Harnessing Generative AI (GenAI) for Automated Feedback in Higher Education: A Systematic Review

Sophia Soomin Lee and Robert L. Moore

University of Florida, USA

Abstract

In this systematic review, we synthesize ten empirical peer-reviewed articles published between 2019 and 2023 that used generative artificial intelligence (GenAI) for automated feedback in higher education. There are significant opportunities and challenges to integrate these tools effectively into learning environments as the demand for timely and personalized feedback grows. We examine the articles based on instructional contexts and system characteristics, identifying critical implementation possibilities for GenAI in automated feedback. Our findings reveal that GenAI provides diverse feedback across various contexts with multiple instructional purposes. GenAI systems can reduce instructor workload by automating routine grading and feedback tasks, allowing educators to focus on more complex teaching responsibilities with augmented capabilities. Additionally, these systems enhance communication, offer cognitive and emotional support, and improve accessibility by creating supportive, stress-free learning environments. Overall, implementing GenAI automated feedback systems improves educational outcomes and creates a more efficient and supportive learning environment for students and instructors. We conclude with future research directions to better integrate GenAI with human instruction by reconsidering instructors' roles, especially in providing feedback to create more effective educational experiences.

Recent advances in generative artificial intelligence (GenAI) have created new opportunities to explore how to integrate this technology into instructional practices. One area where GenAI has potential is through streamlining and scaling up instructor feedback. Particularly in online learning environments, providing personalized and formative feedback to learners can be challenging as course sizes increase without a comparable increase in instructors. In addition, there is a high demand from students for quality feedback (Moore et al., 2023; Mulliner & Tucker, 2017) and its utility in many aspects of learning, including increasing motivation (Koenka et al., 2021), promoting self-regulated learning (Lim et al., 2021), and enhancing students' academic achievement (Cai et al., 2023). Feedback enhances students' ability by guiding their learning process, and there are three central proposed mechanisms by which it does so (Shute, 2008). First, formative feedback signals a gap between learners' current performance and desired performance, thus reducing uncertainty about their current level. Second, formative feedback can help reduce learners' cognitive load via personalized feedback providing scaffolding. Lastly, feedback provides learners with helpful information for correction when it is specific enough to address learners' misunderstandings. Thus, a concerted effort has been made to automate feedback to increase the amount (Bälter et al., 2013) and enhance the quality and timing (Van der Kleij et al., 2015). Evidence for the effectiveness of automated feedback and high student satisfaction (Bayerlein, 2014) and the advancements of Natural Language Processing techniques has spurred interest in studying automated feedback (Deeva et al., 2021; Yan et al., 2024).

The Promises and Concerns of Using Generative AI for Feedback in Higher Education

Given the labor-intensiveness of providing quality and timely feedback, there has been consistent interest in integrating technology into instructional practices. AI has been at the forefront of the trend, with the acceleration further fueled by the impact of the COVID-19 pandemic. AI in Education (AIEd), which refers to the application of AI technologies to support and enhance educational practices, has garnered attention from educational researchers and educators around the world with unique expectations around using it to enhance learning, teaching, assessment, and administration (Chiu et al., 2023). The interest in AI was amplified with the introduction of ChatGPT in 2022, intensifying global curiosity and attention toward AI applications in education.

Generative AI uses deep learning models to produce human-like content, such as images and text, in response to complex prompts, including languages, instructions, and questions (Bozkurt & Bae, 2024; Lim et al., 2024). Examples include ChatGPT and Claude, which can create personalized and interactive learning experiences that enhance students' learning outcomes (Swindell et al., 2024). GenAI can function similarly to a personal assistant through language manipulation and generation capabilities (Bozkurt & Bae, 2024). While there are several categories of GenAI, the focus has been on text generation, especially in higher education. The term "Large Language Model" (LLM) refers to generative AI models that utilize extensive pre-trained text data to produce human-like text content (Yan et al., 2024). These systems leverage LLMs to comprehend and generate language, thereby playing a crucial role in educational settings (Bozkurt & Bae, 2024). Picciano (2024) explains that LLMs, such as

ChatGPT, are trained on vast datasets to predict and generate coherent, contextually appropriate language outputs. This capability facilitates various educational tasks, including essay writing and providing personalized feedback. GenAI harnesses Natural Language Processing (NLP) techniques to utilize large language models to understand natural language patterns and generate human-like text, facilitating automatization.

In the context of feedback, GenAI enables what could be conceptualized as AI-generated feedback (Banihashem et al., 2024; Farrokhnia et al., 2023), and it is expected to facilitate effective feedback practices (Katz et al., 2023). AI-generated feedback shares the features of timeliness and abundance with other automated feedback systems. However, using pre-trained language models such as BART and GPT-based models, AI-generated feedback is expected to provide more personalized and qualified feedback for more complex tasks (Dai et al., 2023) because it does not require specialized training to adapt to different tasks. Also, integrating AI-generated feedback creates opportunities for real-time collaboration (Yan et al., 2024) and interactive learning in online discussions (Lin et al., 2024), which often leads to increased student engagement in learning tasks (Michel-Villarreal et al., 2023; Smolansky et al., 2023). Providing feedback on essay writing (Chieu et al., 2023; Farrokhnia et al., 2023) and language learning (Barrot, 2023) has been especially prevalent.

There are naturally some concerns specific to AI-generated feedback, considering the heavy reliance on writing tasks, especially in higher education settings. Main concerns have been centered around accuracy, reliability, and plagiarism while using a GenAI tool, which could be problematic in the context of student learning and academic integrity (Michel-Villarreal et al., 2023; Moore et al., 2023; Swindell et al., 2024). Especially regarding accuracy and reliability, past research raised concerns over the possibility of students receiving inappropriate feedback, often leading to decreased tool use (Jasin et al., 2023). Also, with the accelerated adoption of AI-based feedback tools in online learning environments and decreasing human touch within the process, AI-generated feedback systems could lead to students' misuse or abuse of the system, especially when institutional guidelines are unclear. Ethical concerns about the potential reinforcement of biases and the impact on human agency and critical thinking skills have also been posed (Moore, 2019; Swindell et al., 2024).

The suspected problems embedded within the AI-generated feedback system could be more deeply understood from an AI-human interaction framework, which posits that AI could play very different roles depending on how it interacts with other components within an educational system (Xu & Ouyang, 2022). However, there is a lack of research on the roles that instructors play within classrooms interacting with GenAI. Also, our understanding of how and to what extent AI that generates text may enhance feedback practices and its capacity to improve feedback's effectiveness, timeliness, and personalization is somewhat limited.

Background

Due to the increasing emphasis on self-regulated and personalized learning coupled with the demanding nature of providing feedback, numerous initiatives have been aimed at creating automated feedback systems. As technology evolves rapidly, there have been significant advancements in how feedback is delivered and utilized within instructional settings (Conrad & Dabbagh, 2015; Elsayed & Cakir, 2023; Pishchukhina & Allen, 2021; Vittorini et al., 2021).

