

## AI-enhanced virtual simulation for vocational engineering education

Yan Li <sup>1,2\*</sup> , Wai Yie Leong <sup>2</sup> 

<sup>1</sup> Heilongjiang Institute of Construction Technology, Harbin, CHINA

<sup>2</sup> INTI International University, Nilai, MALAYSIA

Received 25 December 2025 • Accepted 21 May 2026

### Abstract

Artificial intelligence (AI)-enhanced virtual simulation has emerged as a transformative approach for advancing vocational engineering education by addressing limitations of traditional workshop-based training. This study develops and evaluates an AI-enhanced virtual simulation system that integrates pedagogical, functional, and technical alignment under the pedagogical-functional-technical mapping framework. A six-week quasi-experimental study involving 80 students in a cement process control course was conducted. The proposed system combines intelligent tutoring, adaptive task sequencing, and real-time performance analytics to personalize learning pathways according to individual skill levels and learning progress. By simulating authentic industrial scenarios, the platform enables learners to practice complex engineering operations safely and repeatedly, while receiving immediate, data-driven feedback on procedural accuracy, efficiency, and decision-making. The framework is pedagogically grounded in experiential and mastery-based learning, functionally aligned with occupational standards, and technically supported by AI models for learner modeling and assessment. The findings highlight the potential of AI-enhanced virtual simulation to improve skill acquisition, learning engagement, and training scalability, offering a robust digital solution for future-ready vocational engineering education. Results indicate statistically significant improvements in learning performance and operational efficiency for the experimental group. These outcomes should be interpreted cautiously given the study's contextual and methodological constraints.

**Keywords:** AI-enhanced simulation, vocational engineering education, formative assessment, instructional design alignment, learning analytics, education quality

## INTRODUCTION

Simulation-based teaching has become a core component of vocational and engineering education, particularly in process-intensive domains such as manufacturing, energy systems, and materials engineering. Virtual simulation environments enable learners to practice complex operations without physical risk or excessive resource consumption, and prior studies have shown benefits in terms of engagement and procedural accuracy (Radianti et al., 2020). However, despite increasing technical sophistication, many simulation platforms continue to exhibit persistent pedagogical limitations.

One widely reported issue is that simulation activities often devolve into repetitive interface

manipulation or “button pressing,” where learners follow prescribed operational steps without engaging in deeper reasoning about process logic or system behavior (García-Robles et al. 2024). Empirical studies indicate that such use of simulation does not necessarily lead to improved conceptual understanding or transferable problem-solving skills. In addition, many existing platforms rely on uniform task sequences and generic feedback, offering limited support for personalized learning or self-regulated practice. These limitations are not purely technical; rather, they reflect a misalignment among instructional objectives, system functions, and technical implementation.

Recent research has explored the integration of artificial intelligence (AI) into simulation systems to address these challenges through intelligent tutoring,

### Contribution to the literature

- This study proposes a PFT-M-based AI-enhanced simulation framework that aligns learning objectives, system functions, and AI implementation.
- This study reports evidence from a six-week quasi-experimental intervention, showing improved learning performance and operational efficiency.
- This study offers a human-centered approach to AI-supported vocational education through formative feedback, adaptive learning, and learning analytics.

automated assessment, and adaptive feedback. While AI-enhanced simulation systems show promise, most existing implementations emphasize algorithmic capabilities—such as classification accuracy or dialogue generation—without explicitly mapping AI functions to curriculum objectives or competency structures. As a result, AI is often introduced as an auxiliary feature rather than embedded as an integral component of pedagogical design. Moreover, few studies systematically examine how AI-supported feedback influences learners' reasoning processes over time or how automated assessment aligns with instructor judgment in authentic classroom settings (Zawacki-Richter et al., 2019).

From an Industry 5.0 perspective, these shortcomings are particularly problematic. Industry 5.0 emphasizes human-centered systems, resilience, and responsible human-machine collaboration rather than efficiency-driven automation alone. In vocational engineering education, this implies that learners should develop competencies such as reflective decision-making under constraints, adaptive responses to system disturbances, and informed interaction with intelligent systems. However, these competencies are rarely operationalized in existing simulation studies. Human-centered design is often reduced to usability considerations, resilience is addressed primarily at the system level, and ethical or responsible AI is seldom examined in relation to learner behavior or assessment criteria. In this study, Industry 5.0 principles are used as a conceptual lens to inform system design priorities rather than as a set of directly measured outcome variables.

A key reason for these limitations lies in the lack of a coherent design framework that explicitly links pedagogical intent, functional modules, and technical implementation. Although various instructional design models and simulation frameworks have been proposed—such as task-based training models, competency-oriented curricula, and adaptive learning architectures—many treat pedagogy and system architecture as loosely coupled layers. Consequently, feedback mechanisms, assessment logic, and AI functions are not systematically aligned with learning objectives, leading to fragmented or superficial educational outcomes.

To overcome this limitation, this study adopts and refines the pedagogical-functional-technical mapping

(PFT-M) framework as a guiding principle for the design of AI-enhanced simulation systems. Unlike general instructional design models, PFT-M requires explicit mapping between pedagogical objectives, corresponding system functions, and concrete technical implementations. In this work, the framework is operationalized in a vocational cement process control course, where learning objectives are translated into AI-supported argumentation, automated process assessment, and data-driven personalized practice. Rather than claiming universal applicability, the study evaluates the practical value of this explicit alignment by examining whether it leads to measurable improvements in learning performance, self-directed learning, and assessment reliability.

Accordingly, the paper addresses the following research questions (RQs):

- RQ1:** How does an AI-enhanced simulation system designed under the PFT-M framework influence students' conceptual understanding and operational performance?
- RQ2:** What role does data-driven, AI-supported feedback play in supporting self-regulated learning and instructional decision-making?
- RQ3:** To what extent can a PFT-M-guided design support Industry 5.0-oriented competencies in a vocational engineering context?

By grounding these questions in a concrete system implementation and empirical evaluation, this study seeks to move beyond generic claims of AI effectiveness and contribute a critically examined design-and-evaluation case to the engineering education literature.

## LITERATURE REVIEW

### Simulation-Based Teaching in Vocational Education

Simulation-based teaching has become a key approach in vocational and engineering programs. Virtual plants, digital twins, and remote laboratories enable students to practice operating procedures and fault handling with reduced risk and cost (Vergara et al., 2022; Zhang & Liu, 2023). Studies have reported benefits such as higher engagement and improved procedural accuracy, particularly in hazardous domains like chemical process control and power systems (Zhai & Wibowo, 2023). At the same time, there is concern that

many simulation activities remain close to “button pressing” and do not sufficiently develop conceptual understanding or transfer of learning to new tasks.

