Quarterly Review
of Distance
Education

Leveraging artificial intelligence to enhance student feedback: An AHP-based decision-making framework

Uttara Deolankar and Amber Batwara
Symbiosis Skills and Professional University, Pune, India

Received 28 December 2025
Revised 11 April 2026
Accepted 22 May 2026

Abstract

Purpose – Artificial intelligence (AI) in student feedback systems transforms educational assessment and support. AI can improve student learning by providing personalised, timely, and actionable feedback using powerful algorithms and data analytics. These systems analyse student performance, behavioural patterns, and even the natural language used in written work to provide insights that educators may otherwise find difficult to identify. AI may also identify learning gaps, recommend specialised educational resources, and forecast future performance, creating a more proactive and adaptive teaching environment. AI in education has changed many aspects of learning, especially student feedback.

Design/methodology/approach – Research is separated into three parts: (1) As expected, published studies, project reports, and academic discussions revealed various performance markers. For the Kirkpatrick Model's four tiers, nine performance indicators were chosen. (2) AHP (Analytical Hierarchy Process) calculated the priority weight and rank of selected performance metrics. To ensure the AHP Model stability, a sensitivity analysis was done. (3) The selected performance indicators using the AHP were compared between two groups: traditional feedback and AI-generated feedback.

Findings – AI-generated feedback in education systems was shown to be practical by analysing the performance of these two groups. The linked research shows that AI-generated feedback provides students with timely, consistent, and personalised support, thereby improving learning. It also helps students to discover their talents and weaknesses.

Originality/value – AI technology and its careful integration into educational systems will boost learning and student behavioural confidence. Thus, AI-powered feedback mechanisms may shape education in the future.

Keywords Artificial intelligence (AI), Analytical hierarchy process (AHP), Comparative analysis, Student feedback, Sensitivity analysis

Paper type Research article

1. Introduction

Artificial intelligence (AI) is a technology used in digital devices, such as computers and mobile phones, to imitate human learning, comprehension, problem-solving, decision-making, creativity, and autonomy. AI aims to mimic human cognitive processes, allowing machines to interact judiciously. AI is rapidly progressing and is progressively integrating into numerous organizations, including healthcare, finance, transportation, education, and entertainment (Duan, Edwards, & Dwivedi, 2019). It can transfigure these fields by increasing efficiency, automating tasks, and enabling innovations. Concerns about data privacy, employment displacement, and the necessity of legislation to guarantee responsible growth are some of the ethical issues that AI brings up.

Some examples include virtual assistants like Siri, Alexa, and facial recognition software. These systems operate within a limited scope and are highly effective at what they are programmed to do, but lack general intelligence (Dizon & Tang, 2020).

"We all need people who will give us feedback. That's how we improve." Bill Gates frequently talks about the need for constant self-improvement. We must have people who will give us honest and constructive feedback (Ponnusamy *et al.*, 2021).

In the present case study, feedback is any response regarding a student's performance or behavior. The purpose of feedback in evaluating the learning process is to improve the student's performance. Student feedback is crucial to the educational process, facilitating

Quarterly Review of Distance Education

e-ISSN: 2169-1266
p-ISSN: 1528-3518
DOI 10.1108/QRDE-12-2025-0029

continuous improvement and personalized Learning. It is a requisite and an indispensable component of the educational system (Leckey & Neill, 2001). It is incorporated into the teaching-learning process to enhance teaching and learning techniques since it instantly impacts obtaining knowledge and skills. Student feedback has a reliable impact on both teaching and Learning. Feedback assists students in understanding the subject matter and guides how to improve student's learning procedures. Students who receive feedback may become more self-aware, confident, and enthusiastic about their studies. Reliable feedback acts as a strengths, weaknesses, opportunities, and challenges (SWOC) analysis given to students to improve their performance. Giving students pertinent feedback can help them improve their academic or fieldwork performance.

However, Traditional feedback methods often face challenges such as delayed responses, generic comments, and the inability to cater to individual student needs effectively. AI technologies offer promising solutions, including natural language processing (NLP), machine learning (ML), and data analytics (Chan & Colloton, 2024).

1.1 Applications of AI in student feedback

AI can analyze student data, produce feedback, and personalize learning tracks. Feedback is the information students receive about their performance, comprehension of the concepts, and behavioral change. Feedback can be formative or summative, positive or negative, specific or general, and verbal or written (Shaik *et al.*, 2022). The feedback procured through the AI tools provides a way to formulate the strengths and weaknesses of the student's performance (Naz & Robertson, 2024). This allows students to improve their performance continuously (Wirtz, Weyerer, & Geyer, 2019).

Automated grading systems utilize AI to evaluate student assignments and exams (Naz & Robertson, 2024). These systems can provide immediate feedback, highlight common errors, and suggest areas for improvement. Tools like Gradescope and EdTech platforms have implemented such systems to streamline grading. Intelligent Tutoring Systems (ITS) employ AI to deliver personalized feedback and guidance to students. ITS can tailor instructional content and feedback to individual needs by analyzing student performance and learning behaviors. Examples include Carnegie Learning and Knewton, which adapt content based on student interactions (Naz & Robertson, 2024). Natural Language Processing (NLP) algorithms can analyze student essays for grammar, coherence, and content quality. These systems, such as Grammarly and Turnitin, offer detailed feedback on writing style, structure, and originality, helping students enhance their writing skills (Hooda, Rana, Dahiya, Rizwan, & Hossain, 2022). AI is used by adaptive learning platforms to gauge student performance and modify feedback as required. Systems like Gradescope, DreamBox, and Smart Sparrow monitor student progress and provide real-time feedback, and give suggestions for further study.

1.2 Benefits of AI-driven feedback

AI-powered feedback systems yield a wealth of data related to student performance. Facilitators and teachers can influence the collected data to understand each student's strengths and weaknesses, monitor progress, and make well-informed instructional decisions. This benefits the student for continual performance improvement (Yu & Canton, 2023). AI systems provide instant feedback, allowing students to address errors and misconceptions promptly. This immediate response is crucial for effective Learning and retention (Afzaal *et al.*, 2021). AI can analyze individual learning patterns and offer tailored feedback, catering to the unique needs of each student. This personalized approach helps address specific weaknesses and promote strengths. AI-driven feedback systems can handle large volumes of student data, making them ideal for large classrooms and online courses. This scalability ensures that all students receive consistent, high-quality feedback (Chan & Colloton, 2024). It eliminates human feedback bias, ensuring evaluations are based solely on performance metrics. This objectivity promotes fairness in the assessment process.

1.3 Limitations

As we discussed the benefits, the other side of the coin represents the limitations. Technical Limitations would be briefed as AI systems rely on vast quantities of data and algorithms, which can be resource-intensive to develop and maintain. Ensuring accuracy and reliability in feedback remains a technical challenge (Wirtz *et al.*, 2019). The next aspect is Data Privacy and Security, meaning AI in education involves collecting and analyzing sensitive student data. Data privacy and security are paramount to protecting student information from misuse, a considerable limitation to the education system (Dönmez, 2024).

