

Investigating the Impact of an Intelligent Learning Companion on Learning Effect and Experience in Analog Circuit Laboratory Instruction

Xinyi Tian^a, Jianwei Li^{a,b}, Yanli Ji^c

^a School of Humanities, Beijing University of Posts and Telecommunications, Beijing 102206, China

^b Beijing Key Laboratory of Network System and Network Culture, Beijing University of Posts and Telecommunications, Beijing 100876, China

^c School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

© Higher Education Press 2026

Abstract Transforming engineering education in the AI era requires an evaluation of new instructional tools and a reconceptualization of the division of labor among teachers, students, and intelligent learning companion systems (ILCSs). This work explores how a retrieval-augmented generation intelligent learning companion can be embedded within a human–AI collaborative teaching model by using an analog circuit laboratory instruction as a case study. A controlled experiment compared traditional teacher-led guidance with system-supported instruction, focusing on three core dimensions: knowledge acquisition, learning effect (cognition, skill, and emotion), and flow experience (cognitive control, immersion and time transformation, loss of self-consciousness, and autotelic experience). The results indicate that while the system showed a limited impact on knowledge acquisition and emotion, it significantly enhanced skills, immersion and time transformation, and autotelic experience. These findings suggest that ILCSs serve as effective complements in practice-oriented engineering education, particularly in terms of providing personalized support and instant feedback strengthening hands-on learning and student engagement. Such companions cannot fully serve as a substitute for teacher-led conceptual scaffolding or emotional guidance. The study’s theoretical contribution lies in emphasizing the importance of role allocation in human–AI collaborative education and offers practical implications for the design of learner-centered, practice-oriented instructional models in intelligent education.

Keywords human–AI collaborative education, intelligent learning companion system, analog circuit laboratory instruction, knowledge acquisition, learning effect, flow experience

1 Introduction

The profound transformation of global industrial structures in the 21st century has intensified the demand for innovative engineering talent, optimizing engineering education a critical issue in higher education (Gürdür Broo et al., 2022). Science, technology, engineering, and mathematics (STEM) education, a cornerstone of economic development, aims at cultivating individuals equipped with 21st-century skills to meet global market demands (Bas & Kiraz, 2025; Jang, 2016). However, global STEM education research has gradually shifted its focus from discipline integration to broader areas such as educational equity, teacher professional development, and higher-order skills. The global preference for this evolutionary path indicates that the educational community has recognized the persistent gap between existing educational frameworks and actual industry needs, especially problem-solving, teamwork, and application competencies (Yang et al., 2023). It is imperative to explore novel instructional paradigms that can support students in deep learning and develop higher-order skills.

With the rapid advancement of AI technologies, human–AI collaborative teaching has emerged as a promising new educational paradigm (Filippi & Motyl, 2024; Kim, 2024a; Kong et al., 2025). This paradigm

Received May 8, 2025; revised October 31, 2025; accepted November 17, 2025

Jianwei Li (✉)

E-mail: lijianwei@bupt.edu.cn

seeks to create learning environments powered by hybrid intelligence, which can offer personalized support and instant feedback, and dynamic instructional guidance to significantly enhance teaching effectiveness. Despite the immense potential of human-AI collaborative education, implementing and validating it in complex, practice-oriented engineering education requires addressing numerous challenges. Research (Fan et al., 2025) suggests that an overreliance on generative AI may lead to metacognitive laziness, which hinders students' self-regulated and deep learning. The existing literature has confirmed the potential of AI-assisted engineering education and emphasized the necessity for effective guidance to help students fully leverage these tools (Pham et al., 2023). These literature collectively indicate that despite human-AI collaboration's promising outlook, empirical research on this collaboration's underlying mechanisms and efficacy remains nascent.

This study aims at addressing this research gap by examining an analog circuit laboratory course as a practical case for investigating the application of an intelligent learning companion based on retrieval-augmented generation (RAG) in engineering education. Students in physical laboratories are more effective at developing their acquisition of practical skills and problem-solving competencies (Tokatlidis et al., 2024), but may also experience a high cognitive load hindering their construction of deep knowledge (Li et al., 2025b).

According to cognitive load theory (Sweller, 1988), the instant and targeted assistance provided by intelligent learning companion systems (ILCSs) is designed to reduce students' extraneous cognitive load during complex problem-solving process. By offloading the mental burden of searching for information, the system aims to free up cognitive resources, thereby facilitating students' construction of deep knowledge and enhancing their learning effects. The proposed system is grounded in flow theory and provides students with personalized support and instant feedback (Csikszentmihalyi, 1990). This balance is expected to enhance students' motivation and engagement, thereby leading to more immersive and satisfying learning experiences.

This study sets knowledge acquisition, learning effect, and flow experience as its core dependent variables to capture multifaceted impacts of the instructional design. It is guided by two overarching objectives: to explore how human-AI collaborative teaching models can be effectively designed and to examine how ILCSs can be applied within the context of complex, practice-oriented engineering education. The study examines RAG-based ILCSs in a physical analog circuit

laboratory and generates empirical evidence to develop efficient, learner-centered teaching models.

2 Literature Review

2.1 | Pedagogical Agents

Early research on pedagogical agents (PAs) has primarily focused on their fundamental functions and persona effect within multimedia learning environments. Researchers (Reeves & Nass, 1996) have observed that people tend to anthropomorphize technology and respond to it as they would to a human—a phenomenon formalized as the media equation theory. This theory has provided a critical psychological foundation for the design of PAs (Baylor & Kim, 2009; Sikström et al., 2022). Subsequently, research has shifted toward understanding how PAs influence learners' cognitive and emotional dimensions. A review of 26 relevant studies found that although PAs can effectively enhance student motivation and learning effects, their effectiveness depends on various design conditions (Heidig & Clarebout, 2011).