Cavalcanti et al. (2023) synthesized 63 articles that used automated feedback systems and evaluated the systems' effectiveness in increasing students' learning outcomes and unburdening instructors' workloads. Their findings were that automatic feedback might be as effective as manual feedback provided by instructors. They also found that while utilizing an automatic feedback approach could improve student outcomes and support instructors, there is still a limited understanding of how instructors integrate these tools into their classrooms. Also, this review is further limited in that most of the studies included were designed to provide feedback for a specific context, which is not generalizable to other contexts.

Deeva et al. (2021) presented a comprehensive classification framework for automated feedback, synthesizing 109 automated feedback systems. According to their framework, educational technologies ("architecture"), the educational settings in which they are applied ("educational context"), the properties of automated feedback they deliver ("feedback"), and the approaches for their design and evaluation ("evaluation") form the core of the systems. Also, they asserted that more attention should be paid to the students to provide more personalized feedback. While this comprehensive review offers a valuable tool for designing and understanding automated feedback systems, it emphasizes the importance of students providing more personalized learning experiences rather than implementing the systems within real-world settings involving different stakeholders.

Banihashem et al.'s (2022) higher-education-focused study provided insights into the potential of learning analytics' use for feedback regarding stakeholders, objectives, data used, and learning analytics methods in practice. The findings did not emphasize the automatic aspect of the learning-analytics-based feedback systems analyzed but provided insights into the underlying objectives of these automatic feedback systems in higher education settings including reflection, personalization, and expected outcomes such as enhanced academic performance, self-regulation, and motivation.

Several reviews focus on the use of Generative AI in education. Bahroun et al. (2023) comprehensively reviewed GenAI in educational settings, including its application in higher education contexts. They found that publications focused on integrating GenAI tools (Chaudhry et al., 2023) and students' acceptance and use of GenAI (Strzelecki, 2023). Some reviews attempt to understand the potential of using GenAI or large language models in educational contexts, although they do not specifically focus on feedback. For instance, Yan et al. (2024) outlined the current usage of LLMs in supporting educational tasks. They reviewed 109 articles and found that, by using many different models of LLMs, including BERT and GPT, important educational tasks such as providing feedback, generating content, and offering recommendations were being automated. While some emphasized teachers as agents of implementing automatic systems boosted by LLMs, they were considered more passive implementers than active designers of learning experiences.

Additionally, Kasneci et al. (2023) outline opportunities for adopting LLMs in education, emphasizing the affordance of LLMs in personalizing learning for individual students. They identified two major developments that made significant advancements in NLP: the use of transformer architecture and the underlying attention mechanism, which augmented past models to understand human language better, and the use of pre-training, which broadened the scope of

tasks that language models could address. However, they also raised concerns about the lack of interpretability and ethical considerations. While the authors mention the opportunities to use LLMs for assessment and evaluation by identifying students' difficulties and providing personalized feedback, empirical evidence was limited to specific cases that adopt these approaches.

In summary, past review literature on automatic feedback systems helps us understand the core components (Deeva et al., 2021) and essential applications of automatic feedback systems (Banihashem et al., 2022), especially in higher education. Furthermore, several reviews on Generative AI or LLMs offer essential implications for augmenting past automatic feedback systems (Kasneci et al., 2023; Yan, 2024), highlighting a research gap at the intersection of these fields.

Purpose

While several prior systematic reviews have focused on automated feedback systems or GenAI, we identified gaps in these reviews. First, previous studies did not focus on feedback using generative AI (Cavalcanti et al., 2023; Deeva et al., 2021). This is important because LLM-based GenAI differs from other automated systems in terms of their affordances (Kasneci et al., 2023; Yan et al., 2024). In addition, previous research has paid limited attention to implementing automated feedback in higher education settings, particularly the role of instructors in designing and implementing automatic feedback (Cavalcanti et al., 2023; Banihashem et al., 2022). Cavalcanti et al. (2023) explored the impact of automated feedback on teachers primarily from the workload perspective. Banihashem et al. (2022) focused on different uses of LA-enabled feedback systems but were limited in understanding instructors' active roles. AI can play very different roles depending on how it interacts with other components within an educational system (Moore et al., 2023; Xu & Ouyang, 2022), signaling a need to pair the emerging interest in GenAI with the interest in implementing automated feedback systems to frame our systematic review. Our systematic review answers the following questions:

1. What are the contexts of GenAI automated feedback systems?
2. What are the characteristics (instructional purpose, format, mechanism, technology) of GenAI automated feedback systems?
3. What are the possibilities for GenAI automated feedback systems in higher education?

Methods

We conducted a systematic review to identify peer-reviewed articles that address specific research questions, following the methodology outlined by Arksey and O'Malley (2005). Our approach adhered to the PRISMA principles (Liberati et al., 2009) for systematic reviews, and the steps we followed are summarized in Figure 1.

Figure 1

Article Selection Process. (Adapted from Liberati et al. (2009))
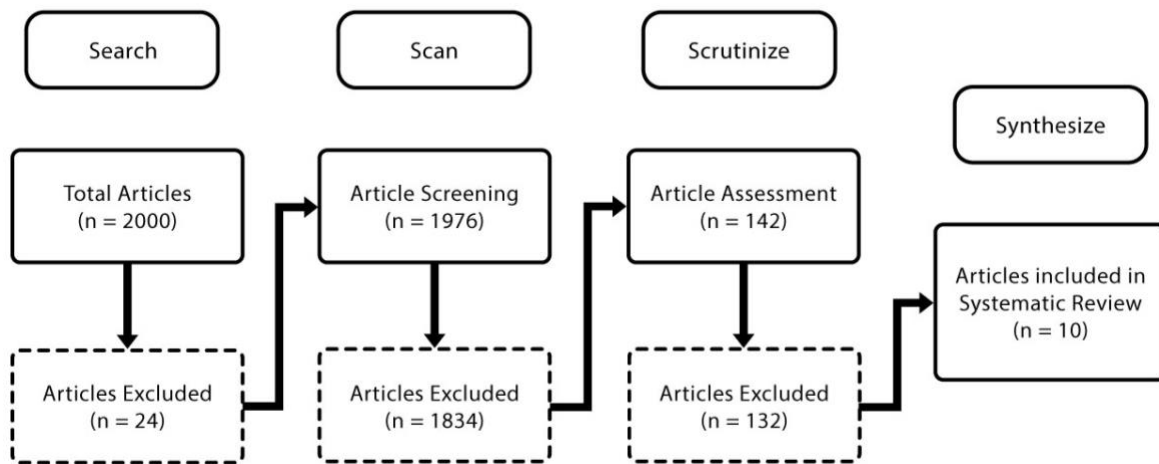
Our search was limited to peer-reviewed articles published in English between 2019 and 2023, utilizing the Academic Search Premier and Education Source databases. These databases are commonly employed for education-focused systematic reviews (Moore et al., 2024; Moore & Miller, 2022; Powers & Moore, 2021). Our search terms were: ("conversational agent*" OR "conversational AI" OR "generative ai" OR "GenAI" OR "chatbot*" OR "generative artificial intelligence" OR "language model*" OR "LLM*" OR "pre*trained language model" OR "natural language processing" OR "NLP") AND ("student*" or "learn*" or "teach*" or "instruct*") AND "feedback" OR "assess*" OR "eval*" AND ("higher education" OR "college" OR "university" or "grad*" OR "undergrad*"). The "*" is added for wildcard searches. Initially, we retrieved 2,000 studies, which were then reduced to 1,976 after removing duplicates for further examination.