Acceptance and Trust by the Educators and students may be skeptical of AI-driven feedback systems. Building trust and demonstrating the effectiveness of these technologies is essential for widespread adoption. Pedagogical Implications AI feedback systems must align with pedagogical goals and not merely focus on technical accuracy (Chan & Colloton, 2024). Ensuring that AI complements human teaching and promotes holistic learning experiences is crucial.

1.4 The traditional student feedback system

Feedback is identified in the works of researchers as an essential element for improving students' learning process. Research shows that a lecturer's feedback creates value when it provides appraisal information and links to further Learning.

The traditional student feedback system, typically used before the integration of digital technology, was more manual and less dynamic than today's systems. Below are the key characteristics and processes of traditional student feedback systems: End-of-term Evaluations: In most cases, feedback was collected at the end of a course or semester. Students often filled out printed questionnaires in class, which asked them to rate various aspects of the course, instructor, and materials. Feedback forms were frequently generic, with the same questions across different classes, departments, and institutions (Shaik *et al.*, 2022). This simplified the process but limited the questions' relevance or specificity for various disciplines or teaching styles. After collecting feedback, the results were manually compiled, which took time. Data would often be entered into spreadsheets, and summary reports would be generated for faculty or department heads. Since the system was manual, there was little to no analysis beyond basic averaging or counting responses. Detailed insights, such as trends over time or correlation between feedback and outcomes, were not easily accessible.

Without the opportunity for more personalized or in-depth responses, the feedback received was often surface-level and less helpful in identifying specific areas for improvement. Because feedback was collected at the end of the semester and took time to process, instructors usually didn't receive the information until well after the course had ended, limiting its immediate usefulness.

1.5 Present scenario: modern methods of student feedback

The present scenario of student feedback systems has evolved significantly with technological advancements and pedagogical practices. Today's systems aim to be more interactive, data-driven, and accessible to students and educators alike. Many institutions use online platforms (e.g. Google Forms, SurveyMonkey, Blackboard) to collect student feedback efficiently. These platforms often allow for anonymity, which encourages more honest responses. Systems like Moodle, Canvas, or Blackboard often integrate feedback collection into course workflows, making it easy for students to provide input and for faculty to analyze results (Dönmez, 2024).

Tools like Kahoot, Mentimeter, or Poll Everywhere allow instructors to gather immediate feedback during lectures, helping to assess understanding in real-time and adjust the lesson accordingly. Some systems enable ongoing feedback rather than just end-of-course evaluations, allowing instructors to change throughout the term. AI-powered systems can analyze student comments for sentiment, highlighting areas of concern or satisfaction. Dashboards and analytics tools provide educators with visual insights into trends in student feedback, helping to track engagement, participation, and satisfaction over time (Wirtz *et al.*, 2019).

These systems analyze individual student performance and provide tailored feedback. Smart Sparrow and Knewton are examples of adaptive learning platforms that modify content in response to student comments and progress. Some systems facilitate peer-to-peer feedback, where students review each other's work, providing diverse perspectives. Many feedback systems are now optimized for mobile devices, making it easier for students to provide feedback. Efforts are made to ensure that feedback systems are accessible to students with disabilities, using standards like WCAG (Web Content Accessibility Guidelines) (Hooda *et al.*, 2022).

Integrating AI in education, especially in feedback systems, can transform the learning process by providing students with more targeted and consistent assistance. AI may assess data from many sources, such as student performance and behavior, to detect learning gaps, recommend tailored resources, and forecast future performance. This study intends to evaluate the efficiency of AI-generated feedback with conventional techniques utilizing the Analytical Hierarchy Process (AHP) approach to rank and prioritize performance metrics, followed by a sensitivity analysis for model stability. The research reveals how AI-driven feedback systems give constant, individualized help, raising students' confidence and enhancing their overall learning experience, and underscores their crucial role in defining future education.

RQ1. How effective is artificial intelligence (AI) in transforming student feedback systems to provide personalized, timely, and actionable feedback, thereby enhancing student learning outcomes?

RQ2. What are each performance indicator's relative weights and rankings?

This study work aims to contribute by addressing the aforementioned two issues. The following section reviews existing research, including performance metrics that will aid in the adoption and integration of AI-generated feedback systems within the educational sector. The following sections include research methodology, research analysis, and results. Subsequently, we address the study's weaknesses and the avenues for further research. Finally, the paper finishes with assertions and citations.

2. Literature review

The review of related literature aims to establish a strong conceptual and empirical foundation for the present study by critically examining prior research in the field. It not only helps in identifying the research gap but also situates the current study within the broader academic discourse.

Artificial Intelligence (AI) has emerged as a transformative force in education, enabling more personalized, adaptive, and efficient learning experiences. Research in Artificial Intelligence in Education (AIEd) highlights the development of customized learning systems through technologies such as social networking platforms, chatbots, expert systems, intelligent tutoring systems, machine learning algorithms, and virtual learning environments. These tools empower educators to tailor instruction to individual learner needs, enhance knowledge acquisition, and support the development of professional competencies (Tapalova & Zhiyenbayeva, 2022).

Extending this perspective of AI-enabled personalization, empirical studies have also examined its impact on specific learning outcomes. For instance, (Kazu & Kuvvetli, 2023) explored the role of AI-driven pronunciation practice in vocabulary acquisition. Using a controlled experimental design with pre-test and post-test groups, their study demonstrated that integrating AI-based pronunciation tools significantly improves long-term vocabulary retention among learners. This finding reinforces the potential of AI not only in content delivery but also in enhancing cognitive learning processes.

In addition to improving learning outcomes, AI has shown significant promise in supporting data-driven educational decision-making. One of the critical challenges faced by educators is the continuous monitoring of students' academic progress and the timely identification of at-risk learners. (Khan, Ahmad, Jabeur, & Mahdi, 2021) addressed this issue by applying machine learning algorithms to analyze student performance data. Their study found that decision tree models were particularly effective in predicting student outcomes and enabling early interventions. This highlights how AI can transform raw educational data into actionable insights, thereby improving both teaching strategies and student support systems.

Collectively, these studies demonstrate the multifaceted role of AI in education, ranging from personalized learning and enhanced retention to predictive analytics and informed decision-making. However, while prior research has extensively explored AI applications in learning and performance monitoring, there remains a need for structured frameworks that integrate these capabilities into student feedback systems.

The above case studies were examined, which led to the present case study, Leveraging Artificial Intelligence to Enhance Student Feedback– Decision-Making Framework.