As the research progressed, the design characteristics of PAs emerged as key factors affecting learning efficacy. Effective cognitive-level PA design accounts for ways of reducing learners' cognitive load. For instance, PAs that provide cues within an instructional animation have been found to help students distinguish relevant information, thereby reducing their cognitive load and improving learning outcomes (Yung & Paas, 2015). Conversely, the mere presence of PAs has been determined to not necessarily increase cognitive load or enhance learning efficacy, suggesting that the agent effect of PAs requires further empirical scrutiny (Schroeder, 2017). Moreover, PAs employing conversational teaching styles could heighten learners' interest and retention albeit potentially at the cost of increased cognitive load (Schroeder, 2017). Research has found that less is more when it comes to nonverbal communication: simple, meaningful expressions are often more effective than complex ones (Lin et al., 2020). Additionally, cognitive load can be conceptualized in terms of intrinsic, extraneous, and germane load. PA design primarily aims to reduce extraneous load while enhancing germane load, thereby optimizing learning. PAs have also been applied to support embodied and discovery-based learning, facilitating the mastery of complex concepts through performance-based feedback and remedial guidance (Abdullah et al., 2017).

The social and emotional roles of PAs have received considerable scholarly attention. For instance, a social cognitive framework considering PAs as

intelligent learning companions illustrated how agents can support learning through modeling and scaffolding based on social-cognitive theory (Kim & Baylor, 2006a). Multi-agent intelligent tutoring systems have demonstrated notable advantages over single-agent systems in promoting students' self-regulated learning by providing timely feedback and helping learners navigate complex tasks more effectively (Martin et al., 2016). Multi-agent systems emerged primarily in the mid-2010s, and their main strengths lay in the supporting collaborative learning and distributed scaffolding rather than in the simple extension of single-agent designs. AI tutors offering empathetic feedback via facial expressions can significantly enhance motivation and improve accuracy in learning tasks (Oker et al., 2020).

The instructional effectiveness of PAs is strongly influenced by their competencies and interaction modes. Higher-competency agents facilitate stronger knowledge application and more positive learner perceptions than lower-competency agents, whereas lower-competency agents may bolster learners' self-efficacy (Kim & Baylor, 2006b & 2016). Moreover, agents adopting a proactive interaction style, such as offering prompts, questions, and real-time feedback, have been shown to improve learners' recall and engagement.

In recent years, the advent of large language models (LLMs) has ushered in a new era of PA research and has marked a transition from pre-scripted agents to highly autonomous, conversational ILCs. One review of collaborative intelligent tutoring systems highlighted the overlap between these systems and PAs, emphasizing the agents' role in fostering collaborative learning (Ubani & Nielsen, 2022).

Emerging studies have examined LLMs as AI tutors, analyzing their educational functions from multiple perspectives (Al-Abri, 2025; Ding et al., 2023; García-Méndez et al., 2025). For example, ChatGPT's role in answering questions, providing writing assistance, and supporting exam preparation has been investigated (Al-Abri, 2025). One study on students' perceptions of using ChatGPT as a physics tutor revealed a correlation between answer accuracy and learners' trust (Ding et al., 2023). Interviews with high school teachers and students have identified the expectations for PAs with enhanced communication and scaffolding competencies (Sikström et al., 2024). These findings underscore the considerable potential of LLM-driven PAs to elevate the quality of agent-learner interactions. In contrast to traditional scripted agents, LLM-based PAs offer real-time, personalized guidance to learners, significantly improving adaptability and autonomy of the agents; however, careful monitoring is required to ensure the accuracy and appropriateness of the generated content.

Despite substantial advancements in the design

and implementation of PAs, several research gaps still remain. Drawing on a systematic review of 75 empirical PA studies, Dai et al. (2022) highlighted inconsistencies across experimental designs and measurement tools, indicating the need for more rigorous studies on the topic, particularly those pertaining to K-12 learners and virtual reality (VR) applications. VR and immersive environments present unique challenges and opportunities for PAs, potentially enhancing experiential engagement and offering novel insights into cognitive and emotional mechanisms (Dai et al., 2022).

In summary, research on PAs has evolved from basic investigations of persona effects and cognitive load through complex socio-emotional interactions and multi-agent systems to the current era featuring research on LLM-driven ILCs. This trajectory enriches the theoretical foundation for designing effective learning technologies and reflects the broader evolution of educational technology from teacher-centered to task-centered and ultimately AI-centered paradigms. Future research should focus on leveraging the advanced AI to construct ILCs that provide high-level communication, empathy, and personalized support for learners.

2.2 | Roles in Human-AI Collaborative Teaching

Although human-AI collaborative teaching model has emerged many years ago and there have been numerous studies on model debugging and teaching effect testing, classroom division of labor between human teachers and AI tutors has still stuck in the theoretical level (Ji et al., 2022). In many studies, the roles split between human teachers and AI tutors in the classroom are mainly divided into three types, selection-execution, teaching-assistance, and emotion-teaching.