Scan

In the next stage, we conducted a detailed scan of abstracts to refine our selection of articles, focusing on using GenAI for feedback. Dissertations and conference proceedings were omitted from consideration. This screening process reduced our initial pool to 142 relevant articles.

Scrutinize

We applied the inclusion and exclusion criteria outlined in Table 1 to assess the full texts of the selected articles. Each article underwent review by both authors, with any conflicts resolved by the first author to achieve consensus. Articles were excluded if they did not meet the specified criteria, and a reason for exclusion was provided in each case. This process led to the removal of 132 articles. The most common reasons for exclusion included a non-higher-education setting (39), not focused on feedback or assessment (34), not empirical (23), and no context for implementation (12). We assessed the articles' quality by using their ranking within SciMago. This technique of using only Q1 or Q2 journals has been used in prior systematic reviews (Bano et al., 2018; Moore & Blackmon, 2022; Moore et al., 2024).

Table 1

Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
Published between 2019-2023 in English	Article was a book chapter, conference proceeding, or dissertation
Peer-reviewed and empirical	Not published in a Q1 or Q2 journal*
Focused on higher education setting	No context for implementation

Source: SCIMago Journal and Country rankings (https://www.scimagojr.com/journalrank.php)

Synthesize

      After completing our screening process, we were left with ten articles. We synthesized the articles included around the three research questions and added additional citations where appropriate.

Results

      Our search focused on empirical peer-reviewed journal articles published between 2019-2023 that analyzed the features of GenAI in providing automated feedback in higher education settings. The articles provided helpful insight into how GenAI is already being used to automate feedback and insights for future implementations.

      RQ1: What are the contexts of GenAI automated feedback systems?
      Our included articles focused on writing tasks in language, business, creative thinking, and STEM contexts (Table 2).

Table 2

Contexts for Included Articles

Context	Included Articles
Language learning	Escalante et al., 2023; J. Li et al., 2023
Writing	X. Li et al., 2023; Wambsganss et al., 2022
Creative thinking	Hu et al., 2023; Neo, 2022
STEM	Hobert & Berens, 2023; Jasin et al., 2023; Lee et al., 2022; Memmert et al., 2023

      Two articles focused on language learning (Escalante et al., 2023; J. Li et al., 2023) and used ChatGPT as a personalized feedback tool for learners. For Escalante et al. (2023), the focus was on learners who were learning English, and ChatGPT was used as a complementary

instructional tool to provide immediate and personalized feedback to learners. J. Li et al. (2023) focused on scaffolding learner support for non-native speakers trying to learn Chinese. Both studies found that the use of ChatGPT gave learners more autonomy in their learning process.

X. Li et al (2023) and Wambsganss et al. (2022) focused on academic writing and how ChatGPT could be used for writing assessments. The X. Li study focused on an undergraduate-level course and used a collection of academic papers to develop an assessment mechanism to expedite learners' feedback. Wambsganss' study focused on improving writing in business courses. Specifically, they explored whether social comparison nudging—a digital nudge that references how other learners have performed on similar tasks—coupled with automated feedback would improve the demonstration of persuasion in a short-form business pitch. Ultimately, they found higher argumentation skills in learners who received automated feedback and social comparison nudges (Wambsganss et al., 2022). An interesting finding from their study is that combining both elements contributed to higher argumentation skills.

Two studies explored how a chatbot could improve learners' writing skills (Hu et al., 2023; Neo, 2022). Hu et al. (2023) integrated a chatbot to provide students with on-demand writing support and feedback, creating personalized assistance. In Neo (2022), the chatbot was integrated to provide real-time support to improve writing confidence. Both studies highlight the effectiveness of chatbots in offering immediate, personalized feedback that fosters self-directed learning and improves writing self-efficacy and proficiency. Lastly, four studies examined chatbots in STEM contexts (Hobert & Berens, 2023; Jasin et al., 2023; Lee et al., 2022; Memmert et al., 2023). The purpose of these chatbots was to engage learners and revise their understanding as if they were engaging with a peer student in courses to learn statistics (Hobert & Berens, 2023), chemistry (Jasin et al., 2023), and public health (Lee et al., 2022).

RQ2: What are the characteristics (instructional purpose, format, mechanism, technology) of GenAI automated feedback systems?

Instructional Purpose

The included studies used GenAI automated feedback systems for personalized learning and various instructional objectives, including collaborative problem-solving, self-regulated learning, and motivation and engagement (Table 3).

Table 3

Instructional Purpose of Included Articles

Purpose	Included Articles
Personalized learning	Escalante et al., 2023; Hobert & Berens, 2023; Hu et al., 2023; Jasin et al., 2023; Memmert et al., 2023; X. Li et al., 2023
Collaborative problem solving	X. Li et al., 2023; Memmert et al., 2023
Self-regulated learning	Hu et al., 2023; Lee et al., 2022 ; Jasin et al., 2023
Enhancing motivation and engagement	Neo, 2022; Jasin et al., 2023; J. Li et al., 2023; Wambsganss et al., 2022

The most significant number of studies found GenAI-enabled feedback promising especially for personalized learning (Escalante et al, 2023; Hobert & Berens, 2023; Hu et al., 2023; Jasin et al., 2023; Memmert et al., 2023; X. Li et al., 2023). In Hu et al.'s (2023) study, students' learning data was collected and analyzed to automatically provide students with appropriate learning support, including suggestions and resources. In Hobert & Berens' (2023) study, the developed digital tutor integrated all learning scenarios for an entire course, providing more comprehensive and individualized student feedback opportunities. This was particularly evident for classes focusing on language learning in which students often demonstrate varying proficiency levels and demand tailored feedback (Escalante et al., 2023; X. Li et al, 2023). The concept of "scaffolding" (Memmert et al., 2023; Jasin et al., 2023) was used, and the term "feedback" was often used interchangeably.