2.1 Discussion on performance indicators

This section discusses the performance indicators, specifically under the four levels of the Kirkpatrick Model (Kirkpatrick, 2010). A widely accepted method for assessing the outcomes of training and educational initiatives is the Kirkpatrick Model. It evaluates both official and informal training techniques and assigns scores based on four levels of criteria: behaviour, learning, results, and reaction. (Smidt, Balandin, Sigafoos, & Reed, 2009).

This research study considers four levels mentioned in the Kirkpatrick Model and categorizes them into nine performance indicators, as shown in Table 1. These nine indicators are about AI-generated feedback. AI-generated feedback will be a guiding torch for the teaching communities to enhance students' feedback, which will ultimately improve the students' performances. For additional reading, a table also contains the details of enablers and their citation.

2.1.1 Reactions. 2.1.1.1 Increase in satisfaction level (I1). The perceived degree of satisfaction and fulfilment that students acquire from their own unique accomplishment is known as student satisfaction. Along with expertise and ideals, job satisfaction is a driving factor in professional behaviour. (Patel, 2018).

Table 1. Finalized performance indicators for the present research

S.No.	Kirkpatrick model levels	Performance indicators	Symbol	References
1	Reactions	Increase in satisfaction level	I1	Hooda *et al.* (2022), Patel (2018)
		Increase in confidence level	I2	Lin, Liu, Poitras, Chang, and Chang (2024), Barak (2010)
2	Learning	Performance	I3	Yu & Canton (2023), Barak (2010)
		Able to demonstrate	I4	Naz & Robertson (2024)
3	Behavior	Apply knowledge	I5	Khan *et al.* (2021), Afzaal *et al.* (2021)
		Government regulations and standards	I6	Selwyn and Bulfin (2016), Barak (2010), Naz and Robertson (2024)
4	Results	Create a new knowledge base	I7	Shaik *et al.* (2022), Keane and Labhrainn (2005)
		Able to judge one's opinion	I8	Khan *et al.* (2021), Dönmez (2024)
		Able to denote similarities and differences	I9	Spector and Ma (2019), Arredondo and Martinez (2010)

2.1.1.2 Increase in confidence level(I2). Increased confidence means you accept and believe in yourself and have a sense of control. One knows the strengths and weaknesses well and has a favorable view of oneself. Student feedback should be specified, which leads to enhanced satisfaction levels.

2.1.2 Learning. 2.1.2.1 Performance (I3). As a cost-effective substitute for human instructors, automatic feedback on student performance has been used for writing courses (Yu & Canton, 2023). AI-generated feedback allows the students to understand their strengths and weaknesses; the strengths would be emphasized while the weaknesses would be worked on to change to strengths.

2.1.2.2 Able to demonstrate(I4). With the feedback from AI-generated tools, students can demonstrate a learning concept independently. This criterion is crucial, as it lets the mind apply the knowledge that was earlier comprehended by the mind. AI replies with, “It is clear you have a strong understanding of the key concepts and theories,” without specifying the exact location where this understanding is demonstrated (Naz & Robertson, 2024).

2.1.3 Behavior. 2.1.3.1 Positive behavioral change (I5). The students expect a positive behavioral change after receiving the feedback. Improvements reflected in the student’s work lead to a positive behavioral change. A positive behavioral change leads to enhancing the satisfaction level of Learning.

2.1.3.2 Apply knowledge (I6). Content or learning matter is taught in the classroom sections, usually in the form of a theory or a principle. The learning content should not remain in the books but be immediately brought into practice to solve daily life problems. If not brought, then the knowledge learned remains in the books. The knowledge learned in the classroom, when brought into practice, is called the application of knowledge (Naz & Robertson, 2024).

2.1.4 Results. 2.1.4.1 Create a new knowledge base(I7). When “Learning” takes place, it is assumed that a new knowledge base is created. Acquiring new knowledge allows us to ponder more deeply about challenges from many viewpoints, polishing skills like reasoning and problem-solving. Hence, creating new knowledge is referred to as a constructivist approach to building knowledge. It is one of the higher-order thinking skills (HOT) in the cognitive domain of Bloom’s Taxonomy. Student feedback should be procured to create a new knowledge base (Shaik *et al.*, 2022).

2.1.4.2 Able to judge one’s opinion(I8). Creating a new knowledge base leads to the formulation of higher-order thinking skills. Feedback should be formulated so that the students can develop their own decisions based on their opinions. Formulating one’s opinion with logic is an essential part of Learning.

2.1.4.3 Able to denote similarities and differences (I9). One of the higher-order thinking skills is the ability to formulate similarities and differences in the taught concepts. When this skill is formulated, there is a greater conceptual understanding of the concepts (Arredondo & Martinez, 2010).

3. Research methodology

This research comprises three tiers, as indicated in Figure 1, to address the gaps in knowledge regarding accepting and integrating AI systems in student feedback. The primary objective of this study is to systematically identify the key performance indicators that influence the acceptance and effectiveness of AI-generated feedback systems. This is achieved through a comprehensive review of existing literature, followed by validation and refinement of these indicators through expert consultation. By integrating theoretical insights with practitioner perspectives, the study ensures that the identified indicators are both conceptually sound and practically relevant.

The second objective is to evaluate and prioritize these identified indicators by assigning relative weights and rankings. This is undertaken using a structured analytical approach grounded in the Kirkpatrick Model, which provides a robust framework for assessing training and evaluation outcomes across multiple levels. By mapping the nine indicators onto this