### Integration of Artificial Intelligence in Simulation Systems

The integration of AI into educational simulation has opened new possibilities for intelligent tutoring, automated assessment, and real-time feedback. AI can analyze learner behaviors, identify misconceptions, and generate personalized guidance through natural language understanding and pattern recognition (Roll & Wylie, 2016).

Recent studies have incorporated AI-driven evaluation modules, yet many of these approaches focus primarily on algorithmic sophistication, neglecting pedagogical alignment with curriculum objectives (Liu et al., 2025). In vocational contexts, where competency-based outcomes are emphasized, the absence of a systematic mapping between instructional goals and AI system functions often limits the educational value of technological innovation.

### Pedagogical and Learning Theories

Constructivist learning theory asserts that learners actively construct knowledge through exploration, reflection, and interaction. Simulation environments embody this principle by enabling learners to test hypotheses and observe outcomes in a safe, virtual space.

Adaptive learning theory further expands this perspective by emphasizing the dynamic adjustment of learning paths based on individual learner data and performance. In simulation teaching, this approach allows the system to tailor scenarios and exercises according to each learner’s progress and difficulties.

Formative assessment theory highlights the importance of continuous feedback and reflection during the learning process (Anderson, 2001). In an AI-enhanced simulation context, formative assessment is implemented through data-driven analytics, automatic grading, and individualized performance reports that support both self-regulation and teacher intervention.

### Research on Frameworks for Simulation Design

Several researchers have attempted to establish structured frameworks that integrate pedagogical principles into simulation system design. Early models focused on task-based training structures, while later frameworks incorporated competency mapping and feedback loops. Nevertheless, most of these designs treat educational, functional, and technical layers separately, which often leads to feature stacking and a lack of coherence between curriculum goals and platform architecture.

The PFT-M framework addresses this issue by providing a unified structure that aligns teaching objectives, system functionalities, and technological implementations. It bridges the gap between instructional theory and practical design, offering a scalable approach for building adaptive and competency-oriented simulation platforms (Segovia et al., 2016).

### Research Gaps and Rationale

Although research on AI-supported education and virtual simulation has progressed rapidly, several gaps remain particularly salient for vocational engineering education:

1. Empirical evaluations of AI-enhanced simulation systems in real teaching environments are still limited.
2. Many systems lack clear alignment between instructional objectives, platform functionalities, and technical implementation.
3. The integration of adaptive learning mechanisms and data-driven formative assessment is often weak or fragmented.

These gaps suggest the need for a systematic design framework that can connect pedagogical intent with system architecture and that embeds continuous performance feedback within authentic simulation tasks.

## RESEARCH DESIGN AND SYSTEM OVERVIEW

This section first introduces the overall research design, followed by the theoretical framework, system architecture, and implementation mechanisms.

### Research Design Overview

This study adopts a quasi-experimental design conducted over six weeks in a vocational cement process control course. A total of 80 students participated and were divided into an experimental group ( $n = 40$ ) and a control group ( $n = 40$ ). The grouping followed existing class arrangements rather than random assignment, which reflects authentic classroom conditions but limits strict causal inference. Baseline equivalence was assumed based on similar academic background and course exposure.

Both groups followed the same syllabus, learning objectives, and instructional schedule. The control group used a conventional simulation platform with instructor-guided procedures, while the experimental group used the AI-enhanced simulation system designed under the PFT-M framework.

The instructional intervention consisted of weekly simulation tasks aligned with course topics, including process configuration, parameter tuning, and fault

**Table 1.** Mapping of pedagogical objectives, functional modules, and technical implementation

Pedagogical	Functional module	Technical implementation
Understand process flow	Scenario-based walkthrough	Unity3D path animation and layered annotations
Master operational skills	Equipment interaction & disassembly	3D object control and step logic
Evaluate system states	DCS task scheduling	Virtual DCS interface and condition scripting
Assess learning outcomes	Quiz & feedback module	Online test engine and xAPI-based data capture

handling. Learning outcomes were evaluated using multiple data sources, including post-test scores, system-generated performance indicators, and behavioral logs.

Data were collected throughout the intervention period and analyzed using descriptive statistics and independent-samples t-tests to compare group performance. Effect sizes (Cohen's *d*) were calculated to assess the magnitude of differences between groups. This study focuses on comparative outcomes rather than establishing causal relationships.

### Theoretical Foundation and Design Methodology

Rapid advances in simulation and AI technology have made it possible to build highly complex learning platforms. Without a clear design method, however, such systems may become collections of loosely related functions with little pedagogical coherence (Leva & Panzani, 2024). In this work, the PFT-M framework is used as the main design guideline.

At the pedagogical layer, course objectives for the cement process control curriculum are analyzed to identify the knowledge, skills, and attitudes to be developed, such as understanding of process flow, parameter coupling, safety constraints, and troubleshooting strategies.

At the functional layer, these objectives are translated into platform functions: scenario-based task presentation, student-driven process design, AI-mediated argumentation and feedback, automatic scoring, and personalized behavior practice management.

At the technical layer, specific technologies are selected to implement these functions. These include a process simulation engine based on real industrial data, natural language processing for dialogue with students, a knowledge graph describing process constraints and fault relationships, and data-logging and analytics components for tracking learner behavior (J. Xu et al., 2023).

By iterating across these layers, the design aims to keep technological choices aligned with educational needs and to avoid redundant features.

### Structure of the Model and Mapping Logic

According to the PFT-M model, the design of a simulation system can be divided into three levels.

The pedagogical level specifies what students should learn, based on curriculum standards and course learning outcomes. For the present course, this includes the ability to read and design process flow diagrams, tune key operating parameters, and reason about the impact of parameter changes on product quality and energy consumption.

The functional level translates these learning goals into concrete system features. Examples include interactive process configuration, real-time visualization of process states, logging of student actions, AI-supported dialogues that probe students' reasoning, and an assessment module that can evaluate both results and intermediate decisions.

The technical level then identifies technologies that can support these functions in a robust and scalable way. The platform uses a modular architecture with a process simulation kernel, an AI interaction module, a learning analytics engine, and a teacher dashboard.

This mapping procedure helps to avoid feature stacking, a design choice that is pedagogically hard to justify (experience API [xAPI], 2023).

It offers an organized approach to align educational requirements with practical technology and design solutions. The layers' fitting together is illustrated in **Table 1**.

Rather than prescribing specific tools, the model guides system design through a structured alignment between curriculum goals, functional tasks, and technological execution. This enhances the platform's utility in a diverse array of disciplines and domains (Alonso-Fernández et al., 2022; Npuls Consortium, 2024; Görzen et al., 2024).

