses of AI. Most educators that are technophobic have embraced AI tools for ease of preparing lesson plans, curating assessments, simplifying the grading process and streamlined administrative tasks (Ajani et al., 2024). Further AI tools have helped the educators to analyze student data and identify the needs for intervention and improvement and thereby making room for personalized and real-time evaluation (Ajani et al., 2024). Educators are also using AI for creating educational escape rooms, mind maps, summarization of length texts, all of which assist student engagement and enhance the quality of educational content transferred to students (Ruiz-Rojas et al., 2024). On the other hand, students are also massively benefitted from AI platforms that offer diverse perspectives and enhance critical thinking through real-time feedback, offer access to guidance that is available all the time and provide content and pace for different learners with individual capacity (Ruiz-Rojas et al., 2024). Students can use AI platforms to assess, analyze and summarize course materials and therefore, create efficient and personalized content for studying and assignments (Reicher et al., 2025). AI has been able to reshape the outcome of education as it enables the identification of varying capacities of students and curate the delivery of knowledge accordingly.

This study will focus on a specific domain of AI, namely Large Language Models (LLMs) and how they support learning. The introduction about the broader application of AI in learning has laid the foundation to identify the more specific application of LLMs in learning for the benefit of an enhanced educational experience.

Relationship between AI and LLMs

An exclusive study on LLMs and their relevance and usage in education must be preceded with an understanding of the relationship between AI and LLMs. LLMs are considered the most advanced natural language processing and generative engines (Xing et al., 2024; see also Davar et al., 2025; Yan et al., 2023). LLMs are specifically distinct from AI from their human-like capabilities that focus on adaptation, creativity, and contextual relevance. Unlike general AI tools, LLMs are used as thought partners offering reasoning, explanations, and enhancing collaboration (Ruiz-Rojas et al., 2024; see also Shen et al., 2025). The journey of AI was founded on rule-based expert systems and search engines and it has currently evolved to LLM enabled Generative AI. While fostering informational access and creativity, LLMs go beyond the tasks of AI and generate new content and knowledge, translate and adapt languages (Nguyen Van Viet et al., see also 2024; Belkina et al., 2025).

Overview of LLMs

Origin and Development of LLMs

The origin of language models was created somewhere between 1950-1960s and the initial models lacked the computational techniques required for effective natural language understanding and processing (Shahzad et al., 2025). The next important development took place between 1980-1990s, when the existing language models started utilizing probabilistic modelling for the prediction of word sequences (Shahzad et al., 2025). Despite this next step in development of LLMs, the models still struggled with correctly capturing context and semantics for effective generation of output. The development of LLMs was significantly shaped and redefined in the mid 2010s with the progressive use of deep learning. The introduction of RNNs, which are one step ahead than the traditional deep learning networks, improved effective sequential predictions and generated conclusions based on prior inputs

(Shahzad et al., 2025). The most distinguished RNN, Transformer models which are associated with LLMs was released in 2017 marking a significant milestone in the development of LLMs. Unlike the traditional RNNs, Transformer models are able to examine long text sequences and perform multiple computational steps simultaneously rather than in a serialized manner (Daver et al., 2025). With the incorporation of transformer models, the Open Ai introduced the first of their series of Generative Pre-Trained models (GPT) in 2018. From GPT-1 to GPT-4, Open AI has introduced further advancements in each model focusing on expanding the number of parameters, improving text generation, including training from human feedback and refining the outputs (Shahzad et al., 2025).

LLMs and their sources of knowledge

LLMs are exposed to massive datasets consisted of books, peer-reviewed journal articles, diverse professional guidelines and handbooks, websites, case databases, and many more sources available online (Newton & Xiromeritis, 2024; see also Lucas et al., 2024; Wu, Dang & Li, 2025; Yan et al., 2023). Once they are exposed, their training is based on deep learning on these corpora, allowing the LLMs to absorb every detail on the datasets including factual knowledge, reasoning patterns, and grammatical conventions across different disciplines. For example, in the health care sector, LLMs are trained on current clinical literature, existing patient records, and guidelines for plausible and sophisticated responses (Lucas et al., 2024). The training of LLMs is not limited to be done on English or Western-centric datasets but a diverse selection of datasets is used including regions or filed specific Open Educational Resources (Alfirević et al., 2024; see also Nguyen Van Viet et al., 2024). One may feel that the foundational knowledge of LLMs is limited and determined by the data in the training datasets at the time of training development. However, it must be noted that the effective responsiveness of the LLMs to carefully structured user prompts make them flexible and adaptable to different contexts and evolving needs of the users (Newton & Xiromeritis, 2024; see also Davar et al., 2025).

LLMs in Education

One simple definition cannot be devoted to the meaning of education as it can include so many aspects related to the facilitation of learning and acquiring knowledge, skills, values and habits. A person can undertake education at different stages of life and education can also take different forms suiting the needs of people. For the ease of relevance of this paper, we are focusing on formal education that is delivered through formal institutes such as primary to high schools, universities, and other educational institutes.

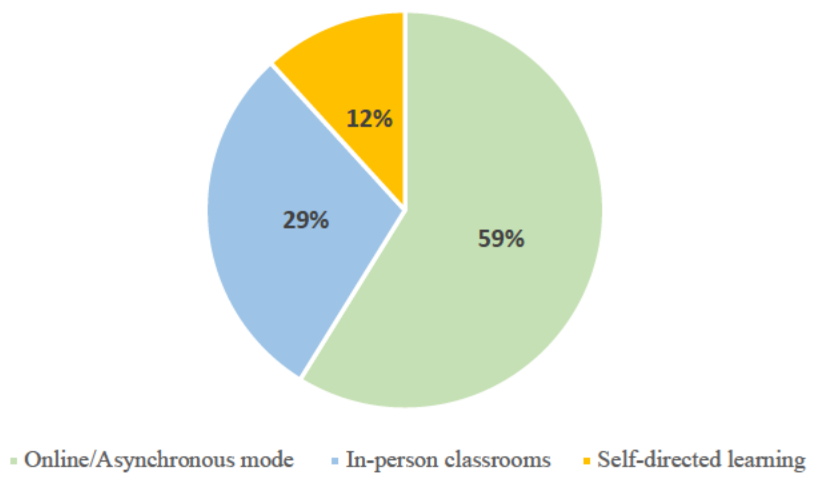
Further, before we can dive into the application of LLMs in education and analyze how they support learning, it is important to understand the objectives of learning to evaluate whether LLMs have proven to be satisfying those objectives. As mentioned before, the analysis of this paper will be based on Bloom's taxonomy and the learning objectives identified through this framework. A core goal of education is considered to be understanding higher-order thinking and to go beyond basic memorization and achieve high levels of understanding and application of knowledge. Accordingly, Bloom's taxonomy has categorized the learning objectives into 6 hierarchical levels; remembering, understanding, applying, analyzing, evaluating and creating. While remembering simply refers to recalling facts that are learnt, understanding means the ability to explain the ideas and concepts that we have learnt. Next, Applying refers to the use of knowledge or applying the knowledge learned to different contexts and for problem solving. Analyzing is the objective of differentiating the facts and concepts learnt and the ability to organize information. Further, evaluating refers to the act of judging the value and effectiveness of what is learnt. Finally, creating means producing new ideas or new work depending on what has been learnt. This paper will focus on how LLMs support learning in formal institutes to achieve the different objectives of education as stipulated by Bloom's taxonomy. However, the research is not restricted to formal or in-person delivery of education but is extended to evaluate the impact of LLMs in modern and more informal methods of education such as online delivery of education, asynchronous educational activities and self-directed learning.

Role of LLMs in different modes of learning

When we discuss the role of LLMs in education one may be tempted to think that it's all about how students get assistance from LLM based AI chatbots such as ChatGpt, Gemini, or Deepseek for their academic writing. However, the use of LLMs in education have gone far beyond this limited and well-known application of AI chatbots. Figure 5 shows the distribution of the reviewed articles based on the educational environment, context or type they have discussed about.

Figure 5

Distribution of reviewed articles based on the learning mode discussed.



Application of LLMs in in-person classrooms

Traditional learning style of in-person or classroom education has been reshaped with the use of LLMs. A study done on Japanese junior high schools has explained how these schools are using LLM powered tools for visualizing and supporting classroom activities. The traditional active reading classroom activities are now being alternated by individual and group activities powered by LMS, e-books and analytics dashboards and daily attendance logs are also captured through the same tools (Horikoshi et al., 2025). Further, learning stream plots have been generated to process multimodal data across various platforms to generate visualizations of timelines of classroom activities (Horikoshi et al., 2025). Such in person classroom activities augmented by LLM platforms have helped the educators to track the errors and discrepancies in their lesson plans and activities and identify the gaps in student performances and issues in student engagement (Horikoshi et al., 2025). It can be also seen that primary school educators, who are traditionally considered to be well based on physical textbooks and restricted to classroom activities, are now taking a shift towards deploying LLM powered chatbots such as ChatGPT for both teaching and assessing student performance (Uğraş et al., 2024). Another study done on the integration of Multimodal LLMs in science education has shown that the multimodal nature of science education which involves not just reading and writing but diagram creation and interpretation, data analysis, visualization of different solutions has been improved and reshaped with LLMs (Bewersdorff et al., 2025). Accordingly, multimodal LLMs are currently used for generating visual explanations that accompany modelling, experiments, and simulations in science lessons (Bewersdorff et al., 2025). These Multimodal LLMs have also enabled real-time feedback and evaluation of the content created through them (Bewersdorff et al., 2025). A similar study has observed the use of ChatGPT to model a drainage of water through a tank as a pre-lab activity which was later used by the students to critically identify the gaps in the model and improve on it through their own innovation (Keith et al., 2025). This study has tried to confirm the role of LLMs not as a replacement for the original work of the students but only as supplementary to guide them and enhance their critical thinking and innovation (Keith et al., 2025).