In the selection-execution model, teachers guide and provide feedback based on students' needs; however, their role shifts. Teachers select teaching tasks that AI performs based on teaching objectives and students' feedback. These tasks include introducing new content, providing learning tasks, and facilitating activities (Fang et al., 2020). In this model, teachers act as organizers, with the main teaching tasks completed by AI (Chiu & Rospigliosi, 2025). This division of labor helps combine AI with teachers' personalized experience, enabling large-scale applications and reducing teachers' pressure. However, the model's drawbacks are obvious. The teaching process is carried out completely by AI, resulting in limited communication between teachers and students, leading to more pressure or anxiety (El Shazly, 2021). The lack of emotional communication is difficult for teachers to obtain accurate feedback, which often leads to the incorrect arrangement of teaching tasks.

The teaching-assistance model has a wide

range of applications and is more easily accepted. In this model, teachers are still the main force in teaching, while AI plays the role of a teacher's assistant to improve their teaching design and helps them expand teaching goals to meet students' various needs (Chiu & Rospigliosi, 2025). By obtaining data analyzed by AI, teachers in this model can improve teaching design and obtain accurate students' feedback, thereby enhancing teaching quality (Holstein & Alevan, 2022; Kim, 2024b). AI can also be trained to find the most suitable ways to assist teachers to intervene and guide students, which will effectively develop human-AI collaborative teaching model and reduce the burden on teachers as much as possible (Cohn et al., 2025).

Given the lack of communication in the selection-execution model, emotion-teaching model has been proposed. In courses that pay more attention to interaction or are highly personalized, such as classes about speaking a foreign language, AI can provide guided learning tasks, such as pronunciation practice, in the assimilation stage so that teachers can spend more time on interaction in the application and expansion stages (Jeon, 2023). This is because when faced with computers, students often find it difficult to arouse learning interest and AI cannot effectively alleviate their anxiety. Students' emotional states should be managed by teachers so that students' anxiety can be kept at a controllable level.

2.3 | Flow Theory

Flow theory, proposed in 1975, is defined as the holistic sensation one experiences when fully engaged in an activity (Csikszentmihalyi, 1975). This sensation typically arises when students perceive a challenge that matches their current competences and is often accompanied by a sense of accomplishment derived from the gradual resolution of problems (Beard, 2015).

According to the flow theory, three core antecedent conditions facilitate the emergence of flow states: first, clearly defined goals; second, real-time feedback; and third, a balance between the perceived challenge of the task and individual competence (Kaye, 2012). The flow theory has seldom been emphasized in STEM education, which can be attributed to the inherently high abstraction level of STEM content, which often exceeds learners' problem-solving competencies. As a result, achieving an optimal balance between challenges and skills becomes difficult in STEM, preventing learners from entering their zone of proximal development (ZPD). In traditional laboratory instruction, teachers are difficult to provide timely and personalized assistance, making immediate feedback largely unattainable. Consequently, learners in laboratory-based STEM environments rarely report entering sustained flow states which are difficult for teachers to

measure systematically.

Students receive analog circuit laboratory instruction toward explicit experimental objectives, an instructional process supported by well-structured laboratory manuals. To address the limitations of traditional laboratory support, this study introduces an ILCS that allows learners to ask questions and receive instant, personalized feedback during experiments. Moreover, this system provides learners with adaptive guidance aligned with their skill levels, ensuring that task difficulty remains appropriately balanced within their ZPD.

The ILCS introduced in this study facilitates the antecedent conditions for a flow experience. Measuring flow experience illuminates learners' psychological engagement with complex laboratory tasks, and provides a meaningful lens through which to evaluate the pedagogical effectiveness of intelligent learning companions in STEM experimental instruction.

3 Research Questions

Building upon current theoretical frameworks and literature in the field, this study seeks to empirically evaluate the effectiveness of human-AI collaborative teaching in engineering education. Using analog circuit laboratory instruction as a practical context, the work focuses on comparing traditional teacher-led guidance supported by ILCS. Accordingly, the study addresses the following three research questions:

(1) RQ1: Is there a significant difference in knowledge acquisition between traditional teacher-led guidance and ILCS-supported instruction?

(2) RQ2: Is there a significant difference in overall learning effect between traditional teacher-led guidance and ILCS-supported instruction?

(3) RQ3: Is there a significant difference in students' flow experience between traditional teacher-led guidance and ILCS-supported instruction?

4 Methods

4.1 | Participants

Thirty students from a electric circuit experiments class were selected to ensure the representativeness of the sample and the reliability of the results. Participants were divided into two groups of 15 students each: an experimental group and a control group.

Participants were all second-year students majoring in electronic science and technology. Twenty-four males and six females were included. The instructor delivered a 20-minute orientation to explain the experimental procedure and highlight essential precautions, thereby ensuring that all students shared a consistent understanding of the upcoming tasks.

Students in the experimental group were then provided with the uniform resource locator of the ILCS and given a brief introduction. During the experiment, students in the experimental group were encouraged to pose questions primarily to the system. If the ILCS was unable to solve problems or provided inaccurate answers, students could then raise their hands to ask the teacher. In contrast, students in the control group relied entirely on the teacher for guidance as they conducted experiments and solved problems. After the experiment, 178 questionnaires were collected, all of which were deemed valid.

4.2 | Instruments

In this study, the developed ILCS was designed to provide accurate, reliable, and personalized instructional support for analog circuit laboratory instruction. An example of the system’s user interface is shown in Figure 1. The system architecture, illustrated in Figure 2, is built upon a RAG framework and incorporates multiple supporting technologies.