### Layers of System Architecture and Function

The simulation platform follows a closed-loop teaching process that includes three main stages: task assignment, task execution, and performance evaluation. It is composed of three core subsystems. The instructor terminal is used to set up training scenarios, assign simulation tasks, and manage assessments. The learner terminal allows students to carry out tasks, interact with equipment, and complete in-system evaluations. The data analytics module records user activities, rebuilds action sequences, and generates performance indicators for further analysis.

These subsystems are connected via network interfaces, Socket for real-time command synchronization, and RESTful API for asynchronous

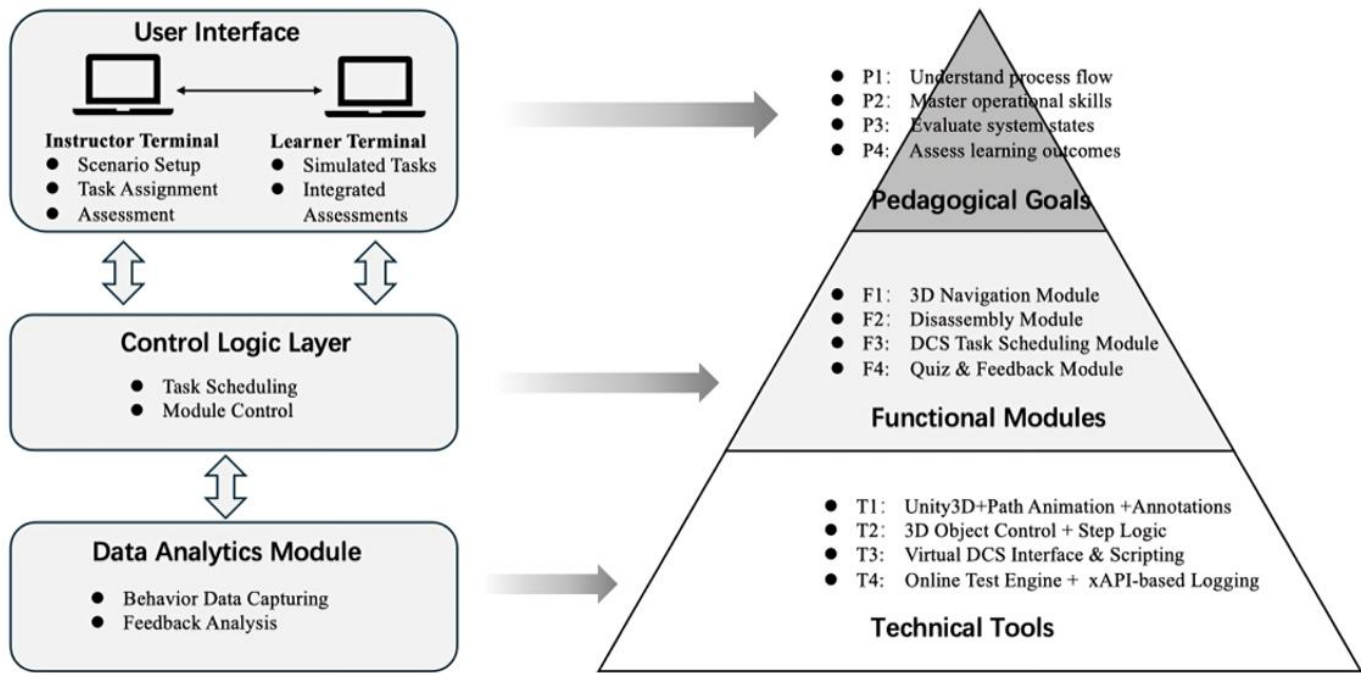


Figure 1. Integrated architecture and PFT-M of the AI-enhanced simulation platform (Source: Authors' own development)

data exchange and feedback delivery. The system architecture is modular and adaptable, allowing for reuse across different course settings with minimal reconfiguration (Lo et al., 2024; Wang et al., 2024).

In Figure 1, the right side of the diagram is represented as a pyramid structure to emphasize the hierarchical dependency among the three layers of the PFT-M framework. Pedagogical goals are positioned at the top to indicate that they define the ultimate direction and evaluation criteria of the system. Functional modules occupy the intermediate layer, translating abstract educational objectives into actionable system features. Technical implementation forms the base layer, providing the necessary infrastructure to support these functions.

This representation highlights that technical choices should be driven by functional requirements, which in turn are derived from pedagogical intent. The pyramid structure is therefore used to illustrate a top-down design logic rather than a bottom-up accumulation of technological features.

### Mechanism for Learning Data Collection and Feedback

To support continuous monitoring and feedback, the platform incorporates a data-driven mechanism with three main components.

**Behavior logging.** During simulation tasks, learner activities such as parameter adjustments, control operations, and task completion states are recorded using the xAPI. Dialogue content from AI-student interactions is also logged, creating a semantic record of

student decision-making and reasoning (Issa et al., 2024; Ngo et al., 2024).

**Data processing.** Behavioral and semantic logs are processed using natural language techniques and knowledge-graph-based reasoning to identify error types, misconceptions, and strategies used by learners. These indicators are cross-checked with instructor evaluations to ensure that automated interpretations are consistent with expert judgement (H. Xu et al., 2024).

**Feedback generation.** Based on these analyses, the system generates formative feedback at different levels: immediate hints during AI dialogues, post-task feedback reports summarizing strengths and weaknesses, and course-level overviews of progress. It also recommends follow-up exercises and adjusts task difficulty to match the learner's evolving competence (Loong & Chang, 2024; Raihan et al., 2025).

This design shifts the emphasis from purely outcome-based assessment to continuous, process-oriented support and lays a foundation for subsequent empirical study.

### AI-Supported Argumentation in Teaching and Learning

To bring the PFT-M framework into everyday classroom practice, the system organizes teaching, learning, practice, and assessment as an AI-supported closed loop.

#### *Learner-centered process design and AI argumentation*

Conventional simulation courses typically follow a sequence in which teachers demonstrate key procedures, students observe, and then practice under teacher

supervision. In the present system, this sequence is deliberately changed. After receiving a production target and basic constraints, students first design their own process schemes in the virtual environment, including choices of raw material proportions, temperature curves, pressure settings, and control strategies.

They then enter a dialogue with the AI learning module. In this dialogue, students are asked to explain why they selected parameter values or operating rules. The AI module interprets these explanations using the domain knowledge graph, compares them with process constraints, and raises questions when inconsistencies or unsafe assumptions are detected (Anwar et al., 2022). Rather than directly providing correct answers, the AI offers stepwise hints and counterexamples. Students modify their designs, run new simulations, and observe system responses until a technically acceptable or improved solution is obtained (Shute & Rahimi, 2017).