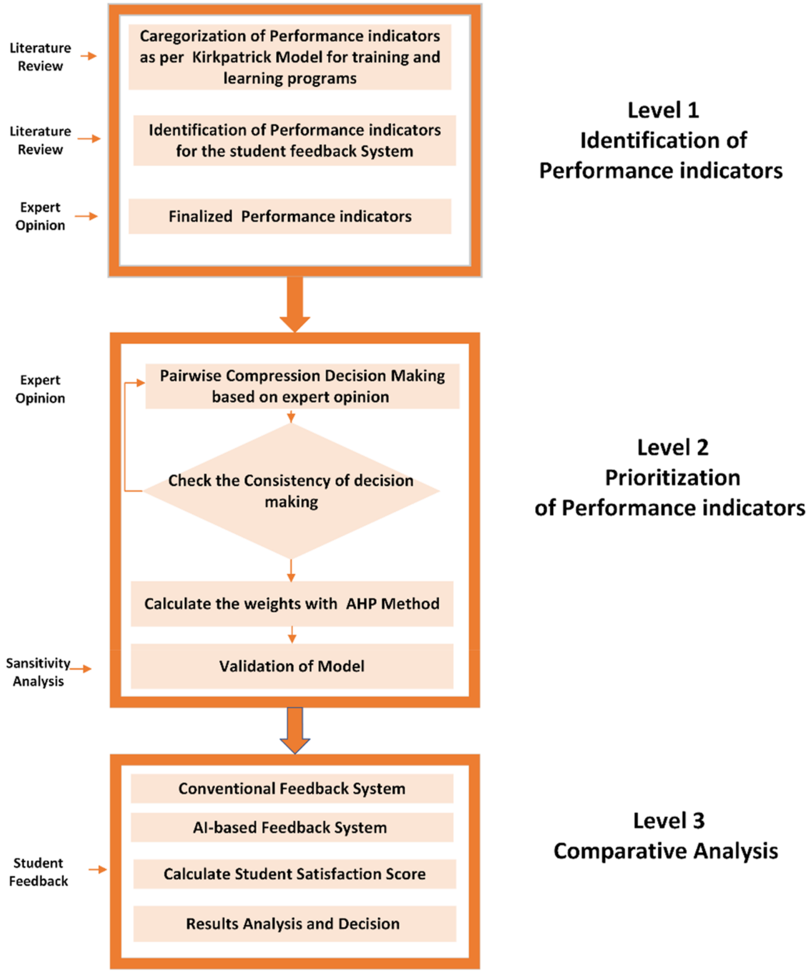

Figure 1. Research flow chart

model, the study seeks to establish their relative significance and interrelationships, thereby offering a systematic basis for decision-making in the implementation and assessment of AI-driven feedback systems.

The first level enumerates "Identification of Performance Indicators". This level is an umbrella of Literature review, Performance indicators, and Expert opinion. The study of various research articles led to a research gap focused on formulating indicators about AI-generated student feedback. The indicators of AI-generated student feedback were formulated with the orientation to the Kirkpatrick evaluation model. An expert's opinion was taken on enhancing the indicators for quality improvement.

The second level of the research study focuses on prioritizing performance indicators. Prioritization was performed logically; pairwise compression decision-making based on expert opinion was performed first. Secondly, the consistency of decision-making was calculated using the AHP method. Thirdly, sensitivity analysis was performed to validate the decision-making model. The third level elaborated on comparative analysis. This included the feedback given to students in the conventional setting and the feedback given to students in an AI-generated manner. GradeScope software was used to generate the AI feedback. A comparative analysis was done on nine parameters. The third level ends with the result analysis and decision, which will be discussed in the next section.

4. Research analysis

Research Analysis was performed in two stages to explore the satisfaction level of students in comparison with the traditional way of feedback and AI-generated feedback approaches.

4.1 AHP (Geometric mean method)

According to Saaty (1980) "AHP is a method of decision-making that divides a challenging MCDM problem into a hierarchy". It is a theory based on the hierarchy and consistency of data, and it judges the data and provides a measure of its importance. AHP incorporates all input from experts into a final choice without requiring pairwise assessments of several options to demonstrate their usability on a number of criteria (Melewar, Alwi, Hsu, & Lin, 2013).

AHP is a modern method for dealing with multiple criteria in decision-making (MCDM), enabling decision-makers to develop preferences that result in the best possible decisions. Due to the issue's intricacy and urgency, we must search for straightforward analytical methods. When evaluating the value assessments of various stakeholders and individual decision-makers, multi-criteria decision analysis (MCDA) techniques might be helpful (Kiker, Bridges, Varghese, Seager, & Linkov, 2005).

4.1.1 AHP methodology. When there are many factors to consider, decisions are made using the Analytical Hierarchy Process (AHP). Its adaptability, ease of computation, and compatibility with other methods have made it the most popular method for supporting multi-criteria decision-making. AHP is a statistical method that builds on the decision maker's knowledge while navigating complex decisions, utilising both quantitative and qualitative data. The creation of paired comparison matrices and the subsequent identification of weighting factors for each attribute are essential components of AHP (Sivakumar, Radha Krishnappa, & Nallanathel, 2020).

Step 1: The hierarchical structure was built like a tree with different levels of performance indicators, as shown in Figure 2.

Step 2: Create a pair-wise comparison matrix for each item on a nine-point importance scale, as shown in Table 4. Saaty's nine-point scale (1–9) to create a pairwise comparison matrix. 1 denotes equal significance, 3 moderately more considerable importance, five vital greater importance, 7 much higher importance, and nine massively greater relevance. The compromised significance levels are shown by 2, 4, 6, and 8. Saaty's scale for pairwise comparison matrices is shown in Table 2.

A matrix was built that compares the aims and criteria in each area, as shown in Table 5 (Paiewise compression matrix). This matrix was developed based on the responses that were gathered from the survey. Table 3 provides an overview of the individuals who took part in the study. A total of seven individuals, ranging from professors to research scholars, are involved in the expert committee. While the viewpoints of the experts were being studied, comparative judgments were integrated by applying the geometric mean to the opinions in order to form the matrix of the relative evaluation. In order to establish the weights for each performance indicator, the geometric mean approach, which was developed by (Buckley, 1985), was used. This technique included determining the geometric mean of the pairwise judgement matrices using the technique.

Step 3: Determine the consistency- The absolute or relative magnitudes of the principal Eigenvector are ascertained. Subsequently, use Equation (1) to get the consistency index (CI) value for each n-dimensional matrix. To ascertain the reliability of the data, compute the

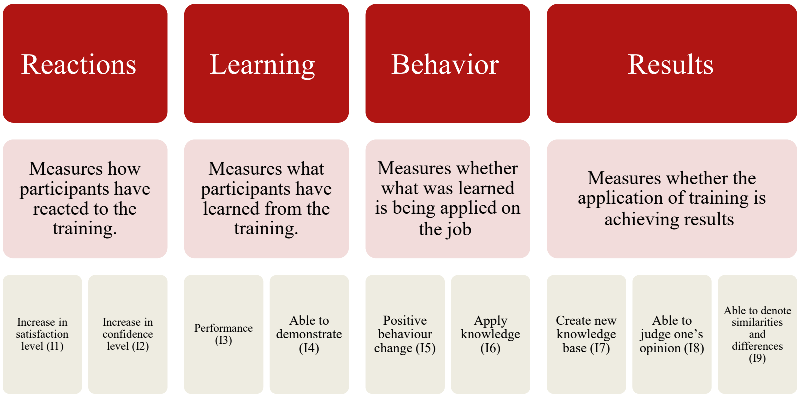

Figure 2. Hierarchical structure for AHP

Table 2. Saaty's scale for AHP

Fuzzy crisp number	Linguistic variable
$1^{\sim}$	Equal importance
$3^{\sim}$	Moderate importance
$5^{\sim}$	Essential or high-importance
$7^{\sim}$	Very high importance
$9^{\sim}$	Extreme importance
$2^{\sim},4^{\sim},6^{\sim},8^{\sim}$	Intermediate values between the two adjacent judgments

