Advances in Educational Technology and Psychology (2026)
Clausius Scientific Press, Canada

DOI: 10.23977/aetp.2026.100415
ISSN 2371-9400 Vol. 10 Num. 4

Teacher–Student–AI Collaborative Teaching and Multi-Dimensional Intelligent Evaluation: A Case Study of the Generative Classroom

Hui Chen[1,a], Yunna Xue[1,b], Liying Gao[1,c], Chonghai Xu[1,d], Zhaoqiang Chen[1,e,*]

[1]School of Mechanical Engineering, Qilu University of Technology (Shandong Academy of Sciences), Jinan, China
[a]chenhui@qlu.edu.cn, [b]snow5821@163.com, [c]gly@qlu.edu.cn, [d]xch@qlu.edu.cn, [e]czq@qlu.edu.cn
**Corresponding author*

Keywords: Generative Artificial Intelligence; Teacher–Student–AI Collaboration; Fundamentals of Mechanical Design; Multi-Dimensional Evaluation

Abstract: Generative artificial intelligence is changing the way teachers and students interact in class, and the teaching subject is expanding from a "teacher–student" pair to a "teacher–student–AI" triad. When such technology is introduced into engineering foundation courses, three questions have to be answered: what teaching responsibilities AI should take, how technological substitution and formalistic collaboration can be avoided, and how an evaluation mechanism suited to the collaborative process can be built. Taking the course Fundamentals of Mechanical Design as its object, this paper constructs a "teacher–student–AI" collaborative teaching framework along three lines, namely role reconstruction, layered interaction and data integration. It designs a human–AI interaction process at the cognitive, analytical and creative levels, and delimits the role of AI according to the principle of "assisting rather than replacing". On this basis, it establishes a whole-process evaluation system covering the three dimensions of knowledge, ability and literacy, explains how learning-process data are transformed into digital portraits and fed back into teaching, and illustrates the correspondence between teaching activities and evaluation evidence with the case of bolt group design. The paper also discusses implementation risks, including the reliability of generated content, data privacy, teachers' workload and evaluation fairness, together with the corresponding safeguards. The study belongs to research on teaching-model construction and implementation design; the quantitative targets mentioned are proposed goals that await verification in subsequent teaching pilots.

1. Introduction

1.1 Research Background

Generative artificial intelligence, represented by large language models, is moving from an auxiliary tool to a participant in teaching. In higher engineering education, with its capacity for natural-language interaction, scheme generation and multimodal output, it is changing the way

course content is organized, the form of classroom interaction and the path of learning evaluation [1,2].

Fundamentals of Mechanical Design is a core foundation course for mechanical engineering majors. It is both theoretical and strongly practical, and it has long faced difficulties that are hard to avoid. Topics such as mechanism motion and shaft-system structure are abstract, and students find it difficult to form a mental picture of motion. Students' foundations differ markedly, and uniform lecturing can hardly accommodate different paces. An assessment system dominated by final examinations can hardly reflect the abilities that students gradually develop in the design process. Generative AI provides the conditions for addressing these problems, but it also raises new questions, such as how the role boundary between human and machine should be drawn, whether collaboration actually takes place, and whether process data can support credible evaluation. The course was chosen for the pilot for several reasons. It has wide coverage, so the experience of reform is easy to transfer to other courses. Knowledge graphs, ability graphs, problem graphs and digital textbooks have already been built, which provides the resource base for AI integration. Its teaching difficulties are fairly typical of similar courses, so the conclusions have some value for wider application.

Existing research on human–AI collaborative teaching mostly stays at the level of conceptual advocacy and model description, and several problems remain unresolved once implementation begins. The role of AI is vaguely defined, often loosely called a teaching assistant or a learning companion without being differentiated by teaching stage and cognitive level, so that it either oversteps and thinks for students or degrades into a demonstration tool [1,3]. The collaboration process lacks an operable design; teachers find it hard to judge when AI should intervene and to what extent, and students seldom develop stable collaborative habits. The evaluation system does not match the form of collaboration. Result-oriented assessment lacks a mechanism for converting process performance into evidence of ability, so that assessment for learning and assessment for improvement are difficult to put into practice [4,5].