The specific pedagogical purpose of adopting the feedback system within a course varies, ranging from facilitating students' collaborative problem-solving (Memmert et al., 2023; X. Li et al., 2023) to supporting students' self-regulated learning (Hu et al., 2023; Lee et al., 2022; Jasin et al., 2023). Memmert et al. (2023) used the system mainly to provide students with soft scaffolding or problem-specific support to facilitate conceptual design development with Design Science Research (DSR). AI was also used to provide feedback to warn the students about failing an online course (Hu et al., 2023) and to help students clarify their understanding during the review process (Jasin et al., 2023; Lee et al., 2022).

Other pedagogical approaches to enhance motivation and engagement, especially in online learning environments, were used via an enhanced amount of interaction (Neo, 2022), communication immediacy (Jasin et al., 2023), empathetic support (Jasin et al., 2023), providing cultural resources and a comfortable environment (J. Li et al., 2023), and social comparison nudging (Wambsganss et al., 2022).

Lastly, in addition to educational needs deriving from pedagogical needs, administrative needs to leverage GenAI to alleviate assessment burden were mentioned prevalently. For instance, X. Li et al. (2023) used the system to provide instructors with information about the composition of knowledge in students' unstructured writing. This enabled instructors to offer more objective feedback, regardless of their writing assessment experience. It was often used in language learning (Escalante et al., 2023; J. Li et al, 2023) and academic writing classes (X. Li et al., 2023; Wambsganss et al., 2022) to serve the purpose where a high level of feedback is indispensable, often leading to teacher burnout.

Format

Feedback often consists of varied formats to serve the purpose of the feedback. While specific details of the feedback provided within each study varied, they could be broadly categorized into three categories: information, course resources, and analysis of student's work, with few of the systems being included in more than one category (Table 4).

Table 4

Feedback Format of Included Articles

Category	Included Articles
Information	Hobert & Berens, 2023; Hu et al., 2023; Jasin et al., 2023; J. Li et al., 2023; Lee et al., 2022; Neo, 2022
Course resources	Hobert & Berens, 2023; Jasin et al., 2023
Analysis of student work	Escalante et al., 2023; J. Li et al., 2023; Memmert et al., 2023; Wambsganss et al., 2022; X. Li et al., 2023

Sales (1993) defined feedback as information presented to the learner, and especially in an online setting, "information" is how feedback is often conceptualized. Of the ten studies, six could be conceptualized as such. Depending on the components or the specific functions of each feedback system represented in each study, the information provided in each study varied. Most common was the information that was provided upon students' requests to help clarify their understanding via hints, reminders, definitions, or explanations (Hobert & Berens, 2023; Jasin et al., 2023, Lee et al., 2022; Neo, 2022). In other cases, more metacognitive information was provided. For instance, Hu et al. (2023) provided information about the pass rates and instructions for self-regulated learning strategies. Two articles provided feedback as suggestions for course resources (Hobert & Berens, 2023; Jasin et al., 2023). Resources recommended included video resources (Jasin et al., 2023) and files and quizzes (Hobert & Berens, 2023). In both cases, course resources were provided in addition to other content, including information.

In five of the articles, the analysis of students' work was provided as feedback. In these cases, the contents of the feedback were often conceptualized as "insights" or "suggestions," implying that the feedback was created based on input from the student that resulted in the

assessment of the submitted work. For instance, in the case of X. Li et al. (2023), Wambsganss et al. (2022), and Escalante et al. (2023), students' writing was evaluated via the tool, creating insights to revise and improve the work. In the case of Memmert et al. (2023), predefined templates were used to generate suggestions to facilitate the design science research process. Lastly, in J. Li et al. (2023), students were asked to use a commercialized GenAI tool to work on their writing tasks, and it could be understood that they were getting both helpful information and analysis since the instructors could not capture how they used the tool. The instructors were not able to access student's accounts and see how exactly students interacted with ChatGPT, thus we had to assume that the student's were getting the information and analysis. An improvement in the research design would have required the students to submit their chatlogs as part of the assessment for that assignment.

Mechanisms

According to Shute (2008), there are three ways to explain how feedback works: signaling a gap, reducing cognitive load, and correcting information. Our analysis showed that these mechanisms could account for how the AI-generated feedback system works (Table 5).

Table 5

Feedback Mechanism of Included Studies

Category	Included Articles
Signaling gap	Escalante et al., 2023; Hu et al., 2023; Wambsganss et al., 2022
Cognitive load	Hobert & Berens, 2023; Lee et al., 2022; Memmert at al., 2023; Neo, 2022
Correcting information	Escalante et al., 2023; Jasin et al, 2023; J. Li et al., 2023; X. Li et al., 2023

One mechanism for feedback systems to be effective is by signaling a gap. For example, in the study by Hu et al. (2023), the system provides information on the probability of students passing the course, along with diagnosis and suggestions to support their self-regulated learning based on students' learning progress. In the study by Wambsganss et al. (2022), the writing was scored based on readability and argumentativeness, highlighting the gap between the ideal and current states, while students were provided with comments on six areas of writing in the study by Escalante et al. (2023).

Feedback could enhance learning by reducing students' cognitive load. Several ways to reduce cognitive load by providing feedback were reflected in the research. One way was promptly providing the information so students could keep learning without interruption. In the cases of Hobert & Berens (2023) and Lee et al. (2022), information needed to continue learning was provided by answering students' questions and providing resources. Another way was by enabling access to advanced knowledge. For instance, Memmert et al. (2023) made suggestions

that would otherwise have been hard to access and integrate into their submitted work. Also, in Neo (2022), depending on students' proficiency level, they were guided to review the contents or advance knowledge. If the evaluation results indicated that students achieved a certain level of proficiency in an area of knowledge, the bot would guide students to advanced knowledge.

Feedback could also enhance students' learning by correcting the most common misunderstandings. J. Li et al. (2023) and Escalante et al. (2023) used the feedback system to provide students with information on their writing, specifically by identifying errors and suggesting the correct usage. In addition, there were some cases where the system seemed to support students' learning via all the mechanisms mentioned above, mainly when AI chatbots were used (Jasin et al., 2023; X. Li et al., 2023).

Technology

Different underlying techniques and tools were employed in each study to embody different instructional or pedagogical designs and uses of the feedback. While all the AI tools in the ten studies used Natural Language Processing techniques and machine learning to generate texts, the specific tools used to implement generative AI varied. They could be classified as commercialized tools (i.e., ChatGPT, Chatlayer) or specially designed tools to serve the particular purpose of the studies. According to the findings of this study, Generative AI embodied an automatic feedback system leveraging varying affordances. Unlike instructor feedback where students can receive feedback only when the instructors have the time to provide it, with an automatic feedback system, students can get feedback whenever they want. This implementation of on-demand feedback leveraged various technological affordances, which we have organized into into three categories: context-generalizing, course-integrating, and interpretive scoring (Table 6).