Table 3. Summary of the participants

Expert	Post	Qualification	No. of experience (Years)
Academicians	Professor	Doctoral Degree	30–45
	Assistant Professor	Master's Degree related field	15
	Research Scholar	Expert in AI	5–8

consistency ratio (CR) using Equation (2) with a reference value known as the random consistency index (R.I.). Presented below are the equations:

$$CI = \frac{(\lambda_{max} - n)}{n - 1} \dots\dots\dots \tag{1}$$

$$CR = \frac{CI}{RI} \dots\dots\dots\dots\dots \tag{2}$$

Table 4 delineates the random indices (R.I.) values for matrices of varying sizes (N) ranging from 1 to 10. The numbers were estimated based on a sample of 500 instances. The variability in R.I. is contingent upon the dimensions of the matrix. To assess the reliability of the AHP model's computations, examine the consistency ratio (C.R.). If the C.R. value was below 0.10, we accepted the result of the AHP model. We adjusted the subjective ratings when they were elevated due to the inconsistency of the AHP output. The permissible range for the consistency

Table 4. Random consistency index values

n	1	2	3	4	5	6	7	8	9	10
RI	0	0	0.58	0.9	1.12	1.24	1.32	1.41	1.45	1.49

ratio (C.R.) is contingent upon the size of the matrix. The value is 0.05 for a 3x3 matrix, 0.08 for a 4x4 matrix, and 0.1 for matrices when n is more than or equal to 5. The matrix assessment is deemed reliable and consistent if the C.R. value is equal to or below these levels. If the C.R. exceeds the permissible limit, it indicates that the matrix is inconsistent, necessitating a reevaluation and enhancement of the assessment procedure.

4.1.2 Results of AHP. The findings of the Analytical Hierarchy Process (AHP) used in this research corroborate the transformational impact of AI on student feedback systems by allocating weights and priority to diverse performance metrics aligned with the four tiers of the Kirkpatrick Model.

The first performance indicator was “Increase in satisfaction level” (I1) at the reaction level, with a weight of 0.247, achieving the first rank as shown in Table 6. This finding aligns with other studies highlighting the significance of student satisfaction in improving engagement and motivation (Hooda *et al.*, 2022). The substantial weight indicates that AI feedback systems significantly enhance student happiness by delivering fast, relevant, and individualized replies. The second measure, “Increase in confidence level” (I2), assigned a weight of 0.179, underscores the beneficial effect of AI on student confidence, corroborated by research by (Lin *et al.*, 2024).

The Learning Level, “Performance” (I3), placed third with a weight of 0.173, underscoring the significance of enhancing student abilities and knowledge using AI feedback systems. Previous research, like (Dönmez, 2024), has shown that AI can tailor itself to individual learning requirements, therefore improving student performance via specific feedback. “Able to demonstrate (I4)”, with a weight of 0.038, signifies the difficulty in quantifying practical applications relative to other indicators, as seen by (Hooda *et al.*, 2022). Under the Behavior Level, the indicator “Apply knowledge” (I5) with a weight of 0.056 highlights the modest relevance of knowledge transfer to practical settings in AI-driven feedback systems. The indicator “Government regulations and standards” (I6), ranked tenth with a weight of 0.027, accords with results showing that compliance has less perceived effect on student feedback outcome (Barak, 2010).

At last, at the Results Level, “Create a new knowledge base” (I7) was given a weight of 0.125, highlighting the significance of developing and distributing new knowledge as a result of AI-generated feedback. Similar results were found by (Selwyn & Bulfin, 2016), who studied AI’s function in supporting knowledge development and exchange. The indicators “Able to judge one’s opinion” (I8) and “Able to denote similarities and differences” (I9), with weights of 0.079 and 0.069, respectively, support the premise that AI feedback fosters critical thinking and analytical abilities (Spector & Ma, 2019). The overall ranking and weights of each performance indicator are represented in Figure 3.

The AHP results validate previously published findings and underscore the transformative potential of AI-generated feedback in education systems. By enhancing satisfaction, confidence, and learning outcomes, AI technologies pave the way for more personalized, adaptive, and operative educational domains (see Table 5).

4.1.3 Sensitivity analysis. Sensitivity analysis is a method used to evaluate the impact of variations in the input variables of a model on the outputs. This method is beneficial when the inputs are uncertain or consist of several subjective judgments, as seen with experts (Punia Sindhu, Nehra, & Luthra, 2016). Sensitivity analysis enables the evaluation of the robustness and reliability of the model’s outcomes by examining the effects of varying these inputs (Batwara, Sharma, Makkar, & Giallanza, 2022). Sensitivity analysis is essential in decision-

Table 5. Pairwise compression matrix

	I1	I2	I3	I4	I5	I6	I7	I8	I9
I1	1	2	3	5	5	6	2	4	4
I2	1/2	1	2	3	3	3	2	5	5
I3	1/3	1/2	1	4	2	6	4	4	5
I4	1/5	1/3	1/4	1	1/4	3	1/3	1/5	1/4
I5	1/5	1/3	1/2	4	1	3	1/5	1/3	1/3
I6	1/6	1/3	1/6	1/3	1/3	1	1/3	1/3	1/4
I7	1/2	1/2	1/4	3	5	3	1	3	4
I8	1/4	1/5	1/4	5	3	3	1/3	1	2
I9	1/4	1/5	1/5	4	3	4	1/4	1/2	1

Note(s): Consistency Ratio (CI) = 0.12

Table 6. Final weights for performance indicators

S.No.	Kirkpatrick model levels	Performance indicators	Symbol	Consistency index	Final weights	Ranking
1	Reactions	Increase in satisfaction level	I1	0.18245	0.247	1
		Increase in confidence level	I2		0.179	2
2	Learning	Performance	I3		0.173	3
		Able to demonstrate	I4		0.038	8
3	Behavior	Apply knowledge	I5		0.056	7
		Government regulations and standards	I6		0.027	9
4	Results	Create a new knowledge base	I7		0.125	4
		Able to judge one’s opinion	I8		0.079	5
		Able to denote similarities and differences	I9		0.069	6

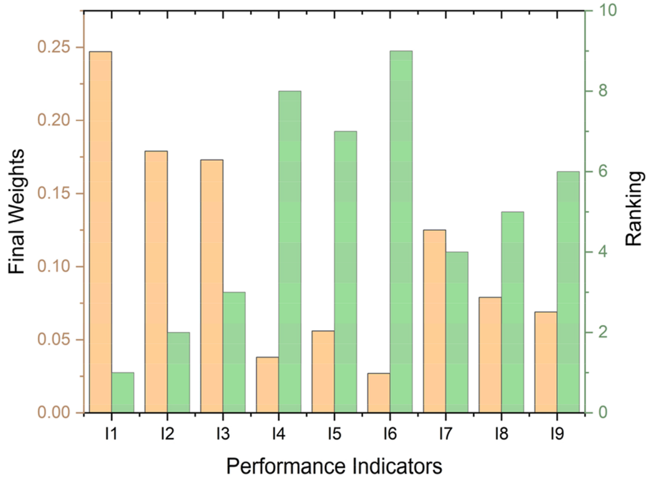

Figure 3. Overall ranking of performance indicators

making models, particularly those using the Analytic Hierarchy Process (AHP). The Analytic Hierarchy Process (AHP) is a Multi-Criteria Decision Analysis (MCDA) methodology designed to assess many criteria according to their relative significance. Nonetheless, since expert assessments often dictate these weights, it is essential to evaluate the impact of changes in these weights on the model's performance. Sensitivity analysis is an effective instrument for mitigating this deficiency by altering the weights to swiftly modify the priority ratings.