In this way, the simulator becomes a space for reasoning and reflection, not just for reproducing fixed procedures. The interaction also produces rich behavioral and semantic data that feed into later analysis (Ifenthaler & Widanapathirana, 2014).

### *Intelligent management of practice and assessment*

In the practice and examination stages, an evaluation module based on AI supports management and feedback:

First, the automated scoring module evaluates classroom practical tasks, formal examinations, and after class exercises by comparing students' operation trajectories and process states with expert reference models (Leva & Panzani, 2024). Both final outcomes and intermediate operations (such as the stability of temperature control or response to alarms) are considered.

Second, the system tracks error prone behavior over time. For each student, it maintains a profile of frequent errors and misconceptions, and records how these change across tasks and sessions (Dias et al., 2020).

Third, the analytics module uses these profiles to assemble personalized behavior training content and practice papers (Chango et al., 2022). Exercises focus on unresolved weak points, while skills that have been mastered receive less emphasis, thus making practice more efficient.

Finally, course level feedback is generated. Learners receive individual reports on their progress, strengths, and remaining difficulties, and instructors can inspect aggregated dashboards showing common error types and learning trajectories for the class.

Through this AI-centered workflow, teaching, learning, practice, and assessment are linked by data and feedback rather than treated as separate stages,

consistent with the PFT-M principle of aligning pedagogy, system functions, and technology.

### **AI-Supported Teaching, Learning, and Assessment Model**

The AI-enhanced system departs from the familiar pattern of "teacher explains-students watch-students practice-teacher corrects." The core mechanisms can be described from two angles: how students design and refine process schemes in dialogue with AI during teaching and learning, and how practice and assessment are handled through automatic scoring and personalized behavior management.

#### *Learner-centered AI argumentation*

In the learning phase, students first design their own process schemes instead of reproducing a teacher's example (Chi & Wylie, 2014). For learner  $i$ , the scheme is written as a decision vector using Eq. (1):

$$\mathbf{x}_i^{(0)} = (x_{i1}^{(0)}, x_{i2}^{(0)}, \dots, x_{im}^{(0)}), \quad (1)$$

where  $x_{ik}$  is the  $k$ -th operating parameter of the cement production line.

For the clinker line used in this course, the vector is instantiated with  $m = 8$  components:

1.  $x_{i1}$ : raw meal feed rate (t/h),
2.  $x_{i2}$ : main burner coal feed rate (kg/s or t/h),
3.  $x_{i3}$ : preheater exit gas temperature setpoint ( $^{\circ}\text{C}$ ),
4.  $x_{i4}$ : rotary kiln rotation speed (rpm),
5.  $x_{i5}$ : ID fan speed or flow rate (% or  $\text{m}^3/\text{h}$ ),
6.  $x_{i6}$ : O<sub>2</sub> concentration setpoint at the kiln outlet (%),
7.  $x_{i7}$ : clinker cooler primary air flow or pressure, and
8.  $x_{i8}$ : cement mill production rate setpoint (t/h).

The feasible region  $\Omega$  is determined by process and safety constraints:

$$g_{\ell}(\mathbf{x}) \leq 0, \quad (2)$$

where  $\ell = 1, \dots, L$ , which impose limits on temperatures, pressures, gas composition and equipment ranges (Graesser et al., 2012).

Once an initial scheme has been proposed, the student explains the reasoning behind  $\mathbf{x}_i^{(t)}$  in natural language. The AI module analyzes this explanation, checks whether constraints are violated and whether the scheme is efficient, and then returns structured feedback. Two scalar functions are used to summarize this judgement. The first measures constraint violation:

$$E(\mathbf{x}_i^{(t)}) = \sum_{\ell=1}^L \max(0, g_{\ell}(\mathbf{x}_i^{(t)})), \quad (3)$$

and the second combines several performance aspects:

$$J(\mathbf{x}_i^{(t)}) = \lambda_1 C_{qual}^{(t)} + \lambda_2 C_{energy}^{(t)} + \lambda_3 C_{stability}^{(t)} \quad (4)$$

with  $\lambda_1, \lambda_2, \lambda_3 \geq 0$  and  $\lambda_1 + \lambda_2 + \lambda_3 = 1$ .

Here,  $C_{qual}^{(t)}$  aggregates deviations of quality indices (for example, free CaO and Blaine fineness) from their targets,  $C_{energy}^{(t)}$  reflects specific heat and power consumption per ton of product, and  $C_{stability}^{(t)}$  penalizes fluctuations in key variables such as kiln hood temperature and main drive current. These choices are illustrative and follow the indicators already used in the course (Masri et al., 2019).

Interaction with the AI module leads to a sequence of revisions:

$$\mathbf{x}_i^{(t+1)} = \mathbf{x}_i^{(t)} + \Delta \mathbf{x}_i^{(t)}, \quad (5)$$

where  $\Delta \mathbf{x}_i^{(t)}$  is the change the learner makes to the eight parameters in response to error prompts, counter-examples and hints. The dialogue continues until:

$$E(\mathbf{x}_i^{(t)}) \leq \varepsilon_E \text{ and } J(\mathbf{x}_i^{(t)}) \leq \varepsilon_J, \quad (6)$$

for small thresholds  $\varepsilon_E, \varepsilon_J > 0$ . At this point the scheme can be regarded as technically acceptable or close to optimal. In this way, students work through the reasoning behind their designs, revise parameters step by step with AI support, and reach a sound operating strategy rather than simply copying a standard answer.

### AI-based assessment and personalized recommendation

Practice and assessment are organized around an automatic scoring engine and a personalized behavior recommendation mechanism instead of uniform manual marking.

For learner  $i$  on task  $j$ , the automatic score  $S_{ij}$  is calculated from four dimensions (Baker & Inventado, 2014):

$$S_{ij} = w_1 C_{ij} + w_2 Sa_{ij} + w_3 St_{ij} + w_4 P_{ij}, \quad (7)$$

with

$$w_1 + w_2 + w_3 + w_4 = 1, \quad (8)$$

where  $w_k \geq 0$ .

Here,  $C_{ij}$  measures correctness against task goals and required process parameters,  $Sa_{ij}$  reflects safety (such as limit violations and alarm handling),  $St_{ij}$  represents process stability, and  $P_{ij}$  refers to procedural compliance including startup and shutdown sequences. The weights  $w_1, \dots, w_4$  are set by the teaching team in line with the competency requirements of vocational practice, with higher weights typically assigned to safety and correctness (Henri et al., 2017).

The course grade  $G_i$  combines theory, conventional practice and AI-based process scores:

$$G_i = \alpha S_i^{theory} + \beta S_i^{practice} + \gamma S_i^{AI}, \quad (9)$$

where  $\alpha + \beta + \gamma = 1$   $S_i^{theory}$  is the written test score,  $S_i^{practice}$  is the instructor's practical score, and  $S_i^{AI}$  is the aggregated automatic score across simulation tasks. This structure makes the assessment basis clear and traceable.

