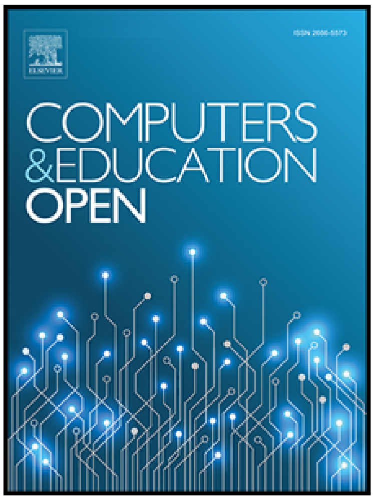
Journal Pre-proof



Unraveling the mechanisms and effectiveness of AI-assisted feedback in education: A systematic literature review

Shen Ba , Lan Yang , Zi Yan , Chee Kit Looi , Dragan Gašević

PII: S2666-5573(25)00043-6
DOI: https://doi.org/10.1016/j.caeo.2025.100284
Reference: CAEO 100284

To appear in: *Computers and Education Open*

Received date: 25 December 2024
Revised date: 5 August 2025
Accepted date: 23 August 2025

Highlights

- Our systematic review (N=129) shows AIFB research surged post-2018, driven by large language models.
- AIFB effectively supports multiple feedback domains (task, process, self-regulation, self) and complexity levels.
- Evidence indicates generally positive impacts on learner perceptions, actions, and outcomes.
- Limited transparency about AI algorithms calls for improved reporting.
- Our new conceptual model synthesizes these insights and guides future AIFB research.

Title:

Unraveling the mechanisms and effectiveness of AI-assisted feedback in education: A systematic literature review

Author names and affiliations:

Shen Ba[a], Email: bas@eduhk.hk

Lan Yang[a], Email: yanglan@eduhk.hk

Zi Yan[a], Email: zyan@eduhk.hk

Chee Kit Looi[a]: cklooi@eduhk.hk

Dragan Gašević[b]: dragan.gasevic@monash.edu

[a]Department of Curriculum and Instruction, The Education University of Hong Kong, Hong Kong SAR, China.

[b]Faculty of Information Technology, Monash University, Clayton, VIC, Australia.

Corresponding author:

Lan Yang, Block D, Department of Curriculum and Instruction, The Education University of Hong Kong, 10 Lo Ping Road, Hong Kong SAR, China. Email: yanglan@eduhk.hk

Abstract

Rapid advancements in Artificial Intelligence (AI) have prompted growing interest in leveraging AI for educational feedback. Yet, the centrality of the learner in this process is often overshadowed by technological excitement, and a broad understanding of AI-assisted feedback (AIFB) in education remains evolving. To address this gap, we conducted a systematic review of 129 peer-reviewed journal articles (2014–2023) based on widely used AI-related search terms to examine how AI, especially generative AI, supports feedback mechanisms and influences learner perceptions, actions, and outcomes. Our analysis identified a sharp rise in AIFB research after 2018, driven by modern large language models. We found that AI tools flexibly cater to multiple feedback foci (task, process, self-regulation, and self) and complexity levels (basic, intermediate, and elaborated). Our findings demonstrate that AIFB can effectively enhance targeted learning outcomes. By employing a transparent and field-aligned methodology, we synthesized recent advances and offers actionable insights for both research and practice. While the focus on widely recognized AI-related search terms ensures strong comparability and relevance, some specialized subfields (e.g., Automated Writing Evaluation), are less prominent in this synthesis. The study also highlights the ongoing need for clearer reporting of underlying AI algorithms. Building on these findings, we propose an original conceptual model that synthesizes current progress and offers a roadmap for future explorations. By illuminating the affordances and constraints of AIFB, we highlight the necessity for transparent methodological reporting and underscores the importance of integrating pedagogical and technological insights to promote meaningful, learner-centered feedback.

Keywords: Artificial intelligence, feedback, education, systematic literature review

1. Introduction

Artificial intelligence (AI) can be defined as "systems that display intelligent behavior by analyzing their environment and taking actions (with some degree of autonomy) to achieve specific goals" (Sheikh et al., 2023, p. 16). Generative AI (GAI) includes specific AI models, such as large language models or generative adversarial networks, that generate new outputs (text, images, or other formats) resembling the data on which they were trained (Sengar et al., 2024; Goodfellow et al., 2014). The rapid evolution of GAI has sparked a wide array of imaginations, expectations, and concerns regarding its prospective impact on education (Hooda et al., 2022; Hwang & Chen, 2023).

On the one hand, GAI is promising for providing learners with immediate feedback, overcoming the restrictions of time, space, and human workload (Ba et al., 2025; Hwang & Chen, 2023). This advancement constitutes a significant step toward educational goals including lifelong learning (Belcadhi, 2016), seamless learning (Wong & Looi, 2011), and equitable access (Jemeli & Fakandu, 2019). On the other hand, there are valid concerns about the accuracy of machine-generated content, the method of information delivery, and how learners interact with GAI (Chan & Hu, 2023; Yang et al., 2024). Even when the content is accurate, the instructional design for content delivery is crucial to stimulate cognitive engagement and foster critical thinking (Ayanwale et al., 2024; Mayordomo et al., 2022). The delivery approach can also affect whether learners utilize GAI tools responsibly and ethically (Hwang & Chen, 2023). Thus, the key challenge is to integrate AI-driven tools while ensuring that learner needs, ethical considerations, and pedagogical value remain paramount.

In this study, we adopt a clearer definition of feedback to underscore its role in shaping learner understanding and performance. Rather than referring to feedback exclusively as a "process," we emphasize that feedback is information presented to learners concerning the accuracy, quality, or effectiveness of their work. This information can correct misunderstandings (Wambsganss et al., 2022), offer actionable advice for skill development

(Jiang et al., 2023), and bolster motivation and confidence (Ouyang et al., 2023), all of which likely lead to a more efficient learning process (Morris et al., 2021).

1.1 Educational feedback

Hattie and Timperley (2007) describe four distinct focus levels of educational feedback. Task-level feedback addresses the correctness or quality of learners' work on a specific task (e.g., highlighting grammatical errors). Process-level feedback considers the strategies and methods learners adopt, often suggesting method adjustments. Self-regulation feedback targets the enhancement of metacognitive and self-regulation skills. Lastly, self-level feedback concerns learners' confidence and self-esteem rather than a specific task.

Educational feedback can also vary in complexity (Fu et al., 2024; Shute, 2008). Drawing on Shute's (2008) taxonomy, feedback can be synthesized into three levels: basic, intermediate, and advanced. *Basic feedback* offers minimal information, such as verifying correctness or providing the right answer without explanation. It supports procedural accuracy but lacks instructional depth. *Intermediate feedback* includes forms like error flagging, attribute isolation, and topic-contingent feedback, which guide learners by drawing attention to task-relevant features without diagnosing specific misconceptions. *Advanced feedback* is highly elaborated and personalized, including response-contingent explanations, strategic hints, and informative tutoring, which actively scaffold learners' reasoning and address underlying misconceptions. This continuum reflects increasing cognitive support and is useful for aligning feedback design with learner needs and instructional goals (see Shute, 2008, pp. 160-161 for details; Based on Shute's work, see also Appendix C for our synthesized descriptions of feedback levels). By synthesizing these dimensions of feedback focus and complexity with AI capabilities, our review aims to clarify how GAI can be harnessed to enrich learning outcomes while maintaining a balanced, learner-centered pedagogical approach.

1.2 AI-assisted feedback

In this review, we define AI-assisted feedback (AIFB) as the use of AI technologies to analyze learning environments, tasks, and learner data in order to provide supportive actions that help learners progress toward specific learning goals. This perspective builds on extensive research into feedback theories, methods, and outcomes, as well as the integration of AI technologies in educational settings. Early intelligent tutoring systems offered feedback based on preprogrammed responses and rigid algorithms, thus delivering only limited adaptivity and personalization (Mousavinasab et al., 2021; Nye et al., 2014). Nonetheless, these pioneering systems laid the groundwork for more advanced approaches utilizing generative AI (GAI). For instance, Hooda et al. (2022) examined AI and machine learning algorithms used to support higher education assessment and feedback, focusing on algorithmic performance and its link to learning outcomes. Similarly, Deeva et al. (2021) proposed a classification framework for systems designed to provide automated feedback, highlighting predominantly technical aspects. Moving into writing-specific domains, Fu et al. (2024) and Shi and Aryadoust (2024) reviewed how automatic writing evaluation can shape writing performance and learners' perceptions of feedback. In digital learning contexts, Bimba et al. (2017), Cavalcanti et al. (2021), and Maier and Klotz (2022) collated information on the generation of automatic, personalized feedback, examining its various purposes and reported effects. Despite these contributions, there is still a lack of cohesive synthesis on how GAI can be strategically integrated into feedback practices in a way that prioritizes both learner-centered and pedagogically sound approaches.

Previous reviews have also largely focused on the technical performance of AI algorithms (Deeva et al., 2021; Hooda et al., 2022) or have been confined to specific learning contexts (Bimba et al., 2017; Cavalcanti et al., 2021; Fu et al., 2024; Law, 2024; Maier & Klotz, 2022; Shi & Aryadoust, 2024). While some studies have explored aspects of learner perceptions or performance (Shi & Aryadoust, 2024), an integrative analysis that synthesizes how AI-assisted feedback (AIFB) is associated with learner perceptions, actions, and outcomes across diverse AI technologies remains limited. Addressing this gap is important for developing a more nuanced and evidence-informed understanding of AIFB's effectiveness from the learner's perspective.

Given the mixed expectations, uncertainties, and concerns surrounding GAI, a retrospective review summarizing recent applications of AI in educational feedback is both timely and necessary. Importantly, the true measure of any technology's value in education lies not in its technical sophistication but in its impact on learners—the ultimate beneficiaries. A review focused on learner experiences offers significant potential to inform and improve the design, implementation, and evaluation of AIFB tools.

2. Aims and objectives

The primary aim of this review is to systematically investigate how AI, particularly GAI, can enhance educational feedback in ways that maintain a strong learner-centered focus. This objective is informed by the *Self-System Model of Motivational Development* (SSMMD) (Connell & Wellborn, 1991; Skinner et al., 2008), which provides a conceptual foundation for examining the interaction between AI-assisted learning environments and learner outcomes. Central to this model are the principles that learning environments directly influence learners' perceptions of competence, autonomy, and relatedness. These perceptions subsequently guide learners' actions and behaviors, ultimately impacting educational outcomes such as performance, knowledge acquisition, and skill development.

Drawing on the SSMMD, this review pursues four interrelated objectives that frame our research questions:

1) To contextualize the landscape of AIFB research by analyzing the bibliometric and methodological profiles of existing studies. This addresses how AIFB functions as an educational environment—a core tenet of the SSMMD.

2) To identify and classify the mechanisms of AIFB across macro (e.g., data sources and AI models), meso (focus levels of feedback), and micro (complexity levels) dimensions. These mechanisms represent the structural components of the AI learning environment that influence learner perceptions.

3) To examine the influence of AIFB on learners' perceptions, actions, and outcomes, in alignment with the SSMMD's proposition that learner behavior is shaped by perceived competence, autonomy, and relatedness within a given environment.

4) To explore the interconnections between AIFB mechanisms and their educational effectiveness, thereby modeling the dynamic and reciprocal pathways outlined in the SSMMD.

These objectives guide a theory-informed synthesis of the empirical literature, aiming to inform educational researchers and practitioners about the nuanced roles of AIFB in shaping learner experiences and outcomes.

3. Methods

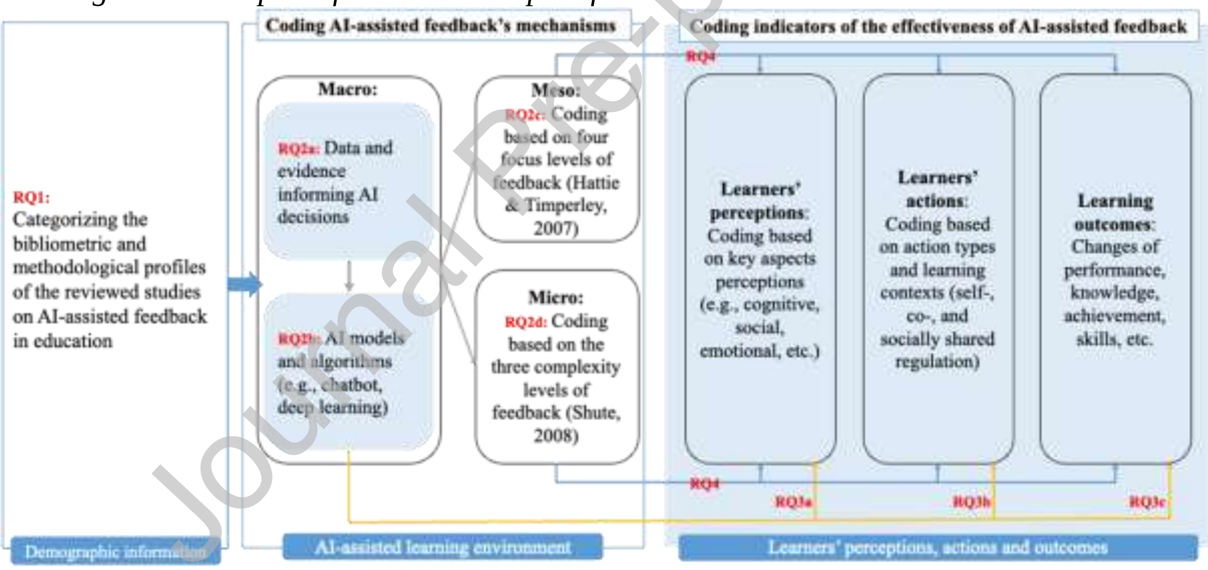
3.1 Research Questions and Rationale

Figure 1 operationalizes the SSMMD as a guiding analytical framework by delineating two interrelated components: the mechanisms of AIFB and the indicators of its effectiveness. On the left, the framework outlines three levels of feedback mechanisms—macro, meso, and micro—coded in alignment with research questions RQ2a to RQ2d. The macro level focuses on the data and algorithms that drive AI decisions (e.g., deep learning and chatbot models), while the meso level draws on established pedagogical feedback constructs (e.g., Hattie & Timperley's four levels), and the micro level refers to task-specific feedback complexity (Shute, 2008). This layered structure reflects the complex characteristics of AIFB as an environmental input that interacts with learners, consistent with the SSMMD's environmental influence dimension.

The right side of Figure 1 aligns with RQ3a to RQ3c and RQ4, mapping how AIFB mechanisms influence learners' perceptions, actions, and outcomes. The framework traces these learner-centered responses that could be shaped by perceptions such as cognitive and emotional responses, actions such as self-regulation and collaboration, and learning outcomes including achievement and skills development. This progression captures the pathways articulated in the SSMMD: from environmental input (AIFB mechanisms) through internal processes (learner perceptions) to behavioral engagement and ultimate academic results. Hence, Figure 1 serves as an integrative conceptual tool to explore how diverse AI feedback mechanisms intersect with learner processes in a manner that is theoretically coherent with the SSMMD model.

Figure 1

An integrated conceptual framework adapted from the SSMMD



The SSMMD model delineates the following core tenets:

1) Environmental influence: The characteristics of educational environments (e.g., AI-assisted feedback) significantly shape learner perceptions.

2) Learner perceptions: Learners develop cognitive, social, and emotional perceptions based on their interactions within the learning environment.

3) Learners' actions: Learners' perceptions directly influence their engagement and regulatory actions, including self-, co-, and socially-shared regulation.

4) Learning outcomes: Learners' actions are the primary determinants of educational outcomes, such as performance improvement and skill development.

Building on these tenets, our review is guided by four overarching research questions (RQs) designed to systematically explore how AI-assisted feedback (AIFB) functions as an environmental input and influences various learner outcomes.

RQ1: What are the bibliometric and methodological profiles of the reviewed studies on AIFB in education?

RQ2: What are the mechanisms of AIFB in education?

- RQ2a: What types of data and evidence inform AIFB?
- RQ2b: What types of AI models and algorithms are employed?
- RQ2c: What are the focus levels of AIFB?
- RQ2d What are the complexity levels of AIFB?

RQ3: What are the effects of AIFB on learners?

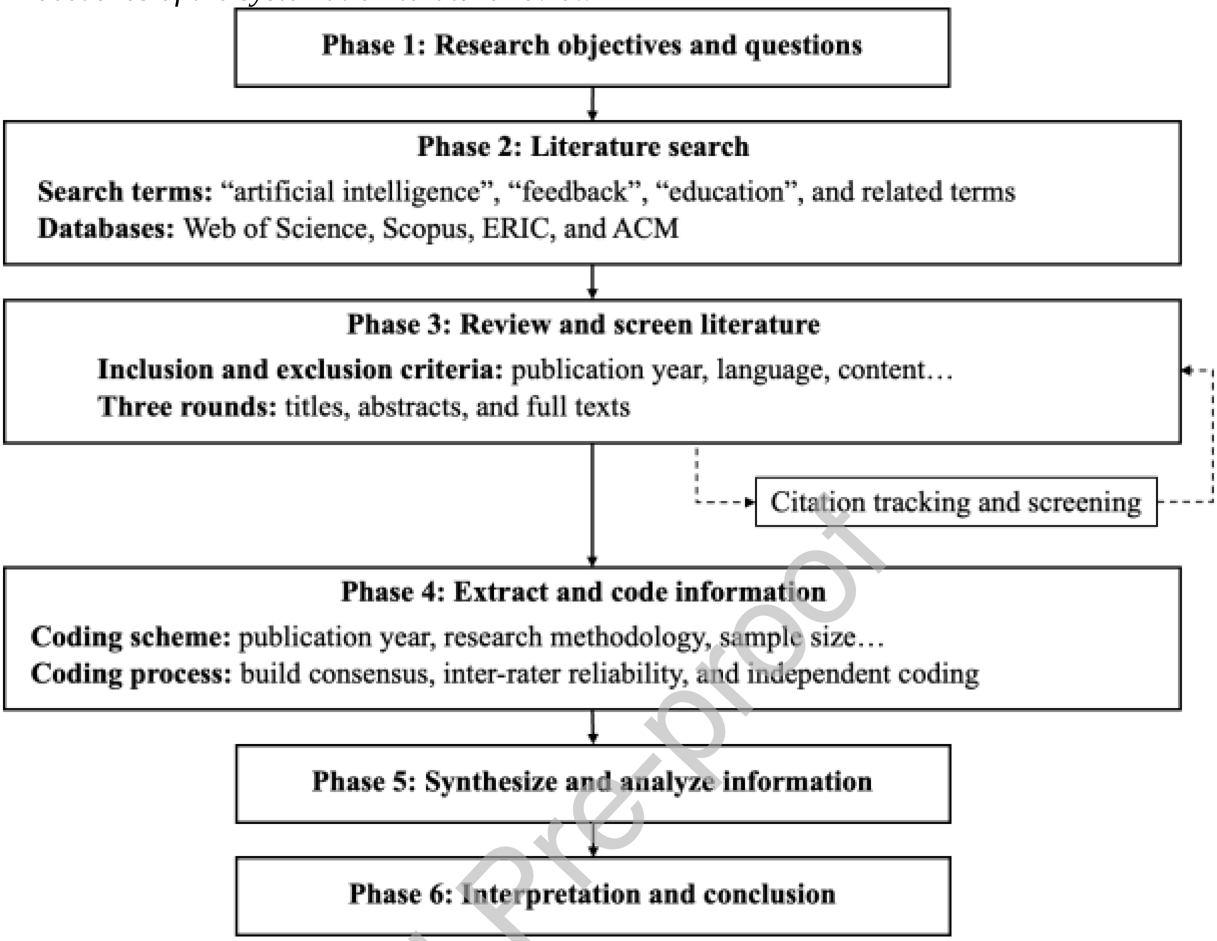
- RQ3a: How is AIFB associated with learner perceptions?
- RQ3b: How is AIFB associated with learner actions?
- RQ3c: How is AIFB associated with learning outcomes?