Application of LLMs in online education

Impact on leaners:

The impact of LLMs is more commonly seen in online education. A cross-cultural study done on 205 participants has observed the use of online quizzes enhanced with gamification and their relation with LLMs. The study has proven that LLMs have been used for the auto generation of the questions, answers, explanations in those gamified quizzes and most importantly, real-time feedback is made possible through LLMs as a part of these quizzes (Issabek et al., 2025). Further, this study has shown that LLMs have supported to curate these quizzes depending on the capacity and performance level of different students and also their demographic and cultural backgrounds. Another vital role played by LLMs to support learning is how they envisioned as interactive conversational agents replacing conversations between students and educators (Blaise Agüera y Arcas, 2022). These LLM powered interactive conversational agents are able to respond to very theoretical and philosophical queries, and provide logical reasoning (Blaise Agüera y Arcas, 2022). An interesting study done on preschool classrooms that are focused on social sciences has found out that these LLM powered interactive conversational agents are used in early childhood education as well to develop language skills, early literacy and numeracy of children through conversations, storytelling, and visuals (Doğan et al., 2025). A similar role is played by AI chatbots that heavily assist the modern time flipped classrooms by helping the students to have access to different resources before class and be ready for the lessons while practicing conversational engagement, language development and brainstorming (Davar et al., 2025).

Impact on educators:

On the other hand, educators are largely using LLMs via GPT models in completely asynchronous and online environments to measure the quality of knowledge and critical thinking skills of students and their cognitive presence phases (Castellanos-Rayes et al., 2025). Discussion posts are a common activity included in the higher education curriculum to improve critical thinking skills, writing skills and student engagement. LLMs which use structured prompts and adapted codebooks such as Practice Inquiry Model are being used to assess hundreds and thousands of discussion posts to classify them into cognitive presence phases (Castellanos-Rayes et al., 2025). Such very specific usage of LLMs focus on critical inquiry and metacognitive learning of higher education (Castellanos-Rayes et al., 2025). A study done exclusively in the context of language education and more specifically on English language learners, who takes English as their second language, has observed the usage of LLM models such as GPT-4, GPT-3.5, Google’s PaLM 2, and Anthropic’s Claude 2 for automated essay scoring (Pack et al., 2024). This study has explained that such automated essay scoring is consequential to explicit instructional prompts by educators that provide detailed scoring rubrics to the LLMs. Accordingly, the authors have confirmed that LLMs have provided not just a grade for the essays but also a detailed reasoning for each score identifying the areas of improvement, gaps in skills and also highlighting the strengths (Pack et al., 2024). This study has also focused on comparing the scoring LLMs to human graders to analyze the benefits of using such automated scoring systems for time savings and more reliable feedback (Pack et al., 2024). Further, the literature has proven that educators are greatly benefitted by the generation of lesson plans by LLMs, creation of diverse assessments and exams that prevent repetition in evaluation models, and facilitating a more convenient and comfortable space for tailored feedback that aims to address the unique drawbacks of each student (Xing et al., 2024; see also Nguyen Van Viet et ak., 2024; Lang et al., 2024).

Application of LLMs in self-directed learning

Self-directed learning is a modern method of learning that offers the learners complete control to manage their educational journey. Accordingly, the learners are expected to identify their goals and objectives in learning, identify and utilize the appropriate resources and strategies for effective learning. This is one of the main areas of education that is heavily regulated by LLMs as LLMs have proven to be very supportive in assisting the learners to organize their self-directed learning schedules. LLMs have been able to democratize access to best suited methods of learning, especially when engaged in self-directed courses. Campbell & Cox (2024) studying the application of generative AI in higher education has established that LLMs are used to carry out the most basic tasks in self-directed learning such as summarizing of lengthy materials, finding out appropriate resources for assignments, generating ideas, curation of formats and structures of assignments, and for receiving of feedback that would critique the both AI and self-generated work and receive potential grades before actual work is submitted. Özçelik & Yangin Ekşi (2024) have also shown the importance of using LLMs such as ChatGPT as learning assistants to cultivate writing skills. A group of researchers have studied a set of radiology postgraduate students in India to assess their awareness and engagement with LLMs in self-directed learning. This area of study is complex and LLMs have proven to be of great assistance in simplifying subject matters, structuring reports, curating differential diagnoses and providing access to a great variety of related textbook materials (Sarangi et al., 2024.). GPT platforms powered by LLMs have offered consistent supplementary support in online and asynchronous environments related to such exclusive and complex areas of study. A literature covering the impact of LLMs in medical education has mentioned how medical students easily have access to up-to-date medical and clinical content and scenarios through LLMs that a considerable amount of time is saved as they do not have to manually refer to hundreds of books to gather the relevant information (Lucas et al, 2024).

Relationship to the theoretical frameworks

In-person classrooms

The above-mentioned applications of LLMs in learning have proven how they directly support the different types of learning introduced by Diana Laurillards' LCF and learning objectives as shown in Bloom's taxonomy.

Generating visualizations in in-person classrooms and using LLMs to create models and diagrams in science education support two types of learning in the LCF. Accordingly, acquiring of knowledge and producing levels are achieved. Using the LLMs to create visualizations help the students to acquire knowledge more conveniently with ease by reducing complications. Further, using LLLMs to create science models reflects the effect of LLMs in producing new work by learners. Turning to the Bloom's taxonomy, these usages of LLMs in physical classrooms. Impact the remembering and understanding skills of learners by helping them to efficiently recall and comprehend large sets of data. Further, by using the guidance of LLMs to produce new work help them to apply the knowledge effectively and achieve initial stages of 'creating' by laying the foundation for learners through initial models and diagrams and allowing them to build on them.

Online education

The usage of LLMs in online education to generate quizzes, their questions and responses and curating the content of quizzes to the different needs of students shows how it supports acquiring and producing. Further, the interactive conversational agents enhance the skills of inquiring and discussing of learners as they build conversations with these robots to express and exchange ideas and seek explanations. Further by using these conversational agents in early childhood educations again supports the acquiring of knowledge of young students through storytelling and visualizations.

The usage of LLMs in online education is also impacting different levels of the Bloom's taxonomy. The conversational agents directly help the learners to comprehend complex information and remember them in simple terms and these conversational engagements help them to analyze educational material and apply them in their day to day studying routines and repetitive questioning and seeking explanations, learners are able to evaluate the responses provided by LLMs and apply their own judgement to come into conclusions.

Further, the educators' usage of LLMs in grading and scoring have also helped to understand and analyze the performance gaps of students. Further, it has helped the educators to evaluate their feedbacks against the feedbacks created by LLMs and create the best matching feedback for the performance of students.

Self-directed learning

In self-directed learning, LLMs are mostly assisting the students in guiding their course journeys and curating easy ways to get through the educational materials. Accordingly, by summarizing lengthy materials, structuring assignments and providing access to resources, LLMs are supporting the acquiring and producing of knowledge in self-directed learning environments. Further, they help the learners to understand course materials in the absence of physical educators, remembering facts, apply the knowledge acquired in assignments and creating their own work through researches, thesis, and responses to exams.

Benefits of using LLMs in learning and education

Following the discussion of the application of LLMs in different modes of learning, the next section is devoted to identifying the benefits highlighted by the current literature of using LLMs in learning and how they support the improvement and enhancement of effectiveness of learning. Even though the usage of LLMs can be diversified across different modes of education such as in-person or classroom learning, online learning, and self-directed learning, the benefits they offer are common to all LLM driven tools. Therefore, this paper will discuss a few common groups under which the benefits of LLMs can be captured.

Personalized Learning Experience

One of the most important features that has been highlighted by most of the literature is how LLMs have offered personalized learning experiences for students with varying skills. This has allowed learners to engage in learning at their own pace which is shaped by their individual needs and prior knowledge. AI assistants powered by LLMs are able to explain any complex text or content in a very simple manner that could be understood by different individuals with varying understanding capacities. Customized learning pathways, explanations, feedback, and AI driven tutors and conversational agents are addressing learning gaps of individuals by offering diversified support in real time (Xing et al., 2024; see also Nguyen Van Viet et al., 2024; Alfirević et al., 2024; Wu et al., 2025; Cecchini et al., 2025; Ruiz-Rojas et al., 2024).

Furthermore, the personalized experience is not limited to students but educators are also benefited by this advantage. LLMS are able to assess the quality of teaching performance, lesson quality, metacognition, and innovation of educators and give feedback so that they are able to identify the gaps in their performances and work on to improve (Fang et al.,2025; see also Belkina et al., 2025).

This benefit enables all levels of the LCF and learning objectives of Bloom's taxonomy. By customizing the educational content to the different needs of the students, the LLMs affect acquiring, inquiring, practicing and discussing stages of student learning and it further supports the achieving of all cognitive levels of the Bloom's taxonomy.

Accessibility, and inclusion

LLMs have been able to overcome many barriers that could not be surpassed by the traditional learning systems. They facilitate different content such as multimodal, multilingual, and translated outputs that can be easily embraced by the learners with disabilities or language barriers (Xing et al., 2024; see also Wu, Dang & Li, 2025; Cecchini et al., 2025; Ruiz-Rojas et al., 2024). LLMS are being used to create alternative formats and adaptive resources for any educational content that cannot be received by individuals with learning disabilities or representing different backgrounds and cultures. By this way, LLMs have been able to uphold equitable education with broadened participation and democratized access which is a vital Sustainable Development Goals (Nguyen Van Viet et al., 2024; see also Ruiz-Rojas et al., 2024b; Nedungadi et al., 2024).