Table 6

Articles by Technology Use

Category	Included Articles
Context-generalizing	Escalante et al., 2023; Lee et al., 2022; Memmert et al., 2023; X. Li et al., 2023
Course-integrating	Hobert & Berens, 2023; Hu et al., 2023; Jasin et al. 2023; Neo, 2022
Interpretive scoring	J. Li. et al, 2023; Wambsganss et al., 2022

The first category of context-generalizable feedback includes tools developed to provide feedback that transcends specific domains and utilizes GenAI to offer suggestions without requiring problem-specific data. Because learning activities in higher education courses often underscore the importance of problem-solving and the domain for the activity is frequently up to students' interests to enhance their motivation, it is frequently challenging for teachers to provide constructive feedback based on in-depth content knowledge. Examples include ChatGPT, T-

Bert, and Chatlayer. Escalante et al. (2023) leveraged GenAI to provide writing assistance for English learners, and Lee et al. (2022) integrated GenAI to improve after-class review in a public health course. X. Li et al. (2023) developed a real-time knowledge-aware academic writing assessment tool, and Memmert et al. (2023) used GenAI to offer learners real-time problem-solving suggestions. These studies demonstrated leveraging GenAI to generate adaptable feedback without needing training data. This broadens the potential for implementation while supporting diverse learner backgrounds and specific needs, creating personalized responses.

The second category includes tools implemented to analyze student learning data and generate comprehensive feedback. Hobert and Berens (2023) developed a digital tutor chatbot to support students in a large-scale formal learning setting, providing continuous guidance throughout the learning process and addressing the need for individualized feedback in extensive lecture courses. Similarly, Hu et al. (2023) implemented an intelligent tutoring robot (ITR) through robotic process automation (RPA) to create an early warning system that offers comprehensive learning support and timely feedback within a course. Jasin et al. (2023) focused on synchronous communication in an online chemistry course. Neo (2022) integrated a chatbot to assist students with their writing. The applications—from large enrollment courses to specific contexts—demonstrate the power and utility of GenAI in providing personalized, course-specific feedback to improve student outcomes.

The third category focuses on tools that provide scores for submitted work to provide score-attached feedback. Examples include interpretable AI, knowledge-aware strategies, and Named Entity Recognition (NER), all used to give rationales and formative suggestions. J. Li et al. (2023) developed an academic writing assessment tool using knowledge-aware strategies and NER to offer rationales and formative suggestions based on scores. Wambsganss et al. (2022) used Interpretable AI to automatically score persuasive writing assignments, providing feedback and social comparison nudging to improve argumentation skills. While both studies focused on writing assessment, J. Li et al. emphasized academic writing with detailed interpretive feedback, whereas Wambsganss et al. incorporated social comparison to enhance persuasive writing quality.

RQ3: What are the possibilities for implementing GenAI automated feedback systems in higher education?

The articles included in our study highlighted several possibilities for implementing a GenAI automated feedback system. We identified categories of transforming instructor roles, enhancing educational dialogues, and cognitive and emotional assistance (Table 7).

Table 7

Opportunities Presented in Included Articles

Category	Included Articles

Transforming instructor roles	Hobert & Berens, 2023; Jasin et al., 2023; J. Li et al., 2023; Memmert et al., 2023; Wambsganss et al., 2022; X. Li et al., 2023
Enhancing educational dialogues	Hobert & Berens, 2023; Hu et al., 2023; Jasin et al., 2023; J. Li et al., 2023
Cognitive and emotional assistance	Escalante et al., 2023; Lee et al., 2022; J. Li et al., 2023; Neo, 2022

GenAI automated feedback systems can potentially transform instructors' roles by reducing instructors' load and augmenting instructors' expertise. The systems can significantly reduce instructor load by automating routine grading and feedback tasks, allowing instructors to focus on more critical and complex aspects of teaching (J. Li et al., 2023; Wambsganss et al., 2022). These systems provide personalized feedback for students, handling basic queries and analyzing their work, enabling instructors to address more detailed and challenging cases. J. Li et al. (2023) and Wambsganss et al. (2022) demonstrated that these systems effectively manage repetitive grading and formative feedback, allowing instructors to engage in higher-level assessments and more personalized student interactions. Furthermore, Hobert & Berens (2023) and Jasin et al. (2023) emphasized that AI tools facilitate better communication and provide real-time feedback, further alleviating the burden on instructors by managing students' basic queries and offering comprehensive support throughout the learning process.

GenAI systems also enable instructors to extend their capabilities beyond traditional limits. For instance, GenAI automated feedback systems augment instructors' capabilities by complementing the specific knowledge needed for constructive learning. Memmert et al. (2023) highlighted the potential of AI to offer broad problem-solving suggestions, extending beyond the expertise of individual instructors. X. Li et al. (2023) emphasized the development of a real-time writing assessment tool that provides generalized feedback applicable across various contexts. Overall, implementing GenAI automated feedback systems enhances teaching efficiency and effectiveness by allowing instructors to dedicate more time and effort to complex and impactful educational tasks. GenAI automated feedback systems can significantly improve communication among different stakeholders within the system, including instructors, teaching assistants, and students (Hobert & Berens, 2023; Hu et al., 2023; Jasin et al., 2023). By reducing instructors' workload, these systems allow educators to provide a more tailored feedback system. In language learning classrooms, AI chatbots provide students with additional opportunities to practice dialogue, acting as knowledgeable entities that complement the instructor's role (J. Li et al., 2023). Hobert & Berens (2023) and Jasin et al. (2023) also emphasized that AI tools facilitate better communication and provide real-time feedback. Hobert and Berens (2023) showcased how digital tutor chatbots enhance communication in large-scale courses by facilitating interactions between students, teaching assistants, and instructors, thereby reducing instructor workload. Similarly, Jasin et al. (2023) demonstrated the effectiveness of chatbots for synchronous communication in an online chemistry course, providing real-time, course-specific feedback. Hu

et al. (2023) developed an early warning system using GenAI to provide timely feedback and support, enhancing communication between students and the automated system. Overall, implementing GenAI systems enhances educational communication and expands the range of support available to students, leading to more effective and comprehensive learning experiences.

Lastly**,** GenAI automated feedback systems offer substantial cognitive and emotional benefits to students. Lee et al. (2022) demonstrated that GenAI systems enhance learning outcomes by offering tailored cognitive support. They provide individualized feedback supporting cognitive and metacognitive development, as evidenced by improved learning outcomes in public health courses (Lee et al., 2022). Escalante et al. (2023) found that AI feedback significantly improves language skills by rephrasing responses to ensure student understanding. These systems act as scaffolding for students' learning, adjusting responses to ensure complete comprehension, which enhances language skills (Escalante et al., 2023). Additionally, GenAI supports self-regulated learning by allowing students to reflect on their progress and plan their next steps, fostering a stress-free environment that encourages engagement and help-seeking behaviors (Neo, 2022).