As indicated in Table 7, the "Increase in satisfaction level" (I1) weight impacts the total weight of other performance indicators and ranges from 0.1 to 0.5 in increments of 0.1. When the "Increase in satisfaction level" (I1) value is adjusted from 0.247 to 0.1, Table 8 provides a rank based on the order in which priorities have changed. Assuming the lead, "Increase in satisfaction level" (I1) changes other performance indicators ' priority weights and rankings.

The top-ranking indicators, such as "Increase in satisfaction level" (I1) and "Increase in confidence level" (I2), continuously kept their places, underlining their dominating influence in AI-driven feedback systems. This stability suggests that these indicators are significant and have a strong, consistent influence on the entire model. Sensitivity analysis also found that lower-ranked indicators, such as "Government regulations and standards" (I6), exhibited a limited effect on the overall priority of the performance model. This means that tiny changes in these indicators have a minimal influence on the overall outcomes, further justifying the model's concentration on core drivers of AI integration.

Table 7. Priority values of performance indicators after varying the weights of "Increase in satisfaction level (I1)"

S.No.	Kirkpatrick model levels	Symbol	Normal	0.1	0.2	0.3	0.4	0.5
1	Reactions	I1	0.247	0.1	0.2	0.3	0.4	0.5
		I2	0.179	0.214	0.191	0.1672	0.143	0.119
2	Learning	I3	0.173	0.207	0.184	0.161	0.138	0.1154
		I4	0.0384	0.045	0.0408	0.035	0.030	0.0255
3	Behavior	I5	0.056	0.068	0.060	0.052	0.045	0.0378
		I6	0.0279	0.0333	0.029	0.0260	0.022	0.0185
4	Results	I7	0.1258	0.1505	0.133	0.117	0.100	0.0836
		I8	0.0798	0.095	0.084	0.074	0.063	0.0530
		I9	0.0697	0.083	0.074	0.064	0.0556	0.0463

Table 8. Ranking of performance indicators after varying the weights of "Increase in satisfaction level (I1)"

S.No.	Kirkpatrick model levels	Symbol	Normal	0.1	0.2	0.3	0.4	0.5
1	Reactions	I1	1	4	1	1	1	1
		I2	2	1	2	2	2	2
2	Learning	I3	3	2	3	3	3	3
		I4	8	8	8	8	8	8
3	Behavior	I5	7	7	7	7	7	7
		I6	9	9	9	9	9	9
4	Results	I7	4	3	4	4	4	4
		I8	5	5	5	5	5	5
		I9	6	6	6	6	6	6

The sensitivity analysis supports the resilience and stability of the AHP model for assessing performance indicators in AI-driven feedback systems, as shown in Figure 4. It validates that the prioritization of indicators remains trustworthy, ensuring that conclusions generated from the model are believable and actionable. This further facilitates the deployment of the AHP-derived insights to achieve fundamental changes in educational outcomes using AI-enhanced feedback mechanisms.

4.2 Comparative analysis

This section includes a comparative discussion that underlines the advantages of AI-generated feedback over conventional feedback systems and corroborates results from earlier published studies in this area. An AI tool, Gradescope, was implemented in the case study to fetch student feedback. Gradescope is a viral tool because of its AI-powered handwriting recognition. This innovative technology allows teachers and instructors to grade short answers quickly and accurately, like fill-in-the-blank questions, True or False statements, and handwritten responses. This AI tool graded hundreds of essays without the tedious task of manually deciphering handwriting. It automatically recognized student's answers and grouped similar responses for efficient review. This saved instructors significant time, ensured consistency, and reduced the risk of error in grading.

Gradescope is an American ed-tech establishment offering online and AI-assisted higher education grading tools. The establishment's grading software provides tools for marking written exams and homework assignments. They were founded in 2014 at Berkeley, with its Headquarters in Berkeley, California, United States.

4.2.1 Quasi-experimental design. The present research design represented a quasi-experimental design. The quasi-experimental design aims to establish a cause-and-effect relationship between an independent and dependent variable: "AI tool: Gradescope" and "student feedback", respectively. The present study was conducted on 160 students studying a learning program on Creative Writing in English. The students in the case study were randomly divided in 2 groups of 80 students in each group.

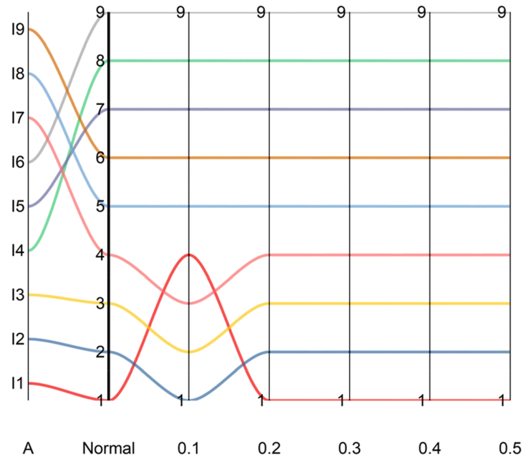

Figure 4. Parallel plot for sensitivity analysis

Both the groups, the experimental group, and the control group, were administered with a test on a taught content for 6 months. The experimental group further experimented with the AI tool to obtain feedback. The control group did not experiment with the AI tool to obtain feedback; the feedback was given traditionally. A comparative analysis was done on the responses procured through the AI tool and traditional method. The responses of both groups were taken through a rating scale, i.e. the questionnaire/post-test, which was quantified and analyzed further.

4.2.2 Post-test only non-equivalent groups design. In this design, participants in one experimental group are exposed to a treatment, a control group is not exposed to the treatment, and then the two groups are compared. In this research study, the treatment was AI-generated feedback to the experimental group. The independent variable is AI-generated feedback, while the dependent variable is scores achieved in the Post-test.

Experimental group → Feedback through AI-generated software →Post-test.

Control group → feedback through traditional way → Post-test.

4.2.3 T- test analysis. A two-sample *T*-test was undertaken to assess students' satisfaction levels across two groups: one getting feedback via conventional techniques (C1) and the other using AI-generated input (C2). The research attempted to analyze whether AI-generated feedback substantially improved student satisfaction. The following outlines the analysis and results.

Variables:

(1) Independent Variable: AI-generated feedback program.