Further, the support of LLMs is available and accessible 24/7 and therefore, providing much flexibility for the learners. In traditional educational systems, the educators are available only during the learning hours and any feedback required after learning hours must wait until the following working day. However, especially for students who are pursuing higher education along with part time jobs and other commitments, such limited access to educators is inconvenient. Therefore, readily available AI tutors or LLMs that provide instant feedback and guidance are greatly benefiting the current trending lifestyles of university students (Sarangi et al., 2025; see also Reicher et al., 2025; Xing et al., 2024; Davar et al., 2025).

By clearing the barriers in educational accessibility, LLMs are allowing a wide range of learners to be included in all levels of the LCF and paving the way for engagement in higher cognitive skills in the Bloom's taxonomy.

Efficiency and scalability

LLMs have assisted the educators to increase efficiency in their routine tasks with the automation of lesson planning, generating of assessments including quizzes, exams, and assignments, grading, and providing feedback. By this way, educators have also been able to handle larger classrooms and focus more on student mentorship, individual attention to students, and curriculum innovation (Wu et al., 2025; see also Belkina et al., 2025; Lang, Triantoro & Sharp, 2024; Davar et al., 2025). Further, through tools such as customizable Open Educational Resource LLMs, instructors can create and consistently update resources (Alfirević et al., 2024; see also Lang et al., 2024).

By improving efficiency and scalability of the educators and their routine tasks, LLMs are supporting the LCF stages related to the educators and thereby impacting acquiring and practicing knowledge and their skills through generating assessment contents, and lesson plans.

Promoting lifelong learning

LLM powered tools and platforms create an independent environment for the learners to pursue their education autonomously. They get the chance to practice exams, quizzes, get scripts for presentations and rehearse, get real-time feedback before any assessments (Wu, Dang & Li, 2025; see also Reicher et al., 2025; Nguyen Van Viet et al., 2024). Lucas et al., (2024) has brought in examples from medical education to show how students rehearse clinical reasoning scenarios and cases with AI conversational agents and using their feedback for improvement. Further, LLMs have helped students to let go of anxiety, frustration or humiliation that is caused by asking 'basic' questions in a classroom setting as LLMs provide immense independence to enquire about the most basic concepts and points until the students have fully understood (Lucas et al., 2024).

Increased motivation, engagement and enjoyment

Students have proven to be more motivated and engaged in using AI powered resources such as gamified quizzes and assessments compared to learning in traditional classrooms. The more tailored and customized format for learning has helped the students to be less stressed as the feedback given through LLMs are clear, curated to the individual skill capacity and therefore, limits comparison and competition with the rest of the learners. As a result, students have experienced increased enjoyment and motivation along with sustained efforts and hard work to achieve their educational goals (Ruiz-Rojas et al., 2024; see also Davar et al., 2025).

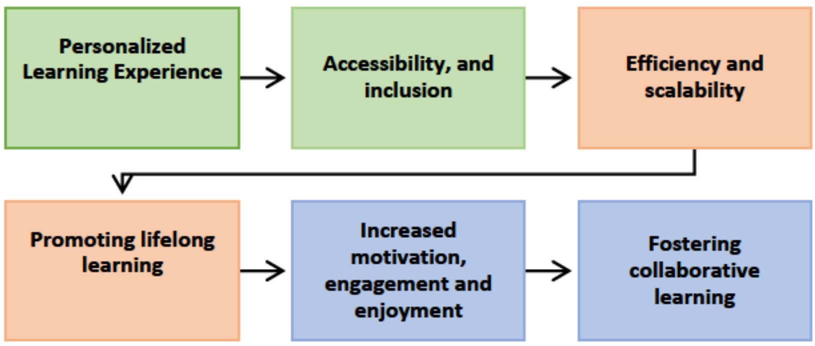
Fostering collaborative learning

LLMs have been able to foster collaborative learning in higher education settings. An exclusive study done on the usage of LLMs by nursing students has shown how LLMs are used by small groups of students for joint construction and refinement of nursing care plans, specifically through shared prompt engineering (Shen et al., 2025). These student groups have cooperated on making, using and updating their prompts to make sure the outputs from the AI fit the care plan better. The teams have collaborated over many steps and discussed the LLM results, comparing them against the textbook materials and relied on different perspectives of each team member to improve the care plans produced by the LLMs (Sgen et al., 2025). These students have stated that they have experienced the value of LLMs not just as a knowledge partner but also as a platform encouraging teamwork and collaborative learning. Newton and Xiromeritis (2024) have also noted that LLMs facilitate collaboration through generation of quizzes, and assessments which are used by students to discuss, and practice critical reasoning. Accordingly, LLMs are portrayed as a 'third collaborator' that gives extra perspective and fills in missing details for the ease of student learning.

This shows how LLMs directly impact the level of collaborating in LCF framework and enhance the space for discussion through facilitating critical engagement through delivering external perspectives in team projects and feeding into group judgements. Furthermore, LLMs impact analyzing and evaluating of knowledge in Bloom's taxonomy through allowing them to judge the content generated by LLMs against their own senses and justifications before coming into conclusions. This further results in creating new ideas amongst peer groups and encouraging collaborative learning.

Figure 6

Benefits of using LLMs in learning and education



Limitations of using LLMs in learning

The increasing development and adoption of LLMs in education has definitely brought in numerous opportunities and benefits that have blurred many of the constraints in the traditional education systems. However, excessive elaboration of the benefits of LLMs might undermine the necessity to critically examine the challenges or limitations in the use of

LLMs in learning environments. The existing literature has elucidated the importance of identifying these limitations in order to ensure that these areas are improved so that they do not compromise the efficiency and integrity of education but complement the same. This section will explore the existing literature's take on the limitations of using LLMs in learning environments and challenges they pose against effective learning outputs. Further, these studies have also identified the ethical considerations surrounding the heavy use of LLMs for educational purposes. This section will therefore shed light on this aspect as well.

Limitations on complexity and clarity of content

A study done on elementary science has explored the scientific validity of responses provided by ChatGPT to measure the quality and sustainability of education. Accordingly, it has found out that even though 94.2% of responses provided by ChatGPT for elementary science questions were valid, only 70.6% responses were clear (Choi, 2025). Further, only 12.8% of responses have been relevant to the curriculum (Choi, 2025). The main issue presented here was that responses were not effectively received by the young learners and the advanced terminology that was used was not proportionate to the age and skill levels of the learners. This issue can be understood as presenting accessibility barriers to the learners and therefore, hindering the effectiveness of learning. LLMs are criticized for producing inaccurate responses or presenting the risk of misconceptions. The same study on elementary science education has found out that the responses generated by ChatGPT involved subtle misconceptions and therefore, demanding the mediation of an educator to prevent the risk of misunderstanding by students (Choi, 2025).

Lack of empirical evidence

Another criticism raised is that responses and content generated by LLMs lack empirical evidence for factual learning improvements and they are greatly considered to be conceptual (Nguyen Van Viet et al., 2024). Further, LLMs are comparatively better in helping learners for 'paper understanding' when the correct information is provided rather than in open-ended 'literature review' tasks where LLMs have proven to be weaker (Wang et al., 2025; see also Yan et al., 2024). As a result, the issue of generating hallucinations or false references has been common (Maci & Anesa, 2025; see also Claman & Sezgin, 2024). Dibrell (2024) has mentioned that the 'Stochastic parrot' phenomenon related to the functionality of LLMs has caused them to produce work that looks credible but often includes many errors and false details. Highlighting the issue of hallucinations, Yan et al (2024) has pointed out how AI detection tools such as Turnitin produce false-positive rates causing unjust decisions that can have negative consequences on student grades and self-perception. Another result of fabrication of data is academic misconduct that happens when students present work in the assumption that research was carried out when it actually was not (Erkılıç & Ibrahim Cifci, 2024).

The issue of over-reliance

In the same study that has observed 48 graduate students and their use of LLMs in higher education, Wang et al (2025) has mentioned that over-reliance or inappropriate adoption of LLMs is a rising issue. Due to mismanagement of time or lack of time and tight deadlines, students tend to rely heavily and completely on tools powered by LLMs. Such over reliance can result in a number of issues. The creativity and critical thinking of students are substantially hindered as the heavy use of LLMs prevent students from expressing authenticity, personal engagement, reflective depth and they produce very generic work that lack individuality (Dibrell, 2024; see also Claman & Sezgin, 2024). Maci & Anesa (2025) has stated that LLM powered automated tools can assist learners and educators but they cannot replace the pivotal role played by the human analyst and that as a result, development of essential human interpretive skills is hindered. Further, when students get used to producing work generated by LLMs, they often lack deep understanding of the learning concepts and indicate very less confidence of the surface level work (Dibrell, 2024).

Lack of digital literacy

Another issue related to the use of LLMs is the lack of digital literacy amongst both students and educators. Most LLM based education tools are currently in lab-testing phases or during early research (Maci and Anesa, 2025). Even though they are already heavily used by learners, they have not been sufficiently tried in real classroom settings. This poses a challenge on the practical efficacy of LLMs and digital illiteracy further feeds the inability of educators and students to effectively recognize the limitations and biases in AI generated content or fix these errors (Maci & Anesa, 2025).