RQ4: What are the associations between AIFB's mechanisms and effects?

3.2 Research design and procedures

Following the *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) statement (Page et al., 2021), this review involved six phases (Figure 2): 1) defining the research objectives and questions; 2) conducting focused searches in major databases; 3) screening articles based on pre-determined inclusion and exclusion criteria and conducting citation tracking; 4) extracting and coding data; 5) synthesizing and analyzing data; and 6) deriving interpretations and conclusions.

Figure 2
Procedures of the systematic literature review



3.3 Inclusion and exclusion criteria

The publication time was set to range from January 2014 to December 2023, aligning with the introduction of generative adversarial networks (GANs), a milestone widely recognized as the foundation of modern generative AI research (Chiu et al., 2023; Goodfellow et al., 2014). Articles were required to be written in English, published in peer-reviewed journals, and to explicitly employ artificial intelligence (AI) technologies, including but not limited to generative adversarial networks (GANs), large language models (e.g., GPT, BERT), machine learning, deep learning, neural networks, intelligent tutors, expert systems, and chatbots, to assist feedback processes in educational contexts. The selection of these search terms was rigorously informed by established practices in the field, as evidenced by recent, high-impact systematic reviews in AI and education (see Cavalcanti et al., 2021; González-Calatayud et al., 2021; Hopcan et al., 2023; Salas-Pilco et al., 2022; Sun et al., 2023; Zhai et al., 2021;

Zhao, 2024). This approach ensured methodological consistency and comparability with the international literature, enabling our findings to be situated within the broader trajectory of AI-in-education research synthesis.

For this review, AI-assisted feedback was defined broadly, encompassing systems that demonstrate intelligent behavior by interpreting data from their environment and taking autonomous or semi-autonomous actions to support pedagogical goals (Sheikh et al., 2023). To operationalize this definition, the search strategy targeted studies that clearly described the use of AI-related technologies in feedback applications, using keywords directly associated with these models in educational contexts (see Appendix A) (Roll & Wylie, 2016; Zawacki-Richter et al., 2019;).

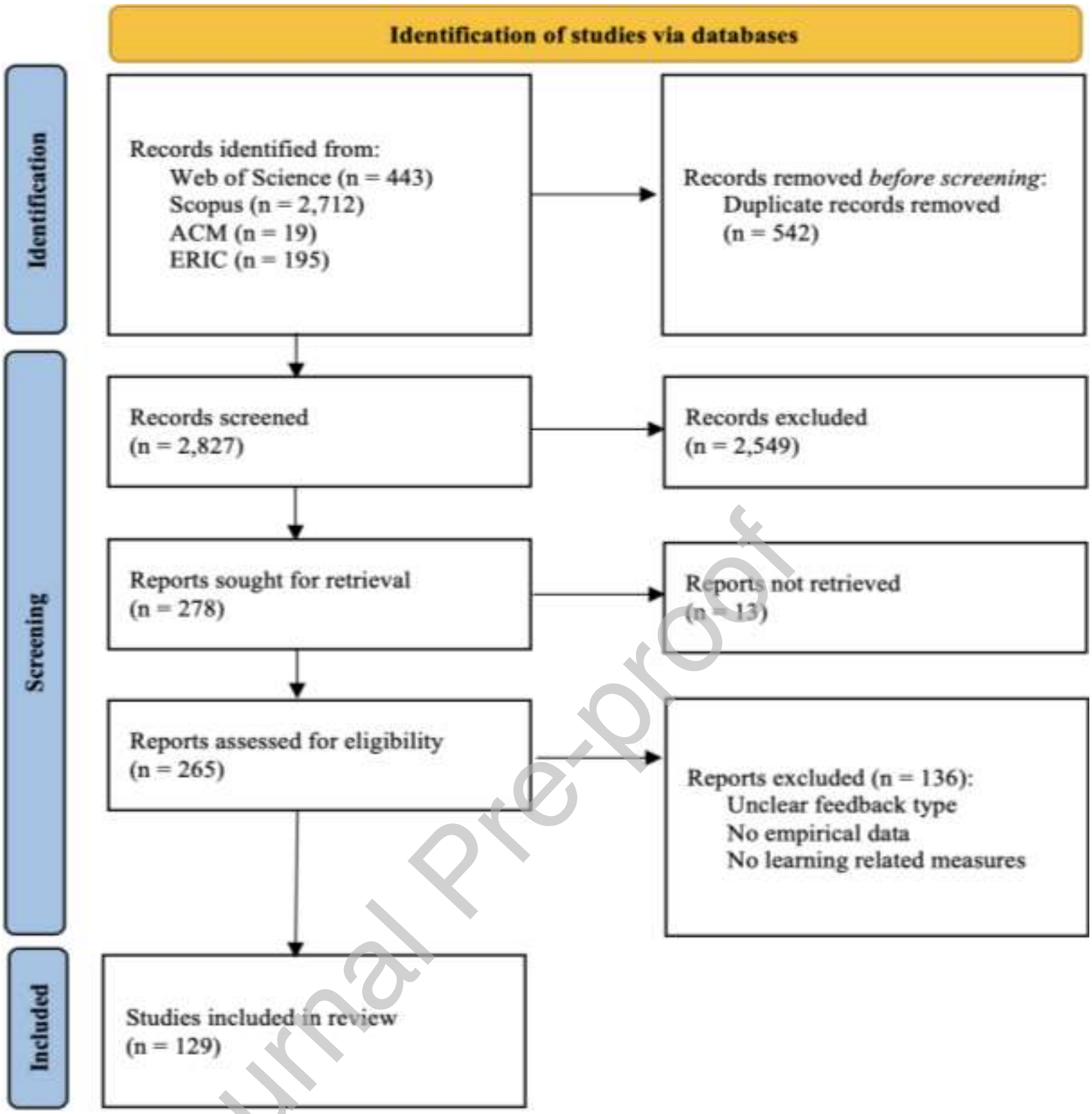
Consequently, studies that focused solely on rule-based, automated, or traditional AI feedback systems (e.g., automated writing evaluation tools or expert systems) were excluded unless they incorporated the AI-related technologies specified above. While this search strategy ensured consistency and replicability, it may have excluded otherwise relevant studies that did not explicitly reference AI-related terminology in their titles, abstracts, or keywords, consistent with the scope and limitations observed in other leading systematic reviews (e.g., Feng, 2022; Stevenson & Phakiti, 2014). Additionally, the articles had to have empirical data on learners' perceptions, actions, and/or outcomes in response to AI-assisted feedback. Articles that did not meet the abovementioned criteria were excluded.

3.4 Searching and screening strategies

This review searched Web of Science, Scopus, ERIC, and ACM, ensuring broad yet methodologically consistent coverage of journal articles across multiple fields (e.g., education, educational technology, computer science, and engineering), guided by the search strategies of established systematic reviews in AI in education (Cavalcanti et al., 2021; Deeva et al., 2021). The search terms were drawn from three categories (“artificial intelligence,” “feedback,” and “education”) as detailed in Appendix A, and were refined to reflect widely accepted AI-related terminology and to emphasize alignment with the inclusion and exclusion criteria described in Section 3.3. Boolean operators were used to combine these terms for systematic and targeted searches in each database. Metadata, including titles, abstracts, and authors, was extracted and stored for all retrieved records.

Following the initial search, this review employed a three-round screening process, systematically filtering the articles by title, abstract, and full text, according to the clearly defined inclusion and exclusion criteria in 3.3. This process (see also Figure 3) was specifically designed to identify studies explicitly describing the use of AI technologies in feedback applications as operationalized in the review, and to exclude those limited to rule-based or non-generative AI systems. To enhance reliability and reduce subjectivity, one researcher performed the screening primarily, while another researcher randomly checked 20% of the articles. Their decisions were consistent for over 90% of the checked articles, and agreement was reached on the rest after discussion. Afterward, citations associated with the remaining articles were examined to minimize the risk of overlooking relevant articles.

Figure 3
PRISMA flow diagram



3.5 Extracting and coding information

Two researchers with considerable experience in AI and education conducted data extraction and coding, following the scheme detailed in Table 1. The researchers utilized established frameworks for certain data elements, including feedback focus categories from Hattie and Timperley (2007) and feedback complexity classifications from Shute (2008). Classification of feedback focus and complexity was based either on direct examination of the feedback content or on the authors' descriptions within the reviewed studies when explicit feedback content was not provided. Meanwhile, codes for perceptions, actions, and outcomes emerged as variables in the reviewed studies and were formed into different categories inductively.

The coders first collectively worked on 5% of the articles to build a baseline consensus. Subsequently, they independently coded a randomly selected 20% of the data, after which, the inter-rater reliability was calculated using Cohen's kappa (Cohen, 1960). Specifically, Cohen's kappa values were calculated for each coding category. The values ranged from 0.81 to 0.87 across categories (see Table 1), indicating substantial to almost perfect agreement. Once an acceptable kappa score was achieved and disagreements were resolved, the coders divided the remaining data equally for independent coding.

Table 1
Data extraction and coding scheme

	Extracted data	Description/Example codes	Cohen's kappa
RQ 1	Publication year	-	-
	Research methodology (Creswell, 2009)	1) quantitative, 2) qualitative, 3) mixed methods.	0.87
	Sample size (Fu et al., 2024)	1) small (size < 30), 2) medium (30 <= size < 50), 3) large (50 <= size <100), 4) very large (size >= 100).	0.82
	Education level	1) elementary, 2) secondary, 3) undergraduate, 4) postgraduate, 5) others.	0.85
	Educational context	e.g., language, science, and personal development.	-
RQ 2	Data and evidence	Types of learning data collected and input to AI models.	0.84
	AI model and algorithm	e.g., chatbot, deep learning, and large language models.	0.81
	Focus level (Hattie & Timperley, 2007)	1) task-level feedback, 2) process-level feedback, 3) self-regulation level feedback, 4) self-level feedback.	0.86
	Complexity level (Shute, 2008)	1) basic feedback, 2) intermediate feedback, 3) elaborated feedback.	0.83
RQ 3	Learner perception	e.g., perceived cognitive development, satisfaction, and technology usability.	0.85
	Learner action	e.g., idea generation, number of revisions, and social interaction.	0.84
	Learning outcome	e.g., test performance, writing skills, and artistic skills.	0.82

3.6 Data analysis

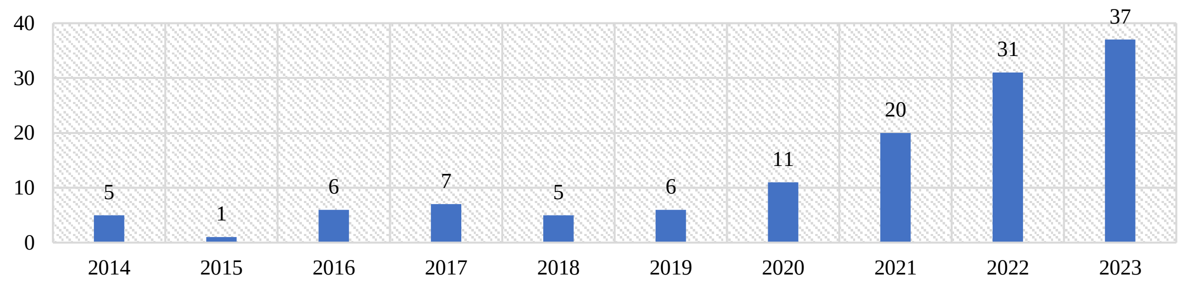
After completing the coding, descriptive and thematic analyses were conducted to answer each research question. For RQ1, bibliometric and methodological profiles were analyzed by categorizing studies based on publication trends, research designs, sample sizes, participant demographics, educational domains, and learning contexts. For RQ2, data modalities, AI models, AIFB focus, and AIFB complexity were examined. Frequencies and distributions were calculated to identify patterns. For RQ3, the associations between AIFB and learner perceptions, actions, and outcomes were explored through categorical analysis of reported findings. Studies were grouped by perception types (e.g., cognitive, social/emotional, technological), action categories (e.g., self-regulation, co-regulation), and learning outcomes (e.g., cognitive achievement, skill development). Finally, RQ4 synthesized associations between AIFB's mechanisms (e.g., feedback levels and complexity) and its reported effects, emphasizing trends such as the alignment of task- and process-level feedback with cognitive and actionable perceptions, as well as the role of intermediate and elaborated feedback in fostering both individual and collaborative learning outcomes.

4. Results

4.1 RQ1: What are the bibliometric and methodological profiles of the reviewed studies on AIFB in education?

The number of publications on AIFB in education has increased drastically, from fewer than 10 annual publications before 2019 to over 35 in 2023 (Figure 4). The reviewed studies exhibited an even distribution between quantitative (n = 61) and mixed methods (n = 60) approaches. Meanwhile, the number of pure qualitative research (n = 8) was found to be much smaller.

Figure 4
Number of publications yearly



The reviewed studies varied in sample sizes, with 31 studies involving 50 to 100 participants ("large"), while 49 had more than 100 ("very large"). "Medium" (30 to 50 participants) and "small" (less than 30 participants) sample size studies were also well represented, with 21 and 24 occurrences, respectively. Besides, three studies incorporated multiple sample groups, and seven did not specify their sample size. Most studies (n = 75) recruited undergraduate students, followed by 15 studies involving postgraduate students and another six having a mix of both. Only a few studies involved elementary (n = 12) and secondary (n = 10) school students. The remaining 11 studies either described their participants generally (e.g., adults) or did not specify.

Around one-third of the studies (n = 44) were conducted in educational domains related to language, communication, and writing. Slightly fewer studies (n = 31) focused on STEM subjects (science, technology, engineering, and mathematics), followed by areas related to personal development, professional skills, and general education (n = 27). Moreover, 16 studies were about health science education, while arts, creativity, and design were the least studied (n = 9).

In terms of learning contexts, the majority of studies were conducted in digital or online environments (n = 66), reflecting the growing adoption of technology-mediated learning platforms. Classroom-based implementations were reported in 19 studies, while 32 studies employed a blended learning approach that combined face-to-face and online components. A smaller number of studies were conducted in laboratory settings (n = 9) or informal learning contexts (n = 3), such as self-directed or extracurricular environments.

4.2 RQ2: What are the mechanisms of AIFB in education?

4.2.1 RQ2a: What types of data and evidence inform AIFB?

The synthesis of data modalities revealed a predominant reliance on textual data (n = 73), such as student essays, conversations with chatbots, and discussion threads. Performance and interaction data (n = 26) was a common input for informing or generating feedback. This data modality consisted of quantifiable metrics such as attendance, quiz scores, and system interaction logs. Audio and visual data, including speech recordings, artwork images, and facial expressions, were considered in 15 studies. Besides, 10 studies modeled movement data for feedback, such as physical gestures and interactions within virtual spaces. Notably, five studies combined multiple data modalities to build AIFB models.

4.2.2 RQ2b: What types of AI models and algorithms are employed?

The summary of AI models and algorithms showcases a range of specialized technologies tailored to assist feedback. Conversational agents, such as IBM Watson (2 studies), Google's Dialogflow (4 studies), CSIEC chatbot, Virtual Operative Assistant, Bilge, and Rasa, were particularly prevalent, with 24 studies employing them to simulate human-like conversations and provide tailored feedback to learners. In this review, conversational agents generally refer to conventional rule-based or retrieval-based systems developed prior to the wide application of GAI tools. Deep learning techniques, including convolutional neural networks, long short-term memory networks (LSTM), and OpenPose, were adopted in 17 studies. This is the same number as those that employed traditional machine learning models, such as support vector machines, decision trees, and hidden Markov models. As a distinct category, large language models (LLMs), including bidirectional encoder representations from transformers (BERT), generative trained transformers (GPT), and ResNet50, were used in 14 studies. Unlike earlier conversational agents, LLMs represent a more advanced class of transformer-based models, which were initially used for natural language processing tasks such as semantic analysis and text classification, rather than interactive conversation. That said, there is some overlap, as four studies in this review specifically employed LLM-based chatbots (e.g., ChatGPT), reflecting the growing integration of GAI. In addition, 12 studies examined automatic writing evaluation systems (IntelliMetric, Pigai, Juku, Willow). As for the remaining studies, 17 investigated specialized AI tools (e.g., iFlyRec), while 28 did not specify the details of AI tools or models.

4.2.3 RQ2c: What are the focus levels of AIFB?

Results have demonstrated two distinct AIFB strategies, namely single-level and multi-level feedback. For single-level feedback, task-level feedback has been more commonly studied (n = 29), followed by process-level feedback (n = 18). Meanwhile, no studies have solely adopted self-regulation level or self-level feedback.

For multi-level feedback, many studies (n = 55) investigated AIFB manifested at both task and process levels. An even more holistic approach was observed in 16 studies, synergizing feedback across three levels (task, process, and self-regulation). The other combinations of feedback levels were less common, including process and self-regulation levels (n = 7) as well as task and self-regulation levels (n = 2).

4.2.4 RQ2d What are the complexity levels of AIFB?

In general, there has been an inclination toward detailed feedback. Specifically, AIFB at the basic level, which typically only includes indicators for correctness, was very rare among the reviewed studies (n = 1). In comparison, 81 studies implemented intermediate feedback, which carries more information such as explanations, hints, and suggestions for improvement. Furthermore, elaborated feedback, consisting of detailed explanations, tailored advice, and constructive scaffolding, was observed in 46 studies.

4.3 RQ3: What are the effects of AIFB on learners?

4.3.1 RQ3a: How is AIFB associated with learner perceptions?

The reviewed studies on learner perceptions in response to AIFB were roughly categorized under cognitive perceptions, social/emotional/motivational/competent perceptions, and usability/technological perceptions (Table 2). The full list of research output (RO) can be found in Appendix B.