To support individualized training, each learner has a time-varying competency profile (Winne, 2017):

$$\mathbf{v}_i^{(t)} = (ER_i^{(1)}(t), ER_i^{(2)}(t), \dots, ER_i^{(D)}(t)), \quad (10)$$

where  $ER_i^{(d)}(t)$  is the error rate of learner  $i$  on competency dimension  $d$  up to opportunity  $t$ , and  $D$  is the number of defined dimensions.

In the present course, the profile is instantiated with  $D = 6$  dimensions:

1. raw meal proportioning and homogenization,
2. preheater and kiln temperature control,
3. kiln safety and alarm handling,
4. ID fan and draft control,
5. clinker cooling and cooler operation, and
6. startup and shutdown procedure compliance.

Each practice task  $r$  is tagged with one or more of these dimensions. For example, a raw meal stabilization task is linked to dimension 1, a kiln temperature control scenario to dimension 2, a combined  $O_2$  and draft control task to dimensions 2 and 4, and an alarm and interlock handling task to dimension 3 and dimension 6. The recommendation priority of task  $r$  for learner  $i$  at opportunity  $t$  is:

$$p_{ir}(t) = \sum_{d \in \mathcal{D}_r} ER_i^{(d)}(t), \quad (11)$$

where  $\mathcal{D}_r$  is the set of dimensions covered by task  $r$ . Tasks with higher  $p_{ir}(t)$  are scheduled earlier, so that practice focuses on weak areas instead of repeating already-mastered skills (Banihashem et al., 2024).

Across the full course, the sequence  $\{\mathbf{v}_i^{(t)}\}_{t=1}^T$  and the trajectory  $\{G_i^{(t)}\}_{t=1}^T$  give a longitudinal view of each learner's performance, misconceptions and progress. On this basis, the AI module can generate reports on learning content, strengths, persistent errors and improvements, supporting continuous guidance and personalized behavior evaluation for both students and teachers.

The following section focuses specifically on methodological alignment and instructional logic, rather than on system description or implementation details.

## INSTRUCTIONAL MECHANISM AND AI IMPLEMENTATION

### Design Rationale and Scope of the PFT-M Framework

The PFT-M framework is employed in this study as a design-oriented alignment mechanism for AI-enhanced simulation systems rather than as a novel

learning theory. Its three-layer structure—pedagogical objectives, functional requirements, and technical implementation—is conceptually intuitive and consistent with established instructional design principles. The value of the framework lies in making the alignment between instructional intent and system functionality explicit, enforceable, and traceable throughout the design process.

In AI-enhanced simulation environments, pedagogical goals can easily become obscured by technical complexity or feature-driven development. To address this risk, the PFT-M framework is used to constrain system design choices in this study: each functional module and AI capability must be directly linked to a clearly defined learning objective. This approach is intended to reduce ad hoc feature accumulation and to ensure that AI components support targeted instructional purposes, particularly those related to reasoning, reflection, and self-directed learning.

### Mapping Pedagogical Objectives to AI-Supported Functions

At the pedagogical level, the target vocational engineering course emphasizes not only procedural accuracy but also reflective decision-making under operational constraints, consistent with human-centered principles associated with Industry 5.0. These objectives are translated into functional requirements, including learner-initiated process design, articulation of reasoning, formative feedback, and iterative revision.

At the functional level, AI is introduced specifically to support formative feedback rather than summative judgment. Learners are required to propose process solutions and explain their decisions, while the AI module provides prompts, hints, and reflective questions when inconsistencies, inefficiencies, or constraint violations are detected. Automated assessment operates continuously during task execution to generate diagnostic information, but final evaluation authority remains with instructors.

At the technical level, these functions are instantiated through a modular architecture combining a simulation engine, an AI reasoning module, learning data logging, and instructor-facing dashboards. The technical implementation is organized around feedback loops rather than linear task completion, reinforcing the formative orientation of the instructional design.

Figure 2 illustrates the AI-supported teaching-learning-assessment loop that operationalizes this mapping under the PFT-M framework.

### Formative Feedback Loop and Human-in-the-Loop Design

As shown in Figure 2, the core instructional mechanism of the system is a formative feedback loop

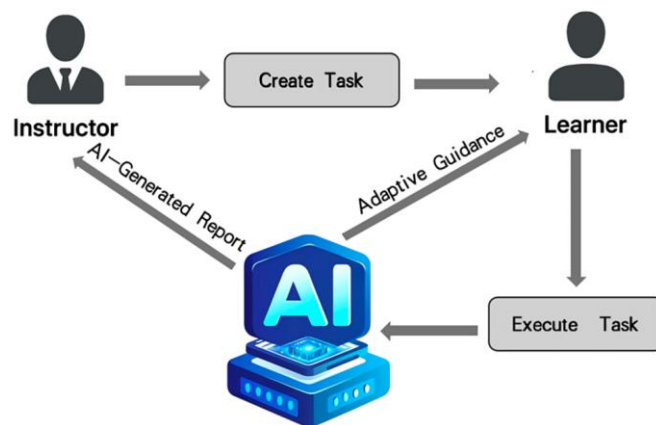


Figure 2. AI-supported teaching-learning-assessment loop illustrating formative feedback and learner revision under the PFT-M framework (Source: Authors' own development)

connecting learners, AI modules, and instructors. Learners first design process solutions and articulate their reasoning. The AI module analyzes these explanations against domain constraints and performance criteria and returns feedback in the form of prompts, warnings, or reflective questions that encourage revision.

Importantly, the system does not enforce a single optimal solution. Multiple solution paths are considered acceptable provided they satisfy safety constraints and remain within defined performance thresholds. AI feedback distinguishes among constraint violations, inefficient but acceptable strategies, and logically inconsistent explanations. Only constraint violations and conceptual inconsistencies trigger corrective feedback, while alternative feasible solutions are acknowledged through reflective prompts.

Instructor involvement is maintained throughout the process. Aggregated views of learner actions, common error patterns, and AI feedback histories are provided through dashboards, allowing instructors to intervene, adjust task parameters, or override automated suggestions when necessary. This human-in-the-loop design ensures that AI functions as a formative support tool rather than a decision authority (Huang, 2023). This human-in-the-loop design also reflects ethical considerations related to transparency and responsibility in AI-supported educational systems.

### AI Reasoning Transparency and Reliability

The AI reasoning module combines a domain-specific knowledge graph with natural language processing techniques to interpret learner explanations. The knowledge graph encodes operational constraints, causal relationships, and expert heuristics derived from process documentation and instructor input. Learner explanations are parsed to identify referenced parameters, causal claims, and justifications, which are then matched against the knowledge graph.