From an emotional support perspective, GenAI automated feedback systems significantly enhance students' accessibility and interaction, particularly those who might not typically seek help (Neo, 2022; Lee et al., 2022). Neo (2022) highlighted the emotional benefits of GenAI systems, noting that they create a comfortable, non-pressured learning atmosphere that encourages students to seek assistance and engage more deeply with their studies. These systems create a comfortable atmosphere that reduces pressure and stress, making students more inclined to seek and benefit from feedback (J. Li et al., 2023). They also facilitate effective communication among students, instructors, and teaching assistants, ensuring timely and relevant feedback (Hobert & Berens, 2023). Neo (2022) and Lee et al. (2022) found that AI chatbots make it easier for students to seek assistance by alleviating stress and creating a supportive environment. J. Li et al. (2023) observed that these systems encourage greater engagement with feedback due to the comfortable learning atmosphere they foster. Hobert & Berens (2023) emphasized that GenAI tools enhance communication among all educational stakeholders, leading to more efficient and effective feedback processes. Overall, GenAI automated feedback systems not only make feedback more accessible and less stressful for students but also streamline communication, making the educational process more effective and supportive for all involved.

Discussion

This section offers insights about designing and implementing GenAI automated feedback systems.

Design of GenAI Automated Feedback Systems

This study reviewed the design of the AI-generated feedback system from various perspectives. As a result, it was found that there are differences in the design depending on the

course's context and the system's characteristics (instructional purpose, content, mechanism, technology) compared to the extant automatic feedback system.

This study addressed instructional purpose in depth compared to previous systematic reviews that analyzed the purpose mostly from functional perspectives. This is because, unlike any automatic feedback system, AI-generated feedback systems are often not explicitly created for educational purposes. Thus, instructors are frequently required to make proactive decisions about how to adopt the system into the course. This study identified various instructional purposes ranging from facilitation of self-regulated learning to collaborative problem-solving, demonstrating the need for specifying the purpose of using the system.

According to Shute (2008), the purpose of feedback can be broadly divided into "directive" and "facilitative." Deeva et al. (2021) presented four purposes of automatic feedback: "corrective," "suggestive," "informative," and "motivational". They asserted that feedback can be of multiple types. Our distinction of purposes for AI-generated feedback is coherent with these distinctions but not entirely the same. This is because various GenAI tools in this study could handle diverse functions based on the learner's request, compared to the task-specific automatic feedback systems reviewed in previous review papers.

Nevertheless, we confirmed that GenAI-automated feedback still follows the feedback mechanism. Specifically, GenAI-automated feedback is also designed and utilized to improve the learner's experience through signaling gaps, reducing cognitive load, and correcting information like previous feedback systems (e.g., Shute, 2008). Furthermore, previous studies did not specify the format in which feedback was presented. This is because, although there may be differences in the specific content of the feedback provided by a single automatic feedback system, there were no differences in the forms. However, according to our analysis, the form of feedback provided by the AI-generated feedback system using GenAI could be broadly categorized into three types: information, analysis, and course resources.

Finally, compared to the methods and techniques used in the existing automatic feedback, the GenAI-automated feedback systems in this study could still be considered valid in the classification of the three feedback generation models mentioned earlier: data-driven model, expert-driven model, and mixed model (Deeva et al., 2021). Providing feedback through prompt engineering could be understood as a mixed model.

<u>Implementation of AI-Generated Feedback Systems</u>

There have been ongoing efforts to automate feedback, and with the advent of GPT models, it has become possible to transcend domains in feedback. Considering the role of the instructor in a learner-centered paradigm, it is vital to understand how AI-generated feedback systems augment their role. A framework delineating the role of AI in educational systems from a Complex Adaptive System perspective was proposed by Xu & Ouyang (2022). According to the authors, AI can play various roles with students and instructors as the leading learning and teaching agents, being seen as 1) a new subject, 2) a direct mediator, and 3) a supplementary

assistant. Our study is coherent with the framework in that GenAI automated feedback systems analyzed in our study also perform these three roles as outlined by Xu & Ouyang (2022). For instance, in some studies, feedback systems act as a dialogue partner, serving as a new subject (e.g., Lee et al., 2022), while in others, they function as direct mediators bridging instructor, teaching assistant, and student roles (e.g., Hobert & Berens, 2023). In many studies, these feedback systems are also viewed as supplementary assistants, answering students' questions or providing emotional support (e.g., Jasin et al., 2023).

However, the studies covered by Xu & Ouyang (2022) and the ten studies analyzed in this research lack a discussion of the role of the instructor. The analysis in this study indirectly or directly mentions the role of the instructor within GenAI-automated feedback systems beyond the typical roles such as designing learning materials or giving lectures. Specifically, the following points were noted:

- Supporting students to use the tool properly (e.g., helping students recognize the limit ations of the tools, preventing unintended usage or overreliance on the system)
- Providing higher-level or detailed feedback (e.g., offering clarification for confused st udents)
- Making final decisions or feedback (e.g., providing a comprehensive evaluation of the learning process and results)
- Prompt engineering (e.g., providing prompts that help students receive better feedbac k)
- Facilitating students' learning (e.g., stimulating and guiding students to further learni ng)

First, guidance from the instructor for using GenAI tools was most frequently mentioned (e.g. Memmert et al., 2023; J. Li et al., 2023). This guidance could be provided on an individual level by the instructor or on an institutional level in the form of guidelines to adhere to. An interesting point is that while many studies designate providing higher-level feedback, which automatic feedback fails to offer as a role of the instructor (e.g. Hu et al., 2023; X. Li et al., 2023), it is unclear what specific feedback they should provide. Considering that some studies assigned the role of complementing the instructor's expertise to AI-generated feedback systems (Memmert et al., 2023), it calls for discussion of what feedback the human instructor should provide that the system cannot. For example, one such role could be making "final decisions" when students cannot receive clear answers through automatic feedback and are confused (e.g. Jasin et al., 2023; J. Li et al., 2023). Alternatively, the instructor could provide long-term feedback that complements students' instant interaction with the AI-generated feedback system. Prompt engineering, as mentioned by J. Li et al. (2023) and Escalante et al. (2023) could also be one aspect of the instructor's role, highlighting collaboration between the instructor and GenAI to improve the quality of feedback provided.

Lastly, Hu et al. (2023) present the facilitator role of the instructor more concretely, emphasizing motivating learners to go beyond course contents by stimulating and guiding them towards further learning. This underscores the need for instructors to play a more active role in

AI-generated feedback systems, where learners have more autonomy than in conventional, domain-specific, and structured systems. Because AI-generated feedback systems offer learners access to more specialized information with lower cognitive loads, instructors can more effectively facilitate students' active learning.

Limitations

Systematic reviews provide insight into the published literature that meets the requirements established by the authors. Our systematic review focused on empirical articles and used a quality assessment filter based on the journal publication. These limitations may have excluded articles that could have added insight to the systematic review. We encourage researchers to use our systematic review to identify future research directions and consider conference proceedings and dissertations, which we excluded from this review. Additionally, generative AI is an emerging concept, so we anticipate that there will be more research in the coming years that will allow a more complete picture to emerge of the landscape of GenAI. Our review did not distinguish between large and small language models (e.g., GPT vs BERT). As the field evolves and research matures in this space, we anticipate that future systematic reviews will want to consider these distinctions.