In the cognitive domain, 21 studies examined outcomes on perceived cognitive development (e.g., perceived skill mastery and understanding), with 20 of them indicating positive associated effects of AIFB. Additionally, 10 out of 14 studies reported favorable perceptions regarding the accuracy and personalization of AIFB. Twelve studies reported an increase in learners' perceived performance. Furthermore, nine out of 10 studies indicated that AIFB did not increase cognitive load among learners. The associated effects of AIFB on learners' perceived metacognitive and higher-order thinking skills was also explored in eight studies, with only one reporting no change.

Regarding social/emotional/motivational/competent perceptions, most studies focused on learner satisfaction (n = 29). Seventeen studies explored learners' motivation and attitude, with positive findings. Thirteen studies also confirmed the positive relationship between AIFB and learners' self-efficacy and confidence. As for perceived emotions, 11 out of 12 studies found increased positive emotions among learners, while six studies demonstrated that AIFB helped reduce negative emotions. Social interaction (e.g., social learning and isolation) was less commonly investigated in only three studies.

AI-assisted and actionable perceptions centered on actionable learning experiences with AI-assisted tools (e.g., attention and engagement), with 21 studies reporting favorable experiences. Technology usability and reliability were assessed in 19 studies, with 17 confirming positive outcomes. The perceived timing of AIFB was assessed and considered suitable by 10 studies. Additionally, five studies explored the potential for AIFB to enhance self-regulated learning.

Table 2
AIFB on learner perceptions

Cognitive perceptions	Cognitive development evaluation (n = 21)	**Increased (n = 20)**: RO16; RO22; RO24; RO32; RO34; RO43; RO49; RO58; RO63; RO65; RO68; RO75; RO78; RO81; RO83; RO90; RO98; RO108; RO116; RO123	**No change (n = 1)**: RO108	-
	Content accuracy and personalization analysis (n = 14)	**Increased (n = 10)**: RO2; RO8; RO14; RO16; RO47; RO80; RO93; RO94; RO111; RO117	**No change (n = 1)**: RO76	**Decreased (n = 3)**: RO53; RO66; RO93
	Performance evaluation (n = 12)	**Increased (n = 12)**: RO4; RO8; RO20; RO43; RO47; RO52; RO60; RO66; RO106; RO112; RO118; RO128	-	-
	Cognitive load analysis (n = 10)	**Increased (n = 1)**: RO61	**No change (n = 4)**: RO26; RO121; RO124; RO126	**Decreased (n = 5)**: RO6; RO12; RO105; RO114; RO118
	Evaluation of metacognitive and higher-order thinking skills (n = 8)	**Increased (n = 7)**: RO30; RO40; RO49; RO68; RO78; RO114; RO116	**No change (n = 1)**: RO37	-
Social, emotional, motivational, and competent perceptions	Satisfaction (n = 29)	**Increased (n = 26)**: RO5; RO10; RO15; RO16; RO17; RO18; RO23; RO24; RO27; RO28; RO43; RO45; RO63; RO65; RO66; RO67; RO70; RO75; RO77; RO84; RO87; RO89; RO90; RO100; RO112; RO118	**No change (n = 2)**: RO37; RO85	**Decreased (n = 1)**: RO58
	Motivation and attitude (n = 17)	**Increased (n = 17)**: RO15; RO23; RO29; RO34; RO40; RO44; RO50; RO56; RO62; RO70; RO75; RO76; RO78; RO93; RO99; RO112; RO114	-	-
	Self-efficacy and confidence (n = 15)	**Increased (n = 13)**: RO15; RO16; RO30; RO56; RO61; RO66; RO73; RO75; RO84; RO89; RO107; RO112; RO116	**No change (n = 2)**: RO37; RO102	-
	Positive emotions (n = 12)	**Increased (n = 11)**: RO4; RO23; RO26; RO32; RO37; RO59; RO70; RO80; RO99; RO119; RO129	**No change (n = 1)**: RO24	-
	Negative emotions (n = 6)	-	-	**Decreased (n = 6)**: RO6; RO12; RO14; RO34; RO41; RO74
	Social interaction (n = 3)	**Increased (n = 3)**: RO24; RO43; RO78	-	-
AI-assisted and actionable perceptions	Actionable learning experience (n = 21)	**Increased (n = 21)**: RO4; RO16; RO18; RO31; RO32; RO43; RO44; RO49; RO50; RO59; RO63; RO70; RO73; RO75; RO79; RO81;RO83; RO85; RO98; RO109; RO114	-	-
	Technology usability and reliability (n = 19)	**Increased (n = 17)**: RO4; RO5; RO9; RO15; RO23; RO27; RO31; RO32; RO35; RO52; RO60; RO72; RO80; RO82; RO84; RO94; RO99	**No change (n = 1)**: RO102	**Decreased (n = 1)**: RO24
	Timing of AIFB (n = 10)	**Increased (n = 10)**: RO6; RO16; RO29; RO32; RO40; RO47; RO84; RO94; RO111; RO117	-	-
	AI-assisted self-regulated learning (n = 5)	**Increased (n = 4)**: RO43; RO61; RO74; RO129	-	**Decreased (n = 1)**: RO39

4.3.2 RQ3b: How is AIFB associated with learner actions?

Around one-third of the reviewed studies captured learners' actions in response to AIFB. Given the locus of control of these actions (i.e., individual, interpersonal, and group), this review categorized them under self-regulation, co-regulation, and socially shared regulation (Hadwin & Oshige, 2011; Zimmerman, 2008), which helped distinguish the context in which changes of actions occurred (Figure 3).

Table 3
AIFB on learners' actions

Self-regulation	Forethought phase	-	-
	Performance phase (n = 12)	Idea/argument/explanation generation (n = 4)	RO36; RO71; RO102; RO115
		Engagement in activities (n = 3)	RO20; RO89; RO101
		Task/assignment completion (n = 3)	RO4; RO9; RO104
		Usage of learning strategies (n = 2)	RO95; RO113; (**No change**)
		Dashboard usage (n = 1)	RO2
	Self-reflection phase (n = 10)	Number/duration of drafts/revisions/attempts (n = 7)	RO55; RO80; RO101; RO102; RO113; RO119; RO127 (**Decreased**)
		Self-reflection and self-correction (n = 3)	RO64; RO80; RO95
Co-regulation	Interaction efficacy (n = 1)		RO18
	Carefulness and truthfulness in peer submission reviews (n = 1)		RO38
	Open-ended questioning (n = 1)		RO85
	Co-regulated behavioral sequences (n = 1)		RO126
	Response frequency and polarity (n = 1)		RO96
Socially shared regulation	Collaborative knowledge building (n = 2)		RO120; RO122
	Social interaction (n = 2)		RO78; RO126
	Information sharing (n = 1)		RO78
	Engagement in discussions (n = 1)		RO124
	Depth and quality of forum posts (n = 1)		RO104
	Constructive behaviors (n = 1)		RO104
	Depth and complexity in knowledge elaboration (n = 1)		RO126
	Knowledge convergence (n = 1)		RO126
	Metacognitive regulation (n = 1)		RO121
	Behavioral transitions (n = 1)		RO121

In the reviewed studies, 22 studies capturing learner actions were within individual learning contexts (i.e., self-regulation). Among these, 12 measures were recorded during the performance phase (e.g., dashboard usage), while 10 measures pertained to the self-reflection phase (e.g., number of revisions). Notably, none of the studies considered learner actions related to the forethought phase.

Regarding co-regulation (n = 5) and socially shared regulation (n = 8), where learning involves interactions among multiple individuals, there have been relatively fewer studies exploring the associated effects of AIFB. The actions examined across these studies also varied. Despite the diversity, most of the reviewed studies indicated positive effects of AIFB on learner actions.

4.3.3 RQ3c: How is AIFB associated with learning outcomes?

Around half of the reviewed studies investigated the association between AIFB and learning outcomes (Table 4). The most commonly examined outcomes were cognitive in nature, with 36 studies evaluating learners' academic achievement. Notably, all but one of these studies reported positive effects, suggesting that AIFB generally contributes to improved learning performance, often assessing conceptual understanding or memory of subject content over time.

Beyond cognitive outcomes, a considerable number of studies explored the role of AIFB in supporting skill development. Writing skills were the most frequently studied (n = 13), particularly in language and communication contexts, where feedback targeted textual coherence, grammar, or argumentation. Speaking skills were examined in five studies, often in pronunciation or oral presentation tasks. For these two skill areas, we also examined the linguistic and geographic contexts in which the studies were conducted, as these factors have important implications for the interpretation and transferability of findings. Among the 13 studies on writing, most involved learners using English as a second language, with studies conducted in Mainland China, Hong Kong, Iran, and Norway. Others took place in English-dominant contexts such as the USA and Canada, while two studies involved German-speaking learners in Germany and Western Europe, and one study used Chinese as the medium of instruction. Similarly, the five studies focusing on speaking skills were largely situated in East Asian contexts, including Taiwan and Mainland China, and all employed English as the target language. These patterns suggest that AIFB systems for writing and speaking are frequently deployed in English as a second language learning environments. Other domain-specific skills included surgical (n = 3), kinematic or motor coordination (n = 3), and artistic skills (n = 3). A small number of studies also addressed additional competencies such as communication skills, punctuation accuracy, and self-diagnosis ability.

Social and collaborative dimensions of learning were addressed in seven studies that examined how AIFB influenced group processes and outcomes, such as the quality of co-constructed artefacts (e.g., knowledge graphs). These studies suggest that AIFB can support not only individual learning gains but also the enhancement of collective cognitive outcomes.

Table 4
AIFB on learning outcomes

Test performance and learning achievement (n = 36)		RO1; RO8; RO9; RO10; RO15; RO17; RO19; RO21; RO22; RO25; RO27; RO28; RO29; RO33; RO36; RO41; RO42 (**No change**); RO45; RO52; RO53; RO54; RO55; RO56; RO67; RO79; RO80; RO87; RO88; RO92; RO97; RO100; RO103; RO105; RO114; RO115; RO129
Acquisition of knowledge (n = 8)		RO18; RO26; RO37 (No change); RO44; RO46; RO72; RO98; RO117
Group knowledge graph scores (n = 7)		RO120; RO121; RO122; RO123; RO124; RO125; RO126
Skills	Writing skills (n = 13)	RO11 (China, English); RO57 (HK, English); RO60 (Canada, English); RO61 (China, English); RO62 (China, English); RO64 (USA, English); RO71 (Norway, English); RO78 (China, Chinese); RO83 (Iran, English); RO86 (Germany, German); RO102 (Western Europe, German); RO106 (China, English); RO107 (USA, English)
	Speaking skills (n = 5)	RO12 (Taiwan, English); RO14 (No change) (Taiwan, English); RO48 **(Increased + No change) (China, English)**; RO58 (Taiwan, English); RO128 (Zou, English)
	Surgical skills (n = 3)	RO25 (**Increased + No change**); RO68; RO69
	Kinematic skills (n = 3)	RO34; RO50; RO51
	Artistic skills (n = 3)	RO7; RO13; RO117
	Critical thinking skills (n = 1)	RO10
	Communication skills (n = 1)	RO3 (**Increased + No change**)
	Punctuation skills (n = 1)	RO99
	Presentation skills (n = 1)	RO77
	Prosocial skills (n = 1)	RO91
Others	Self-diagnosis accuracy (n = 1)	RO86 (**No change**)
	Comprehension abilities (n = 1)	RO110
	Simulation training effectiveness (n = 1)	RO82

4.4 RQ4: What are the associations between AIFB's mechanisms and effects?

This review first explored the associations between feedback foci and learner perceptions (Figure 5). Overall, there has been a predominant trend of employing task-level and process-level feedback to support most learning perceptions. Specifically, task-level feedback was frequently associated with cognitive development evaluation (n = 20), satisfaction (n = 24), and actionable learning experience (n = 17). Meanwhile, the number of studies leveraging process-level feedback was relatively higher regarding perceived performance evaluation (n = 10), cognitive load analysis (n = 10), and negative emotions (n = 5). While self-regulation level feedback was less common overall, it was proportionally more prominent in studies examining perceived cognitive load (20%) and self-regulated learning (25%), compared to its frequency across the full set of outcomes.

Figure 5

The focus levels of AIFB on learner perceptions

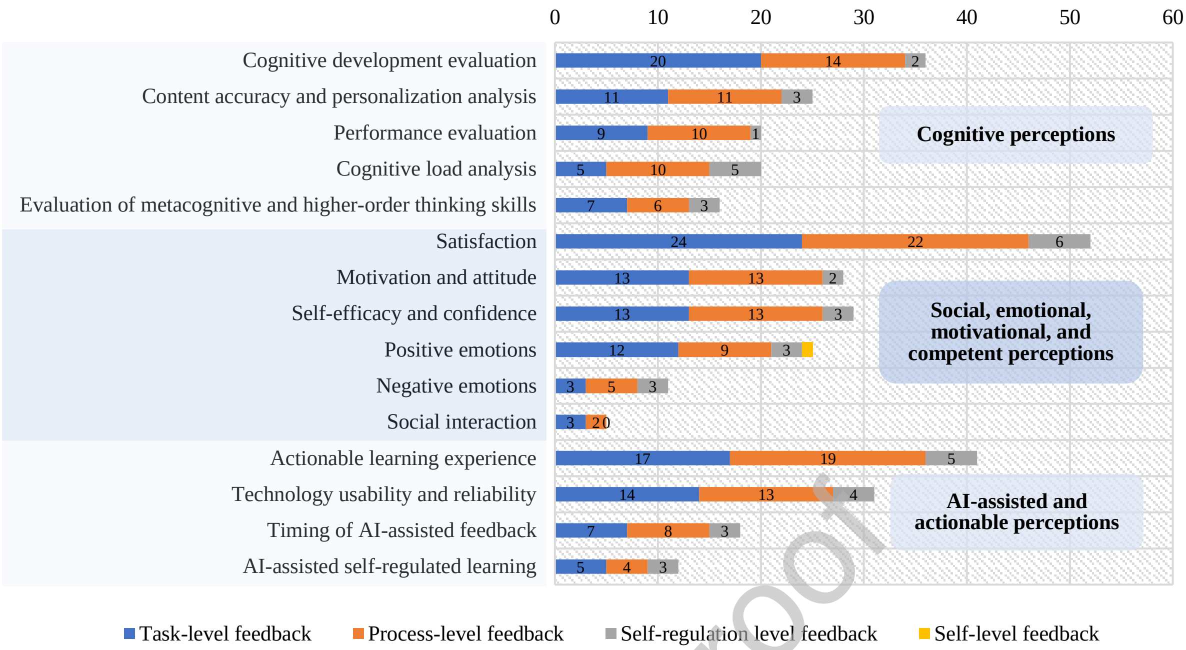


Figure 6 presents associations between various feedback complexity and learner perceptions. Under the cognitive dimension, the reviewed studies more often adopted intermediate feedback, except for cognitive load. Studies commonly took cognitive load into account when using AI-assisted elaborated feedback (n = 8).

Regarding social/emotional/motivational/competent perceptions, the distribution for intermediate feedback was found to be similar across perception variables. However, a deviation was observed for negative emotions, which were more frequently associated with elaborated feedback.

As for AI-assisted and actionable perceptions, the percentage of intermediate feedback was high for actionable learning experiences (n = 16). Meanwhile, the usage of intermediate and elaborated feedback was roughly equal among other perception measures in this dimension.

Figure 6

The complexity levels of AIFB on learner perceptions

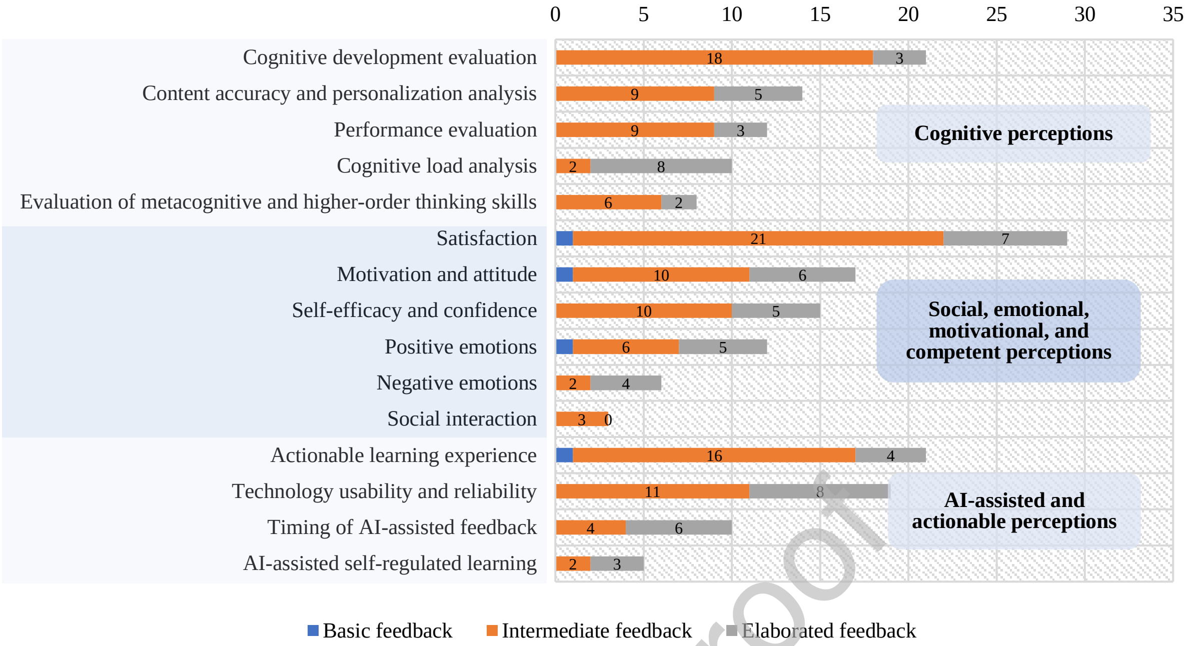
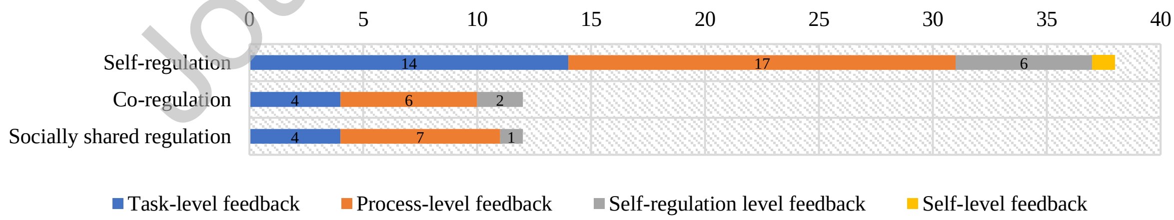


Figure 7 presents the associations between feedback focus levels and learner actions. Process-level feedback has been adopted the most frequently across scenarios (self-regulation n = 17, co-regulation n = 6, socially shared regulation n = 7). Meanwhile, task-level feedback was 2 to 3 occurrences less frequent on average. A small proportion of studies have employed AI-assisted self-regulation level feedback.