**Table 2.** Statistical comparison between experimental and control groups

Metric	Experimental group	Control group	t-value	p-value	Cohen's d
Post-test score (mean $\pm$ standard deviation)	86.3 $\pm$ 7.4	66.7 $\pm$ 8.1	3.91	< 0.01	0.87
Task time (min)	12.4	16.1	3.10	< 0.01	0.65
Error rate (%)	8.2	12.5	2.57	< 0.05	0.52

To assess interpretive reliability, AI-identified reasoning issues were compared with instructor judgments on a representative subset of interaction records. The agreement rate reached 87%, indicating that the AI module provides sufficiently accurate interpretations for formative feedback purposes. Discrepancies were primarily associated with incomplete learner verbalization rather than incorrect conceptual reasoning. These findings underscore that the AI module offers approximate, pedagogically useful interpretations rather than definitive diagnoses of learner understanding.

### Learning Data Collection for Formative Feedback

To support AI-assisted formative feedback and instructional oversight, the system incorporates a structured learning data collection mechanism aligned with the pedagogical objectives defined under the PFT-M framework. The purpose of learning data collection in this study is not to infer latent cognitive states or to predict learning outcomes, but to provide observable indicators that inform feedback, reflection, and instructional decision-making during simulation-based activities.

During task execution, the system records learner interactions at multiple levels, including parameter adjustments, control actions, task progression states, and time-stamped event sequences. In addition, textual data generated through learner-AI dialogue are logged to capture explanations, justifications, and revisions articulated by learners. All interaction records are stored as xAPI-compatible events to ensure consistency and traceability across tasks and sessions.

Collected data are processed to generate descriptive indicators relevant to formative feedback, such as the frequency of constraint violations, stability of operational adjustments, response latency to system prompts, and recurrence of specific error types. These indicators are used to trigger AI-supported prompts, reflective questions, or hints during task execution, enabling learners to identify inconsistencies and iteratively revise their decisions. At the instructional level, aggregated data support instructor awareness and targeted intervention without conflating data availability with claims of learning effectiveness.

## RESULTS

The results reported in this section reflect outcomes associated with the integrated AI-enhanced simulation system and should not be attributed to any single

component, such as AI dialogue, automated scoring, or personalized task recommendation. They should be interpreted as system-level outcomes rather than as direct causal validation of the PFT-M framework or any individual system component.

The post-test used in this study was developed based on course learning objectives and included both conceptual understanding and applied problem-solving items. The test consisted of 20 items with a total score of 100. Reliability analysis yielded a Cronbach's alpha of 0.82, indicating acceptable internal consistency.

Operational performance indicators, including task completion time and error rate, were automatically recorded by the simulation system. Error rate was defined as the proportion of incorrect operations relative to total required steps in each task.

### Learning Outcomes and Operational Efficiency

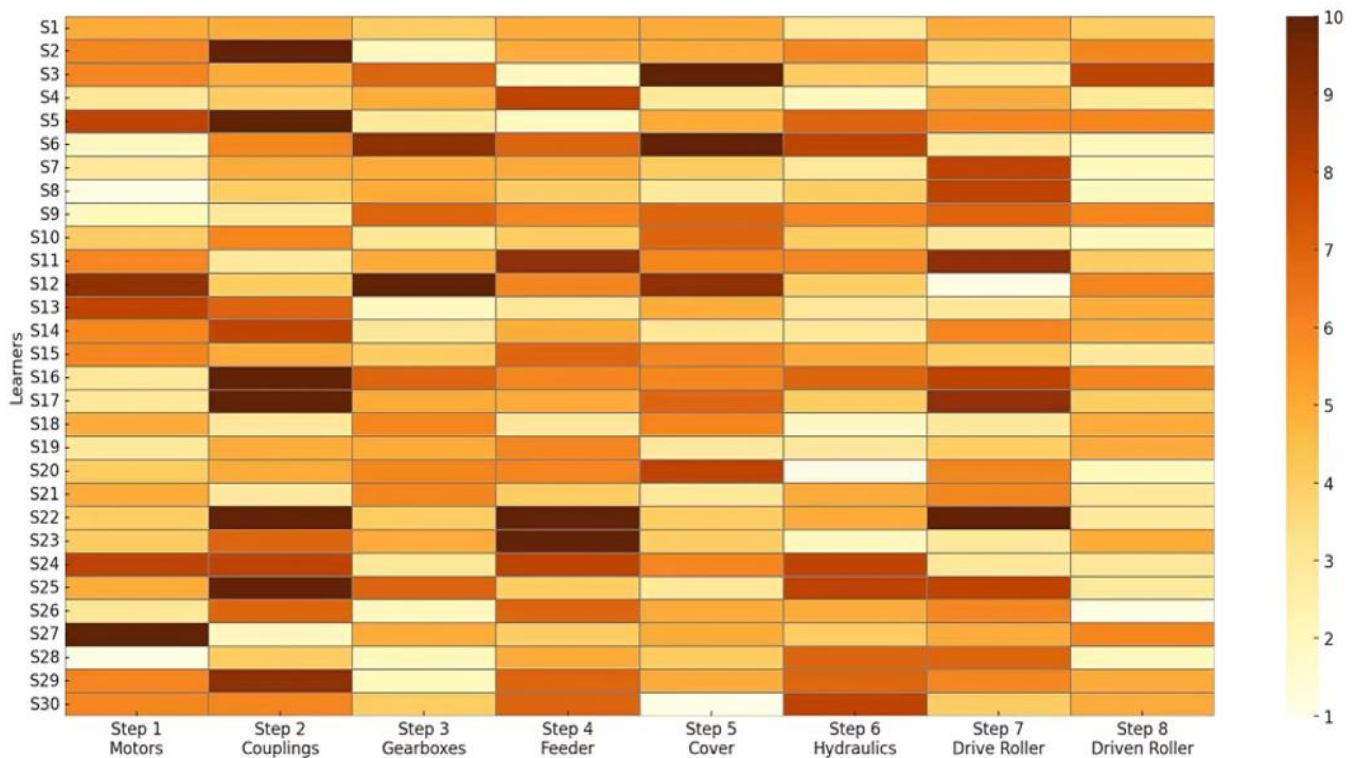
Learning outcomes were evaluated through post-test performance, task completion time, and operational error rates. Comparisons between the experimental group and the control group indicate that students using the AI-enhanced simulation system achieved higher post-test scores, completed tasks more efficiently, and exhibited lower error rates.