LangChain framework is at the core of the ILCS and serves as a flexible development environment, enabling the integration and orchestration of multiple

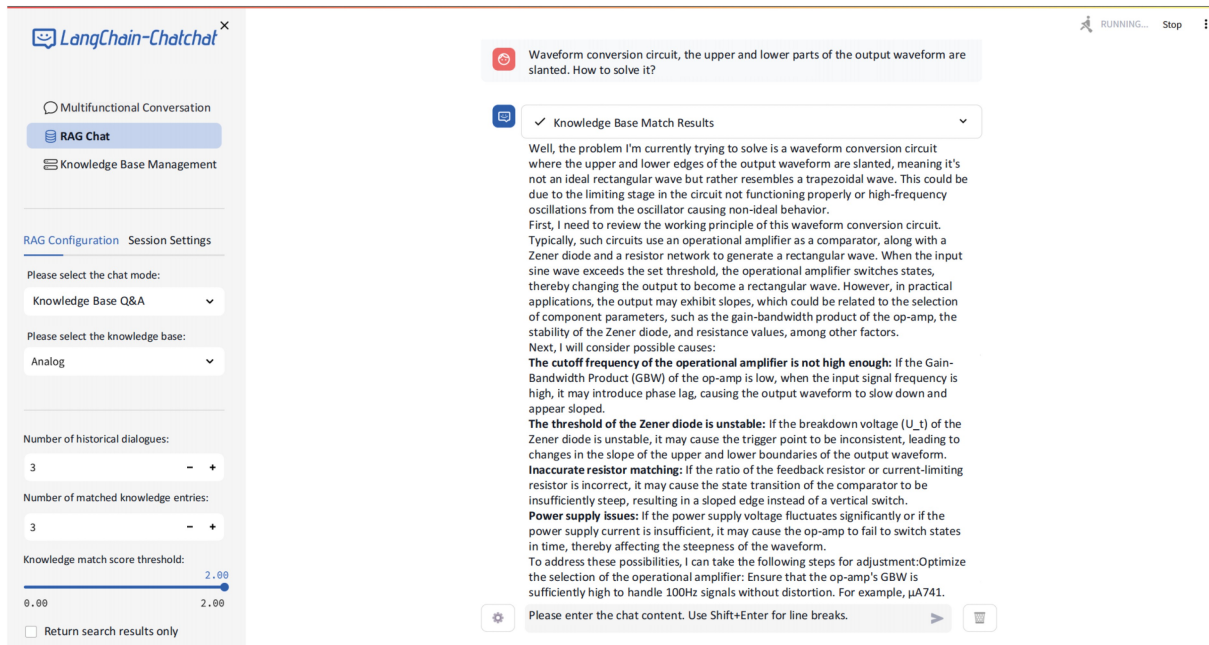


Figure 1 Personalized retrieval-augmented generation (RAG)-based intelligent learning companion system.

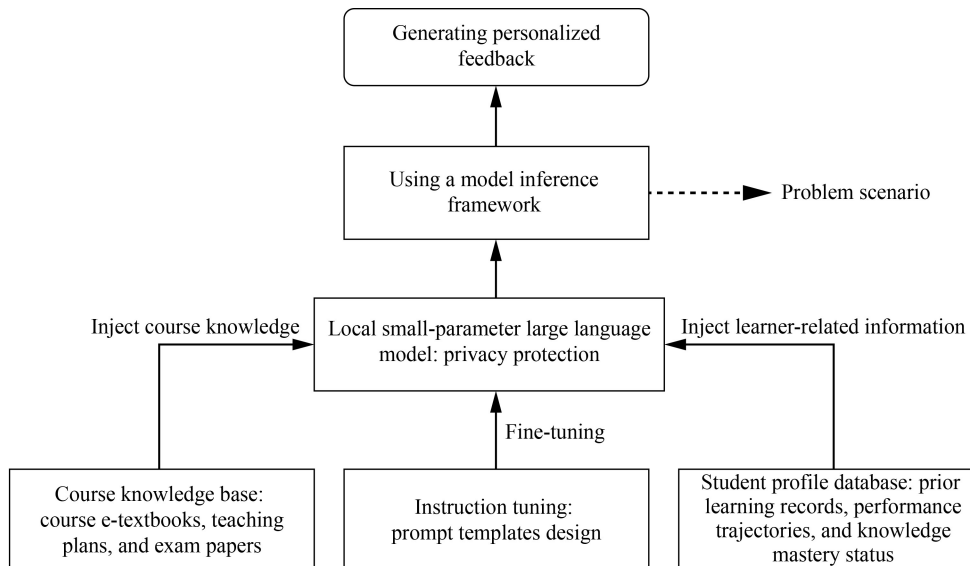


Figure 2 Architecture of the personalized retrieval-augmented generation-based intelligent learning companion system.

natural language processing modules. The ILCS performs two simultaneous retrieval processes when a student submits a query, as shown in Figure 2. It searches the course knowledge base, which consists of course e-textbooks, teaching plans, and exam papers, to extract relevant knowledge. The ILCS queries the student profile database with of prior learning records, performance trajectories, and knowledge mastery status to capture individualized learning needs and contextual information.

The retrieved materials and learner-related information are then combined with the student’s original query to form an enriched prompt processed by a locally deployed, small-parameter LLM to ensure both efficient answer delivery and privacy protection. The ILCS carries out its inference process through the XInference framework with the Qwen2-7B model serving as the system’s backbone for generating responses during laboratory sessions. The integration of RAG ensures that the generated answers are both contextually relevant and tailored to individual learning conditions.