Limitations and biases in training data

Turning towards the ethical concerns surrounding the use of LLMs in education, one of the most vital issues is how learning equity can be restricted due to limitations in training datasets of LLMs. Nguyen Van Viet et al (2024) has mentioned that training datasets of LLMs are considered to be globally unbalanced and they are trained by and large on English data and from western, educated, industrialized, rich and democratic countries. This limitation on training has therefore marginalized learners representing other cultural backgrounds from having equitable access and deriving the expected and most effective outputs from the LLMs (Nguyen Van Viet et al., 2024). Further, the deployment and maintenance of LLMs is expensive and therefore, institutions with less resources find it difficult to have access to high level LLMs (Guizani et al., 2025). This again creates educational inequalities. Another result of such bias in training data is the representation of systemic inequalities and underrepresentation of the minority (Dibrell, 2024). The cost of modifying or updating model-generated content is high and therefore, this drawback hampers accuracy in the output created by LLMSs (Guizani et al., 2025).

Data privacy concerns

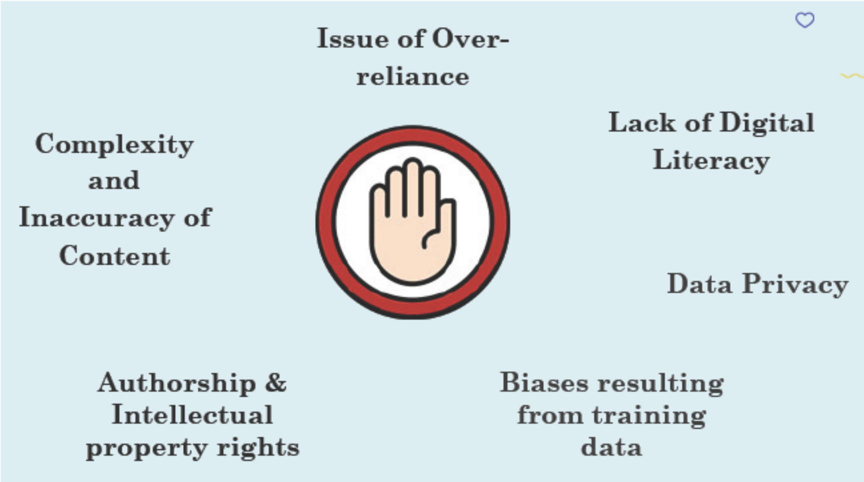
Another common ethical consideration is the privacy related to using student data for fine-tuning of LLMs (Claman & Sezgin, 2024). According to Yan et al. (2024) if LLMs are fine-tuned using private data of students without any strong anonymization, details of personal identities can leak through future responses produced by LLMs.

Issues related to authorship

When researchers or academic writers depend on LLM powered tools for content generation, few writers can submit the same work for multiple journals, probably minimally altered (Erkılıç & Ibrahim Cifci, 2024). This happens due to the duplication of work done by LLMs and slicing or spilling one work or project into several publications. This is considered a substantial violation of publication ethics (Erkılıç & Ibrahim Cifci, 2024). Another important ethical concern surrounding the use of LLMs is the ambiguities related to authorship and attribution of academic work. AI powered LLMs are not able to fulfill the requirements of scientific authorship and their ability to generate content blurs academic records (Erkılıç & Ibrahim Cifci, 2024). Accordingly, they end up removing credit from actual contributors or incorrectly assign authorship (Erkılıç & Ibrahim Cifci, 2024). This is also tied to the issue of unclear intellectual property rights of content generated by LLMs through training on proprietary materials (Erkılıç & Ibrahim Cifci, 2024; see also Guizani et al., 2025).

Figure 7

Limitations of using LLMs



Recommendations to mitigate the effects of limitations

The existing literature has introduced several recommendations that can be implemented to address the limitations and challenges posed by LLMs.

Introducing relevant socio-legal frameworks

Matijevic et al (2024) has performed a systematic literature review to construct a socio-legal framework for the usage of LLMs and their application in education. This study has identified the dire need for a regulatory framework to govern how LLMs are used in education given the socio-legal challenges resulting from their application. According to the authors the majority of AI governance is based on non-binding guidelines and therefore results in inconsistent enforcement (Matijevic et al., 2024). Furthermore, only a few countries such as the USA, China and EU states have copyright and data laws in place for AI. Accordingly, the authors have recommended introducing data governance laws that will focus on GDPR compliance, anonymizing student data and obtaining consent for use of personal data (Matijevic et al (2024). Another solution recommended for this is to deploy LLM systems within institution-controlled environments or local/closed environments in order to mitigate the unauthorized use of data and leaks (Yan et al., 2024).

Fine-tuning LLMs

Further, for the biases presented in the outputs of LLMs stemming from the biases and social inequalities embedded in the training data sets, the existing literature has recommended to fine-tune LLMs in a way to match the local languages, different cultural context and literacy levels (Choi, 2025). It is also proposed to use adaptive language processing which will allow the LLMs to tailor their outputs to the varying backgrounds of the learners so that the social achievement gaps will not be reinforced through the responses of LLMs (Choi, 2025). Another solution to this same issue is to calibrate the LLMS for local academic standards to prevent content generation focused on one-size-fits-all style (Nguyen Van Viet et al., 2024). This will also help educators from different regions and demographics to get the optimal assistance from LLMs. To address the issue of hallucinations, Jacobsen & Weber (2023) has recommended critically evaluating the outputs generated by LLMs and refine the prompts iteratively to mitigate the errors.

Developing prompt engineering skills

Another important solution to mitigate the negative impact of LLMs is to enhance the prompt engineering skills of both educators and students. The output of the LLMs is greatly dependent on the prompt we send to them and if the prompt is poorly constructed, it can lead to inaccurate or irrelevant results (Lee & Palmer, 2025). Wang et al. (2021) has introduced the CRISPE framework that outlines the basis to craft precise prompts. Accordingly, Capacity and role refers to defining the role of the LLM in which we want them to operate and therefore, confirm whether we want them to act as a copywriter, researcher, or translator. Next is ‘I’ for insight which refers to providing the necessary information and contextual background to the LLM for the reason of relevancy. Next, ‘S’ stands for Statement and this means the core of the prompt. This refers to what we ask the LLM to do and it is

recommended to be very specific to get a more precise response. Next is Personality and this reflects on the personality of the LLMs or how we want the LLMs to respond. For example, do we require the LLM to provide a professional response or casual friendship? This will enhance personalization in outputs and individuality. The final part is Experiment which requires the user to ask LLMs to provide multiple responses or examples and therefore, having the option to choose the best response or combine few different responses.
A similar approach is using the AIPROMT framework for refining prompts to achieve clarity and specificity (Korzynski et al., 2023). Another step that can be taken to enhance prompt engineering skills is to include the same in academic curriculum and train the students on this. Accordingly, prompt engineering skills of learners should be observed and assessed to make sure that they are well positioned to obtain optimal results from the use of LLMs. Knoth et al. (2024) has suggested a quantitative prompt quality score to assess the clarity, context and restrictions in prompts written by learners.

Policies to enhance academic integrity

To address the issues related to authorship and academic integrity, the educational institutions must implement policies to robust authorship and practice attribution habits. Accordingly, learners must be encouraged to clearly distinguish between contributions by human brain and AI tools and disclose any assistance taken from LLMs in their assignments and publications (Erkılıç & Cifci, 2025; see also Dibrell, 2024). Further, assessments must be modelled in such a way that they cannot be easily outsourced to LLMs. To do this, assessments must be created in a way to focus on emphasizing synthesis, analysis, and real-world examples and applications so that learners cannot simply follow responses provided by LLMs (Wang et al., 2025). Further, introducing clear institutional policies to guide the students in the usage of AI will mitigate any overreliance and force the students to balance their involvement with the integration of LLMs in their work (Klar, 2025).

Enhancing digital literacy

Another solution to improve the usage of LLMs amongst the learners is to enhance their information literacy education. Bećirović et al. (2025) has mentioned that by fostering vetting skills and teaching the learners to analyze the validity of output generated by LLMs against trusted sources and therefore be informed and trained to identify any hallucinations produced by LLMs. It should be taught to the students that LLMs must only be used as debaters or moderators but not as mere answer providers. Learners must be skilled to properly justify, critique or build upon the content generated through LLMs (Reicher et al., 2025).

Increase empirical studies

One of the focal issues related to the development of LLMs is the lack of empirical evidence related to their impact in actual classroom settings. Therefore, it is recommended for higher education institutes and high schools to carry out open and real-world pilot studies that will foster the sharing of methods and outcomes that need to develop a solid knowledge of how LLMs work (Dibrell, 2024; see also Yan et al., 2024). Further, involving the main stakeholders, students and educators in the process of designing and evaluation will enhance the user perspective and improve the identification of real-world issues related to usage that will pave the way for efficient and practical solutions (Filippi & Motyl, 2024).

The future of LLMs in Learning; Emerging trends and possible trends

The educational sector is looking ahead to identify the emerging trends that would define the role of LLMs in learning in the future. LLMs are already creating a powerful transformative effect in education even though most of these LLMs are still in their early stages of development. Therefore, if the creators of these models can address the limitations and drawbacks hindering the optimal performance of these models with effective solutions, the power of LLMs and what they can do is beyond one's imagination.

Parakh et al. (2024), speaking of the future role of LLMs in education, has mentioned that models such as Llama 3.1, CodeLlama, and Mistral can be used for retrieval-augmented generation (RAG) systems and vector databases to enhance personalized learning experiences through more targeted and adaptive content delivery. The aim of such developments is to encourage engagement and knowledge retention of learners by enhancing the role of LLMs as both content generators and effective real-time feedback agents (Parakh et al., 2024).