Figure 7

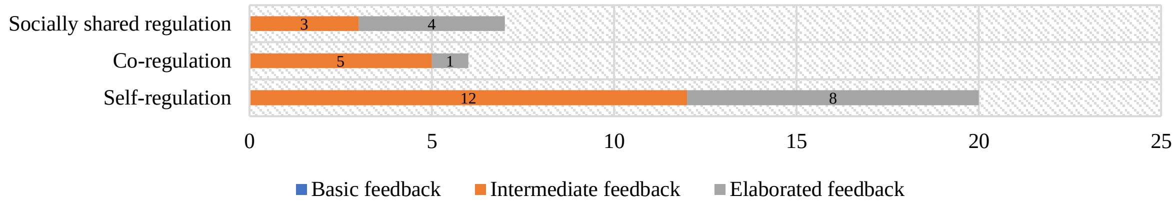
The focus levels of AIFB on learner actions



Regarding AIFB's complexity levels (Figure 8), the reviewed studies mainly examined the association between intermediate feedback and learners' actions within self-regulation and co-regulation contexts. In socially shared regulation scenarios, studies employing intermediate feedback (n = 3) and elaborated feedback (n = 4) were almost even. None of the reviewed studies considered basic feedback for altering learner actions.

Figure 8

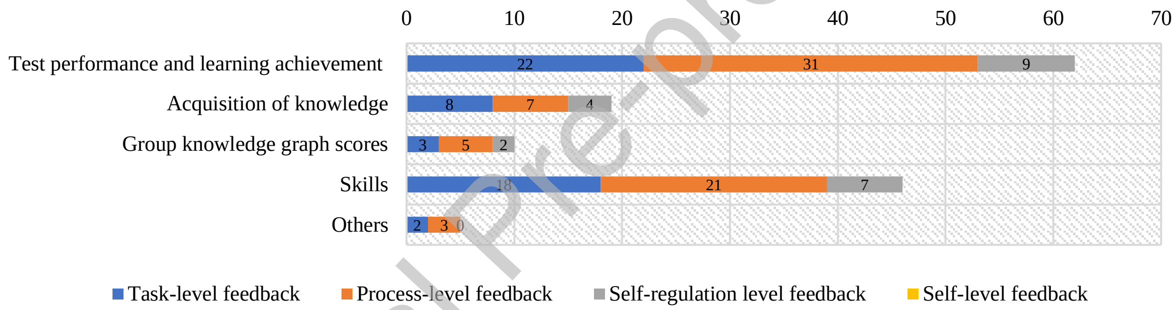
The complexity levels of AIFB on learner actions



For AIFB foci and learning outcomes, Figure 9 depicts a similar distribution among the four focus levels. Feedback was primarily manifested at task and process levels. In terms of self-regulation level feedback, it was adopted more frequently for test performance and learning achievement.

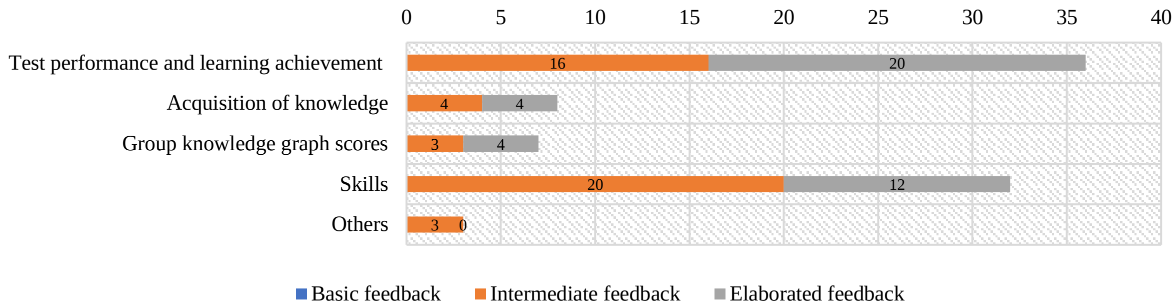
Figure 9

The focus levels of AIFB on learning outcomes



As for AIFB complexity and learning outcomes, intermediate and elaborated feedback were implemented almost equally to enhance various outcomes except for skill development (Figure 10). The proportion of intermediate feedback is larger than elaborated feedback for fostering various skills.

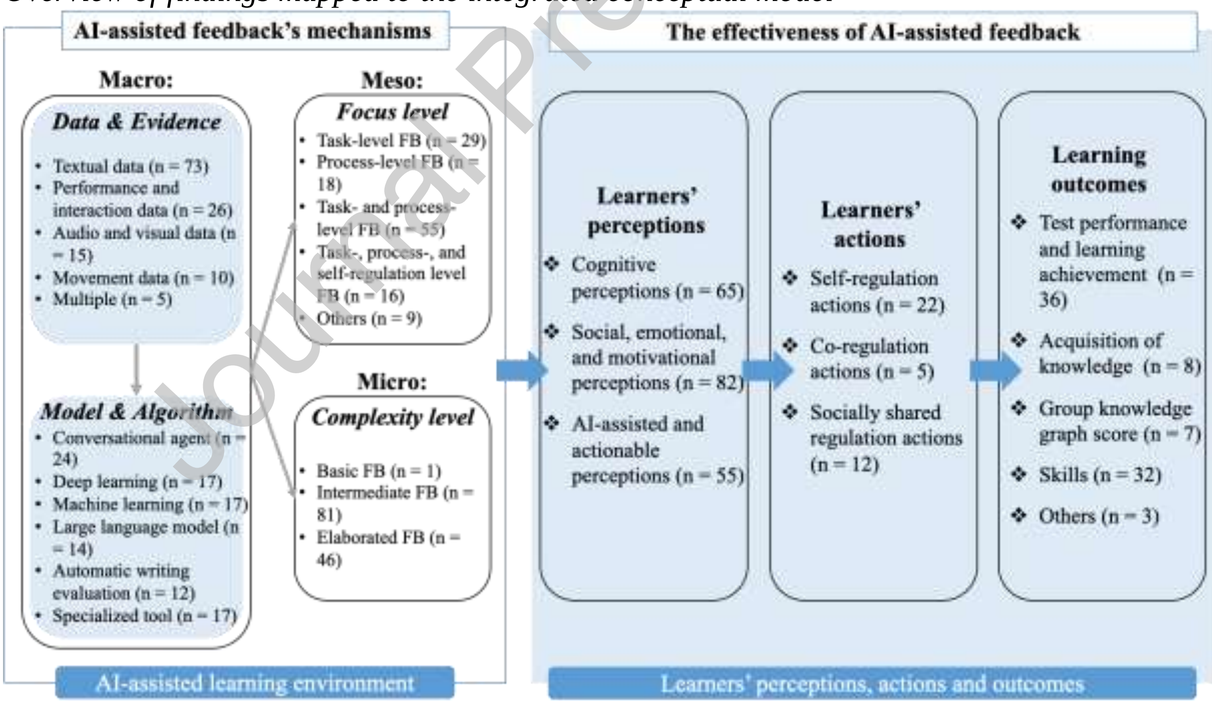
Figure 10
The complexity levels of AIFB on learning outcomes



5. Discussion

This review revealed meaningful findings regarding the mechanisms and effectiveness of AIFB. Based on the conceptual model (Figure 1), findings were categorized and summarized accordingly (Figure 11). This section discusses the findings and their implications.

Figure 11
Overview of findings mapped to the integrated conceptual model



5.1 Bibliometric and methodological profiles

The growth of publications is consistent with observations by Fu et al. (2024) on automatic writing evaluation feedback and the general development of AI in education (Chen et al., 2022). One explanation could be that the evolution of AI technologies motivated explorations of educational applications. Several modern large language models were proposed around 2018, such as generative pre-training (GPT 1.0) (Radford et al., 2018) and bidirectional encoder representation from transformers (BERT) (Devlin et al., 2019).

Methodology-wise, the broad adoption of quantitative (n = 61) and mixed methods (n = 60) approaches echoes the data-driven nature of AIFB. Quantitative data, such as logs and test scores, are not only crucial for AI modeling but may also be parts of the research results. While doing so is technologically straightforward, it calls for attention and rigor in terms of research ethics (Guan et al., 2023).

Although there have been few pure qualitative studies, the substantial number of mixed methods research indicates that qualitative insights into the “how” and “why” behind the effectiveness of AIFB have been valued. With GAI’s advanced capabilities for in-depth interaction and communication, qualitative methods such as content analysis can be crucial for revealing human-AI interaction dynamics (Wang et al., 2023). That being said, with increasing data generated, qualitative analysis may also benefit from automatic approaches (Ba et al., 2023; Shaffer et al., 2016).

The review revealed that more than half of the studies involved large (50 or more participants) or very large (100 or more participants) samples. This prevalence of large samples is attributable not only to the primary adoption of quantitative methods but also to the integration of digital platforms. Since AIFB needs to be hosted by platforms that can process data, analyze patterns, and generate data-based insights, the deployment of such platforms could make AIFB more accessible and scalable.

As for participants’ education levels, the finding is not surprising that 90 studies recruited undergraduate or graduate students. Besides the factor that tertiary students are more autonomous and accessible, another factor might be the suitability of AI technologies for younger students. When designing AIFB for less mature learners, especially in the GAI era, researchers need to consider both the potential benefits and risks, such as younger learners' abilities in discerning GAI hallucinations (Williams et al., 2023).

The distribution of studies across subject fields is relatively even, except for slightly more studies on language, communication, and writing (n = 44). This disparity may be attributed to the development of large language models, which enable sophisticated analyses of text data and the generation of accurate and constructive feedback in language-related disciplines (Xiao & Zhi, 2023).

5.1.1 Data and evidence informing AIFB

This review identified the predominant use of textual data (n = 73) in AIFB, largely due to advancements in natural language processing techniques (Shaik et al., 2022). The reliance on textual input may also stem from many learning outcomes being inherently textual or translatable into text, such as written reports and discussion transcripts (Smith et al., 2020).

Other reviewed studies indicated a shift towards a multimodal approach in AI applications (e.g., Fazlollahi et al., 2023; Yamamoto et al., 2023; Zou et al., 2023). This approach is crucial for developing AI systems that offer personalized, context-aware feedback and accommodate various learning activities.

5.1.2 AI models and algorithms

The results on computational models and algorithms varied widely regarding both techniques employed and their descriptions. While some studies provided detailed descriptions of the AI models (e.g., long short-term memory model), others only mentioned a broad category (e.g., deep learning) or simply the software (e.g., Grammarly). This inconsistency poses challenges when comparing AI models across studies.

Future studies are recommended to provide systematic and specific descriptions of AI models and algorithms. This specificity is particularly worthwhile as we enter the GAI era, where each iteration and model can present very different capabilities (e.g., ChatGPT 3.5 vs. ChatGPT 4) (Farhat et al., 2023).

5.1.3 Focus levels of AIFB

Regarding AIFB focus levels, some studies concentrated on a single level (n = 37), while others incorporated multiple levels (n = 55). Among studies implementing single-level feedback, task level was the most frequent, followed by process level. This distribution may be explained by the complexity behind data collection and modeling. Specifically, providing task-level feedback is relatively straightforward, where the AI models are confined to identifying errors and offering corrective suggestions. Conversely, process-level feedback requires the AI models to not only process learners' action data but also analyze the associations hidden within actions (Cerezo et al., 2020). Such analyses entail a deeper understanding of how various actions collectively reflect learners' cognitive, social, and other states.

Despite these challenges, many recent studies have integrated multiple feedback levels (e.g., Chaiprasurt et al., 2022; Ouyang et al., 2023; Zheng et al., 2023). By combining task, process, and self-regulation feedback, these studies highlight AI's potential to help stakeholders build a well-rounded understanding of the learning processes.

5.1.4 Complexity levels of AIFB

The scarcity of studies using basic feedback, coupled with the prevalence of intermediate feedback, suggests a shift towards offering learners more than right-or-wrong indicators. Intermediate feedback, which includes not only correctness indicators but also some explanations, hints, and/or suggestions, signifies a balance between providing learners with basic information and overwhelming them with non-personalized details. That is to say, while details in feedback can be informative, excessive details may also distract learners from the core feedback (Ferguson, 2011). The capacity to deliver detailed yet personalized feedback is highly reliant on the affordance of AI models.

In this sense, the adoption of elaborated feedback in 46 studies reflects the fast-growing capabilities of AI to provide detailed, tailored, and constructive feedback. Such feedback aligns with constructivist learning theories that emphasize the importance of active learner engagement and the construction of knowledge through personalized learning experiences (Bada & Olusegun, 2015). Elaborated feedback demonstrates the potential of AI to serve as interactive and responsive agents, capable of adapting to individual learner needs and fostering higher-order thinking skills.

5.2 Associations between AIFB and learner outcomes

5.2.1 Association with learner perceptions

The synthesis of learner perceptions towards AIFB has revealed primarily positive results across dimensions. The distribution among cognitive perceptions has indicated a focus on perceived learning gains (i.e., cognitive development evaluation ($n = 21$) and performance evaluation $n = 12$)). Meanwhile, AIFB would not increase learners' perceived cognitive load in most cases. This suggests that quality AIFB can reduce uncertainty and confusion about a topic and facilitate cognitive processing (Wang et al., 2019). Furthermore, the results have demonstrated a need to continue exploring the associated effects of AIFB on fostering metacognitive and higher-order thinking skills.

On social/emotional/motivational/competent perceptions, learner satisfaction emerged as the primary outcome, with 26 studies that highlighted learners' favorable responses to learning with AIFB. The enhancement of motivation, attitude, self-efficacy, and confidence has further underscored the potential of AIFB in creating a supportive and worthwhile learning environment. Moreover, studies measuring learning emotions indicated that AIFB effectively relieves negative emotions and promotes positive ones. This finding confirms the good practices of AIFB in maintaining a positive stance (Ryan & Henderson, 2018). The perceived social interaction remains underexplored, suggesting a potential gap in understanding how AIFB influences group dynamics and collaboration effectiveness.

Regarding AI-assisted and actionable perceptions, the positive actionable learning experiences across studies have confirmed that AI-assisted tools can actively engage learners. Moreover, the favorable perceptions of technology usability and reliability have indicated that current AI-assisted educational systems are accessible and stable, which is crucial for future applications in educational practices. Studies on the timing of AIFB have further highlighted the advantage of AI in providing immediate feedback (Ba et al., 2023).

5.2.3 Association with learner actions

Learner actions were the focus in only a small subset of the reviewed studies, with a tendency to focus on individual learning scenarios rather than social or collaborative learning. This tendency may partially be attributed to the importance of self-regulated learning, which is pivotal across various learning activities (Booth et al., 2018). In that sense, despite the popularity of collaborative learning at all education levels, there is a notable lack of studies examining AIFB in group-based learning contexts. This scarcity may stem from the complexities associated with the dynamic nature of collaborative learning, with multiple learners sharing information and exchanging ideas (Cress et al., 2019). It remains to be understood whether current or forthcoming large language models can effectively manage these complexities and provide high-quality feedback in collaborative learning settings.

5.2.4 Association with learning outcomes

The reviewed studies shed light on several outcome dimensions. A majority of the studies (n = 36) concluded that AIFB was positively associated with test performance and learning achievement. The findings of these studies have suggested that the benefits of AIFB are broadly applicable to various educational settings.

The reviewed studies also extensively explored the development of various skills facilitated by AIFB. The notable emphasis on writing skills is likely due to the objective metrics for writing evaluation and the clear delineation of improvement areas where AI can offer precise and actionable feedback (Deeva et al., 2021; Fu et al., 2024). In more specialized domains, studies investigating the association of AIFB with surgical, kinematic, and artistic skills have reflected the adaptability of AI feedback mechanisms to diverse and practical skill sets. Nevertheless, more evidence is necessary to better understand the effectiveness of AIFB in enhancing various skills.

5.3 Associations between AIFB mechanisms and effects

5.3.1 Associations between feedback focus and learner perceptions

According to Figure 5, task-level and process-level feedback assisted by AI have been predominantly used to shape learner perceptions. The frequent association of task-level feedback with cognitive development evaluation, satisfaction, and actionable learning experience suggests that task-specific feedback has often been considered suitable for enhancing both the quality of learning and learners' engagement with the content (Butler & Woodward, 2018). Meanwhile, AI-assisted process-level feedback has been diversely associated with performance evaluation, cognitive load analysis, and negative emotions, highlighting the feasibility of employing AI to scaffold the learning process by breaking down complex information and assisting learners through challenging tasks. Despite the relatively sparse focus on self-regulation feedback assisted by AI, it is shown to be significantly tied to managing cognitive load and fostering self-regulated learning. The complexity of implementing this type of feedback with AI might contribute to its sparse usage (Molenaar, 2022). Addressing these barriers in future studies could enrich our understanding of how AI can support sustained learning efforts.

5.3.2 Associations between feedback complexity and learner perceptions

The observed trend has indicated a preference for intermediate feedback with regard to cognitive perceptions, except for cognitive load, where there is a shift towards elaborated feedback (Figure 6). This finding has indicated that researchers have been cautious when employing detailed feedback, which might require additional cognitive processing.

A similar inclination towards intermediate feedback remains for social/emotional/motivational/competent perceptions. However, the domain of negative emotions stands out, with more studies employing elaborate feedback. Combined with the association between negative emotions and process-level feedback, the results have suggested that existing studies tended to address learners' negative emotions with more intricate and detailed feedback assisted by AI. Moreover, while that feedback might not be directly emotional, its level of detail and clarity could help learners decompose complex questions and thus reduce their negative emotions (Lang et al., 2022; Wang et al., 2019).

5.3.3 Associations between feedback focus and learner actions

The linkages between the focus levels of AIFB and learner actions have presented a slight predilection for process-level feedback. This suggests that AIFB targeting learning approaches and strategies might be considered more impactful on individual learner actions compared to other focus levels. Despite the critical role of self-regulation in guiding learner actions (Yamada et al., 2017), AIFB operationalized at the self-regulation level has been underrepresented, pointing to a potential research gap. This gap may stem from the complexities involved in designing and delivering self-regulation level feedback with AI (Molenaar, 2022).

5.3.4 Associations between feedback complexity and learner actions

The reviewed studies have mainly employed intermediate AIFB to adjust learner actions in self-regulation and co-regulation scenarios (Figure 8). This trend may represent a common practice of feedback designs in educational settings (García-Jiménez et al., 2015; Labuhn et al., 2010). Notably, elaborated feedback has been adopted more often for socially shared regulation actions. While the overall number of studies in this category is small, the relative prevalence of elaborated feedback echoes a trend of utilizing detailed and nuanced feedback assisted by AI in collaborative learning contexts (Guasch et al., 2013; Han et al., 2021).