The proposed RAG-enhanced ILCS demonstrates notable advantages over conventional LLM-based systems, including higher content relevance, more accurate responses, and more efficient allocation of instructional resources. The evaluation results from the analog circuit laboratory scenario indicate that the system achieved a knowledge coverage rate of 96.1%, an answer accuracy of 99.4%, and a personalized feedback rate of 67.2%.

4.3 | Data Collection

In this section, we first introduce three questionnaires used as research instruments and then describe the procedures for data collection and the evaluation of their reliability and validity.

In this study, three separate questionnaires were administered to assess students’ knowledge acquisition, learning effect, and flow experience before and after the laboratory instruction. As shown in Table 1, each questionnaire includes three basic demographic questions, including student identification, questionnaire name, and number of question.

Three distinct questionnaires were employed to evaluate students’ knowledge acquisition, learning

effect, and flow experience of analog circuits and electronic measurement before and after laboratory instruction.

The first questionnaire, knowledge acquisition questionnaire, consisted of multiple-choice questions. It was used to evaluate students’ mastery of relevant concepts and applied knowledge before and after the laboratory instruction. The items covered both fundamental principles and applied tasks pertinent to the course, enabling a comparison between the pretest and the posttest to quantify knowledge acquisition. The test was reviewed by experts in electrical engineering and educational technology to ensure its reliability and validity. The evaluation focused on the following three aspects:

(1) Content relevance: test comprehensively covering concepts and applications relevant to the course;

(2) Construct validity: items assessing foundational knowledge acquisition and applied skills, aligning with the key learning outcomes of the course;

(3) Clarity and precision: test items reviewed for clear and precise wording, aligning with the course objectives. A total of 60 valid responses were collected for this questionnaire.

The second questionnaire, learning effect scale, was designed to assess students’ achievements during the experimental learning process. In addition to basic demographic information, the scale included 13 items across three dimensions: cognition, skill, and emotion. All items adopted a five-point Likert scale (1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree), with higher scores indicating stronger performance in the corresponding dimension. The scale was adapted from the framework proposed by Chen (2022) and modified to suit the context of laboratory-based instruction. A total of 60 valid responses were obtained.

The third questionnaire, flow experience scale, aimed to measure students’ psychological states and flow experiences during experimental learning. It was developed based on the Flow in Education scale v.2 and covered four dimensions: cognitive control, immersion and time transformation, loss of self-consciousness, and autotelic experience (Heutte et al., 2021). A total of 58 valid responses to the third questionnaire were received, and expert evaluations confirmed the test’s reliability and validity in measuring students’ knowledge and skills in analog circuit laboratory instruction.

To assess the reliability and validity of the learning effect scale, internal consistency was examined using Cronbach’s alpha, which was measured to be 0.985, indicating excellent reliability ($\alpha > 0.7$). Construct validity was evaluated using the Kaiser-Meyer-Olkin (KMO) test and Bartlett’s test of sphericity. The KMO value was 0.890, suggesting high

Table 1 Questionnaire content

Student identification	Questionnaire name	Number of question
1	Knowledge acquisition questionnaire	8
2	Learning effect scale	16
3	Flow experience scale	15

sampling adequacy. The Bartlett's test of sphericity yielded a chi-square value of 1,332.716 with 78 degrees of freedom that was statistically significant ($p < 0.001$), indicating the scale's appropriateness for factor analysis (see Table 2). The Cronbach's alpha for the flow experience scale was 0.973, reflecting strong internal consistency. The KMO value was 0.894 and Bartlett's sphericity test produced a chi-square value of 960.596 with 66 degrees of freedom ($p < 0.001$), further supporting the scale's validity.

4.4 | Procedure

This study adopted a pretest–posttest control group experimental design to compare the effects of traditional teacher-led instruction and a RAG-based ILCS on students' learning effect and experiences. Thirty students were enrolled in an electric circuit experiments class who were randomly assigned to a control group (group A) and an experimental group (group B), each of which had 15 students.

Prior to the experiment, the teacher delivered a 20-minute orientation to explain the experimental procedure and highlight essential precautions, thereby ensuring that all students shared a consistent understanding of the upcoming tasks. Following this, all participants received a 5-minute introduction of the ILCS. This session was designed to familiarize students with the system's core functions and usage, providing necessary initial guidance for the experimental group while ensuring that the control group also had a basic understanding of the system for subsequent questionnaire completion. Following this session, all participants completed the pretest, which consisted of a knowledge acquisition questionnaire and validated scales measuring learning effect and flow experience. The pretest and posttest instruments contained identical items to guarantee the comparability of results.

The two groups received different forms of

instructional support during the experimental phase. Students in the control group relied exclusively on teacher guidance to complete the experiments. When these students encountered difficulties, they raised their hands to request assistance and received individualized, one-on-one support from the teacher. In contrast, students in the experimental group were required to first seek assistance from the ILCS. Only when the system failed to provide a satisfactory response did they turn to their teacher for help. Throughout the process, the research team systematically recorded the number of questions posed to the teacher in each group, indicating how different guidance modalities affected students' dependency on and autonomy in learning. Students in the experimental group used the ILCS solely for task execution and problem-solving during the experiment rather than for the targeted posttest preparation, thereby minimizing any potential testing effects.

All students took the posttest after completing the experiment, consisted of a knowledge acquisition questionnaire and validated scales measuring learning effect and flow experience. By analyzing the pretest–posttest results together with the recorded frequency of teacher interventions, the study aimed to provide a comprehensive evaluation of the impact of different instructional approaches on knowledge acquisition, learning effect, and flow experience. The experimental procedure is illustrated in Figure 3.