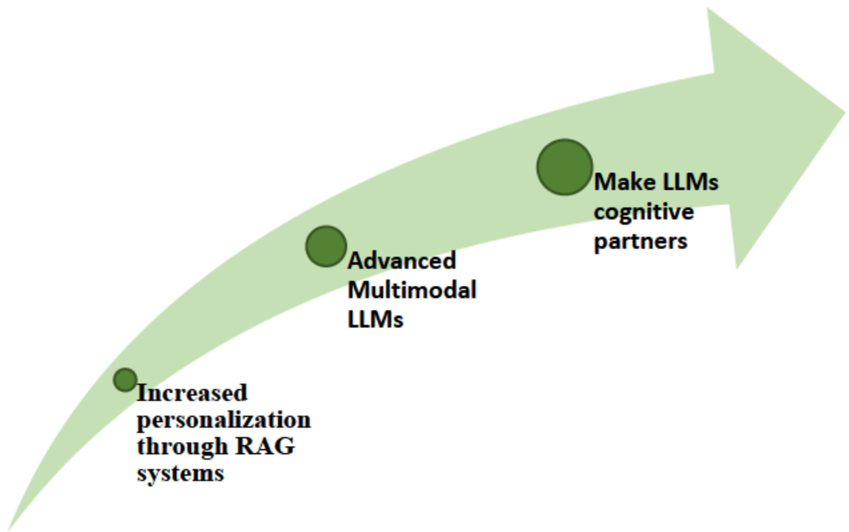
Another emerging trend of LLMs is their potential ability to create a rich learning experience through multimodal LLMs. Al-Dokhny et al. (2024) has discussed the potential use of multimodal LLMs mediated by acceptance frameworks such as Task-technology Fit, AIDUA that will consider the hedonic motivation, social influence, expectancy of effort and willingness to adopt. Such advancements will result in sustainable improvements in higher education and align with continuing customization. This study has empirically confirmed that multimodal LLMs are able to refine personalized education through exclusive fit of technology to the task and user context (Al-Dokhny et al., 2024). It has also been discussed that it is likely to adopt new systems based on adaptive learning and automated 'sensing' tools for common use of research and practice (Whitehead et al., 2025).

Another potential role that is envisioned for the LLMs is the role to function as cognitive partners rather than mere information sources. Expanding the application of LLMs to stakeholder and collaborative project-based assignments show considerable potential of LLMs improving group work and enhancing collaborative education by facilitating social-emotional skills acting as interactive partners (Cain, 2025). Further, it has been identified that some models like GPT variants have started to shift from labor-intensive manual coding to scalable, just-in-time analytics and therefore extending access to advanced learning analytics even to non-experts (Castellanos-Reyes, 2025).

The reviewed studies have emphasized collective evidence to envision a remarkable transformation of the role played by LLMs in learning. These potential advancements are setting new benchmarks in personalized education featured by collaborative learning, real-time feedback and other advancements. The developers should however navigate the potential drawbacks resulting from such possible advancements and keep developing these models to address any factors that hinder their optimal usage.

Figure 8

Emerging trends for LLMs



Conclusion

In the current context of education, learning has taken the shape of many methods. It is not restricted to the traditional approach of in-person learning which happens primarily through in-person interactions between educators and learners. However, with the evolution of technology and perspective of learning, the methods of learning has extended beyond the traditional system to many other methods such as online learning, and self-directed learning. These new learning systems are consisted of a number of new features that give rise to different activates that are undertaken to achieve the objectives of learning. The development of technology and more specifically, artificial intelligence has heavily supported these learning activates by integrating advancements to the mode of delivery and receipt of education. AI has reshaped the learning systems, especially through LLMs, it's one of the dominant counterparts. This study has conducted a systematic overview of the existing literature on how LLMs support education and the analysis has proven a wide-range of applications of LLMs in educational context. These LLMs have introduced many new features that give rise to numerous benefits. The existing literature has identified how LLMs have reshaped learning and education by bringing personalization, engagement, collaboration, enhancement of efficiency to not only learning experiences but also to the role of the educators. Collaborative experiences and real-time feedback have been heavily highlighted by the current studies as valuable outputs of integrating LLMs in education. However, the existing researchers have been able to identify drawbacks and limitations associated with the application of LLMs in learning. These limitations can result from many reasons but mainly due to the fact that these models are still in the early stages of development and the lack of empirical evidences in practice. Ethical concerns related the operation of the LLMs are also requires substantial attention. The current studies have not failed to highlight the appropriate solutions for these drawbacks and emphasized the importance of increasing the deployment of LLMs in actual scenarios including physical classrooms and personal experiences to gain real-time evaluation. By this way, empirical evidence can be collected to measure the success of the role played by LLMs in learning.

It is vital to understand by both learners and educators that LLMs are an efficient tool to supplement and enhance the learning experience. By understanding how these LLMs work and trained, the learners will realize that they are not supernatural creatures that offer what you need. Rather, they are modelled and trained to support the learning process and human mediation plays important role in achieving the optimal results out of the LLMs. Therefore, it must be strongly understood that LLMs are here to support the learners and educators but not replace any of them or any part of the learning process. They are here to be developed and to be used for attaining the highest results of education through supplementing of human cognitive skills.

References

Adams, N. E. (2015). Bloom's taxonomy of cognitive learning objectives. *Journal of the Medical Library Association*, *103*(3), 152–153. https://doi.org/10.3163/1536-5050.103.3.010

Agüera y Arcas, B. (2022). Do large language models understand us? *Dædalus, 151*(2), 183–197. https://doi.org/10.1162/DAED_a_01909

Ajani, O. A., Akintolu, M., & Afolabi, S. O. (2024). The emergence of artificial intelligence in higher education: Prospects and challenges of AI. International Journal of Research in Business and Social Science, 13(8), 157–165. https://doi.org/10.20525/ijrbs.v13i8.3507

Al-Dokhny, A., AlHarthi, H. M., Alalwan, N., & Alkhalaf, S. (2024). Can Multimodal Large Language Models Enhance Performance Benefits Among Higher Education Students? *Sustainability*, 16(10780), 1–28.

Alfirević, N., Gojević, A., Blažević, N., & Crnjac Milić, D. (2024). Custom-Trained Large Language Models as Open Educational Resources. *Information*, 15(2), 89. https://doi.org/10.3390/info15020089

Baltaci, K., Herrmann, M., & Turkmen, A. (2024). Integrating Artificial Intelligence into Electrical Engineering Education: A Paradigm Shift in Teaching and Learning.

Bećirović, S., Polz, E., & Tinkel, I. (2025). Exploring students' AI literacy and its effects on their AI output quality, self-efficacy, and academic performance. *Smart Learning Environments, 12*(29). https://doi.org/10.1186/s40561-025-00384-3

Belkina, M., Daniel, S., Nikolic, S., Haque, R., Lyden, S., Neal, P., Grundy, S., & Hassan, G. M. (2025). Implementing generative AI (GenAI) in higher education: A systematic review of case studies. *Computers and Education: Artificial Intelligence, 8*, 100407. https://doi.org/10.1016/j.caeai.2025.100407

Bewersdorff, E., Sandmann, A., Ulbricht, C., Schreiber, F., Greefrath, G., & Rumann, S. (2025). Taking the next step with generative artificial intelligence: The transformative role of multimodal large language models in science education. *Education and Information Technologies, 30*(2), 1205–1234. https://doi.org/10.1007/s10639-025-12345-x

Bucea-Manea-Tonis, R., et al. (2022). Artificial Intelligence Potential in Higher Education Institutions: Enhanced Learning Environment in Romania and Serbia. *Sustainability, 14*, 5842.

Cain, W. (2025). Cognitive partners in design: Using generative AI for active online learning in a graduate-level course. *Educational Research: Theory and Practice*, 36(1), 42–51.

Campbell, L. O., & Cox, T. D. (2024). Utilizing generative AI in higher education teaching and learning. *Journal of the Scholarship of Teaching and Learning, 24*(4), 162–173. https://doi.org/10.14434/josotl.v24i4.36575

Castellanos-Reyes, D., Olesova, L., & Sadaf, A. (2025). Transforming online learning research: Leveraging GPT large language models for automated content analysis of cognitive presence. *International Review of Research in Open and Distributed Learning, 26*(2), 98–115. https://doi.org/10.19173/irrodl.v26i2.7850

Cecchini, S., Ceccacci, S., Giaconi, C., & Del Bianco, N. (2025). AI chatbots in education: Challenges and opportunities. *Information, 16*(3), 235. https://doi.org/10.3390/info16030235

Choi, Y. (2025). Exploring the scientific validity of ChatGPT's responses in elementary science for sustainable education. *Sustainability, 17*, 2962. https://doi.org/10.3390/su17072962

Claman, D., & Sezgin, E. (2024). Artificial intelligence in dental education: Opportunities and challenges of large language models and multimodal foundation models. *JMIR Medical Education, 10*, e52346. https://doi.org/10.2196/52346

Davar, N. F., Dewan, M. A. A., & Zhang, X. (2025). AI chatbots in education: Challenges and opportunities. *Information, 16*(3), 1-25. https://doi.org/10.3390/info16030235

Davar, P., Tang, K.-Y., & Wang, F. (2025). Evaluating simulated teaching audio performance using retrieval-augmented generation and large language models. *Information*, 16(4), 264. https://doi.org/10.3390/info16040264

Dibrell, D. (2024). Ethical considerations of using AI in writing pedagogy. *Sociological Imagination, 60*(1), 20–24.