Conclusion

While artificial intelligence, specifically GenAI, is not a new concept, it is a rapidly evolving research area. In this systematic review, we synthesized the current literature on GenAI in educational contexts, focusing on articles that explored how GenAI can be used for automated feedback.

The first research question found that the contexts for GenAI automated feedback systems were not only limited to writing and language learning but also included creative thinking and STEM learning. Several aspects of the feedback systems were addressed for the second research question. First, the most widely used instructional rationale used to build the system was personalizing students' learning, which aligns with the findings from Banihashem et al. (2022). In addition, a GenAI-based feedback system was used to facilitate collaborative problem-solving (Memmert et al., 2023; X. Li et al., 2023), self-regulated learning (Hu et al., 2023; Lee et al., 2022; Jasin et al., 2023), and enhance motivation and engagement (Jasin et al., 2023; J. Li et al., 2023; Neo, 2022; Wambsganss et al., 2022). In addition to instructional purposes, assessment burden was commonly mentioned (Escalante et al., 2023; Wambsganss et al., 2022; X. Li et al., 2023), which aligns with the purpose of automated feedback systems reflected in the research by Cavalcanti et al. (2023). Next, the feedback format could be categorized into information, course resources, and analysis of student work, which often overlapped. While past research on (automated) feedback often conceptualized feedback as "information" (Sales, 1993) and focused on properties of feedback such as learner control (Deeva et al., 2021), we found that GenAI enabled providing many different formats or types of feedback possible within one system, expanding the boundary of feedback. In terms of feedback mechanisms, all three were still viable. While corrective feedback was the most common in automated feedback systems (Deeva et al., 2021), it was found that feedback systems utilizing

GenAI were not only limited to correcting information but were also used to signal gaps and reduce cognitive load in this study. Specifically, providing information that students need to continue learning without being interrupted (Hobert & Berens, 2023; Lee et al., 2022; Memmert et al., 2023) or to advance learning (Neo, 2022) was found to be heavily dependent on the capacity of GenAI to create text easily. Lastly, three categories of affordances stemming from different underlying technologies were identified: AI chatbots, learning analytics-based systems, and automatic scoring-based systems. For the last research question, in addition to unburdening instructors, augmenting instructors' capabilities and providing emotional support and cognitive support for the students were identified as potential possibilities for integrating feedback systems utilizing GenAI.

GenAI has many potential applications within educational contexts, and our interest in focusing on feedback is just one part of a complex puzzle of research angles. Ultimately, the goal of integrating GenAI is to scaffold learners better. We anticipate that this systematic review will be the first of many to explore using GenAI and automated feedback.

Declarations

References

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

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

Banihashem, S. K., Noroozi, O., van Ginkel, S., Macfadyen, L. P., & Biemans, H. J. (2022). A systematic review of the role of learning analytics in enhancing feedback practices in higher education. *Educational Research Review*, 37, Article 100489. https://doi.org/10.1016/j.edurev.2022.100489

Banihashem, S. K., Kerman, N. T., Noroozi, O., Moon, J., & Drachsler, H. (2024). Feedback sources in essay writing: peer-generated or AI-generated feedback? *International Journal of Educational Technology in Higher Education*, *21*(1), 23. https://doi.org/10.1186/s41239-024-00455-4

Bano, M., Zowghi, D., Kearney, M., Schuck, S., & Aubusson, P. (2018). Mobile learning for science and mathematics school education: A systematic review of empirical evidence. *Computers & Education*, *121*, 30–58. https://doi.org/10.1016/j.compedu.2018.02.006

Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, 100745. https://doi.org/10.1016/j.asw.2023.100745

Bayerlein, L. (2014). Students' feedback preferences: how do students react to timely and automatically generated assessment feedback? *Assessment & Evaluation in Higher Education, 39*(8), 916-931. https://doi.org/10.1080/02602938.2013.870531

Bälter, O., Enström, E., & Klingenberg, B. (2013). The effect of short formative diagnostic web quizzes with minimal feedback. *Computers & Education, 60*(1), 234-242. https://doi.org/10.1016/j.compedu.2012.08.014

Bozkurt, A., & Bae, H. (2024). May the force be with you Jedi: Balancing the light and dark sides of generative AI in the educational landscape. *Online Learning Journal, 28*(2), 1-6. https://doi.org/10.24059/olj.v28i2.4563

Cai, Z., Gui, Y., Mao, P., Wang, Z., Hao, X., Fan, X., & Tai, R. H. (2023). The effect of feedback on academic achievement in technology-rich learning environments (TREs): A meta-analytic review. *Educational Research Review*, *39*, 100521. https://doi.org/10.1016/j.edurev.2023.100521

Cavalcanti, A. P., Barbosa, A., Carvalho, R., Freitas, F., Tsai, Y. S., Gašević, D., & Mello, R. F. (2021). Automatic feedback in online learning environments: A systematic literature

review. *Computers and Education: Artificial Intelligence*, 2, 100027. https://doi.org/10.1016/j.caeai.2021.100027

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

Chiu, T. K., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *4*, 100118. https://doi.org/10.1016/j.caeai.2022.100118

Conrad, S. S., & Dabbagh, N. (2015). Examining the factors that influence how instructors provide feedback in online learning environments. *International Journal of Online Pedagogy and Course Design*, *5*(4), 47–66. https://doi.org/10.4018/IJOPCD.2015100104

Deeva, G., Bogdanova, D., Serral, E., Snoeck, M., & De Weerdt, J. (2021). A review of automated feedback systems for learners: Classification framework, challenges and opportunities. *Computers & Education, 162*, 104094. https://doi.org/10.1016/j.compedu.2020.104094

Elsayed, S., & Cakir, D. (2023). Implementation of assessment and feedback in higher education. *Acta Pedagogia Asiana*, *2*(1), 34–42. https://doi.org/10.53623/apga.v2i1.170

Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, *20*(1), 57. https://doi.org/10.1186/s41239-023-00425-2

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 1–15. https://doi.org/10.1080/14703297.2023.2195846

Hobert, S., & Berens, F. (2023). Developing a digital tutor as an intermediary between students, teaching assistants, and lecturers. *Educational Technology Research and Development*, 1-22. https://doi.org/10.1007/s11423-023-10293-2

Hu, Y. H., Fu, J. S., & Yeh, H. C. (2023). Developing an early-warning system through robotic process automation: Are intelligent tutoring robots as effective as human teachers? *Interactive Learning Environments*, 1-14. https://doi.org/10.1080/10494820.2022.2160467