4.5 | Data Analysis Method

All quantitative analyses were conducted using IBM SPSS Statistics 27. The study employed multiple statistical methods to examine the effects of the ILCS on students' knowledge acquisition, learning effect, and flow experience.

Knowledge acquisition was assessed using a pretest–posttest control group design. Posttest scores were analyzed through analysis of covariance

Table 2 Reliability and validity of the questionnaires

Name of questionnaire	Testing item	Cronbach's alpha	KMO	Bartlett's sphericity test		
				chi-square	d.f.	Sig.
Learning effect scale	Cognition	0.961	0.890	1332.716	78	$p < 0.001$
	Skill	0.975				
	Emotion	0.940				
	Total	0.985				
Flow experience scale	Cognition	0.946	0.894	960.596	66	$p < 0.001$
	Time distortion	0.930				
	Focus	0.917				
	Emotion	0.948				
	Total	0.973				

Notes. KMO: Kaiser-Meyer-Olkin, d.f.: degree of freedom, Sig.: significance.

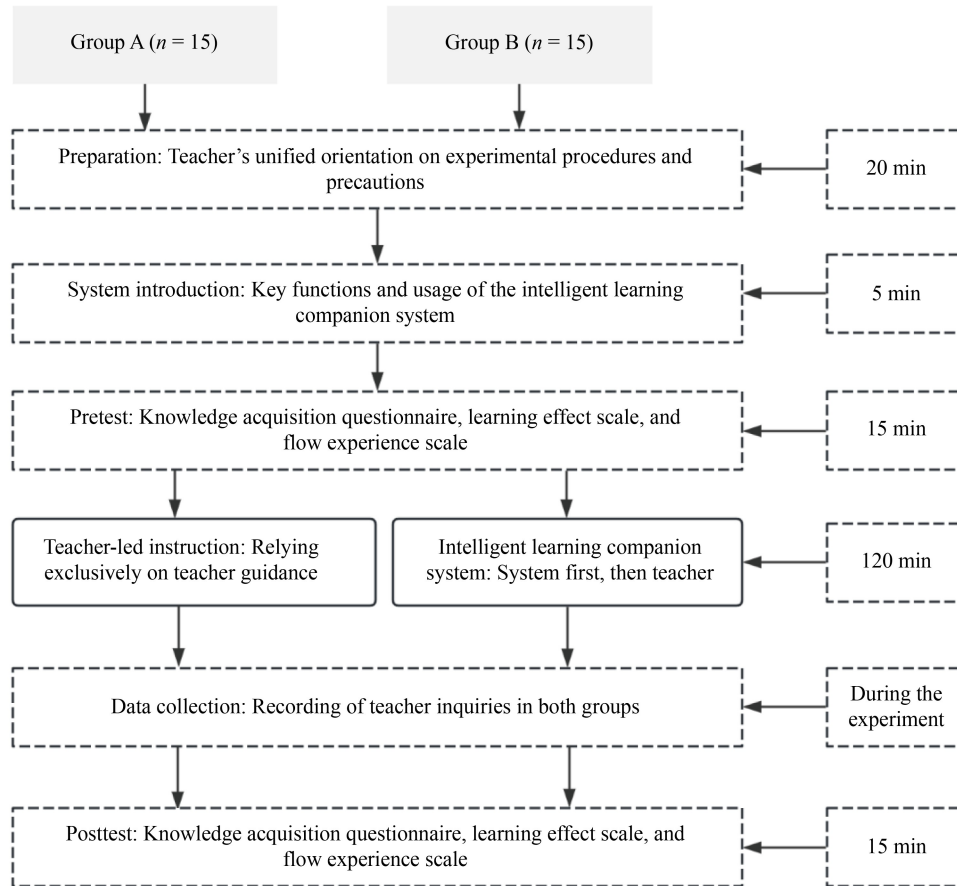


Figure 3 Experimental procedure for evaluating the impact of instructional approaches on knowledge acquisition, learning effect, and flow experience.

(ANCOVA), with pretest scores included as covariates to account for baseline differences between the experimental and control groups. The partial eta squared was calculated to evaluate the practical significance of the observed effects in accordance with Cohen's (2013) guidelines.

The students' learning effect were evaluated across three dimensions: cognition, skill, and emotion. Independent-sample *t*-tests were conducted to compare the performances of the experimental and control groups. Cohen's *d* was calculated to provide an estimate of the effect size and to interpret the practical relevance of any differences between the students in each group.

Flow experience was assessed in four dimensions: cognitive control, immersion and time transformation, loss of self-consciousness, and autotelic experience. Each dimension was analyzed using independent-sample *t*-tests and the effect size of each group was computed to evaluate the practical implications of group differences.

Semi-structured interviews were conducted with a subset of five students from the experimental group to collect qualitative feedback on their experi-

ences using the ILCS. The interview transcripts were then analyzed using ChatGPT-5 to facilitate the thematic coding of relevant experiences, such as problem-solving effectiveness, system usability, and areas for improvement. The insights derived from this qualitative analysis were subsequently integrated with quantitative findings to comprehensively evaluate the system's effectiveness.