Doğan, D., Yıldız, S., & Özkan, S. (2025). Effectiveness of artificial intelligence practices in the teaching of social sciences: Pre-school education. *Educational Technology & Society, 28*(3), 77–92.

Eber, P. A., & Parker, T. S. (2007). Assessing Student Learning: Applying Bloom's Taxonomy. *Human Service Education*, *27*(1), 45–53.

Erkılıç, E., & Cifci, I. (2025). Ethical issues and violations in using chatbots for academic writing and publishing: The answers from ChatGPT. *Journal of Multidisciplinary Academic Tourism, 10*(1), 111–120. https://doi.org/10.31822/jomat.2025-10-1-111

Filippi, S., & Motyl, B. (2024). Large language models (LLMs) in engineering education: A systematic review and suggestions for practical adoption. *Information*, *15*(6), 1-19.. https://doi.org/10.3390/info15060345

Guizani, I., Mehmood, R., Leclercq, E., Abuadbba, A., Song, H., & Yasar, A. (2025). A systematic literature review to implement large language model in higher education: Issues and solutions. *Discover Education, 5*(1), 12–34. https://doi.org/10.xxxx/discedu.2025.xxxxx

Horikoshi, K., Ikeda, K., Shibata, Y., & Takahashi, T. (2025). Learning stream plot: Classroom activity visualization with daily learning log data. *Journal of Smart Learning Environments, 12*(15), 1–23.

Imran, M., & Almusharraf, N. (2024). Google Gemini as a Next Generation AI Educational Tool: A Review of Emerging Educational Technology. *Smart Learning Environments, 11:22*.

Issabek, A., Oliveira, W., Hamari, J., & Bogdanchikov, A. (2025). The effects of demographic factors on learners' flow experience in gamified educational quizzes. *Smart Learning Environments, 12*(25), 1–20. https://doi.org/10.1186/s40561-025-00378-1

Keith, J. R., Deng, S., & Carmichael, P. (2025). Harnessing generative AI in chemical engineering education through critical engagement. *Chemical Engineering Education, 59*(1), 4–13.

Klar, M. (2025). Using ChatGPT is easy, using it effectively is tough? A mixed methods study on K-12 students' perceptions, interaction patterns, and support for learning with generative AI chatbots. *Smart Learning Environments, 12*(32). https://doi.org/10.1186/s40561-025-00385-2

Lang, M., Triantoro, T., & Sharp, J. H. (2024). Large language models as "educational assistant": Automatic writing of teaching cases for IS/business/IT education. *Journal of Information Systems Education, 35*(2), 133–144.

Lee, D., & Palmer, E. (2025). Prompt engineering in higher education: A systematic review to help inform curricula. *International Journal of Educational Technology in Higher Education, 22*(7). https://doi.org/10.1186/s41239-025-00503-7

Lucas, H. C., Upperman, J. S., & Robinson, J. R. (2024). A systematic review of large language models and their implications in medical education. *Medical Education, 58*(11), 1276–1285. https://doi.org/10.1111/medu.15402

Maci, S. M., & Anesa, P. (2025). The impact of AI on discourse analysis: Challenges and opportunities. *International Journal of Language Studies, 19*(2), 1–12. https://doi.org/10.5281/zenodo.15250758

Mezak Matijevic, M., Pisker, B., & Dokic, K. (2024). Constructing a socio-legal framework proposal for governing large language model usage and application in education. *Social Sciences, 13*(9), 479. https://doi.org/10.3390/socsci13090479

Nedungadi, P., Raman, R., Deng, X., & Bianchi, B. (2024). Generative AI for achieving SDG 4: Equity and quality in education. *Smart Learning Environments, 11*(1), 5. https://doi.org/10.1186/s40561-024-00336-5

Newton, P., & Xiromeritis, M. (2024). ChatGPT performance on multiple choice question examinations in higher education: A pragmatic scoping review. *Assessment & Evaluation in Higher Education, 49*(6), 781–798. https://doi.org/10.1080/02602938.2023.2299059

Nguyen Van, V., Luong Thi Minh, H., Nguyen, V., Mong Quoc, T., Bui Anh, T., Le Anh, T., Phan Thi, C., & Nguyen Huu, K. (2024). Revolutionizing education: An extensive analysis of large language models integration. *International Research Journal of Science, Technology, Education, and Management, 4*(4), 10–21. https://doi.org/10.5281/zenodo.14744029

Özçelik, N. P., & Yangın Ekşi, G. (2024). Cultivating writing skills: The role of ChatGPT as a learning assistant—A case study. *Smart Learning Environments, 11*(10), 1–20. https://doi.org/10.1186/s40561-024-00296-8

Pack, A., Barrett, A., & Escalante, J. (2024). Large language models and automated essay scoring of English language learner writing: Insights into validity and reliability. *Computers and Education: Artificial Intelligence, 6*, 100234. https://doi.org/10.1016/j.caeai.2024.100234

Parakh, D., Lodha, P., & Subalalitha, C. N. (2024). Adaptive Personalized Learning System with Generative AI. *Frontiers in Health Informatics*, 13(8), 2612–2637.

Praveena, T., & Anupama, K. (2025). Machine learning meets language learning: The transformative potential of artificial intelligence in English language instruction. Human Research in Rehabilitation, 15(1), 133–144. https://doi.org/10.21554/hrr.042511

Reicher, H., Frenkel, Y., Lavi, M. J., Nasser, R., Ran-milo, Y., Sheinin, R., Shtaif, M., & Milo, T. (2025). A generative AI-empowered digital tutor for higher education courses. Information, 16(4), 264. https://doi.org/10.3390/info16040264

Ruiz-Rojas, L. I., Salvador-Ullauri, L., & Acosta-Vargas, P. (2024). Collaborative working and critical thinking: Adoption of generative artificial intelligence tools in higher education. Sustainability, 16(13), 5367. https://doi.org/10.3390/su16135367

Sarangi, S., Choudhury, S. S., Agarwal, A., Thakker, A., & Cherian, J. (2024). Exploring radiology postgraduate students' engagement with large language models for educational purposes: A study of knowledge, attitudes, and practices. *Journal of European CME, 13*(1), 224–235. https://doi.org/10.1080/21614083.2024.1234567

Shahzad, T., Mazhar, T., Tariq, M. U., Ahmad, W., Ouahada, K., & Hamam, H. (2025). A comprehensive review of large language models: Issues and solutions in learning environments. *Discover Sustainability, 6*, Article 27. https://doi.org/10.1007/s43621-025-00815-8

Shen, M., Shen, Y., Liu, F., & Jin, J. (2025). Prompts, privacy, and personalized learning: Integrating AI into nursing education—a qualitative study. *BMC Nursing, 24*, Article 470. https://doi.org/10.1186/s12912-025-03115-8

Sun, X., Yang, Y., & Du, Q. (2025). Advancing total quality management in higher education through AI and big data analytics. Total Quality Management & Business Excellence. Advance online publication. https://doi.org/10.1080/14783363.2025.2495103

Uğraş, M., Soylu, Y., & Zerengök, D. (2024). ChatGPT-supported education in primary schools: Teachers' perspectives. *Eğitimde Nitel Araştırmalar Dergisi (Journal of Qualitative Research in Education), 12*(52), 317–338. https://doi.org/10.14689/issn.2148-2624.1.12c.52m.13

Wang, J., Huang, C., Yan, S., Xie, W., & He, D. (2025). When young scholars cooperate with LLMs in academic tasks: The influence of individual differences and task complexities. *International Journal of Human–Computer Interaction, 41*(8), 4624–4639. https://doi.org/10.1080/10447318.2024.2352919

Wu, Q., Dang, Y., & Li, F. (2025). Systematic review of generative artificial intelligence (GenAI) in higher education: Opportunities, challenges, and practical implications. *Information, 16*(3), 235. https://doi.org/10.3390/info16030235

Xing, W., Lan, A., Crossley, S., Yu, Z., Denny, P., Nixon, N., & Stamper, J. (2024). Revolutionizing education: An extensive analysis of large language models integration. *International Research Journal of Science, Technology, Education, and Management, 4*(4), 10–21. https://doi.org/10.5281/zenodo.14744029

Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2023). Practical and ethical challenges of large language models in education: A systematic literature review. *British Journal of Educational Technology, 54*(5), 91–108.