Jasin, J., Ng, H. T., Atmosukarto, I., Iyer, P., Osman, F., Wong, P. Y. K., Pua C. Y., & Cheow, W. S. (2023). The implementation of chatbot-mediated immediacy for synchronous communication in an online chemistry course. *Education and Information Technologies*, *28*(8), 10665-10690. https://doi.org/10.1007/s10639-023-11602-1

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

Katz, A., Wei, S., Nanda, G., Brinton, C., & Ohland, M. (2023). Exploring the efficacy of ChatGPT in analyzing student teamwork feedback with an existing taxonomy. *arXiv preprint arXiv:2305.11882*. https://doi.org/10.48550/arXiv.2305.11882

Koenka, A. C., Linnenbrink-Garcia, L., Moshontz, H., Atkinson, K. M., Sanchez, C. E., & Cooper, H. (2021). A meta-analysis on the impact of grades and comments on academic motivation and achievement: A case for written feedback. *Educational Psychology*, *41*(7), 922-947. https://doi.org/10.1080/01443410.2019.1659939

Li, J., Ren, X., Jiang, X., & Chen, C. H. (2023). Exploring the use of ChatGPT in Chinese language classrooms. *International Journal of Chinese Language Teaching*, *4*(3). https://doi.org/10.46451/ijclt.20230303

Li, X., Ouyang, F., Liu, J., Wei, C., & Chen, W. (2023). Examining the effects of a real-time, knowledge-aware tool for academic writing assessment. *Journal of Educational Computing Research*, *61*(6), 1143-1174. https://doi.org/10.1177/07356331221136889

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., Clarke, M., Devereaux, P J, Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *Annals of Internal Medicine, 151(*4). https://doi.org/10.1371/journal.pmed.1000100

Lim, L. A., Gasevic, D., Matcha, W., Ahmad Uzir, N. A., & Dawson, S. (2021, April). Impact of learning analytics feedback on self-regulated learning: Triangulating behavioural logs with students' recall. In *LAK21: 11th International Learning Analytics and Knowledge Conference* (pp. 364-374). https://doi.org/10.1145/3448139.3448174

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

Lin, X., Luterbach, K., Gregory, K. H., & Sconyers, S. E. (2024). A case study investigating the utilization of ChatGPT in online discussions. *Online Learning Journal, 28*(2), 124-149. https://doi.org/10.24059/olj.v28i2.4407

Memmert, L., Tavanapour, N., & Bittner, E. (2023). Learning by doing: Educators' perspective on an illustrative tool for AI-generated scaffolding for students in conceptualizing design

science research studies. *Journal of Information Systems Education*, *34*(3), 279-292. https://aisel.aisnet.org/jise/vol34/iss3/3

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

Moore, R. L. (2019). The role of data analytics in education: Possiblities and limitations. In B. Khan, J. R. Corbeil, & M. E. Corbeil (Eds.), *Responsible analytics and data mining in education: Global perspectives on quality, support, and decision making* (pp. 101–118). Routledge. https://doi.org/10.4324/9780203728703-8

Moore, R. L., & Blackmon, S. J. (2022). From the learner's perspective: A systematic review of MOOC learner experiences (2008–2021). *Computers & Education*, *190*, 1–15. https://doi.org/10.1016/j.compedu.2022.104596

Moore, R. L., Hwang, W., & Moses, J. D. (2024). A systematic review of mobile-based microlearning in adult learner contexts. *Educational Technology & Society*, *27*(1), 137–146. https://doi.org/10.30191/ETS.202401_27(1).SP02

Moore, R. L., Jiang, S., & Abramowitz, B. (2023). What would the matrix do? A systematic review of K-12 AI learning contexts and learner-interface interactions. *Journal of Research on Technology in Education*, *55*(1), 7–20. https://doi.org/10.1080/15391523.2022.2148785

Moore, R. L., & Miller, C. N. (2022). Fostering cognitive presence in online courses: A systematic review (2008-2020). *Online Learning*, *26*(1), 130–149. https://doi.org/10.24059/olj.v26i1.3071

Mulliner, E., & Tucker, M. (2017). Feedback on feedback practice: Perceptions of students and academics. *Assessment & Evaluation in Higher Education, 42*(2), 266-288. https://doi.org/10.1080/02602938.2015.1103365

Neo, M. (2022). The Merlin Project: Malaysian students' acceptance of an AI chatbot in their learning process. *Turkish Online Journal of Distance Education*, *23*(3), 31-48. https://doi.org/10.17718/tojde.1137122

Picciano, A. G. (2024). Graduate teacher education students use and evaluate ChatGPT as an essay-writing tool. *Online Learning*, 28(2), 1-20. https://olj.onlinelearningconsortium.org/index.php/olj/index

Pishchukhina, O., & Allen, A. (2021). Supporting learning in large classes: Online formative assessment and automated feedback. *2021 30th Annual Conference of the European*

Association for Education in Electrical and Information Engineering (EAEEIE), 1–4. https://doi.org/10.1109/EAEEIE50507.2021.9530953

Powers, F. E., & Moore, R. L. (2021). When failure is an option: A scoping review of failure states in game-based learning. *TechTrends*, *65*(4), 615–625. https://doi.org/10.1007/s11528-021-00606-8

Sales, G. C. (1993). Adapted and adaptive feedback in technology-based instruction. *Interactive instruction and feedback*, *14*, 159-175.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, *78*(1), 153-189. https://doi.org/10.3102/0034654307313795

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

Strzelecki, A. (2023). To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interactive learning environments*, 1-14. https://doi.org/10.1080/10494820.2023.2209881

Swindell, A., Greeley, L., Farag, A., & Verdone, B. (2024). Against artificial education: Towards an ethical framework for generative artificial intelligence (GenAI) use in education. *Online Learning Journal, 28*(2), 7-28. https://doi.org/10.24059/olj.v28i2.4438

Van der Kleij, F. M., Feskens, R. C., & Eggen, T. J. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Review of Educational Research, 85*(4), 475–511. https://doi.org/10.3102/0034654314564881

Vittorini, P., Menini, S., & Tonelli, S. (2021). An AI-Based system for formative and summative assessment in data science courses. *International Journal of Artificial Intelligence in Education*, *31*(2), 159–185. https://doi.org/10.1007/s40593-020-00230-2

Wambsganss, T., Janson, A., & Leimeister, J. M. (2022). Enhancing argumentative writing with automated feedback and social comparison nudging. *Computers & Education*, *191*, 104644. https://doi.org/10.1016/j.compedu.2022.104644

Xu, W., & Ouyang, F. (2022). A systematic review of AI role in the educational system based on a proposed conceptual framework. *Education and Information Technologies*, *27*(3), 4195–4223. https://doi.org/10.1007/s10639-021-10774-y

Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2024). Practical and ethical challenges of large language models in education: A

systematic scoping review. *British Journal of Educational Technology*, *55*(1), 90–112. https://doi.org/10.1111/bjet.13370