5 Data Analysis Results

5.1 | Knowledge Acquisition Assessment

This study aimed to investigate the impact of an ILCS on students' knowledge acquisition. As shown in Table 3, the independent-sample *t*-test for pretest knowledge scores indicated no significant differences between the experimental group and control group ($p = 0.641$). Descriptive statistics revealed that the experimental group's score on the posttest ($M = 4.85$, $SD = 0.56$, here M denoting mean value and SD indicating standard deviation) was slightly higher than that of the control

group (M = 4.36, SD = 1.15), suggesting a preliminary positive trend. However, an ANCOVA controlling for pretest scores indicated that students' group membership on posttest knowledge acquisition was not statistically significant ($F(1, 24) = 1.620$, here F means F -test, $p = 0.215$), as shown in Tables 4 and 5. In Table 5, group denotes the effect of variables in ANCOVA.

While the statistical results were not significant, according to Cohen's guidelines, the calculated effect size (partial $\eta^2 = 0.063$) represents a medium effect, suggesting that the intervention may have had a meaningful practical impact on knowledge acquisition, even if statistical significance was not reached. The failure to achieve significance is potentially attributable to the limited statistical power resulting from the small sample size ($n = 15$, per group). Thus, this non-significant finding should be interpreted as a lack of sufficient evidence to confirm effectiveness rather than as a definitive negation of its effect.

5.2 | Learning Effect Analysis

Prior to the intervention, an independent-sample t -test was conducted on the pretest data to ensure baseline equivalence between the control (group A) and experimental (group B) groups. As shown in Table 6, the t -test results indicated no statistically significant differences ($p > 0.050$) between the two groups in terms of

Table 3 Descriptive statistics and independent-sample t -test for the pretest knowledge acquisition

Group	M	SD	t -value	p -value	Cohen's d
A	3.860	1.027	-0.471	0.641	-0.180
B	4.080	1.382			

Notes. M: mean value, SD: standard deviation.

Table 4 Descriptive statistics of the posttest knowledge acquisition

Group	M	SD
A	4.360	1.151
B	4.850	0.555
Total	4.590	0.930

Notes. M: mean value, SD: standard deviation.

Table 5 ANCOVA results for the posttest knowledge acquisition (after controlling for pretest scores)

Source	Type III sum of squares	d.f.	Mean square	F	Sig.	Partial η^2
Pretest score	2.147	1	2.147	2.746	0.111	0.103
Group	1.269	1	1.269	1.623	0.215	0.063
Error	18.760	24	0.782			

Notes. d.f.: degree of freedom, Sig.: significance.

the cognition ($t = -0.616$, $p = 0.543$), skill ($t = 0.921$, $p = 0.365$), and emotion ($t = -0.777$, $p = 0.444$). This confirms that both groups had a comparable baseline level prior to the intervention.

An independent-sample t -test was conducted across three dimensions—cognition, skill, and emotion—to examine the impact of different instructional approaches on students' learning effect. The test's results are summarized in Table 7.

The analysis revealed no statistically significant differences between the control and experimental groups in the cognition and emotion dimensions. Specifically, the control group scored 3.400 and the experimental group 3.640 in the cognition dimension ($t = -1.335$, $p = 0.193$), with a medium effect size (Cohen's $d = -0.480$), suggesting a practically meaningful difference despite the non-significant p -value. The scores for the emotion dimension were 3.620 (control group) and 3.600 (experimental group) ($t = 0.131$, $p = 0.897$), with a very small effect size (Cohen's $d = 0.040$), indicating minimal practical difference.

A significant difference was observed in the skill dimension. The experimental group supported by the ILCS, scored 3.880—a significantly higher score than the control group's score of 3.500 ($t = -2.709$, $p = 0.011$), with a large effect size (Cohen's $d = -0.97$), reflecting the system's substantial practical impact on students' skill acquisition. This suggests that the use of the ILCS provides students with notable advantages in

Table 6 Independent-sample t -test results for the pretest learning effect analysis of experimental and control groups

Dimension	Group	M	SD	t -value	p -value	Cohen's d
Cognition	A	3.310	0.390	-0.616	0.543	-0.230
	B	3.400	0.400			
Skill	A	3.110	0.340	0.921	0.365	0.340
	B	3.020	0.210			
Emotion	A	3.110	0.390	-0.777	0.444	-0.280
	B	3.220	0.390			

Notes. M: mean value, SD: standard deviation.

Table 7 Independent-sample t -test results for the posttest learning effect analysis of experimental and control groups

Dimension	Group	M	SD	t -value	p -value	Cohen's d
Cognition	A	3.400	0.590	-1.335	0.193	-0.480
	B	3.640	0.390			
Skill	A	3.500	0.460	-2.709*	0.011	-0.970
	B	3.880	0.300			
Emotion	A	3.620	0.600	0.131	0.897	-0.280
	B	3.600	0.260			

Notes. M: mean value, SD: standard deviation, * $p < 0.050$.

skill acquisition during analog circuit laboratory instruction.

5.3 | Flow Experience Analysis

An independent-sample *t*-test was conducted on the pretest data for flow experience (see Table 8) prior to the evaluation of the intervention's effectiveness. The results indicated no statistically significant differences ($p > 0.050$) between the groups in the dimensions of cognitive control ($t = -0.426, p = 0.673$), immersion and time transformation ($t = 0.373, p = 0.712$), loss of self-consciousness ($t = -0.699, p = 0.49$), and autotelic experience ($t = -0.375, p = 0.711$), confirming that both groups possessed a comparable baseline level of flow experience before the intervention.

Independent-sample *t*-tests were conducted across four dimensions, including cognitive control, immersion and time transformation, loss of self-consciousness, and autotelic experience, to evaluate the effectiveness of the ILCS in enhancing flow experience. The test results are presented in Table 9.