Yu, Y. F., Mai, N., & Hin, H. S. (2020). Enhancing Students' Collaborative Learning Experiences with Laurillard's Conversational Framework. *International Journal of Creative Multimedia*, *1*(Special Issue 1), 152–167. https://doi.org/10.33093/ijcm.2020.X1.13

Appendix 1

Keywords, their synonyms and search strings used for the search of academic papers

Key words	Synonyms	Search Srting	# of articles	# of relevant articles
AI in education	AI tools in education, machine learning applications for schools	("artificial intelligence" OR AI OR "AI tools" OR "machine learning") AND (educat* OR school* OR teach* OR learn* OR classroom OR "higher education" OR universit*)	80	6
LLM in eductaion	LLM in education and schools	("LLM" OR "large language model" OR "large language models") AND (educat* OR school* OR teach* OR learn* OR classroom OR universit*)	16,041	50
intelligent turoring systems based on AI and LLM	AI and LLM tutoring platforms, personalized AI/ LLM tutors	("intelligent tutoring system*" OR "ITS" OR "AI tutor*" OR "LLM tutor*" OR "AI tutoring platform*" OR "LLM tutoring platform*" OR "personalized AI tutor*" OR "personalized LLM tutor*") AND ("artificial intelligence" OR "AI" OR "large language model*" OR "LLM" OR "generative AI" OR "ChatGPT" OR "GPT-3" OR "GPT-4") AND (educat* OR school* OR teach* OR learn* OR classroom* OR	82,000	10

		universit* OR "higher education" OR "K-12")		
Natural Language Processing in Eductaion	Natural Language Processing for Learning, Language AI in Education	("Natural Language Processing" OR "NLP" OR "Language AI" OR "Text AI") AND (educat* OR learn* OR teach* OR school* OR universit* OR "higher education" OR "K-12" OR classroom*)	30,000	10
Virtual Learning Environments through AI	Digital classroms, online learning spaces	("Virtual Learning Environment*" OR "VLE" OR "Digital Classroom*" OR "Online Learning Space*" OR "AI-powered Learning Environment*" OR "Intelligent Learning Platform*") AND ("Artificial Intelligence" OR "AI" OR "Machine Learning" OR "Adaptive Learning" OR "Intelligent Tutoring System*" OR "Generative AI" OR "LLM" OR "ChatGPT") AND (educat* OR learn* OR teach* OR school* OR universit* OR "K-12" OR "higher education"	161	8

		OR "distance learning" OR "e-learning")		
Virtual Learning Environments through LLM	LLM based digital classroms and online learning spaces	(("Virtual Learning Environment*" OR "VLE" OR "LLM-based Digital Classroom*" OR "LLM-powered Online Learning Space*") AND ("Large Language Model*" OR "LLM" OR "Generative AI" OR "ChatGPT" OR "GPT-3" OR "GPT-4" OR "Transformer Model*") AND (educat* OR learn* OR teach* OR school* OR universit* OR "K-12" OR "higher education" OR "distance learning"))	1	1
Student engagement with AI	AI driven student motivation, learner engagement with technology	("student engagement" OR "learner engagement" OR "student motivation") AND ("artificial intelligence" OR "AI" OR "AI-driven" OR "adaptive learning") AND (educat* OR learn* OR school* OR universit*)	1,121	12

Student engagement through LLMs	LLMs driven student motivation, learner engagement through LLMs	LLM-driven student motivation, learner engagement through LLMs ("student engagement" OR "learner engagement") AND ("large language model*" OR "LLM" OR "ChatGPT" OR "generative AI") AND (educat* OR learn* OR classroom*)	0	0
Collaborative learning with AI	AI supported group learning, digital peer collaboration	("collaborative learning" OR "group learning" OR "peer collaboration") AND ("artificial intelligence" OR "AI" OR "AI-supported") AND (educat* OR school* OR "higher education")	573	5
Collaborative learning through LLMs	Group learning based on LLMs, digital peer collaboration	("collaborative learning" OR "group learning") AND ("large language model*" OR "LLM" OR "GPT-4") AND (educat* OR universit* OR "online learning")	5,812	8
Knowledge Retention with AI	Memory enhancement tools, AI assisted memorization	("knowledge retention" OR "memory enhancement" OR "memorization") AND ("artificial intelligence" OR "AI" OR "adaptive learning") AND (educat* OR learn* OR school*)	191	2
Knowledge Retention with LLMs	LLM supported memory enhancement tools, LLMs assisted memorization	("knowledge retention" OR "memory enhancement") AND ("large language model*" OR "LLM" OR "generative AI") AND (educat* OR learn* OR "higher education")	3	9

Critical Thinking Development with AI	Higher-order cognitive skills, reasoning skills with AI	("critical thinking" OR "higher-order cognitive skills" OR "reasoning skills") AND ("artificial intelligence" OR "AI" OR "machine learning") AND (educat* OR school* OR universit*)	1,330	11
Critical Thinking Development with LLMs	Higher-order cognitive skills enahnced by LLMs, reasoning skills with LLMs	("critical thinking" OR "reasoning skills") AND ("large language model*" OR "LLM" OR "ChatGPT") AND (educat* OR "higher education")	434	5
Assessment Automation with AI	AI-graded assignments, smart testing systems	("assessment automation" OR "automated grading" OR "smart testing") AND ("artificial intelligence" OR "AI" OR "machine learning") AND (educat* OR school* OR universit*)	71	4
Assessment automation with LLM	Grading assignments through LLMs, smart testing sytems	("automated grading" OR "assessment automation") AND ("large language model*" OR "LLM" OR "GPT-4") AND (educat* OR "higher education")	2	0
Personalized Feedback through AI	Adaptive feedback, instant AI feedback	("personalized feedback" OR "adaptive feedback" OR "instant feedback") AND ("artificial intelligence" OR "AI" OR "machine learning") AND (educat* OR learn* OR school*)	294	20
Personalized Feedback through LLMs	Adaptive feedback and instant feedback from LLMs	("personalized feedback" OR "adaptive feedback") AND ("large language model*" OR "LLM" OR	15	10

		"ChatGPT") AND (educat* OR universit* OR "online learning")		
AI-enhanced Creativity	Creative idea generation by AI, AI support for innovation	("creativity" OR "creative idea generation" OR "innovation") AND ("artificial intelligence" OR "AI" OR "generative AI") AND (educat* OR learn* OR school*)	33,017	12
LLM enhanced creativity	Creative idea generation through LLMs, LLM support for innovation	("creativity" OR "idea generation") AND ("large language model*" OR "LLM" OR "GPT-4") AND (educat* OR "higher education")	5273	6
AI-powered Curriculum Design	AI in syllabus development, intelligent course design	("curriculum design" OR "syllabus development" OR "course design") AND ("artificial intelligence" OR "AI" OR "machine learning") AND (educat* OR school* OR universit*)	228	0
LLM-powered Curriculum Design	Syllabus development and course designing through LLMs,	("curriculum design" OR "syllabus development") AND ("large language model*" OR "LLM" OR "generative AI") AND (educat* OR "higher education")	18	12
Learning Analytics via AI	Educational data mining, AI for student performance tracking	("learning analytics" OR "educational data mining" OR "student performance") AND ("artificial intelligence" OR "AI" OR "predictive analytics") AND (educat* OR school* OR universit*)	930	20

Learning anlytics via LLMs	Educational data mining, student performance tracking through LLMs	("learning analytics" OR "student performance") AND ("large language model*" OR "LLM" OR "generative AI") AND (educat* OR "higher education")	46	10
Educational Chatbots through AI	AI conversation agents, virtual learning assistants	("educational chatbot*" OR "virtual assistant*" OR "conversational agent*") AND ("artificial intelligence" OR "AI" OR "NLP") AND (educat* OR learn* OR school*)	915	25
LLM based educational chatbots	LLM based converstaion agents, virtual learning assistants	("educational chatbot*" OR "conversational agent*") AND ("large language model*" OR "LLM" OR "ChatGPT") AND (educat* OR "online learning")	44	12
Competency-based Education with AI	Skills mastery tracking, personalized skill development	("competency-based education" OR "skills mastery" OR "skill development") AND ("artificial intelligence" OR "AI" OR "adaptive learning") AND (educat* OR school* OR universit*)	425	10
Competency-based Education with LLMs	Skills mastery tracking, personalized skill development	("competency-based education" OR "skills mastery") AND ("large language model*" OR "LLM" OR "generative AI") AND (educat* OR "higher education")	3	0
AI for Skill Development	Soft skills AI training, hard skills AI learning	("skill development" OR "soft skills" OR "hard skills") AND ("artificial intelligence" OR "AI" OR "machine learning") AND	517	18

		(educat* OR learn* OR school*)		
LLMs for skill Development	Soft and hard skills trainiong and learning through LLMS	("skill development" OR "soft skills" OR "hard skills") AND ("large language model*" OR "LLM" OR "GPT-4") AND (educat* OR "vocational training")	12	2
Microlearning with AI	Bite-sized learning modules, AI-driven short lessons	("microlearning" OR "bite-sized learning" OR "short lessons") AND ("artificial intelligence" OR "AI" OR "adaptive learning") AND (educat* OR learn* OR school*)	8	0
Microlearning with LLMs	Bite-sized learning modules, LLM-driven short lessons	("microlearning" OR "bite-sized learning") AND ("large language model*" OR "LLM" OR "ChatGPT") AND (educat* OR "mobile learning")	0	0
Virtual Reality and AI Learning	Immersive learning technologies, VR/AR blended education	("virtual reality" OR "VR" OR "augmented reality" OR "immersive learning") AND ("artificial intelligence" OR "AI" OR "machine learning") AND (educat* OR school* OR universit*)	5,121	2
Virtual Relaity and LLM learning	Immersive learning technologies, VR/AR blended education	("virtual reality" OR "VR" OR "augmented reality") AND ("large language model*" OR "LLM" OR "generative AI") AND (educat* OR "higher education")	15	6