5.3.5 Associations between feedback focus and learning outcomes

The examination of AIFB focus and learning outcomes has revealed a consistent pattern in the wide deployment of task- and process-level feedback. This close-to-equal distribution between the two focus levels has indicated that both the specificities of tasks and the underlying learning processes were considered crucial for enhancing academic performance (Wisniewski et al., 2020). Meanwhile, this distribution may also represent a shift from task- to process-level AIFB as a result of the development of AI technologies (Zhai et al., 2021).

5.3.6 Associations between feedback complexity and learning outcomes

The complexity of AIFB in relation to learning outcomes has highlighted a balanced application of intermediate and elaborated feedback across most outcome metrics. This observation may be explained by the fact that both types of feedback could be beneficial for promoting academic achievement. Nevertheless, it remains unclear which complexity level of AIFB may be more effective in promoting learning outcomes. Furthermore, the balanced distribution shifts when focusing particularly on skill development, where intermediate feedback has been employed more often than elaborated feedback. This shift suggests that learners may better assimilate and apply skills when feedback is concise and actionable rather than excessively detailed (Ferguson, 2011). This trend has further emphasized the consideration required to calibrate the complexity of AIFB to the nature of the targeted learning outcome.

5.4 Implications

The ongoing advancement of AI in education presents both opportunities and challenges regarding the immediacy, focus, complexity, and personalization of AIFB. Building on a seminal review on formative feedback (Shute, 2008), many studies have advocated for the importance of timely and actionable feedback to boost learners' metacognition and engagement (Morris et al., 2021; Zhan et al., 2022). However, a major challenge remains regarding the difficulty for a single instructor to effectively support a large group of learners (Ba et al., 2023). With AI's data processing affordance, the reviewed studies have consistently supported AIFB's efficiency. Future efforts may focus on fine-tuning the content and delivery formats of AIFB.

The four levels of focus carry diverse roles and are appropriate for different learning tasks, goals, and contexts (Hattie & Timperley, 2007). While a majority of the reviewed studies have examined AI-assisted task- and process-level feedback, how AI may assist with other feedback levels awaits to be examined. Moreover, it is an emerging and promising path to explore the potential of multilevel feedback.

The results examined via the lens of Shute (2008) in this review indicated that AI could dynamically adapt the complexity of feedback. While this study took consideration of Shute's work, it also went beyond Shute's relatively static categorization by showing that feedback can be fluid and evolve during learning based on ongoing assessments of learner performance and engagement assisted by AI tools.

Hence, personalization is central to AIFB. Given particular learning objectives (i.e., perceptions, actions, and outcomes), researchers and practitioners may thoughtfully consider the type of learning data they can collect (macro), the type of AI models available to them (macro), the areas of improvement they focus on (meso), and the level of feedback complexity that is necessary. By synergizing the above considerations with the advancements of GAI, it is promising to establish a learner-centered and AIFB approach for future learning.

To sum up, conceptually, we made a unique contribution to combine feedback focus levels (task, process, self-regulation, and self; Hattie & Timperley, 2007) with feedback complexity (basic, intermediate, elaborated; Shute, 2008), providing an integrated model of feedback (meso and micro levels) in examining AIFB's mechanisms. This extended conceptual model (Figure 1) has suggested that the effectiveness of feedback would be contingent not just on the focus of the feedback but also on its complexity. This integrated conceptual model contributes to a more nuanced understanding of how different types of feedback interact with perceptions, actions, and outcomes in learning (see Figure 11). The findings of our study have also highlighted the need for dynamic feedback systems in the age of AI that adapt both the focus level and complexity of feedback based on real-time assessments of learners' understanding of task requirements and task complexity in digital contexts.

Practically, our findings encourage the adoption of AI tools in educational settings to explore the mechanisms of AIFB. For instance, AI tools can be used to analyze learners' responses and automatically adjust the complexity and focus of feedback to address individual-based needs. This could lead to more personalized learning experiences and potentially higher engagement and achievement. Our findings on task and process-level feedback in AI applications have also practical implications for software developers and educational technologists aiming to create AI that supports deeper learning.

The integration of AI in feedback systems represents a key area of innovation in education. The conceptual model developed in this study provides a foundation for future research and the development of dynamic feedback systems that can adapt both focus levels and complexity in real time, addressing the evolving needs of learners in increasingly complex digital contexts. As AI continues to advance, the potential for anticipatory feedback, where systems predict and address learning gaps before they manifest, represents a critical direction for the field. This highlights the ongoing relevance of this review, as it informs the design of next-generation AI tools that support personalized, predictive, and adaptive learning experiences, contributing to deeper engagement and enhanced outcomes in education.

5.5 Limitations

This review has several limitations to acknowledge. First, to ensure the quality and credibility of the included studies, the review was restricted to peer-reviewed journal articles. While this criterion enhances methodological rigor, it may have excluded relevant and innovative work found in conference proceedings, technical reports, or preprints, particularly given the rapid development of AI in education. Second, the search strategy relied on a defined set of keywords related to AI, feedback, and education. Consistent with established practices in leading systematic reviews in AI and education (e.g., Cavalcanti et al., 2021; González-Calatayud et al., 2021; Hopcan et al., 2023; Salas-Pilco et al., 2022; Sun et al., 2023; Zhai et al., 2021; Zhao, 2024), our approach prioritized broad and widely recognized AI-related terminology, rather than subfield-specific terms such as "automated feedback" or "automated writing evaluation" (AWE). This methodological decision was made to ensure methodological consistency, credibility, and comparability within the scholarly community. Nevertheless, we acknowledge that this approach may have led to the underrepresentation of studies specifically focused on AWE and similar subfields, as well as studies that did not explicitly reference AI-related keywords in their titles, abstracts, or indexing. As a result, certain relevant research domains within AI-assisted feedback may be less visible in the current synthesis. Third, given the substantial variations in research designs, feedback implementations, learning contexts, and outcome measures, this review did not aim to draw generalizable conclusions about the overall effectiveness of AIFB. Instead, its primary goal was to synthesize how AIFB has been implemented and what kinds of processes and outcomes have been associated with its use. Accordingly, any interpretations of impact should be understood as descriptive and exploratory rather than causal. Finally, the review may be subject to publication bias, as studies reporting significant or positive findings are more likely to be published in peer-reviewed venues, whereas null or negative results may remain unpublished or appear in less prominent sources. This could lead to an overrepresentation of favorable outcomes in the evidence base.

5.6 Future directions of research

Based on the aforementioned limitations and insights derived from our findings on AIFB, as summarized in the conclusion section, we propose three significant directions for future research as follows: First, there is a critical need for standardized reporting of AI models and algorithms to enhance transparency and replicability in feedback research. Exploring the advantages and limitations of different AI tools available for educators and researchers worldwide could potentially lead to more tailored feedback mechanisms. Future studies should also aim to optimize AIFB, assessing how different feedback strategies (e.g., different combinations of feedback focus and complexity levels) affect various learning outcomes and how they can be best implemented to balance cognitive load and enhance learners' engagement without overwhelming them. Second, future research could endeavor to broaden the methodological approaches employed in investigating AIFB, integrating qualitative analyses alongside the prevalent quantitative and mixed methods. This would enrich or provide a deeper understanding of the nuanced impacts of AI feedback on learning. For measures of learning outcomes, greater emphasis should be placed on incorporating objective measures, such as assessments of actual cognitive skills, engagement, and learning behaviors, to complement self-report data. This would help mitigate the inherent limitations of self-reported perceptions and provide a more well-rounded understanding of the effects of AIFB. Additionally, research needs to extend to a broader range of educational levels and diverse subject fields to validate the effectiveness of AIFB across various learning stages. Aside from textual data, it is also meaningful for researchers to expand the range of data sources used in AI decision-making, including audio, video, and physiological data. Doing so would enrich the analysis, offering a more comprehensive perspective on learners' interactions and responses to AI tools. Thirdly, future systematic reviews should consider the inclusion of subfield-specific terminologies, such as "automated feedback" and "automated writing evaluation" (AWE), as well as related domains. This would contribute to a comprehensive coverage of the rapidly evolving landscape of AI-assisted feedback and address potential gaps identified in the current synthesis, particularly the underrepresentation of AWE-focused studies. Fourthly, Long-term studies are essential to assess the impacts of AIFB on learning perceptions, actions, and skill development (for skill development, there could also be different levels [e.g., low vs. high] to explore further), helping to understand how learners' interactions with AI influence their outcomes. Additionally, there is a need to focus on how AIFB can be designed to help learners manage their cognitive and emotional loads effectively. Future research can also explore the relationships between learners' perceptions of AIFB and their learning actions, particularly in contexts of self-regulation and collaborative learning, to better understand the efficacy and applicability of AI in empowering feedback for maximizing productive learning.

References

Adamson, D., Dyke, G., Jang, H., & Rosé, C. P. (2014). Towards an agile approach to adapting dynamic collaboration support to student needs. *International Journal of Artificial Intelligence in Education, 24*(1), 92–124. https://doi.org/10.1007/s40593-013-0012-6

Afzaal, M., Zia, A., Nouri, J., & Fors, U. (2023). Informative feedback and explainable AI-based recommendations to support students' self-regulation. *Technology, Knowledge and Learning, 29*(1), 331–354. https://doi.org/10.1007/s10758-023-09650-0

Ajjawi, R., Boud, D., Henderson, M., & Molloy, E. (2019). Improving feedback research in naturalistic settings. In M. Henderson, R. Ajjawi, D. Boud, & E. Molloy (Eds.), *The impact of feedback in higher education: Improving assessment outcomes for learners* (pp. 245–265). Springer International Publishing. https://doi.org/10.1007/978-3-030-25112-3_14

Ali, R., Hoque, E., Duberstein, P., Schubert, L., Razavi, S., Kane, B., Silva, C., Daks, J., Huang, M., & Van Orden, K. (2021). Aging and engaging: A pilot randomized controlled trial of an online conversational skills coach for older adults. *American Journal of Geriatric Psychiatry, 29*(8), 804–815. https://doi.org/10.1016/j.jagp.2020.11.004

Ayanwale, M. A., Adelana, O. P., Molefi, R. R., Adeeko, O., & Ishola, A. M. (2024). Examining artificial intelligence literacy among pre-service teachers for future classrooms. *Computers and Education Open, 6*, 100179-. https://doi.org/10.1016/j.caeo.2024.100179

Ba, S., Hu, X., Stein, D., & Liu, Q. (2023). Assessing cognitive presence in online inquiry-based discussion through text classification and epistemic network analysis. *British*

Journal of Educational Technology, 54(1), 247–266. https://doi.org/10.1111/bjet.13285

Ba, S., Zhan, Y., Huang, L., & Lu, G. (2025). Investigating the impact of ChatGPT-assisted feedback on the dynamics and outcomes of online inquiry-based discussion. *British Journal of Educational Technology*, 00, 1–25. https://doi.org/10.1111/bjet.13605

Bada, S. O., & Olusegun, S. (2015). Constructivism learning theory: A paradigm for teaching and learning. *Journal of Research & Method in Education, 5*(6), 66–70. https://doi.org/10.9790/7388-05616670

Belcadhi, L. C. (2016). Personalized feedback for self assessment in lifelong learning environments based on semantic web. *Computers in Human Behavior, 55*, 562–570. https://doi.org/10.1016/j.chb.2015.07.042

Benotti, L., Martinez, M., & Schapachnik, F. (2018). A tool for introducing computer science with automatic formative assessment. *IEEE Transactions on Learning Technologies, 11*(2), 179–192. https://doi.org/10.1109/TLT.2017.2682084

Bernius, J. P., Krusche, S., & Bruegge, B. (2022). Machine learning based feedback on textual student answers in large courses. *Computers and Education: Artificial Intelligence, 3*. https://doi.org/10.1016/j.caeai.2022.100081

Bimba, A. T., Idris, N., Al-Hunaiyyan, A., Mahmud, R. B., & Shuib, N. L. B. M. (2017). Adaptive feedback in computer-based learning environments: A review. *Adaptive Behavior, 25*(5), 217–234. https://doi.org/10.1177/1059712317727590

Biswas, U., & Bhattacharya, S. (2023). ML-based intelligent real-time feedback system for blended classroom, *Education and Information Technologies*, 1–29. https://doi.org/10.1007/s10639-023-11949-5

Bonneton-Botte, N., Fleury, S., Girard, N., Le Magadou, M., Cherbonnier, A., Renault, M., Anquetil, E., & Jamet, E. (2020). Can tablet apps support the learning of handwriting?

An investigation of learning outcomes in kindergarten classroom. *Computers & Education, 151*. https://doi.org/10.1016/j.compedu.2020.103831

Booth, A., Hennessy, E., & Doyle, O. (2018). Self-regulation: Learning across disciplines. *Journal of Child and Family Studies, 27*(12), 3767–3781. https://doi.org/10.1007/s10826-018-1202-5

Butler, A. C., & Woodward, N. R. (2018). Chapter One—Toward consilience in the use of task-level feedback to promote learning. In K. D. Federmeier (Ed.), *Psychology of Learning and Motivation*, *69*, 1–38. Academic Press. https://doi.org/10.1016/bs.plm.2018.09.001

Castellano, M. S., Contreras-McKay, I., Neyem, A., Farfán, E., Inzunza, O., Ottone, N. E., del Sol, M., Alario-Hoyo, C., Alvarado, M. S., & Tubbs, R. S. (2023). Empowering human anatomy education through gamification and artificial intelligence: An innovative approach to knowledge appropriation. *Clinical Anatomy, 13*(1), 12–24. https://doi.org/10.1002/ca.24074

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*, Article 100027. https://doi.org/10.1016/j.caeai.2021.100027

Cerezo, R., Bogarín, A., Esteban, M., & Romero, C. (2020). Process mining for self-regulated learning assessment in e-learning. *Journal of Computing in Higher Education, 32*(1), 74–88. https://doi.org/10.1007/s12528-019-09225-y

Chaiprasurt, C., Amornchewin, R., & Kunpitak, P. (2022). Using motivation to improve learning achievement with a chatbot in blended learning. *World Journal on Educational Technology: Current Issues, 14*(4), 1133–1151. https://doi.org/10.18844/wjet.v14i4.6592

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education, 20*(1), 43. https://doi.org/10.1186/s41239-023-00411-8

Chang, C. Y., Kuo, S. Y., & Hwang, G. H. (2022). Chatbot-facilitated nursing education. *Educational Technology & Society, 25*(1), 15–27. https://www.jstor.org/stable/48647027

Chen, B., Bao, L., Zhang, R., Zhang, J., Liu, F., Wang, S., & Li, M. (2023). A multi-strategy computer-assisted EFL writing learning system with deep learning incorporated and its effects on learning: A writing feedback perspective. *Journal of Educational Computing Research, 61*(8), 60–102. https://doi.org/10.1177/07356331231189294

Chen, C. H., Koong, C. S., & Liao, C. (2022). Influences of integrating dynamic assessment into a speech recognition learning design to support students' English speaking skills, learning anxiety and cognitive load. *Educational Technology & Society, 25*(1), 1–14. https://www.jstor.org/stable/48647026

Chen, S. Y., Lin, P. H., & Chien, W. C. (2022). Children's digital art ability training system based on ai-assisted learning: a case study of drawing color perception. *Frontiers in Psychology*, *13*, Article 823078. https://doi.org/10.3389/fpsyg.2022.823078

Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two decades of artificial intelligence in education: Contributors, collaborations, research topics, challenges, and future directions. *Educational Technology & Society, 25*(1), 28–47. https://www-jstor-org.ezproxy.eduhk.hk/stable/48647028

Chen, Y. C. (2022). Effects of technology-enhanced language learning on reducing EFL learners' public speaking anxiety. *Computer Assisted Language Learning,* 1–25. https://doi.org/10.1080/09588221.2022.2055083

Chiu, M. C., Hwang, G. J., Hsia, L. H., & Shyu, F. M. (2022). Artificial intelligence-supported art education: A deep learning-based system for promoting university students' artwork appreciation and painting outcomes. *Interactive Learning Environments,* 1–19. https://doi.org/10.1080/10494820.2022.2100426

Chiu, T. K. F., 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

Cho, K., Foo, Y. M., Dalziel, B., & Hu, W. (2022). Chatbot-mediated learning of cardiac auscultation. *Internal Medicine Journal, 52*(12), 2176–2180. https://doi.org/10.1111/imj.15971

Chrysafiadi, K., Troussas, C., & Virvou, M. (2022). Personalised instructional feedback in a mobile-assisted language learning application using fuzzy reasoning. *International Journal of Learning Technology, 17*(1), 53–76. https://doi.org/10.1504/IJLT.2022.123676

Chrysafiadi, K., Virvou, M., Tsihrintzis, G. A., & Hatzilygeroudis, I. (2023). Evaluating the user's experience, adaptivity and learning outcomes of a fuzzy-based intelligent tutoring system for computer programming for academic students in Greece. *Education and Information Technologies, 28*(6), 6453–6483. https://doi.org/10.1007/s10639-022-11444-3

Chu, Y. S., Yang, H. C., Tseng, S. S., & Yang, C. C. (2014). Implementation of a model-tracing-based learning diagnosis system to promote elementary students' learning in mathematics. *Journal of Educational Technology & Society, 17*(2), 347–357. https://www.jstor.org/stable/jeductechsoci.17.2.347

Cobos, R., & Ruiz-Garcia, J. C. (2021). Improving learner engagement in MOOCs using a learning intervention system: A research study in engineering education. *Computer Applications in Engineering Education, 29*(4), 733–749. https://doi.org/10.1002/cae.22316

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20,* 37–46. https://doi.org/10.1177/001316446002000104

Cress, U., Rosé, C. P., Law, N., & Ludvigsen, S. (2019). Investigating the complexity of computer-supported collaborative learning in action. *International Journal of Computer-Supported Collaborative Learning, 14*(2), 137–142. https://doi.org/10.1007/s11412-019-09305-2

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Sage Publications, Inc.