(2) Dependent Variable: Satisfaction level of students.

Hypotheses:

(1) Null Hypothesis (Ho): AI-generated feedback will not influence the students' satisfaction level.

(2) Alternative Hypothesis (H1): AI-generated feedback will increase the satisfaction level of the students' feedback.

The P-value of 0.000 is less than the standard significance threshold of 0.05, suggesting that the null hypothesis (Ho) may be rejected. This suggests a statistically significant difference in satisfaction ratings between the two groups. The mean satisfaction level for the group getting AI-generated feedback (C2) was considerably greater (Mean = 4.402) compared to the group receiving conventional input (C1), with a mean of 1.759 as per Table 9. The enormous negative T-Value (−21.44) also illustrates the considerable variation in satisfaction levels. The 95% confidence interval for the difference in means (−2.921, −2.364) does not include zero, further showing a substantial difference in satisfaction levels between the two groups.

The findings reveal that AI-generated feedback dramatically improves the satisfaction level of pupils compared to conventional feedback techniques. This data confirms the alternative hypothesis (H1), indicating the usefulness of AI-driven feedback in boosting students' educational experience and happiness. The graphical depiction enumerates the

Table 9. T- test analysis

	N	Mean	StDev	SE mean
Traditional methods (C1)	9	1.759	0.355	0.12
AI-generated feedback (C2)	9	4.402	0.102	0.034

Note(s): *T*-value: 21.44 *p*-value: 0.000 Degrees of Freedom (DF): 9

difference, conveying an impression that students in the experimental group were happy with the AI-produced feedback, as shown in Figure 5.

5. Conclusion

In today's era, technology has entered almost all walks of life, like businesses, corporations, industries, and educational organizations. Educational organizations are no exception in the usage of technology. Technology is used in curriculum development, industry collaboration, teaching-learning, and evaluation processes. Evaluation is a crucial process guiding students' performances to greater heights. The evaluative remarks should reflect on the weaknesses and the strengths of the student's performance. Teachers/facilitators focused on traditional feedback approaches, which could consume much time and energy. The feedback remarks could often be monotonous, leading to ineffective teaching-learning and evaluation processes. Now, with the technology event, these remarks are AI-generated, which has many advantages. AI-generated feedback procures a treasure of data related to student performance. Facilitators and teachers can influence the collected data to understand each student's strengths and weaknesses, monitor progress, and make well-informed instructional decisions. This benefits the student for continual improvement in their performance.

Students receive personal feedback, helping them to understand their strengths and weaknesses and further areas of improvement. AI-generated feedback is delivered instantly, allowing students to address mistakes quickly and continue waiting for teacher review remarks. Unlike human grading, which may vary due to fatigue or biases, AI-generated feedback provides uniform student feedback, promoting fairness and objectivity. With AI-generated feedback, teachers can manage larger classes with a lot of effectiveness. AI-generated input encourages students to self-correct, fostering independent Learning and problem-solving skills. By AI-generated feedback, which happens automatically, teachers and facilitators can focus on more complex teaching aspects like mentoring, counselling those in need, and improving their teaching-learning processes with novel and creative ideas.

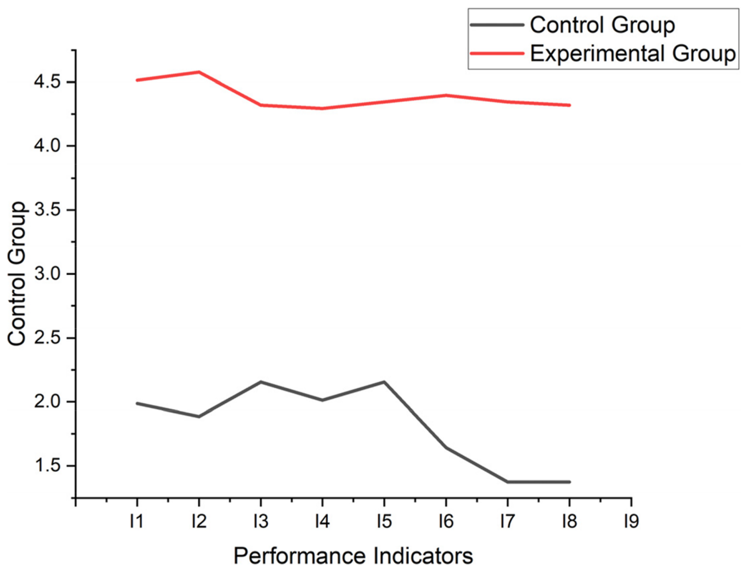

Figure 5. Comparative analysis of student satisfaction

In conclusion, AI-generated feedback can be interactive and adaptive, making the learning experience more engaging and responsive to students' progress in real-time.

6. Limitations and future scope

The results of the research study should be cautiously generalised for the following reasons. The students' encounters with AI-generated feedback were confined to "Gradescope," an AI tool for producing feedback. The participants chosen for the research study were from an advanced educational program, namely the "Creative Writing" curriculum. The Analytic Hierarchy Process (AHP) is an effective instrument for factor assessment; nevertheless, the number of components that may be analysed concurrently is restricted. Consequently, we were compelled to focus on the paramount performance metric that affects students' feedback mechanisms. Nevertheless, supplementary indicators are likely influencing the implementation of feedback systems, and we advocate for more quantitative and qualitative studies to investigate these factors. We advocate for further study to collect data from a larger sample size to empirically validate our hypotheses. The subjective character of the expert panel's judgements is a significant limitation of the research. The findings may have been affected by any bias present in the questionnaire replies. Additionally, other multi-criteria decision-making (MCDM) methods may be used to enhance the fuzzy technique.

Consequently, future research should explore the application of a wider range of AI tools across diverse learning programs and disciplines. Furthermore, the performance indicators identified in this study are broadly applicable and are not confined to any specific subject area, type of institution, or geographical context. This presents a valuable opportunity for future studies to validate and refine these indicators across varied educational settings and cultural environments, thereby enhancing their generalizability and robustness.

Institutional review board statement

This study involved human participants through interviews and surveys. Prior informed consent was obtained from all participants before data collection. The study did not involve any procedures that could cause physical or emotional harm, and therefore, formal Institutional Review Board (IRB) approval was not required under the applicable guidelines.

Originality statement

This manuscript presents an original contribution to the field of education and e-learning by integrating Artificial Intelligence (AI) with the Analytic Hierarchy Process (AHP) to develop a structured decision-making framework for enhancing student feedback.

References

Afzaal, M., Nouri, J., Zia, A., Papapetrou, P., Fors, U., Wu, Y., . . ., & Weegar, R. (2021). Explainable AI for data-driven feedback and intelligent action recommendations to support students self-regulation. *Frontiers in Artificial Intelligence*, *4*, 1–20. doi: 10.3389/frai.2021.723447.