Gamification with AI	AI-based educational games mechanics in learning	("gamification" OR "educational game*" OR "game-based learning") AND ("artificial intelligence" OR "AI" OR "adaptive learning") AND (educat* OR school* OR universit*)	646	0
Gamification through LLMs	LLM-based educational games and mechanics in learning	("gamification" OR "game-based learning") AND ("large language model*" OR "LLM" OR "GPT-4") AND (educat* OR "higher education")	1	0
Ethics in AI Education	Responsible AI use, ethical considerations in educational AI	("ethics" OR "responsible AI" OR "ethical considerations") AND ("artificial intelligence" OR "AI" OR "machine learning") AND (educat* OR school* OR universit*)	17,330	0
Ethics in LLM supported education	Responsible use of LLMs, ethical consideration of LLMs in education	("ethics" OR "responsible AI") AND ("large language model*" OR "LLM" OR "generative AI") AND (educat* OR "higher education")	494	10
Bias and Fairness in Educational AI	Algorithmic bias in education, fairness in AI assessments	("bias" OR "fairness" OR "algorithmic bias") AND ("artificial intelligence" OR "AI" OR "machine learning") AND (educat* OR school* OR universit*)	16,495	10
Bias and Fairness in Educational LLMs	Algorithmic bias in education, fairness in LLM assessments	("bias" OR "fairness" OR "algorithmic bias") AND ("large language model*" OR "LLM" OR "ChatGPT") AND (educat* OR "higher education")	402	19

Multimodal Learning with AI	AI combining text, Images, and audio for education	("multimodal learning" OR "multimedia learning") AND ("artificial intelligence" OR "AI" OR "computer vision" OR "speech recognition") AND (educat* OR school* OR universit*)	1,684	23
Multimodal Learning with LLMs	LLMs combining texts, images and audio for education	("multimodal learning" OR "multimedia learning") AND ("large language model*" OR "LLM" OR "GPT-4 Vision") AND (educat* OR "higher education")	36	5
Equity and Inclusion in AI Education	Accessible AI learning, inclusive EdTech	("equity" OR "inclusion" OR "accessible learning") AND ("artificial intelligence" OR "AI" OR "machine learning") AND (educat* OR school* OR universit*)	10,150	22
Equity and Inclusion in LLMs supported learning	Accessible learning with LLMs, inclusive EdTech	("equity" OR "inclusion" OR "accessible learning") AND ("large language model*" OR "LLM" OR "generative AI") AND (educat* OR "higher education")	162	7
Teacher-AI Collaboration	Human-AI co-teaching models educators using AI tools	("teacher collaboration" OR "co-teaching" OR "educator tools") AND ("artificial intelligence" OR "AI" OR "machine learning") AND (educat* OR school* OR universit*)	37	9
Teachers and LLMs collaboration	Human-LLM co-teaching models, educators using LLMs	("teacher collaboration" OR "co-teaching") AND ("large language model*" OR "LLM" OR "ChatGPT") AND (educat* OR "higher education")	4	0

Appendix 2

Summary of the selected peer reviewed academic articles for the research

Authors	Year	Method	Contribution
Adams, N. E.	2015	Conceptual/Review	Explains Bloom's taxonomy of cognitive learning objectives.
Agüera y Arcas, B.	2022	Conceptual/Perspective	Explores whether large language models truly "understand" human language.
Ajani, O. A. et al.	2024	Review/Conceptual	Discusses the prospects and challenges of AI's emergence in higher education.
Al-Dokhny, A. et al.	2024	Empirical/Survey	Investigates how multimodal LLMs enhance performance benefits among higher education students.
Alfirević, N. et al.	2024	Conceptual/Case Study	Explores custom-trained LLMs as Open Educational Resources (OER).
Baltaci, K. et al.	2024	Conceptual/Review	Discusses AI integration in electrical engineering education as a paradigm shift.

Bećirović, S. et al.	2025	Empirical	Examines students' AI literacy and its effects on AI output quality, self-efficacy, and academic performance.
Belkina, M. et al.	2025	Systematic Review (Case Studies)	Provides a systematic review of case studies on implementing generative AI in higher education.
Bewersdorff, E. et al.	2025	Conceptual/Review	Discusses the transformative role of multimodal LLMs in science education.
Bucea-Manea-Tonis, R. et al.	2022	Empirical/Case Study	Explores AI potential for enhancing learning environments in higher education institutions in Romania and Serbia.
Cain, W.	2025	Case Study/Empirical	Examines using generative AI as cognitive partners for active online learning in a graduate course.
Campbell, L. O., & Cox, T. D.	2024	Review/Conceptual	Discusses the utilization of generative AI in higher education teaching and learning.
Castellanos-Reyes, D. et al.	2025	Empirical/Methodological	Explores leveraging GPT LLMs for automated content analysis of cognitive presence in online learning discussions.

Cecchini, S. et al.	2025	Review/Conceptual	Outlines challenges and opportunities of AI chatbots in education.
Choi, Y.	2025	Empirical/Content Analysis	Investigates the scientific validity of ChatGPT's responses in elementary science.
Claman, D., & Sezgin, E.	2024	Review/Conceptual	Discusses opportunities and challenges of LLMs and multimodal models in dental education.
Davar, N. F. et al.	2025	Review/Conceptual	Examines challenges and opportunities of AI chatbots in education (similar to Cecchini, distinct reference).
Davar, P. et al.	2025	Empirical/Experimental	Evaluates simulated teaching audio performance using retrieval-augmented generation (RAG) and LLMs.
Dibrell, D.	2024	Conceptual/Ethical Analysis	Discusses ethical considerations of AI use in writing pedagogy, including issues of authenticity.
Doğan, D. et al.	2025	Empirical	Studies the effectiveness of AI practices in pre-school social sciences education.

Eber, P. A., & Parker, T. S.	2007	Conceptual/Practical Guide	Provides guidance on assessing student learning by applying Bloom's Taxonomy.
Erkılıç, E., & Cifci, I.	2025	Conceptual/Ethical Analysis	Examines ethical issues and violations of using chatbots for academic writing and publishing.
Filippi, S., & Motyl, B.	2024	Systematic Review	Provides a systematic review of LLMs in engineering education with practical adoption suggestions.
Guizani, I. et al.	2025	Systematic Literature Review	Reviews issues and solutions for implementing LLMs in higher education.
Horikoshi, K. et al.	2025	Empirical/Developmental	Presents "Learning stream plot" for classroom activity visualization using daily learning log data.
Imran, M., & Almusharraf, N.	2024	Review/Conceptual	Reviews Google Gemini as a next-generation AI educational tool.
Issabek, A. et al.	2025	Empirical	Studies the effects of demographic factors on learner flow experience in gamified educational quizzes.

Keith, J. R. et al.	2025	Conceptual/Pedagogical	Discusses harnessing generative AI in chemical engineering education through critical engagement.
Klar, M.	2025	Mixed Methods Study	Explores K-12 students' perceptions, interaction patterns, and learning support with generative AI chatbots.
Lang, M. et al.	2024	Conceptual/Applied Research	Examines LLMs as "educational assistants" for automatic writing of teaching cases.
Lee, D., & Palmer, E.	2025	Systematic Review	Systematic review on prompt engineering in higher education to inform curricula.
Lucas, H. C. et al.	2024	Systematic Review	Systematic review of LLMs and their implications in medical education.
Maci, S. M., & Anesa, P.	2025	Conceptual/Review	Discusses the impact of AI on discourse analysis, highlighting challenges and opportunities.
Mezak Matijevic, M. et al.	2024	Conceptual/Legal Analysis	Proposes a socio-legal framework for governing LLM usage and application in education.

Nedungadi, P. et al.	2024	Conceptual/Review	Examines generative AI's role in achieving SDG 4 (quality education, equity).
Newton, P., & Xiromeritis, M.	2024	Scoping Review	Reviews ChatGPT's performance on multiple-choice examinations in higher education.
Nguyen Van, V. et al.	2024	Review/Conceptual	Provides an extensive analysis of LLM integration revolutionizing education.
Özçelik, N. P., & Yangın Ekşi, G.	2024	Case Study	Case study on ChatGPT's role as a learning assistant for cultivating writing skills.
Pack, A. et al.	2024	Empirical/Quantitative	Examines LLMs for automated essay scoring of English language learner writing, focusing on validity and reliability.
Parakh, D. et al.	2024	Conceptual/System Design	Proposes an adaptive personalized learning system with generative AI.
Praveena, T., & Anupama, K.	2025	Conceptual/Review	Discusses the transformative potential of AI in English language instruction.
Reicher, H. et al.	2025	Conceptual/Developmental	Proposes a generative AI-empowered digital tutor for higher education courses.

Ruiz-Rojas, L. I. et al.	2024	Empirical/Survey	Examines the adoption of generative AI tools for collaborative working and critical thinking in higher education.
Sarangi, S. et al.	2024	Empirical/Survey	Studies radiology postgraduate students' engagement with LLMs for educational purposes.
Shahzad, T. et al.	2025	Comprehensive Review	Provides a comprehensive review of LLM issues and solutions in learning environments.
Shen, M. et al.	2025	Qualitative Study	Qualitative study on integrating AI into nursing education, focusing on prompts, privacy, and personalized learning.
Sun, X. et al.	2025	Conceptual/Review	Discusses advancing total quality management in higher education through AI and big data analytics.
Uğraş, M. et al.	2024	Empirical/Qualitative	Explores primary school teachers' perspectives on ChatGPT-supported education.
Wang, J. et al.	2025	Empirical/Qualitative	Examines how young scholars cooperate with LLMs in academic tasks,

			focusing on individual differences and task complexities.
Wu, Q. et al.	2025	Systematic Review	Systematic review of generative AI in higher education, covering opportunities, challenges, and implications.
Xing, W. et al.	2024	Extensive Analysis/Review	Provides an extensive analysis of LLM integration revolutionizing education.
Yan, L. et al.	2023	Systematic Literature Review	Focuses on practical and ethical challenges of LLMs in education.
Yu, Y. F. et al.	2020	Conceptual/Framework Application	Discusses enhancing collaborative learning experiences using Laurillard's Conversational Framework.