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,* Article 104094. https://doi.org/10.1016/j.compedu.2020.104094

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*. arXiv. https://doi.org/10.48550/arXiv.1810.04805

Divekar, R., Drozdal, J., Chabot, S., Zhou, Y., Su, H., Chen, Y., Zhu, H., Hendler, J., & Braasch, J. (2022). Foreign language acquisition via artificial intelligence and extended reality: Design and evaluation. *Computer Assisted Language Learning, 35*(9), 2332–2360. https://doi.org/10.1080/09588221.2021.1879162

Easterday, M. W., Aleven, V., Scheines, R., & Carver, S. M. (2017). Using tutors to improve educational games: A cognitive game for policy argument. *Journal of the Learning Sciences*, *26*(2), 226–276. https://doi.org/10.1080/10508406.2016.1269287

Ebadi, S., & Amini, A. (2022). Examining the roles of social presence and human-likeness on Iranian EFL learners' motivation using artificial intelligence technology: A case of CSIEC chatbot. *Interactive Learning Environments,* 1–19. https://doi.org/10.1080/10494820.2022.2096638

Ericsson, E., & Johansson, S. (2023). English speaking practice with conversational AI: Lower secondary students' educational experiences over time. *Computers and Education: Artificial Intelligence, 5*, Article 100164. https://doi.org/10.1016/j.caeai.2023.100164

Farhat, F., Chaudry, B., Nadeem, M., Sohail, S., & Madsen, D. (2023). *Evaluating AI models for the national pre-medical exam in India: A head-to-head analysis of ChatGPT-3.5, GPT-4, and Bard (Preprint)*. https://doi.org/10.2196/preprints.51523

Fazlollahi, A. M., Yilmaz, R., Winkler-Schwartz, A., Mirchi, N., Ledwos, N., Bakhaidar, M., Alsayegh, A., & Del Maestro, R. F. (2023). AI in surgical curriculum design and unintended outcomes for technical competencies in simulation training. *JAMA Network Open*, *6*(9), Article e2334658. https://doi.org/10.1001/jamanetworkopen.2023.34658

Fazlollahi, A., Bakhaidar, M., Alsayegh, A., Yilmaz, R., Winkler-Schwartz, A., Mirchi, N., Langleben, I., Ledwos, N., Sabbagh, A., Bajunaid, K., Harley, J., & Del Maestro, R. (2022). Effect of artificial intelligence tutoring vs expert instruction on learning simulated surgical skills among medical students a randomized clinical trial. *JAMA Network Open*, *5*(2). https://doi.org/10.1001/jamanetworkopen.2021.49008

Feng, H.-H. (2022). Genre-based AWE system for engineering graduate writing: Development and evaluation. Language Learning & Technology, 26(2), 58-. https://doi.org/10.125/73479

Ferguson, P. (2011). Student perceptions of quality feedback in teacher education. *Assessment & Evaluation in Higher Education*, *36*(1), 51–62. https://doi.org/10.1080/02602930903197883

Fernandez-Aleman, J., Lopez-Gonzalez, L., Gonzalez-Sequeros, O., Jayne, C., Lopez-Jimenez, J., Carrillo-de-Gea, J., & Toval, A. (2016). An empirical study of neural network-based audience response technology in a human anatomy course for pharmacy students. *Journal of Medical Systems*, *40*(4). https://doi.org/10.1007/s10916-016-0440-6

Fernandez-Aleman, J., Lopez-Gonzalez, L., Gonzalez-Sequeros, O., Jayne, C., Lopez-Jimenez, J., & Toval, A. (2016). The evaluation of i-SIDRA - a tool for intelligent feedback—In a course on the anatomy of the locomotor system. *International journal of medical informatics*, *94*, 172–181. https://doi.org/10.1016/j.ijmedinf.2016.07.008

Fidan, M., & Gencel, N. (2022). Supporting the instructional videos with chatbot and peer feedback mechanisms in online learning: The effects on learning performance and intrinsic motivation. *Journal of Educational Computing Research*, *60*(7), 1716–1741. https://doi.org/10.1177/07356331221077901

Forsberg, E., Ziegert, K., Hult, H., & Fors, U. (2016). Assessing progression of clinical reasoning through virtual patients: An exploratory study. *Nurse Education in Practice*, *16*(1), 97–103. https://doi.org/10.1016/j.nepr.2015.09.006

Fu, Q.-K., Zou, D., Xie, H., & Cheng, G. (2024). A review of AWE feedback: Types, learning outcomes, and implications. *Computer Assisted Language Learning*, *37*(1–2), 179–221. https://doi.org/10.1080/09588221.2022.2033787

Fu, S., Gu, H., & Yang, B. (2020). The affordances of AI-enabled automatic scoring applications on learners' continuous learning intention: An empirical study in China. *British Journal of Educational Technology, 51*(5), 1674–1692. https://doi.org/10.1111/bjet.12995

Fui, C. S., & Lian, L. H. (2018). How do students perceived computerized feedback as effective? *Eurasia Journal of Mathematics, Science and Technology Education, 14*(6), 2669–2682. https://doi.org/10.29333/ejmste/90265

Furlan, R., Gatti, M., Mene, R., Shiffer, D., Marchiori, C., Levra, A., Saturnino, V., Brunetta, E., & Dipaola, F. (2021). A natural language processing-based virtual patient simulator and intelligent tutoring system for the clinical diagnostic process: Simulator development and case study. *JMIR Medical Informatics, 9*(4). https://doi.org/10.2196/24073

García-Jiménez, E., Gallego-Noche, B., & Gómez-Ruíz, M. Á. (2015). Feedback and self-regulated learning: How feedback can contribute to increase students' autonomy as learners. *Sustainable Learning in Higher Education: Developing Competencies for the Global Marketplace*, 113–130. https://doi.org/10.1007/978-3-319-10804-9_9

Glushkova, A., & Manitsaris, S. (2018). Gesture recognition and sensorimotor learning-by-doing of motor skills in manual professions: A case study in the wheel-throwing art of pottery. *Journal of Computer Assisted Learning, 34*(1), 20–31. https://doi.org/10.1111/jcal.12210

González-Calatayud, V., Prendes-Espinosa, P., & Roig-Vila, R. (2021). Artificial intelligence for student assessment: A systematic review. *Applied Sciences*, *11*(12), 5467.

Gonzalez-Carrillo, C., Restrepo-Calle, F., Ramirez-Echeverry, J., & Gonzalez, F. (2021). Automatic grading tool for Jupyter notebooks in artificial intelligence courses. *Sustainability, 13*(21). https://doi.org/10.3390/su132112050

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *arXiv*. https://doi.org/10.48550/arXiv.1406.2661

Guan, X., Feng, X., & Islam, A. Y. M. A. (2023). The dilemma and countermeasures of educational data ethics in the age of intelligence. *Humanities and Social Sciences Communications*, *10*(1), 1–14. https://doi.org/10.1057/s41599-023-01633-x

Guasch, T., Espasa, A., Alvarez, I. M., & Kirschner, P. A. (2013). Effects of feedback on collaborative writing in an online learning environment. *Distance Education*, *34*(3), 324–338. https://doi.org/10.1080/01587919.2013.835772

Guerrero-Roldan, A., Rodriguez-Gonzalez, M., Baneres, D., Elasri-Ejjaberi, A., & Cortadas, P. (2021). Experiences in the use of an adaptive intelligent system to enhance online learners' performance: A case study in economics and business courses. *International Journal of Educational Technology in Higher Education, 18*(1). https://doi.org/10.1186/s41239-021-00271-0

Hadwin, A., & Oshige, M. (2011). Self-regulation, coregulation, and socially shared regulation: Exploring perspectives of social in self-regulated learning theory. *Teachers College Record, 113*(2), 240–264. https://doi.org/10.1177/016146811111300204

Han, J., Kim, K. H., Rhee, W., & Cho, Y. H. (2021). Learning analytics dashboards for adaptive support in face-to-face collaborative argumentation. *Computers & Education, 163*, Article 104041. https://doi.org/10.1016/j.compedu.2020.104041

Han, J., Park, J., & Lee, H. (2022). Analysis of the effect of an artificial intelligence chatbot educational program on non-face-to-face classes: A quasi-experimental study. *BMC Medical Education, 22*(1). https://doi.org/10.1186/s12909-022-03898-3

Han, Y., Wu, W., Liang, Y., & Zhang, L. (2023). Peer grading eliciting truthfulness based on autograder. *IEEE Transactions on Learning Technologies, 16*(3), 353–363. https://doi.org/10.1109/TLT.2022.3216946

Harati, H., Sujo-Montes, L., Tu, C.-H., Armfield, S. J. W., & Yen, C.-J. (2021). Assessment and learning in knowledge spaces (Aleks) adaptive system impact on students' perception and self-regulated learning skills. *Education Sciences*, *11*(10). https://doi.org/10.3390/educsci11100603

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Henderson, M., Ajjawi, R., Boud, D., & Molloy, E. (2019). Designing Feedback for Impact. In M. Henderson, R. Ajjawi, D. Boud, & E. Molloy (Eds.), *The impact of feedback in higher education* (pp. 267–285). Springer International Publishing AG. https://doi.org/10.1007/978-3-030-25112-3_15

Hooda, M., Rana, C., Dahiya, O., Rizwan, A., & Hossain, M. S. (2022). Artificial intelligence for assessment and feedback to enhance student success in higher education. *Mathematical Problems in Engineering, 2022*, 1–19. https://doi.org/10.1155/2022/5215722

Hopcan, S., Polat, E., Ozturk, M. E., & Ozturk, L. (2023). Artificial intelligence in special education: A systematic review. *Interactive Learning Environments, 31*(10), 7335–7353.

Hsia, L., Hwang, G., & Hwang, J. (2023). AI-facilitated reflective practice in physical education: An auto-assessment and feedback approach. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2023.2212712

Hsu, T., Huang, H., Hwang, G., & Chen, M. (2023). Effects of incorporating an expert decision-making mechanism into chatbots on students? Achievement, enjoyment, and anxiety. *Educational Technology & Society*, *26*(1), 218–231. https://doi.org/10.30191/ETS.202301_26(1).0016

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

Huang, W., Hew, K. F., & Gonda, D. E. (2019). Designing and evaluating three chatbot-enhanced activities for a flipped graduate course. *International Journal of Mechanical Engineering and Robotics Research*, *8*(5), 813–818. https://doi.org/10.18178/ijmerr.8.5.813-818

Hwang, G.-J., & Chen, N.-S. (2023). Editorial position paper: Exploring the potential of generative artificial intelligence in education: Applications, challenges, and future research directions. *Educational Technology & Society*, *26*(2). https://www.jstor.org/stable/48720991

Ijaz, K., Bogdanovych, A., & Trescak, T. (2017). Virtual worlds vs books and videos in history education. *Interactive Learning Environments*, *25*(7), 904–929. https://doi.org/10.1080/10494820.2016.1225099

Ingkavara, T., Panjaburee, P., Srisawasdi, N., & Sajjapanroj, S. (2022). The use of a personalized learning approach to implementing self-regulated online learning. *Computers and Education: Artificial Intelligence*, *3*. https://doi.org/10.1016/j.caeai.2022.100086

Jemeli, C. M., & Fakandu, A. M. (2019). Equitable access to education and development in a knowledgeable society as advocated by UNESCO. *Educational Research and Reviews*, *14*(6), 200–205. https://doi.org/10.5897/ERR2018.3647

Jeon, J. (2021). Chatbot-assisted dynamic assessment (CA-DA) for L2 vocabulary learning and diagnosis. *Computer Assisted Language Learning, 36*(7), 1338–1364. https://doi.org/10.1080/09588221.2021.1987272

Jeon, J. (2022). Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2021.2021241

Jiang, M., Jong, M., Lau, W., Chai, C., & Wu, N. (2023). Exploring the effects of automatic speech recognition technology on oral accuracy and fluency in a flipped classroom. *Journal of Computer Assisted Learning, 39*(1), 125–140. https://doi.org/10.1111/jcal.12732

Kamardeen, I. (2014). Adaptive e-tutorial for enhancing student learning in construction education. *International Journal of Construction Education and Research, 10*(2), 79–95. https://doi.org/10.1080/15578771.2012.756437

Kamel, A., Liu, B., Li, P., & Sheng, B. (2019). An investigation of 3D human pose estimation for learning tai chi: A human factor perspective. *International Journal of Human–Computer Interaction, 35*(4–5), 427–439. https://doi.org/10.1080/10447318.2018.1543081

Kamnardsiri, T., Hongsit, L.-O., Khuwuthyakorn, P., & Wongta, N. (2017). The effectiveness of the game-based learning system for the improvement of American Sign Language using kinect. *Electronic Journal of E-Learning, 15*(4), 283–296. https://10.1080/10447318.2018.1543081

Kochmar, E., Vu, D. D., Belfer, R., Gupta, V., Serban, I. V., & Pineau, J. (2022). Automated data-driven generation of personalized pedagogical interventions in intelligent tutoring systems. *International Journal of Artificial Intelligence in Education, 32*(2), 323–349. https://doi.org/10.1007/s40593-021-00267-x

Koć-Januchta, M. M., Schönborn, K. J., Tibell, L. A., Chaudhri, V. K., & Heller, H. C. (2020). Engaging with biology by asking questions: Investigating students' interaction and learning with an artificial intelligence-enriched textbook. *Journal of Educational Computing Research, 58*(6), 1190–1224. https://doi.org/10.1177/0735633120921581

Labuhn, A. S., Zimmerman, B. J., & Hasselhorn, M. (2010). Enhancing students' self-regulation and mathematics performance: The influence of feedback and self-evaluative standards. *Metacognition and Learning, 5*(2), 173–194. https://doi.org/10.1007/s11409-010-9056-2

Lang, Y., Xie, K., Gong, S., Wang, Y., & Cao, Y. (2022). The impact of emotional feedback and elaborated feedback of a pedagogical agent on multimedia learning. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.810194

Law, L. (2024). Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. *Computers and Education Open, 6*, 100174-. https://doi.org/10.1016/j.caeo.2024.100174

Leddo, J., Kindi, R., Bhandarkar, S., Chadeva, N., Ganotra, K., Jayakumar, P., & Somaiya, Y. (2018). Next-generation interactive, educational television: Using artificial intelligence and the internet to customize instruction to student learning needs. *Journal of Educational Multimedia and Hypermedia, 27*(1), 103–114. https://www.learntechlib.org/primary/p/178442/

Lee, H.-S., Gweon, G.-H., Lord, T., Paessel, N., Pallant, A., & Pryputniewicz, S. (2021). Machine learning-enabled automated feedback: Supporting students' revision of scientific arguments based on data drawn from simulation. *Journal of Science Education and Technology, 30*(2), 168–192. https://doi.org/10.1007/s10956-020-09889-7

Lee, Y.-F., Hwang, G.-J., & Chen, P.-Y. (2022). Impacts of an AI-based chatbot on college students' after-class review, academic performance, self-efficacy, learning attitude, and motivation. *Educational Technology Research and Development, 70*(5), 1843–1865. https://doi.org/10.1007/s11423-022-10142-8

Lei, C.-U., Man, K. L., & Ting, T. O. (2014). Using learning analytics to analyze writing skills of students: A case study in a technological common core curriculum course. *IAENG International Journal of Computer Science, 41*(3), 193–197. https://www.iaeng.org/IJCS/issues_v41/issue_3/IJCS_41_3_05.pdf

Li, K.-C., Chang, M., & Wu, K.-H. (2020). Developing a task-based dialogue system for English language learning. *Education Sciences, 10*(11), 1–20. https://doi.org/10.3390/educsci10110306

Li, W. (2022). Analysis of piano performance characteristics by deep learning and artificial intelligence and its application in piano teaching. *Frontiers in Psychology, 12*, Article 751406. https://doi.org/10.3389/fpsyg.2021.751406

Lin, M., & Chang, D. (2020). Enhancing post-secondary writers' writing skills with a chatbot: A mixed-method classroom study. *Journal of Educational Technology & Society*, *23*(1), 78–92. https://www.jstor.org/stable/26915408

Liu, C., Hou, J., Tu, Y., Wang, Y., & Hwang, G. (2021). Incorporating a reflective thinking promoting mechanism into artificial intelligence-supported English writing environments. *Interactive Learning Environments, 31*(9), 5614–5632. https://doi.org/10.1080/10494820.2021.2012812

Liu, C., Liu, S., Hwang, G., Tu, Y., Wang, Y., & Wang, N. (2023). Engaging EFL students' critical thinking tendency and in-depth reflection in technology-based writing contexts: A peer assessment-incorporated automatic evaluation approach. *Education* *and Information Technologies, 28*(10), 13027–13052. https://doi.org/10.1007/s10639-023-11697-6

Liu, J., Liu, X., & Yang, C. (2022). A study of college students' perceptions of utilizing automatic speech recognition technology to assist English oral proficiency. *Frontiers in Psychology, 13,* Article 1049139. https://doi.org/10.3389/fpsyg.2022.1049139

Liu, M., Li, Y., Xu, W., & Liu, L. (2017). Automated essay feedback generation and its impact on revision. *IEEE Transactions on Learning Technologies, 10*(4), 502–513. https://doi.org/10.1109/TLT.2016.2612659

Loftus, N., & Narman, H. S. (2023). Use of machine learning in interactive cybersecurity and network education. *Sensors, 23*(6). https://doi.org/10.3390/s23062977

Lu, X. (2019). An empirical study on the artificial intelligence writing evaluation system in China CET. *Big Data, 7*(2), 121–129. https://doi.org/10.1089/big.2018.0151

Ma, Z.-H., Hwang, W.-Y., & Shih, T. K. (2020). Effects of a peer tutor recommender system (PTRS) with machine learning and automated assessment on vocational high school students' computer application operating skills. *Journal of Computers in Education, 7*(3), 435–462. https://doi.org/10.1007/s40692-020-00162-9

Mahrous, A., Botsko, D. L., Elgreatly, A., Tsujimoto, A., Qian, F., & Schneider, G. B. (2023). The use of artificial intelligence and game-based learning in removable partial denture design: A comparative study. *Journal of Dental Education, 87*(8), 1188–1199. https://doi.org/10.1002/jdd.13225

Maier, U., & Klotz, C. (2022). Personalized feedback in digital learning environments: Classification framework and literature review. *Computers and Education: Artificial Intelligence, 3,* Article 100080. https://doi.org/10.1016/j.caeai.2022.100080

Mayordomo, R. M., Espasa, A., Guasch, T., & Martínez-Melo, M. (2022). Perception of online feedback and its impact on cognitive and emotional engagement with feedback.