**Table 2** presents the statistical comparison between the experimental and control groups. Independent-samples t-tests indicate that the experimental group achieved significantly higher post-test scores ( $t = 3.91$ ,  $p < 0.01$ ), shorter task completion time ( $t = 3.10$ ,  $p < 0.01$ ), and lower error rates ( $t = 2.57$ ,  $p < 0.05$ ). Cohen's d values suggest medium to large effect sizes, indicating that these differences are both statistically significant and practically meaningful in the context of vocational training.

### Reliability of AI-Based Assessment

To examine the reliability of automated assessment, AI-generated scores were compared with instructor ratings for a set of operational tasks. Correlation analysis indicates a strong positive relationship between AI-based scores and instructor evaluations, suggesting that the automated assessment mechanism provides consistent and interpretable performance indicators suitable for formative use.

These findings support the use of AI-generated scores as a supplement to instructor judgment rather than as a replacement, reinforcing the human-in-the-loop assessment design adopted in the system.



**Figure 3.** Heatmap of learner interaction density highlighting high-cognitive-load stages during the simulation task (Source: Authors' own development based on simulation log data)

### Behavioral Patterns and Learning Trajectories

Learner interaction data were further analyzed to examine behavioral patterns during simulation tasks. **Figure 3** presents a heatmap of learner interaction density, highlighting task stages associated with higher levels of interaction and revision activity. These stages correspond to operations requiring greater coordination, parameter tuning, or constraint awareness.

The observed concentration of interactions at these stages suggests that the AI-supported formative feedback effectively directed learner attention toward cognitively demanding aspects of the task. Rather than indicating uniform engagement, the behavioral patterns reflect differentiated interaction aligned with task complexity.

### CONCLUSION AND FUTURE WORK

This study presents the design and classroom implementation of an AI-enhanced virtual simulation system for vocational engineering education, guided by the PFT-M framework. The system integrates AI-supported feedback, automated assessment, and learning analytics to support process-oriented learning and skill development.

The six-week quasi-experimental results show that students using the proposed system achieved higher post-test scores, completed tasks more efficiently, and made fewer operational errors than those using a conventional simulation platform. In addition, the

strong agreement between AI-generated scores and instructor evaluations suggests that the system can provide reliable support for formative assessment. Behavioral data further indicate increased engagement in cognitively demanding task stages, suggesting that AI-supported feedback helped guide attention and iterative revision.

With respect to the RQs, the findings show that the proposed system is associated with improved learning performance (**RQ1**), supports self-regulated learning through feedback and revision (**RQ2**), and demonstrates potential to foster reflective and human-centered competencies aligned with Industry 5.0 (**RQ3**). However, the third aspect was not directly measured and requires further empirical validation.

These findings should be interpreted with caution. Due to the quasi-experimental design, differences between groups, such as prior knowledge, motivation, and engagement, cannot be fully controlled. Although both groups followed the same syllabus and were taught by the same instructors, unobserved factors may have influenced the results. In addition, the relatively short duration of the intervention means that novelty effects cannot be excluded.

Another limitation is that the paper focuses on a single course in cement process control, which may limit the generalizability of the findings. The statistical analysis is limited to basic comparisons and does not capture more complex learning processes or interactions among variables.

Despite these limitations, the paper provides practical evidence that aligning pedagogical goals, system functions, and technical implementation can support more effective simulation-based learning. The PFT-M framework offers a structured approach for integrating AI into vocational education with an emphasis on learning processes rather than only outcomes.

Future work should extend this research in several directions. First, longitudinal and within-subject studies are needed to examine how learners' skills develop over time. Second, studies across multiple courses and institutions would help evaluate generalizability. Third, more advanced analytical methods, such as learning trajectory modeling, could provide deeper insights into how AI-supported feedback influences learning. Finally, further work is needed to improve the transparency and interpretability of AI systems in educational settings.

**Author contributions:** YL: conceptualization, methodology, software, investigation, data curation, formal analysis, writing – original draft, visualization; WYL: supervision, validation, writing – review & editing, project administration. Both authors agreed with the results and conclusions.

**Funding:** No funding source is reported for this study.

**Ethical statement:** This study involved an educational intervention in a vocational engineering course. Participation was voluntary, and informed consent was obtained before data collection. No sensitive personal data were collected, and all data were anonymized for analysis and reporting. According to institutional requirements and the nature of the study, formal ethics committee approval was not required.

**AI statement:** Generative AI tools were used only for language refinement during manuscript preparation. All scientific content, data interpretation, and conclusions were verified by the authors, who take full responsibility for the manuscript.

**Declaration of interest:** No conflict of interest is declared by the authors.

**Data sharing statement:** Data supporting the findings and conclusions are available upon request from the corresponding author.

## REFERENCES

- Alonso-Fernández, C., Calvo-Morata, A., Freire, M., Martínez-Ortiz, I., & Fernández-Manjón, B. (2022). Game learning analytics: Blending visual and data mining techniques to improve serious games and to better understand player learning. *Journal of Learning Analytics*, 9(3), 32-49. <https://doi.org/10.18608/jla.2022.7633>
- Anderson, L. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of education objectives*. Longman.
- Anwar, A., Haq, I. U., Mian, I. A., Shah, F., Alroobaea, R., Hussain, S., Ullah, S. S., & Umar, F. (2022). Applying real-time dynamic scaffolding techniques during tutoring sessions using intelligent tutoring systems. *Mobile Information Systems*. <https://doi.org/10.1155/2022/6006467>
- Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. Larusson, & B. White (Eds.), *Learning analytics* (pp. 61-75). Springer. [https://doi.org/10.1007/978-1-4614-3305-7\\_4](https://doi.org/10.1007/978-1-4614-3305-7_4)
- Banihashem, S. K., Dehghanzadeh, H., Clark, D., Noroozi, O., & Biemans, H. J. (2024). Learning analytics for online game-based learning: A systematic literature review. *Behaviour & Information Technology*, 43(12), 2689-2716. <https://doi.org/10.1080/0144929X.2023.2255301>
- Chango, W., Lara, J. A., Cerezo, R., & Romero, C. (2022). A review on data fusion in multimodal learning analytics and educational data mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 12(4), Article e1458. <https://doi.org/10.1002/widm.1458>
- Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219-243. <https://doi.org/10.1080/00461520.2014.965823>
- Dias, S. B., Hadjileontiadou, S. J., Diniz, J., & Hadjileontiadis, L. J. (2020). DeepLMS: A deep learning predictive model for supporting online learning in the COVID-19 era. *Scientific Reports*, 10, Article 19888. <https://doi.org/10.1038/s41598-020-76740-9>
- García-Robles, P., Cortés-Pérez, I., Nieto-Escámez, F. A., García-López, H., Obrero-Gaitán, E., & Osuna-Pérez, M. C. (2024). Immersive virtual reality and augmented reality in anatomy education: A systematic review and meta-analysis. *Anatomical Sciences Education*, 17(3), 514-528. <https://doi.org/10.1002/ase.2397>
- Görzen, S., Heinemann, B., & Schroeder, U. (2024). Towards using the xAPI specification for learning analytics in virtual reality. In *Proceedings of LAK 2024 Workshops*.
- Graesser, A. C., Conley, M. W., & Olney, A. (2012). Intelligent tutoring systems. In K. R. Harris, S. Graham, T. Urdan, A. G. Bus, S. Major, & H. L. Swanson (Eds.), *APA educational psychology handbook, vol. 3. Application to learning and teaching* (pp. 451-473). American Psychological Association. <https://doi.org/10.1037/13275-018>
- Henri, M., Johnson, M. D., & Nepal, B. (2017). A review of competency-based learning: Tools, assessments, and recommendations. *Journal of Engineering Education*, 106(4), 607-638. <https://doi.org/10.1002/jee.20180>
- Huang, L. (2023). Ethics of artificial intelligence in education: Student privacy and data protection. *Science Insights Education Frontiers*, 16(2), 2577-2587. <https://doi.org/10.15354/sief.23.re202>