1.2 Research Approach

To address these problems, this paper takes the generative classroom of Fundamentals of Mechanical Design as its carrier and attempts three tasks. First, it constructs a "teacher–student–AI" collaborative teaching ecology, defines the layered responsibilities of the different actors, and delimits the boundary of collaboration with the principle of "assisting rather than replacing". Second, it designs a three-level human–AI interaction process at the cognitive, analytical and creative levels, so that collaboration becomes an executable teaching procedure rather than a slogan. Third, it establishes an evaluation system in the three dimensions of knowledge, ability and literacy, links process data to digital portraits and then to teaching feedback, and offers an operable scheme through the case of bolt group design.

The research materials include teaching design texts, classroom performance records, student design work and human–AI interaction logs. In the subsequent pilot stage, pre- and post-tests, learning behavior analysis and interviews with teachers and students will be used to examine the feasibility of the framework and the reliability and validity of the evaluation indicators. The contribution of this paper lies in refining the rather general notion of "collaboration" into three interlocking parts, namely roles, process and evaluation, and in attempting to establish a correspondence among process data, evaluation indicators and demonstrated ability, so as to provide a testable design scheme for AI-integrated teaching in engineering foundation courses. As the study is still at the stage of model construction and implementation design, its judgments about effectiveness and its quantitative targets await verification in teaching pilots.

2. Theoretical Basis

The introduction of generative AI expands the teaching subject from a binary teacher–student structure to a ternary "teacher–student–machine" structure, yet whether AI can play its part still depends on human subjectivity [6]. The role that AI assumes is therefore not fixed; it should be adjusted according to the nature of the cognitive task. The point of a generative classroom is not to let AI give answers directly, but to use authentic design problems to move students towards engineering thinking through repeated cycles of design, verification and revision. This is also the pedagogical basis for the principle of "assisting rather than replacing".

Process-oriented evaluation advocates the continuous collection of evidence and the provision of timely feedback during learning, while developmental evaluation attaches more importance to the role of evaluation in promoting learners' growth [4]. Process data such as dialogue records, scheme iterations and operation sequences were difficult to collect in the past; generative AI makes them recordable and analyzable, and thus provides the technical conditions for process-oriented and developmental evaluation [5]. Abundant data, however, do not amount to credible evaluation, and how to establish an effective correspondence among indicators, evidence and ability remains a difficulty to be overcome.

Taken as a whole, the academic community has reached considerable consensus on "human–AI collaboration" [5,6], and practical explorations such as digital portraits and intelligent evaluation have appeared in engineering education [3,5]. Two shortcomings remain. First, collaboration mechanisms stay largely at the level of model advocacy and lack a layered interaction process matched to the cognitive levels of specific courses. Second, evaluation research concentrates on technical implementation and has not adequately answered the evaluation-studies question of how process data can be turned into credible evidence of ability. This paper seeks to make some contribution on these two points.

3. Construction of the "Teacher–Student–AI" Collaborative Teaching Framework

3.1 Overall Framework

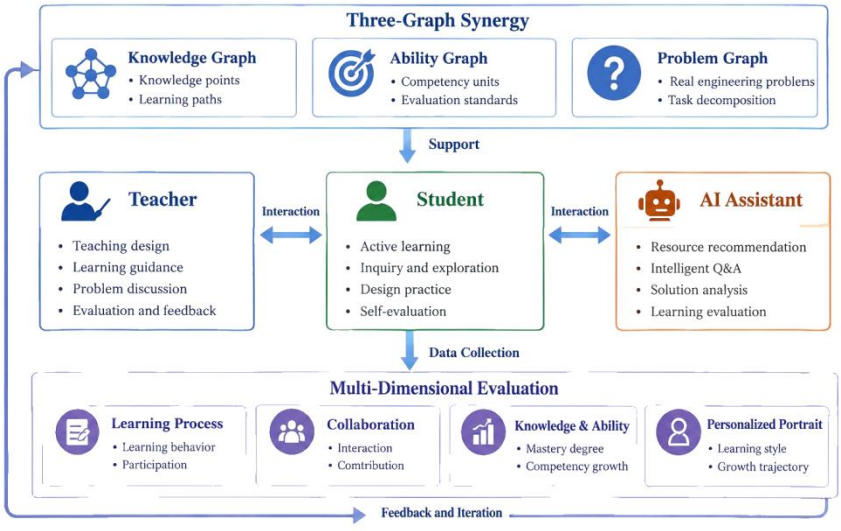


Figure 1. Overall framework of the "teacher–student–AI" collaborative teaching model

The framework proposed in this paper consists of three parts, namely role reconstruction, layered interaction and data integration. Role reconstruction answers who does what, layered interaction

answers how it is done, and data integration answers how teaching is evaluated and improved. Linked by authentic design tasks, the three form an operating chain of teacher guidance, student construction, AI support and data feedback, as shown in Figure 1.