Education and Information Technologies, *27*(6), 7947–7971. https://doi.org/10.1007/s10639-022-10948-2

Mirchi, N., Bissonnette, V., Yilmaz, R., Ledwos, N., Winkler-Schwartz, A., & Del Maestro, R. (2020). The virtual operative assistant: An explainable artificial intelligence tool for simulation-based training in surgery and medicine. *PloS ONE, 15*(2). https://doi.org/10.1371/journal.pone.0229596

Mokmin, N. A. M. (2020). The effectiveness of a personalized virtual fitness trainer in teaching physical education by applying the artificial intelligent algorithm. *International Journal of Human Movement and Sports Sciences*, *8*(5), 258–264. https://doi.org/10.13189/saj.2020.080514

Molenaar, I. (2022). The concept of hybrid human-AI regulation: Exemplifying how to support young learners' self-regulated learning. *Computers and Education: Artificial Intelligence, 3*, Article 100070. https://doi.org/10.1016/j.caeai.2022.100070

Mørch, A. I., Engeness, I., Cheng, V. C., Cheung, W. K., & Wong, K. C. (2017). EssayCritic: Writing to learn with a knowledge-based design critiquing system. *Educational Technology and Society, 20*(2), 213–223. https://www.jstor.org/stable/90002176

Morris, R., Perry, T., & Wardle, L. (2021). Formative assessment and feedback for learning in higher education: A systematic review. *Review of Education*, *9*(3), Article e3292. https://doi.org/10.1002/rev3.3292

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

Nalli, G., Culmone, R., Perali, A., & Amendola, D. (2023). Online tutoring system for programming courses to improve exam pass rate. *Journal of E-Learning and Knowledge Society, 19*(1), 27–35. https://doi.org/10.20368/1971-8829/1135704

Nazari, N., Shabbir, M. S., & Setiawan, R. (2021). Application of Artificial Intelligence powered digital writing assistant in higher education: Randomized controlled trial. *Heliyon, 7*(5). https://doi.org/10.1016/j.heliyon.2021.e07014

Nelekar, S., Abdulrahman, A., Gupta, M., & Richards, D. (2022). Effectiveness of embodied conversational agents for managing academic stress at an Indian University (ARU) during COVID-19. *British Journal of Educational Technology, 53*(3), 491–511. https://doi.org/10.1111/bjet.13174

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

Neumann, A. T., Arndt, T., Köbis, L., Meissner, R., Martin, A., de Lange, P., Pengel, N., Klamma, R., & Wollersheim, H.-W. (2021). Chatbots as a tool to scale mentoring processes: Individually supporting self-study in higher education. *Frontiers in Artificial Intelligence, 4*, Article 668220. https://doi.org/10.3389/frai.2021.668220

Nye, B. D., Graesser, A. C., & Hu, X. (2014). Autotutor and family: A review of 17 years of natural language tutoring. *Grantee Submission, 24*, 427–469. https://doi.org/10.1007/s40593-014-0029-5

Ochoa, X., & Dominguez, F. (2020). Controlled evaluation of a multimodal system to improve oral presentation skills in a real learning setting. *British Journal of Educational Technology, 51*(5), 1615–1630. https://doi.org/10.1111/bjet.12987

Ouyang, F., Wu, M., Zheng, L., Zhang, L., & Jiao, P. (2023). Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course. *International Journal of Educational Technology in Higher Education, 20*(1). https://doi.org/10.1186/s41239-022-00372-4

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Pérez-Marín, D., Hijón-Neira, R., & Santacruz, L. (2016). Active learning through collaborative knowledge building using an automatic free-text scoring system in a b-learning environment. *Behaviour and Information Technology, 35*(7), 572–585. https://doi.org/10.1080/0144929X.2016.1171398

Perikos, I., Grivokostopoulou, F., & Hatzilygeroudis, I. (2017). Assistance and feedback mechanism in an intelligent tutoring system for teaching conversion of natural language into logic. *International Journal of Artificial Intelligence in Education, 27*(3), 475–514. https://doi.org/10.1007/s40593-017-0139-y

Prescott, J., Ogilvie, L., & Hanley, T. (2023). Student therapists' experiences of learning using a machine client: A proof-of-concept exploration of an emotionally responsive interactive client (ERIC). *Counselling and Psychotherapy Research*. https://doi.org/10.1002/capr.12685

Qi, D., Ryason, A., Milef, N., Alfred, S., Abu-Nuwar, M. R., Kappus, M., De, S., & Jones, D. B. (2021). Virtual reality operating room with AI guidance: Design and validation of a fire scenario. *Surgical Endoscopy, 35*(2), 779–786. https://doi.org/10.1007/s00464-020-07447-1

Rad, H. S., Alipour, R., & Jafarpour, A. (2023). Using artificial intelligence to foster students' writing feedback literacy, engagement, and outcome: A case of Wordtune application. *Interactive Learning Environments*, 1-21. https://doi.org/10.1080/10494820.2023.2208170

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving language understanding by generative pre-training*. https://openai.com/index/language-unsupervised/

Rodriguez-Arrastia, M., Martinez-Ortigosa, A., Ruiz-Gonzalez, C., Ropero-Padilla, C., Roman, P., & Sanchez-Labraca, N. (2022). Experiences and perceptions of final-year nursing students of using a chatbot in a simulated emergency situation: A qualitative study. *Journal of Nursing Management, 30*(8), 3874–3884. https://doi.org/10.1111/jonm.13630

Røed, R. K., Baugerud, G. A., Hassan, S. Z., Sabet, S. S., Salehi, P., Powell, M. B., Riegler, M. A., Halvorsen, P., & Johnson, M. S. (2023). Enhancing questioning skills through child avatar chatbot training with feedback. *Frontiers in Psychology, 14*. https://doi.org/10.3389/fpsyg.2023.1198235

Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. International Journal of Artificial Intelligence in Education, 26(2), 582–599. https://doi.org/10.1007/s40593-016-0110-3

Ryan, T., & Henderson, M. (2018). Feeling feedback: Students' emotional responses to educator feedback. *Assessment & Evaluation in Higher Education, 43*(6), 880–892. https://doi.org/10.1080/02602938.2017.1416456

Sailer, M., Bauer, E., Hofmann, R., Kiesewetter, J., Glas, J., Gurevych, I., & Fischer, F. (2023). Adaptive feedback from artificial neural networks facilitates pre-service teachers' diagnostic reasoning in simulation-based learning. *Learning and Instruction, 83*. https://doi.org/10.1016/j.learninstruc.2022.101620

Sáiz-Manzanares, M. C., Marticorena-Sánchez, R., Martín-Antón, L. J., González Díez, I., & Almeida, L. (2023). Perceived satisfaction of university students with the use of chatbots as a tool for self-regulated learning. *Heliyon, 9*(1). https://doi.org/10.1016/j.heliyon.2023.e12843

Salas-Pilco, S. Z., Xiao, K., & Hu, X. (2022). Artificial intelligence and learning analytics in teacher education: A systematic review. *Education Sciences, 12*(8), 569.

Samarakou, M., Fylladitakis, E. D., Karolidis, D., Früh, W.-G., Hatziapostolou, A., Athinaios, S. S., & Grigoriadou, M. (2016). Evaluation of an intelligent open learning system for engineering education. *Knowledge Management and E-Learning, 8*(3), 496–513. https://doi.org/10.34105/j.kmel.2016.08.031

Sengar, S.S., Hasan, A.B., Kumar, S. *et al.* Generative artificial intelligence: a systematic review and applications. *Multimed Tools Appl* (2024). https://doi-org.ezproxy.eduhk.hk/10.1007/s11042-024-20016-1

Shaffer, D. W., Collier, W., & Ruis, A. R. (2016). A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *Journal of Learning Analytics, 3*(3), 9–45. https://doi.org/10.18608/jla.2016.33.3

Shah, C., Davtyan, K., Nasrallah, I., Bryan, R. N., & Mohan, S. (2023). Artificial intelligence-powered clinical decision support and simulation platform for radiology trainee education. *Journal of Digital Imaging, 36*(1), 11–16. https://doi.org/10.1007/s10278-022-00713-9

Shaik, T., Tao, X., Li, Y., Dann, C., McDonald, J., Redmond, P., & Galligan, L. (2022). A review of the trends and challenges in adopting natural language processing methods for education feedback analysis. *IEEE Access, 10*, 56720–56739. https://doi.org/10.1109/ACCESS.2022.3177752

Shaikh, S., Yayilgan, S. Y., Klimova, B., & Pikhart, M. (2023). Assessing the usability of ChatGPT for formal English language learning. *European Journal of Investigation in Health, Psychology and Education, 13*(9), 1937–1960. https://doi.org/10.3390/ejihpe13090140

Sheikh, H., Prins, C., & Schrijvers, E. (2023). Artificial intelligence: Definition and background. In H. Sheikh, C. Prins, & E. Schijvers (Eds.), *Mission AI: The new system technology* (pp. 15–41). Springer International Publishing. https://doi.org/10.1007/978-3-031-21448-6

Shi, H., & Aryadoust, V. (2024). A systematic review of AI-based automated written feedback research. *ReCALL*, 1–23. https://doi.org/10.1017/S0958344023000265

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

Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and disaffection in the classroom: Part of a larger motivational dynamic? *Journal of Educational Psychology, 100*(4), 765–781. https://doi.org/10.1037/a0012840

Smith, G. G., Haworth, R., & Žitnik, S. (2020). Computer science meets education: Natural language processing for automatic grading of open-ended questions in eBooks. *Journal of Educational Computing Research, 58*(7), 1227–1255. https://doi.org/10.1177/0735633120927486

Stefanidis, K., Psaltis, A., Apostolakis, K. C., Dimitropoulos, K., & Daras, P. (2019). Learning prosocial skills through multiadaptive games: A case study. *Journal of Computers in Education, 6*(1), 167–190. https://doi.org/10.1007/s40692-019-00134-8

Steif, P., Fu, L., & Kara, L. (2016). Providing formative assessment to students solving multipath engineering problems with complex arrangements of interacting parts: An intelligent tutor approach. *Interactive Learning Environments, 24*(8), 1864–1880. https://doi.org/10.1080/10494820.2015.1057745

Stevenson, M., & Phakiti, A. (2014). The effects of computer-generated feedback on the quality of writing. *Assessing Writing, 19*(2014), 51–65. https://doi.org/10.1016/j.asw.2013.11.007

Stojanov, A. (2023). Learning with ChatGPT 3.5 as a more knowledgeable other: An autoethnographic study. *International Journal of Educational Technology in Higher Education, 20*(1). https://doi.org/10.1186/s41239-023-00404-7

Sun, L., Yin, C., Xu, Q., & Zhao, W. (2023). Artificial intelligence for healthcare and medical education: A systematic review. *American Journal of Translational Research, 15*(7), 4820.

Sweidan, S. Z., Abu Laban, S. S., Alnaimat, N. A., & Darabkh, K. A. (2021). SIAAA-C: A student interactive assistant android application with chatbot during COVID-19 pandemic. *Computer Applications in Engineering Education, 29*(6), 1718–1742. https://doi.org/10.1002/cae.22419

Tanana, M., Soma, C., Srikumar, V., Atkins, D., & Imel, Z. (2019). Development and evaluation of ClientBot: Patient-like conversational agent to train basic counseling skills. *Journal of Medical Internet Research, 21*(7). https://doi.org/10.2196/12529

Tärning, B., & Silvervarg, A. (2019). "I didn't understand, I'm really not very smart"—How design of a digital tutee's self-efficacy affects conversation and student behavior in a digital math game. *Education Sciences, 9*(3). https://doi.org/10.3390/educsci9030197

Taskiran, A., & Goksel, N. (2022). Automated feedback and teacher feedback: Writing achievement in learning English as a foreign language at a distance. *Turkish Online*

Journal of Distance Education, 23(2), 120–139. https://doi.org/10.17718/tojde.1096260

Vanichvasin, P. (2021). Chatbot development as a digital learning tool to increase students' research knowledge. *International Education Studies, 14*(2), 44–53. https://doi:10.5539/ies.v14n2p44

Vázquez-Cano, E., Mengual-Andrés, S., & López-Meneses, E. (2021). Chatbot to improve learning punctuation in Spanish and to enhance open and flexible learning environments. *International Journal of Educational Technology in Higher Education, 18*, 1–20. https://doi.org/10.1186/s41239-021-00269-8

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

Vollmer, A.-L., Mühlig, M., Steil, J. J., Pitsch, K., Fritsch, J., Rohlfing, K. J., & Wrede, B. (2014). Robots show us how to teach them: Feedback from robots shapes tutoring behavior during action learning. *PLoS ONE, 9*(3). https://doi.org/10.1371/journal.pone.0091349

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

Wang, D., Han, H., Zhan, Z., Xu, J., Liu, Q., & Ren, G. (2015). A problem solving oriented intelligent tutoring system to improve students' acquisition of basic computer skills. *Computers & Education, 81*, 102–112. https://doi.org/10.1016/j.compedu.2014.10.003

Wang, Q., Rose, C., Ma, N., Jiang, S., Bao, H., & Li, Y. (2022). Design and application of automatic feedback scaffolding in forums to promote learning. *IEEE Transactions on Learning Technologies, 15*(2), 150–166. https://doi.org/10.1109/TLT.2022.3156914

Wang, X., Liu, Q., Pang, H., Tan, S. C., Lei, J., Wallace, M. P., & Li, L. (2023). What matters in AI-supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis. *Computers & Education, 194*, 104703. https://doi.org/10.1016/j.compedu.2022.104703

Wang, X., Zhang, L., & He, T. (2022). Learning performance prediction-based personalized feedback in online learning via machine learning. *Sustainability*, *14*(13). https://doi.org/10.3390/su14137654

Wang, Z. (2022). Computer-assisted EFL writing and evaluations based on artificial intelligence: A case from a college reading and writing course. *Library Hi Tech, 40*(1), 80–97. https://doi.org/10.1108/LHT-05-2020-0113

Wang, Z., Gong, S.-Y., Xu, S., & Hu, X.-E. (2019). Elaborated feedback and learning: Examining cognitive and motivational influences. *Computers & Education, 136*, 130–140. https://doi.org/10.1016/j.compedu.2019.04.003

Williams, R., Ali, S., Devasia, N., DiPaola, D., Hong, J., Kaputsos, S. P., Jordan, B., & Breazeal, C. (2023). AI + ethics curricula for middle school youth: Lessons learned from three project-based curricula. *International Journal of Artificial Intelligence in Education, 33*(2), 325–383. https://doi.org/10.1007/s40593-022-00298-y

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology, 10*. https://doi.org/10.3389/fpsyg.2019.03087

Wolf, R. R., & Wolf, A. B. (2023). Using AI to evaluate a competency-based online writing course in nursing. *Online Learning Journal, 27*(3), 41–69. https://doi.org/10.24059/olj.v27i3.3974

Wong, L.-H., & Looi, C.-K. (2011). What seams do we remove in mobile-assisted seamless learning? A critical review of the literature. *Computers & Education, 57*(4), 2364–2381. https://doi.org/10.1016/j.compedu.2011.06.007

Xiao, Y., & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages, 8*(3). https://doi.org/10.3390/languages8030212

Xie, T., Liu, R., Chen, Y., & Liu, G. (2021). MOCA: A motivational online conversational agent for improving student engagement in collaborative learning. *IEEE Transactions on Learning Technologies, 14*(5), 653–664. https://doi.org/10.1109/TLT.2021.3129800

Xu, Y., Wang, D., Collins, P., Lee, H., & Warschauer, M. (2021). Same benefits, different communication patterns: Comparing children's reading with a conversational agent vs. A human partner. *Computers and Education, 161*. https://doi.org/10.1016/j.compedu.2020.104059

Yamada, M., Shimada, A., Okubo, F., Oi, M., Kojima, K., & Ogata, H. (2017). Learning analytics of the relationships among self-regulated learning, learning behaviors, and learning performance. *Research and Practice in Technology Enhanced Learning, 12*(1), 13. https://doi.org/10.1186/s41039-017-0053-9

Yamamoto, S., Tobe, Y., Tawatsuji, Y., & Hirashima, T. (2023). In-process feedback by detecting deadlock based on EEG data in exercise of learning by problem-posing and its evaluation. *Research and Practice in Technology Enhanced Learning, 18*. https://doi.org/10.58459/rptel.2023.18028

Yang, E., & Dorneich, M. C. (2018). Affect-aware adaptive tutoring based on human–automation etiquette strategies. *Human Factors*, *60*(4), 510–526. https://doi.org/10.1177/0018720818765266

Yang, H., Gao, C., & Shen, H. (2023). Learner interaction with, and response to, AI-programmed automated writing evaluation feedback in EFL writing: An exploratory study. *Education and Information Technologies*, 1–22. https://doi.org/10.1007/s10639-023-11991-3

Yang, Q., Lian, L., & Zhao, J. (2023). Developing a gamified artificial intelligence educational robot to promote learning effectiveness and behavior in laboratory safety courses for undergraduate students. *International Journal of Educational Technology in Higher Education*, *20*(1), 18. https://doi.org/10.1186/s41239-023-00391-9

Yang, Z., Wu, J. G., & Xie, H. (2024). Taming Frankenstein's monster: Ethical considerations relating to generative artificial intelligence in education. *Asia Pacific Journal of Education*, *00*, 1–14. https://doi.org/10.1080/02188791.2023.2300137

Yannier, N., Hudson, S. E., & Koedinger, K. R. (2020). Active learning is about more than hands-on: A mixed-reality AI system to support STEM education. *International Journal of Artificial Intelligence in Education*, *30*(1), 74–96. https://doi.org/10.1007/s40593-020-00194-3

Yu, E. (2023). Intelligent enough? Artificial intelligence for online learners. *Journal of Educators Online*, *20*(1). https://doi.org/10.9743/JEO.2023.20.1.16

Yu, S.-J., Hsueh, Y.-L., Sun, J. C.-Y., & Liu, H.-Z. (2021). Developing an intelligent virtual reality interactive system based on the ADDIE model for learning pour-over coffee brewing. *Computers and Education: Artificial Intelligence, 2*. https://doi.org/10.1016/j.caeai.2021.100030

Zahid Iqbal, M., & Campbell, A. G. (2023). AGILEST approach: Using machine learning agents to facilitate kinesthetic learning in STEM education through real-time touchless hand interaction. *Telematics and Informatics Reports, 9*. https://doi.org/10.1016/j.teler.2022.100034

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0

Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., Liu, J.-B., Yuan, J., & Li, Y. (2021). A review of artificial intelligence (AI) in education from 2010 to 2020. *Complexity, 2021*, Article e8812542. https://doi.org/10.1155/2021/8812542

Zhan, Y., Wan, Z. H., & Sun, D. (2022). Online formative peer feedback in Chinese contexts at the tertiary Level: A critical review on its design, impacts and influencing factors. *Computers & Education*, *176*, 104341. https://doi.org/10.1016/j.compedu.2021.104341

Zhang, Z., & Xu, L. (2022). Student engagement with automated feedback on academic writing: A study on Uyghur ethnic minority students in China. *Journal of Multilingual and Multicultural Development*, 1–14. https://doi.org/10.1080/01434632.2022.2102175

Zhao, C. (2024). AI-assisted assessment in higher education: A systematic review. *Journal of Educational Technology and Innovation, 6*(4).