- Ifenthaler, D., & Widanapathirana, C. (2014). Development and validation of a learning analytics framework: Two case studies using support vector machines. *Technology, Knowledge and Learning*, 19(1), 221-240. <https://doi.org/10.1007/s10758-014-9226-4>
- Issa, M., Faraj, M., AbiGhannam, N. (2024). Exploring ChatGPT's ability to classify the structure of literature reviews in engineering research articles. *IEEE Transactions on Learning Technologies*, 99, 1-10. <https://doi.org/10.1109/TLT.2024.3409514>
- Leva, A., & Panzani, G. (2024). A software suite for online practice and autonomous activity in a process control course. *IFAC-PapersOnLine*, 58(26), 19-24. <https://doi.org/10.1016/j.ifacol.2024.10.264>
- Liu, C., Wang, G.-C., & Wang, H.-F. (2025). The application of artificial intelligence in engineering education: A systematic review. *IEEE Access*, 13, 17895-17910. <https://doi.org/10.1109/ACCESS.2025.3532595>
- Lo, C. K., Hew, K. F., & Jong, M. S.-Y. (2024). The influence of ChatGPT on student engagement: A systematic review and future research agenda. *Computers & Education*, 219, Article 105100. <https://doi.org/10.1016/j.compedu.2024.105100>
- Loong, C. N., & Chang, C.-C. (2024). Control knowledge tracing: Modeling students' learning dynamics from a control-theory perspective. *Computers and Education: Artificial Intelligence*, 7, Article 100292. <https://doi.org/10.1016/j.caeai.2024.100292>
- Masri, N., Sultan, Y. A., Akkila, A. N., Almasri, A., Ahmed, A., Mahmoud, A. Y., Zaqout, I., & Abu-Naser, S. S. (2019). Survey of rule-based systems. *International Journal of Academic Information Systems Research*, 3(7), 1-22.
- Ngo, T. T. A., Tran, T. T., An, G. K., & Nguyen, P. T. (2024). ChatGPT for educational purposes: Knowledge management factors and continued use. *IEEE Transactions on Learning Technologies*, 17, 1341-1352. <https://doi.org/10.1109/TLT.2024.3383773>
- Npuls Consortium. (2024). Onwards to success with learning analytics: An xAPI implementation guide. *Npuls Consortium*. [https://npuls.nl/\\_assets/2cbc8f9f-5e03-4e62-9c95-2722f46f0161/Magazine-Npuls-learning-analytics-EN.pdf](https://npuls.nl/_assets/2cbc8f9f-5e03-4e62-9c95-2722f46f0161/Magazine-Npuls-learning-analytics-EN.pdf)
- Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, 147, Article 103778. <https://doi.org/10.1016/j.compedu.2019.103778>
- Raihan, N., Siddiq, M. L., Santos, J. C. S., & Zampieri, M. (2025). Large language models in computer science education: A systematic literature review. In *Proceedings of the 56<sup>th</sup> ACM Technical Symposium on Computer Science Education* (pp. 938-944). ACM. <https://doi.org/10.1145/3641554.3701863>
- 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>
- Segovia, M. G., Segura, M. G., & Noriega, H. R. (2016). Análisis del desarrollo de un material multimedia orientado al manejo higiénico de los alimentos [Analysis of the development of multimedia material focused on hygienic food handling]. *Didáctica, Innovación y Multimedia*, (33).
- Shute, V. J., & Rahimi, S. (2017). Review of computer-based assessment for learning in elementary and secondary education. *Journal of Computer Assisted Learning*, 33(1), 1-19. <https://doi.org/10.1111/jcal.12172>
- Vergara, D., Fernandez-Arias, P., Extremera, J., Davila, L. P., & Rubio, M. P. (2022). Educational trends post COVID-19 in engineering: Virtual laboratories. *Materials Today: Proceedings*, 49(Part 1), 155-160. <https://doi.org/10.1016/j.matpr.2021.07.494>
- Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, 252(Part A), Article 124167. <https://doi.org/10.1016/j.eswa.2024.124167>
- Winne, P. H. (2017). Learning analytics for self-regulated learning. In C. Lang, G. Siemens, A. F. Wise, D. Gasevic, & A. Merceron (Eds.), *Handbook of learning analytics* (pp. 241-249). SOLAR. <https://doi.org/10.18608/hla17.021>
- xAPI. (2023). Experience API (xAPI). *xAPI*. <https://xapi.com/overview/>
- Xu, H., Gan, W., Qi, Z., Wu, J., & Yu, P. S. (2024). *Large language models for education: A survey*. arXiv. <https://doi.org/10.48550/arXiv.2405.13001>
- Xu, J., Zhu, J., Gan, J., & Li, C. (2023). Development and application of virtual simulation teaching system for mechanical disassembly based on Unity3D. In *Proceedings of the 2023 International Conference on Information Education and Artificial Intelligence*. ACM. <https://doi.org/10.1145/3660043.3660163>
- 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, Article 30. <https://doi.org/10.1186/s41239-019-0171-0>

Zhai, C., & Wibowo, S. (2023). A systematic review on artificial intelligence dialogue systems for enhancing English as foreign language students' interactional competence in the university. *Computers and Education: Artificial Intelligence*, 4, Article 100134. <https://doi.org/10.1016/j.caeai.2023.100134>

Zhang, N., & Liu, Y. (2023). Design and implementation of virtual laboratories for higher education sustainability: A case study of Nankai University. *Frontiers in Education*, 8. <https://doi.org/10.3389/feduc.2023.1322263>

<https://www.ejmste.com>