3.2 Reconstruction of Roles and Responsibilities

In the new collaborative relationship, teachers change from knowledge transmitters to guides of engineering thinking, students from passive recipients to active constructors, and AI from a question-answering tool to an aid for honing thinking. The responsibilities of the three parties need to be specified for each teaching stage, so that neither overlap nor gaps occur (Table 1).

Table 1. Role positioning and division of responsibilities

Subject	Role positioning	Before class	In class	After class
Teacher	Guide of engineering thinking; instructional designer	Set objectives, diagnose learning, review AI resources	Organize discussion, correct misconceptions, guide values	Review data, provide tiered tutoring, optimize design
Students	Active inquirers; constructors of ability	Prepare independently, raise questions	Decide on schemes, explore collaboratively, verify iteratively	Reflect and summarize, transfer and apply
AI	Intelligent learning companion; design checking tool	Push preparatory resources, generate an initial portrait	Visualize calculations, generate alternative schemes, record the process	Grade assignments and answer questions, push targeted resources, update the portrait

The responsibility boundary of AI can be summarized as “assisting rather than replacing”. Wherever engineering judgment, value trade-offs and innovative decisions are involved, students should take the lead, and AI only provides support in calculation, checking and scheme comparison.

3.3 Layered Human–AI Interaction Process

Following the progression of cognitive tasks from lower to higher order, this paper divides human–AI interaction into three levels (Table 2). The three levels form a spiral progression rather than a simple linear sequence. Students enter the analytical level after consolidating the foundations at the cognitive level, and the creative level after accumulating decision-making experience at the analytical level. Teachers adjust the level dynamically according to learning conditions, so that collaboration moves gradually from human-led, machine-assisted work towards parallel human–AI work.

The bottom line of collaboration is the educational mission itself. This paper draws three boundaries for AI intervention. Cognitively, AI must not replace students in the design reasoning that they should experience themselves. Evidentially, calculations and conclusions generated by AI must be traceable and verifiable, so that “hallucinations” do not mislead judgment. Ethically, students’ design work should indicate the extent of AI involvement, and teachers are responsible for reviewing AI-generated resources. Together, the three boundaries are meant to make technology serve people and to prevent it from replacing human judgment.

Table 2. Design of the layered human–AI interaction process

Level	Cognitive objective	Role of AI	Role of students	Typical activities
Cognitive	Understand basic concepts and calculations	Undertake basic calculation and visualization	Master operating logic and principles	Mechanism motion simulation, parameter calculation
Analytical	Scheme comparison and engineering decisions	Generate multi-scheme comparisons and checking results	Make engineering judgments and trade-offs	Transmission scheme comparison, strength-check decisions
Creative	Integrated design and innovation	Serve only as a checking tool and feedback source	Lead design and iteration	Overall machine design, innovative improvement

4. Design of the Multi-Dimensional Intelligent Evaluation System

4.1 Evaluation Philosophy

This paper proposes an evaluation framework in three dimensions, namely knowledge, ability and literacy. The knowledge dimension concerns conceptual understanding and calculation accuracy. The ability dimension concerns design decisions, scheme iteration and engineering expression. The literacy dimension concerns collaborative contribution, reflective awareness and engineering ethics. Evaluation runs through the whole process before, during and after class, combining process-oriented evaluation with necessary summative evaluation, so that students' overall performance is not judged by final results alone [4,5].

4.2 Evaluation Index System

On the basis of this philosophy, an observable, collectable and quantifiable evaluation index system is constructed (Table 3). The design follows the principle of evidential traceability, and every observation point corresponds to an explicit data source.