Zheng, L., Fan, Y., Huang, Z., & Gao, L. (2023). Impacts of three approaches on collaborative knowledge building, group performance, behavioural engagement, and socially shared regulation in online collaborative learning. *Journal of Computer Assisted Learning*. https://doi.org/10.1111/jcal.12860

Zheng, L., Long, M., Chen, B., & Fan, Y. (2023). Promoting knowledge elaboration, socially shared regulation, and group performance in collaborative learning: An automated assessment and feedback approach based on knowledge graphs. *International Journal of Educational Technology in Higher Education, 20*(1), 46. https://doi.org/10.1186/s41239-023-00415-4

Zheng, L., Long, M., Niu, J., & Zhong, L. (2023). An automated group learning engagement analysis and feedback approach to promoting collaborative knowledge building, group performance, and socially shared regulation in CSCL. *International Journal of Computer-Supported Collaborative Learning, 18*(1), 101–133. https://doi.org/10.1007/s11412-023-09386-0

Zheng, L., Niu, J., & Zhong, L. (2022). Effects of a learning analytics-based real-time feedback approach on knowledge elaboration, knowledge convergence, interactive relationships and group performance in CSCL. *British Journal of Educational Technology, 53*(1), 130–149. https://doi.org/10.1111/bjet.13156

Zheng, L., Niu, J., Long, M., & Fan, Y. (2023). An automatic knowledge graph construction approach to promoting collaborative knowledge building, group performance, social interaction and socially shared regulation in CSCL. *British Journal of Educational Technology, 54*(3), 686–711. https://doi.org/10.1111/bjet.13283

Zheng, L., Zhong, L., & Niu, J. (2022). Effects of personalised feedback approach on knowledge building, emotions, co-regulated behavioural patterns and cognitive load in online collaborative learning. *Assessment & Evaluation in Higher Education, 47*(1), 109–125. https://doi.org/10.1080/02602938.2021.1883549

Zheng, L., Zhong, L., Niu, J., Long, M., & Zhao, J. (2021). Effects of personalized intervention on collaborative knowledge building, group performance, socially shared metacognitive regulation, and cognitive load in computer supported collaborative learning. *Educational Technology and Society, 24*(3), 174–193. https://www.jstor.org/stable/27032864

Zhu, M., Liu, O., & Lee, H. (2020). The effect of automated feedback on revision behavior and learning gains in formative assessment of scientific argument writing. *Computers & Education, 143*. https://doi.org/10.1016/j.compedu.2019.103668

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*(1), 166–183. https://doi.org/10.3102/0002831207312909

Zou, B., Du, Y., Wang, Z., Chen, J., & Zhang, W. (2023). An investigation into artificial intelligence speech evaluation programs with automatic feedback for developing EFL learners' speaking skills. *Sage Open, 13*(3). https://doi.org/10.1177/21582440231193818

Zou, S. (2017). Designing and practice of a college English teaching platform based on artificial intelligence. *Journal of Computational and Theoretical Nanoscience, 14*(1), 104–108. https://doi.org/10.1166/jctn.2017.6133

Appendix A. Search term list

Artificial Intelligence	**Feedback**	**Education**
Machine learning	Recommendation	Teaching
Deep learning		Instruction
Neural network		Teacher
Intelligent tutor		Instructor
Expert system		Learner
Chatbot		Student

Search query (formatted for the Web of Science)

((((((TS=("artificial intelligence")) OR TS=("machine learning")) OR TS=("deep learning")) OR TS=("neural network")) OR TS=("intelligent tutor")) OR TS=("expert system")) OR TS=("chatbot")
AND
(TS=("feedback")) OR TS=("recommendation")
AND
((((((TS=("education")) OR TS=("teaching")) OR TS=("instruction")) OR TS=("teacher")) OR TS=("instructor")) OR TS=("learner")) OR TS=("student")

Appendix B. The reviewed research output (RO)

RO (Year)	Author (Last names ranked from A to Z) (See full citations via the reference list)
RO1 (2014)	Adamson, D.; Dyke, G.; Jang, H.; Rosé, C. P.
RO2 (2023)	Afzaal, M.; Zia, A.; Nouri, J.; Fors, U.
RO3 (2021)	Ali, R.; Hoque, E.; Duberstein, P.; Schubert, L.; Razavi, S. Z.; Kane, B.; Silva, C.; Daks, J. S.; Huang, M. G.; Van Orden, K.
RO4 (2018)	Benotti, L.; Martinez, M. C.; Schapachnik, F.
RO5 (2022)	Bernius, J. P.; Krusche, S.; Bruegge, B.
RO6 (2023)	Biswas, U.; Bhattacharya, S.
RO7 (2020)	Bonneton-Botte, N.; Fleury, S.; Girard, N.; Le Magadou, M.; Cherbonnier, A.; Renault, M.; Anquetil, E.; Jamet, E.
RO8 (2023)	Castellano, M. S.; Contreras-McKay, I.; Neyem, A.; Farfán, E.; Inzunza, O.; Ottone, N. E.; del Sol, M.; Alario-Hoyo, C.; Alvarado, M. S.; Tubbs, R. S.
RO9 (2022)	Chaiprasurt, C.; Amornchewin, R.; Kunpitak, P.
RO10 (2022)	Chang, C. Y.; Kuo, S. Y.; Hwang, G. H.
RO11 (2023)	Chen, B. B.; Bao, L. N.; Zhang, R.; Zhang, J. Y.; Liu, F.; Wang, S.; Li, M. J.
RO12 (2022)	Chen, C. -H.; Koong, C. -S.; Liao, C.
RO13 (2022)	Chen, S. Y.; Lin, P. H.; Chien, W. C.
RO14 (2022)	Chen, Y. C.
RO15 (2022)	Chiu, M. C.; Hwang, G. J.; Hsia, L. H.; Shyu, F. M.
RO16 (2022)	Cho, K.; Foo, Y. M.; Dalziel, B.; Hu, W.
RO17 (2022)	Chrysafiadi, K.; Troussas, C.; Virvou, M.
RO18 (2023)	Chrysafiadi, K.; Virvou, M.; Tsihrintzis, G. A.; Hatzilygeroudis, I.
RO19 (2014)	Chu, Y. S.; Yang, H. C.; Tseng, S. S.; Yang, C. C.
RO20 (2021)	Cobos, R.; Ruiz-Garcia, J. C.
RO21 (2022)	Divekar, R. R.; Drozdal, J.; Chabot, S.; Zhou, Y. L.; Su, H.; Chen, Y.; Zhu, H. M.; Hendler, J. A.; Braasch, J.
RO22 (2017)	Easterday, M. W.; Aleven, V.; Scheines, R.; Carver, S. M.
RO23 (2022)	Ebadi, S.; Amini, A.
RO24 (2023)	Ericsson, E.; Johansson, S.
RO25 (2023)	Fazlollahi, A. M.; Yilmaz, R.; Winkler-Schwartz, A.; Mirchi, N.; Ledwos, N.; Bakhaidar, M.; Alsayegh, A.; Del Maestro, R. F.
RO26 (2022)	Fazlollahi, A. M.; Bakhaidar, M.; Alsayegh, A.; Yilmaz, R.; Winkler-Schwartz, A.; Mirchi, N.; Langleben, I.; Ledwos, N.; Sabbagh, A. J.; Bajunaid, K.; Harley, J. M.; Del Maestro, R. F.
RO27 (2016)	Fernandez-Aleman, J. L.; Lopez-Gonzalez, L.; Gonzalez-Sequeros, O.; Jayne, C.; Lopez-Jimenez, J. J.; Carrillo-de-Gea, J. M.; Toval, A.
RO28 (2016)	Fernandez-Aleman, J. L.; Lopez-Gonzalez, L.; Gonzalez-Sequeros, O.; Jayne, C.; Lopez-Jimenez, J. J.; Toval, A.
RO29 (2022)	Fidan, M.; Gencel, N.
RO30 (2016)	Forsberg, E.; Ziegert, K.; Hult, H.; Fors, U.
RO31 (2020)	Fu, S. X.; Gu, H. M.; Yang, B.
RO32 (2018)	Fui, C. S.; Lian, L. H.
RO33 (2021)	Furlan, R.; Gatti, M.; Mene, R.; Shiffer, D.; Marchiori, C.; Levra, A. G.; Saturnino, V.; Brunetta, E.; Dipaola, F.
RO34 (2018)	Glushkova, A.; Manitsaris, S.
RO35 (2021)	Gonzalez-Carrillo, C. D.; Restrepo-Calle, F.; Ramirez-Echeverry, J. J.; Gonzalez, F. A.
RO36 (2021)	Guerrero-Roldan, A. E.; Rodriguez-Gonzalez, M. E.; Baneres, D.; Elasri-Ejjaberi, A.; Cortadas, P.

RO37 (2022)	Han, J. W.; Park, J.; Lee, H. N.
RO38 (2023)	Han, Y.; Wu, W.; Liang, Y.; Zhang, L.
RO39 (2021)	Harati, H.; Sujo-Montes, L.; Tu, C. -H.; Armfield, S. J. W.; Yen, C. -J.
RO40 (2023)	Hsia, L. H.; Hwang, G. J.; Hwang, J. P.
RO41 (2023)	Hsu, T. C.; Huang, H. L.; Hwang, G. J.; Chen, M. S.
RO42 (2023)	Hu, Y. H.; Fu, J. S.; Yeh, H. C.
RO43 (2019)	Huang, W.; Hew, K. F.; Gonda, D. E.
RO44 (2017)	Ijaz, K.; Bogdanovych, A.; Trescak, T.
RO45 (2022)	Ingkavara, T.; Panjaburee, P.; Srisawasdi, N.; Sajjapanroj, S.
RO46 (2021)	Jeon, J.
RO47 (2022)	Jeon, J.
RO48 (2023)	Jiang, M. Y. C.; Jong, M. S. Y.; Lau, W. W. F.; Chai, C. S.; Wu, N.
RO49 (2014)	Kamardeen, I.
RO50 (2019)	Kamel, A; Liu, B. W.; Li, P.; Sheng, B.
RO51 (2017)	Kamnardsiri, T.; Hongsit, L. -O.; Khuwuthyakorn, P.; Wongta, N.
RO52 (2022)	Kochmar, E.; Vu, D. D.; Belfer, R.; Gupta, V.; Serban, I. V.; Pineau, J.
RO53 (2020)	Koc-Januchta, M. M.; Schonborn, K. J.; Tibell, L. A. E.; Chaudhri, V. K.; Heller, H. C.
RO54 (2018)	Leddo, J.; Kindi, R.; Bhandarkar, S.; Chadeva, N.; Ganotra, K.; Jayakumar, P.; Somaiya, Y.
RO55 (2021)	Lee, H. -S.; Gweon, G. -H.; Lord, T.; Paessel, N.; Pallant, A.; Pryputniewicz, S.
RO56 (2022)	Lee, Y. -F.; Hwang, G. -J.; Chen, P. -Y.
RO57 (2014)	Lei, C. -U.; Man, K. L.; Ting, T. O.
RO58 (2020)	Li, K. -C.; Chang, M.; Wu, K. -H.
RO59 (2022)	Li, W. Y.
RO60 (2020)	Lin, M. P. C.; Chang, D.
RO61 (2021)	Liu, C. C.; Hou, J. R.; Tu, Y. F.; Wang, Y. M.; Hwang, G. J.
RO62 (2023)	Liu, C. C.; Liu, S. J.; Hwang, G. J.; Tu, Y. F.; Wang, Y. M.; Wang, N. N.
RO63 (2022)	Liu, J. X.; Liu, X. H.; Yang, C.
RO64 (2017)	Liu, M.; Li, Y.; Xu, W. W.; Liu, L.
RO65 (2023)	Loftus, N.; Narman, H. S.
RO66 (2019)	Lu, X. X.
RO67 (2020)	Ma, Z. -H.; Hwang, W. -Y.; Shih, T. K.
RO68 (2023)	Mahrous, A.; Botsko, D. L.; Elgreatly, A.; Tsujimoto, A.; Qian, F.; Schneider, G. B.
RO69 (2020)	Mirchi, N.; Bissonnette, V.; Yilmaz, R.; Ledwos, N.; Winkler-Schwartz, A.; Del Maestro, R. F.
RO70 (2020)	Mokmin, N. A. M.
RO71 (2017)	Mørch, A. I.; Engeness, I.; Cheng, V. C.; Cheung, W. K.; Wong, K. C.
RO72 (2023)	Nalli, G.; Culmone, R.; Perali, A.; Amendola, D.
RO73 (2021)	Nazari, N.; Shabbir, M. S.; Setiawan, R.
RO74 (2022)	Nelekar, S.; Abdulrahman, A.; Gupta, M.; Richards, D.
RO75 (2022)	Neo, M.

RO76 (2021)	Neumann, A. T.; Arndt, T.; Köbis, L.; Meissner, R.; Martin, A.; de Lange, P.; Pengel, N.; Klamma, R.; Wollersheim, H. -W.
RO77 (2020)	Ochoa, X.; Dominguez, F.
RO78 (2023)	Ouyang, F.; Wu, M.; Zheng, L. Y.; Zhang, L. Y.; Jiao, P. C.
RO79 (2016)	Pérez-Marín, D.; Hijón-Neira, R.; Santacruz, L.
RO80 (2017)	Perikos, I.; Grivokostopoulou, F.; Hatzilygeroudis, I.
RO81 (2023)	Prescott, J.; Ogilvie, L.; Hanley, T.
RO82 (2021)	Qi, D.; Ryason, A.; Milef, N.; Alfred, S.; Abu-Nuwar, M. R.; Kappus, M.; De, S.; Jones, D. B.
RO83 (2023)	Rad, H. S.; Alipour, R; Jafarpour, A.
RO84 (2022)	Rodriguez-Arrastia, M.; Martinez-Ortigosa, A.; Ruiz-Gonzalez, C.; Ropero-Padilla, C.; Roman, P.; Sanchez-Labraca, N.
RO85 (2023)	Røed, R. K.; Baugerud, G. A.; Hassan, S. Z.; Sabet, S. S.; Salehi, P.; Powell, M. B.; Riegler, M. A.; Halvorsen, P.; Johnson, M. S.
RO86 (2023)	Sailer, M.; Bauer, E.; Hofmann, R.; Kiesewetter, J.; Glas, J.; Gurevych, I.; Fischer, F.
RO87 (2023)	Sáiz-Manzanares, M. C.; Marticorena-Sánchez, R.; Martín-Antón, L. J.; González Díez, I.; Almeida, L.
RO88 (2016)	Samarakou, M.; Fylladitakis, E. D.; Karolidis, D.; Früh, W. -G.; Hatziapostolou, A.; Athinaios, S. S.; Grigoriadou, M.
RO89 (2023)	Shah, C.; Davtyan, K.; Nasrallah, I.; Bryan, R. N.; Mohan, S.
RO90 (2023)	Shaikh, S.; Yayilgan, S. Y.; Klimova, B.; Pikhart, M.
RO91 (2019)	Stefanidis, K.; Psaltis, A.; Apostolakis, K. C. ; Dimitropoulos, K. ; Daras, P.
RO92 (2016)	Steif, P. S.; Fu, L. T.; Kara, L. B.
RO93 (2023)	Stojanov, A.
RO94 (2021)	Sweidan, S. Z.; Abu Laban, S. S.; Alnaimat, N. A.; Darabkh, K. A.
RO95 (2019)	Tanana, M. J.; Soma, C. S.; Srikumar, V.; Atkins, D. C.; Imel, Z. E.
RO96 (2019)	Tärning, B.; Silvervarg, A.
RO97 (2022)	Taskiran, A.; Goksel, N.
RO98 (2021)	Vanichvasin, P.
RO99 (2021)	Vazquez-Cano, E.; Mengual-Andres, S.; Lopez-Meneses, E.
RO100 (2021)	Vittorini, P.; Menini, S.; Tonelli, S.
RO101 (2014)	Vollmer, A. -L.; Mühlig, M.; Steil, J. J.; Pitsch, K.; Fritsch, J.; Rohlfing, K. J.; Wrede, B.
RO102 (2022)	Wambsganss, T.; Janson, A.; Leimeister, J. M.
RO103 (2015)	Wang, D. Q.; Han, H; Zhan, Z. H.; Xu, J.; Liu, Q. B.; Ren, G. J.
RO104 (2022)	Wang, Q.; Rose, C. P.; Ma, N.; Jiang, S. Y.; Bao, H. G.; Li, Y. Y.
RO105 (2022)	Wang, X. Z.; Zhang, L. J.; He, T.
RO106 (2022)	Wang, Z. J.
RO107 (2023)	Wolf, R. R.; Wolf, A. B.
RO108 (2023)	Xiao, Y.; Zhi, Y.
RO109 (2021)	Xie, T; Liu, R. B.; Chen, Y. J.; Liu, G. P.
RO110 (2021)	Xu, Y.; Wang, D.; Collins, P.; Lee, H.; Warschauer, M.
RO111 (2023)	Yamamoto, S.; Tobe, Y.; Tawatsuji, Y.; Hirashima, T.
RO112 (2018)	Yang, E.; Dorneich, M. C.
RO113 (2023)	Yang, H. Z.; Gao, C.; Shen, H. Z.
RO114 (2023)	Yang, Q. F.; Lian, L. W.; Zhao, J. H.

RO115 (2020)	Yannier, N.; Hudson, S. E.; Koedinger, K. R.
RO116 (2023)	Yu, E.
RO117 (2021)	Yu, S. -J.; Hsueh, Y. -L.; Sun, J. C. -Y.; Liu, H. -Z.
RO118 (2023)	Zahid Iqbal, M.; Campbell, A. G.
RO119 (2022)	Zhang, Z.; Xu, L.
RO120 (2023)	Zheng, L.; Niu, J.; Long, M.; Fan, Y.
RO121 (2021)	Zheng, L.; Zhong, L.; Niu, J.; Long, M.; Zhao, J.
RO122 (2023)	Zheng, L. Q;. Fan, Y. C.; Huang, Z. C.; Gao, L.
RO123 (2023)	Zheng, L. Q.; Long, M. L.; Chen, B. D.; Fan, Y. C.
RO124 (2023)	Zheng, L. Q.; Long, M. L.; Niu, J. Y.; Zhong, L.
RO125 (2022)	Zheng, L. Q.; Niu, J. Y.; Zhong, L.
RO126 (2022)	Zheng, L. Q.; Zhong, L.; Niu, J. Y.
RO127 (2020)	Zhu, M. X.; Liu, O. L.; Lee, H. S.
RO128 (2023)	Zou, B.; Du, Y.; Wang, Z.; Chen, J.; Zhang, W.
RO129 (2017)	Zou, S.

Appendix C. Synthesized definitions/descriptions of feedback levels based on Shute (2008)

Feedback level	Synthesized definition/description by the authors of this systematic review	Examples	Shute (2008) Pages
Basic feedback	Provides minimal information, typically indicating correctness without explanation. This level supports recognition but not understanding or repair of errors.	- Verification (right/wrong) - Correct response - Try again	160–161
Intermediate feedback	Offers moderate elaboration that highlights errors or revisits relevant content but stops short of full diagnostic explanation. It may address task features or provide cues without deep personalization.	- Error flagging - Attribute isolation - Topic contingent	160–161
Advanced (Elaborated) feedback	Provides detailed, targeted information aimed at addressing learners' misconceptions or guiding their reasoning. Often includes scaffolding, strategic hints, and personalized feedback without necessarily giving the answer.	- Response contingent - Hints/cues/prompts - Bugs/misconceptions - Informative tutoring	160–161

Note. No feedback in Shute's (2008, pp. 160-161) Table was omitted intentionally as it is absence of feedback.

Acknowledgments

Funding: This work was supported by the Start-up Research Grant [grant number RG 50/2023-2024R], the Teaching Development Grant [grant number T-24-07], and partially supported by an internal research fund [grant number 04377] awarded by the Education University of Hong Kong. This work was also supported by the Joint Centre for Artificial Intelligence Research Scheme [grant number 02189].

Declaration of interests

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

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