Arredondo, F., & Martinez, E. (2010). Learning and adaptation of a policy for dynamic order acceptance in make-to-order manufacturing. *Computers & Industrial Engineering*, *58*(1), 70–83. doi: 10.1016/j.cie.2009.08.005.

Barak, M. (2010). Motivating self-regulated learning in technology education. *International Journal of Technology and Design Education*, *20*(4), 381–401. doi: 10.1007/s10798-009-9092-x.

Batwara, A., Sharma, V., Makkar, M., & Giallanza, A. (2022). An empirical investigation of green product design and development strategies for eco industries using kano model and fuzzy AHP. *Sustainability*, *14*(14), 8735. doi: 10.3390/su14148735.

Buckley, J. J. (1985). Fuzzy hierarchical analysis. *Fuzzy Sets and Systems*, *17*(3), 233–247. doi: 10.1016/0165-0114(85)90090-9.

Chan, C. K. Y., & Colloton, T. (2024). The future of artificial intelligence in education. *Generative AI in Higher Education*, *23*(1), 240–255. doi: 10.4324/9781003459026-7.

Dizon, G., & Tang, D. (2020). Intelligent personal assistants for autonomous second language learning: An investigation of Alexa. *JALT CALL Journal*, *16*(2), 107–120. doi: 10.29140/jaltcall.v16n2.273.

Dönmez, M. (2024). AI-based feedback tools in education : A comprehensive bibliometric analysis study. *11*(4), 622–646, 10.21449/ijate.1467476.

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. *International Journal of Information Management*, *48*(January), 63–71. doi: 10.1016/j.ijinfomgt.2019.01.021.

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*. doi: 10.1155/2022/5215722.

Kazu, İ. Y., & Kuvvetli, M. (2023). The influence of pronunciation education via artificial intelligence technology on vocabulary acquisition in learning English. *International Journal of Psychology and Educational Studies*, *10*(2), 480–493. doi: 10.52380/ijpes.2023.10.2.1044.

Keane, E., & Labhrainn, M. (2005). Obtaining student feedback on teaching & course quality. *Briefing Paper*, *2*(April), 1–19. Available from: http://www.nuigalway.ie/celt/documents/evaluation_ofteaching.pdf

Khan, I., Ahmad, A. R., Jabeur, N., & Mahdi, M. N. (2021). An artificial intelligence approach to monitor student performance and devise preventive measures. *Smart Learning Environments*, *8*(1), 17. doi: 10.1186/s40561-021-00161-y.

Kiker, G. A., Bridges, T. S., Varghese, A., Seager, P. T. P., & Linkov, I. (2005). Application of multicriteria decision analysis in environmental decision making. *Integrated Environmental Assessment and Management*, *1*(2), 95–108. doi: 10.1897/IEAM_2004a-015.1.

Kirkpatrick, J. (2010). *An Introduction to the New World Kirkpatrick Model*. Newnan, GA: Kirkpatrick Partners, LLC.

Leckey, J., & Neill, N. (2001). Quantifying quality: The importance of student feedback. *Quality in Higher Education*, *7*(1), 19–32. doi: 10.1080/13538320120045058.

Lin, M. P. C., Liu, A. L., Poitras, E., Chang, M., & Chang, D. H. (2024). An exploratory study on the efficacy and inclusivity of AI technologies in diverse learning environments. *Sustainability*, *16*(20), 8992. doi: 10.3390/su16208992.

Melewar, T. C., Alwi, S., Hsu, P. F., & Lin, F. L. (2013). Developing a decision model for brand naming using Delphi method and analytic hierarchy process. *Asia Pacific Journal of Marketing and Logistics*, *25*(2), 187–199. doi: 10.1108/13555851311314013.

Naz, I., & Robertson, R. (2024). Exploring the feasibility and efficacy of ChatGPT3 for personalized feedback in teaching. *Electronic Journal of e-Learning*, *22*(2), 98–111. doi: 10.34190/ejel.22.2.3345.

Patel, N. K. (2018). Effect of integrated feedback on classroom climate of secondary school teachers. *International Journal of Evaluation and Research in Education*, *7*(1), 65. doi: 10.11591/ijere.v7i1.11146.

Ponnusamy, P., Ghias, A. R., Yi, Y., Yao, B., Guo, C., & Sarikaya, R. (2021). Feedback-based self-learning in large-scale conversational AI agents. *AI Magazine*, *42*(4), 43–56. doi: 10.1609/aaai.12025.

Punia Sindhu, S., Nehra, V., & Luthra, S. (2016). Recognition and prioritization of challenges in growth of solar energy using analytical hierarchy process: Indian outlook. *Energy*, *100*, 332–348. doi: 10.1016/j.energy.2016.01.091.

Saaty, T. L. (1980). *The analytic hierarchy process: Planning, priority setting, resource allocation*. McGraw-Hill.

Selwyn, N., & Bulfin, S. (2016). Exploring school regulation of students' technology use – rules that are made to be broken?. *Educational Review*, *68*(3), 274–290. doi: 10.1080/00131911.2015.1090401.

Shaik, T., Tao, X., Li, Y., Dann, C., McDonald, J., Redmond, P., & Galligan, L. (2022). The effectiveness of Learning to improve students' higher-order thinking skills. *IEEE Access*, *10*, 56720–56739. doi: 10.1109/ACCESS.2022.3177752.

Sivakumar, V. L., Radha Krishnappa, R., & Nallanathel, M. (2020). Drought vulnerability assessment and mapping using Multi-Criteria decision making (MCDM) and application of Analytic Hierarchy process (AHP) for Namakkal District, Tamilnadu, India, *Materials Today: Proceedings*, *xxxx*, 1592–1599. doi: 10.1016/j.matpr.2020.09.657.

Smidt, A., Balandin, S., Sigafoos, J., & Reed, V. A. (2009). The Kirkpatrick model: A useful tool for evaluating training outcomes. *Journal of Intellectual and Developmental Disability*, *34*(3), 266–274. doi: 10.1080/13668250903093125.

Spector, J. M., & Ma, S. (2019). Inquiry and critical thinking skills for the next generation: From artificial intelligence back to human intelligence. *Smart Learning Environments*, *6*(1), 8. doi: 10.1186/s40561-019-0088-z.

Tapalova, O., & Zhiyenbayeva, N. (2022). Artificial intelligence in education: AIEd for personalised learning pathways. *Electronic Journal of e-Learning*, *20*(5), 639–653. doi: 10.34190/ejel.20.5.2597.

Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). The future of artificial intelligence in education. *International Journal of Public Administration*, *42*(7), 596–615. doi: 10.1080/01900692.2018.1498103.

Yu, E., & Canton, S. (2023). Intelligent enough ? Artificial intelligence. *Journal Of Education Online*, *10*(1), 1–14.

Corresponding author
Amber Batwara can be contacted at: amberbatwara@gmail.com