Table 3. Multi-dimensional intelligent evaluation index system of knowledge, ability and literacy

Dimension	Observation points	Data sources	Evaluation method
Knowledge	Conceptual understanding, calculation accuracy	Pre-study quizzes, assignments, AI grading records	Automatic scoring plus teacher sampling
Ability	Rationality of schemes, number and quality of iterations, engineering expression	Design logs, revision traces, design reports	Rubric scoring plus AI-assisted analysis
Literacy	Collaborative contribution, depth of reflection, standards of AI use	Collaboration platform records, reflection logs, dialogue records	Peer assessment plus teacher assessment

As for weighting, process-oriented and summative evaluation should each account for a certain proportion, which changes with course objectives. The specific proportions need to be determined in the pilot through reliability and validity tests; this paper does not preset fixed values.

4.3 Feedback of Evaluation Results to Teaching

The purpose of evaluation is not merely to assign scores but to drive improvement. This paper designs an operating mechanism in which monitoring, feedback and iteration are linked. The system keeps recording the learning process and generates digital portraits for individuals and for the class. Teachers adjust the teaching pace and tiered strategies according to the portraits, and students adjust their learning paths according to the diagnostic results. The adjusted teaching then produces new data and enters the next round of monitoring, so that assessment for learning, for teaching and for improvement becomes part of everyday practice.

4.4 Reliability, Validity and Fairness Safeguards

Whether multi-dimensional evaluation holds up depends on three aspects. The first is whether the correspondence between indicators and evidence is clear; every observation point needs an explicit data source and scoring rubric, and the amount of data must not be used in place of the level of ability. The second is dual review by human and machine; the initial evaluation given by AI must be reviewed by teachers, and disputed samples must be arbitrated manually so as to reduce algorithmic bias. The third is cross-validation of multiple sources; design logs, work quality and peer assessment are compared with one another to prevent distortion caused by a single data source. In addition, channels for appeal and data correction should be established to protect students' legitimate rights.

5. A Case of Process Evaluation in Bolt Group Design

5.1 Teaching Objectives and Instructional Design

Bolt group design is a comprehensive application in the connection-parts section of Fundamentals of Mechanical Design. It involves force analysis, strength checking and structural layout, and it is strongly oriented towards engineering practice. There are three teaching objectives. Students should master the method of force analysis for bolt groups, be able to complete strength checking and structural optimization, and develop an awareness of safety and standardization in design and checking.

Before class, the teacher releases an authentic engineering scenario, such as a reducer housing connection, and AI pushes preparatory resources and generates an initial learning portrait. In class, the teacher organizes discussion of the schemes. At the cognitive level, students use AI to complete force calculations and to visualize the results; at the analytical level, they compare the checking results of different layout schemes and make engineering decisions; at the creative level, they lead structural optimization while AI only provides checking feedback. After class, AI grades assignments and pushes targeted exercises, and the teacher provides tiered tutoring based on the portraits.

5.2 Implementation of Digital-Portrait-Driven Process Evaluation

Around the bolt group design task, process observation points and corresponding evidence requirements are set (Table 4). Evaluation no longer looks only at the final drawing; it follows the four stages of force analysis, strength checking, scheme iteration and reflection, and traces students' performance throughout the process.

5.3 Case Analysis and Discussion

In this case, collaboration and multi-dimensional evaluation support each other. AI undertakes repetitive work such as calculation, grading and recording, teachers concentrate on scheme discussion and the correction of misconceptions, and students lead engineering decisions and scheme iteration. Process evidence makes ability growth, which is otherwise hard to observe, visible, and evaluation thus takes on a developmental function. At the same time, the accuracy of AI grading and the authenticity of dialogue records still need teacher review, and the credibility of evaluation depends on whether dual review is actually carried out.

Table 4. Observation points and evidence for process evaluation of bolt group design

Stage	Observation points	Forms of evidence	Evaluators
Force analysis	Whether load simplification and the force model are correct	Analysis drafts, AI dialogue records	Teacher and AI
Strength checking	Whether the checking formula and the calculation are accurate	Calculation sheets, AI grading results	AI initial evaluation plus teacher review
Scheme iteration	Whether the layout is optimized and the iterations are justified	Design logs, revision traces	Teacher and peers
Reflection	Whether methods and an awareness of standards are formed	Reflection logs	Teacher

6. Implementation Outcomes, Risks and Safeguards

In terms of expected outcomes, four kinds of change can be observed. Participation in classroom interaction increases; the proportion of learning tasks in virtual–real integrated scenarios reaches more than 40% (a proposed project target rather than a verified outcome); process evaluation covers the main design tasks; and digital portraits are able to present students' staged growth trajectories. Whether these targets are met needs to be judged through teaching process data, pre- and post-test comparisons, and reliability and validity tests of the evaluation instruments.

Risks in implementation are concentrated in four areas. The first is AI hallucination, for which a pre-review mechanism of AI generation, teacher review and classroom use should be established. The second is data privacy; process data should follow the principle of minimum necessity and be properly anonymized and authorized. The third is teachers' workload; the tool chain and the division of labor between human and machine need to be optimized so that reform does not add clerical work. The fourth is evaluation fairness; collection bias should be corrected through cross-validation of multiple sources and manual review.

Safeguards involve four levels, namely organization, technology, institution and capacity. Organizationally, the course reform is incorporated into the key tasks of the school, with teaching arrangements and platform resources guaranteed. Technologically, knowledge graphs, AI teaching assistants and data analysis modules are deployed on the smart teaching platform. Institutionally, standards for reviewing AI resources, for data use and for evaluation appeals are established. In terms of capacity, teachers receive training in AI-based teaching, which helps them move from tool users to instructional designers.

7. Conclusion

Targeting the generative classroom of Fundamentals of Mechanical Design, this paper constructs a "teacher–student–AI" collaborative teaching framework consisting of role reconstruction, layered interaction and data integration, and proposes a three-level interaction process at the cognitive,

analytical and creative levels together with the principle of “assisting rather than replacing”. On this basis, it designs an evaluation system in the dimensions of knowledge, ability and literacy, outlines the generation path of digital portraits and the operating mechanism of monitoring, feedback and iteration, and illustrates its application through the bolt group design case. It should be noted that the effectiveness of collaboration depends not on how many AI tools are used, but on whether the responsibility boundaries among teachers, students and AI are clear, and on whether process evidence is traceable and reviewable.

The limitations of this paper are also evident. The study has mainly completed the construction of a teaching model and the design of a case, and it still lacks systematic empirical data from real classrooms. Subsequent work will rely on teaching pilots to collect learning processes, design work and feedback from teachers and students, to test the reliability and validity of the evaluation indicators, and to compare the effects of different human–AI collaboration strategies on learning outcomes, engineering decision-making ability and learning experience.

Acknowledgements

This work is supported by the Interdisciplinary Integration Courses 2025 at Qilu University of Technology (Shandong Academy of Sciences)("Mechanical Design"–A Course Integrating Scientific Research and Education, "Fundamentals of Mechanical Design"–An Artificial Intelligence (AI+) Course).

References

[1] Wang Xingwei. Building a New Ecology of “Teacher–Student–Machine” Collaborative Teaching[EB/OL]. People’s Daily Online – Education Channel, 2026-07-27. https://edu.people.com.cn/BIG5/n1/2026/0727/c1006-40768621.html.

[2] China Education Daily. How Applied Universities Can Reshape a New Ecology of Digital-Intelligent Education at the Start of the “15th Five-Year Plan”[EB/OL]. China Education News, 2026-07-06. http://www.jyb.cn/rmtxwwyyq/jyxx1306/202607/t20260706_2111498978.html.

[3] China Education Daily. Reshaping the New Form of Future Classrooms in Universities[EB/OL]. China Education News, 2026-06-24. http://www.jyb.cn/rmtzgjyb/202606/t20260624_2111494380.html.

[4] Guangming Daily. Xiangtan University: Using “Digital Intelligence” to Break Through the Dilemmas of Engineering Education[N/OL]. Guangming Daily, 2025-08-25. https://epaper.gmw.cn/gmrb/html/2025-08/25/nw.D110000gmrb_20250825_5-10.htm.

[5] Xinhua Net. Research on Teaching Reform of Programming Courses Empowered by Generative AI[EB/OL]. Xinhua Net, 2026-05-09. https://www.news.cn/book/20260509/e70421d5a61043db881b25673fc94a90/c.html.

[6] Yi Kaiyu, Han Xibin. From Blended Teaching to Human–AI Collaborative Teaching: A New Teaching Form under the Transformation of Generative Artificial Intelligence[EB/OL]. National Institute of Education Sciences[2026-10-21]. https://www.cnaes.edu.cn/post/30679.