Table 8 Independent-sample *t*-test results for the pretest flow experience of experimental and control groups

Dimension	Group	M	SD	<i>t</i> -value	<i>p</i> -value	Cohen's <i>d</i>
Cognitive control	A	3.730	0.520	-0.426	0.673	-0.160
	B	3.820	0.620			
Immersion and time transformation	A	3.670	0.500	0.373	0.712	0.140
	B	3.600	0.470			
Loss of self-consciousness	A	3.420	0.910	-0.699	0.490	-0.260
	B	3.620	0.630			
Autotelic experience	A	3.690	0.460	-0.375	0.711	-0.140
	B	3.760	0.510			

Notes. M: mean value, SD: standard deviation.

Table 9 Independent-sample *t*-test results for the posttest flow experience of experimental and control groups

Dimension	Group	M	SD	<i>t</i> -value	<i>p</i> -value	Cohen's <i>d</i>
Cognitive control	A	3.780	0.410	-1.772	0.087	-0.650
	B	4.090	0.540			
Immersion and time transformation	A	3.530	0.370	-4.049***	< 0.001	-1.480
	B	4.270	0.590			
Loss of self-consciousness	A	3.730	0.620	-1.458	0.156	-0.530
	B	4.070	0.630			
Autotelic experience	A	3.890	0.480	-2.190*	0.037	-0.800
	B	4.290	0.520			

Notes. M: mean value, SD: standard deviation, * $p < 0.050$, *** $p < 0.001$.

The analysis indicated that there were no statistically significant differences between the control and experimental groups in the cognitive control and loss of self-consciousness dimensions. Specifically, in the cognitive control dimension, the control group obtained a mean score of 3.780, whereas the experimental group scored 4.090 ($t = -1.772, p = 0.087$) with a medium effect size (Cohen's $d = -0.650$), suggesting the presence of a practically meaningful difference despite the non-significant *p*-value. The scores in the loss of self-consciousness dimension were 3.730 (control group) and 4.070 (experimental group) ($t = -1.458, p = 0.156$) with a medium effect size (Cohen's $d = -0.530$), indicating that the intervention may still have practical relevance.

Conversely, statistically significant differences were observed in the immersion and time transformation and autotelic experience dimension. The experimental group scored significantly higher than the control group ($t = -4.049, p < 0.001$) in the immersion and time transformation dimension. The effect size was determined to be very large effect (Cohen's $d = -1.480$), reflecting a measurable and substantial practical impact. In the autotelic experience dimension, the experimental group outperformed the control group ($t = -2.190, p = 0.037$), with a large effect size (Cohen's $d = -0.800$), indicating a significant enhancement in students' enjoyment.

These results suggest that the RAG-based ILCS effectively enhances students' sense of immersion and enjoyment during the analog circuit laboratory learning process and that the reported effect sizes provide additional insight into the practical significance of these differences.

5.4 | Thematic Analysis of Student Interview Responses

Semi-structured interviews were conducted with a subset of five students from the experimental group to obtain qualitative insights into students' experiences with the ILCS. The interview protocol consisted of three main questions:

(1) "Could ILCS help you solve problems encountered during the experiment?"

(2) "How many problems were solved with the assistance of the system? Please describe the types or content of these problems."

(3) "What suggestions do you have for improving the system in terms of usability or functionality?"

The interview transcripts were first analyzed using ChatGPT-5 to facilitate thematic coding and to identify recurring patterns in students' responses. The preliminary themes generated by the system were subsequently reviewed, verified, and refined manually to ensure accuracy and reliability. Three major themes,

each of which had two subthemes, emerged from this combined automated and manual analysis, as summarized in Table 10. The first theme, effectiveness in solving foundational knowledge problems, reflects how the system supported foundational experimental knowledge and facilitated autonomous learning. The second theme, limitations in handling operational problems, highlights the system’s constraints in addressing complex hands-on tasks and limitations imposed by the single-turn dialogue mode. The third theme, functional enhancement needs, captures students’ suggestions for broader functionalities, including file and image support and improvement to practical usability.

P1 to P5 in Table 10 denote the supporting evidence five students provided. Table 10 features specific feedback provided by each interviewee. The thematic analysis of student interviews detailed above offers qualitative insights relevant to research questions. For RQ1, which focuses on knowledge acquisition, the theme, effectiveness in solving foundational knowledge problems, suggests that students generally perceived the ILCS as helpful for resolving foundational experimental issues and supporting autonomous knowledge acquisition. These findings indicate that the system can potentially facilitate students’ understanding of core concepts, although its effectiveness may vary depending on the type and complexity of the problems encountered.

For RQ2, which concerns overall learning effect, the themes, limitations in handling operational problems and functional enhancement needs, highlight certain constraints in the ILCS’s support for complex hands-on tasks, as well as students’ suggestions for usability and functionality improvements. While the system appears to assist with foundational experimental

problem-solving, these qualitative insights imply that the system’s impact on broader learning effect may be limited by operational challenges and interaction designs.

For RQ3, which covers students’ flow experiences, the same themes indicate that practical limitations and interface constraints could affect students’ engagement and immersion during experiments. Students’ requests for the system include broader functionalities indicate that enhancements in support and usability could potentially improve the quality of the learning experience.

Overall, the above thematic analysis provides a nuanced, qualitative perspective on students’ interactions with the ILCS, highlighting both its supportive role in foundational knowledge acquisition and the areas in which its design may limit learners’ broader learning experiences.

6 Discussion

6.1 | Knowledge Acquisition

The ANCOVA results indicated no statistically significant differences between experimental group ($M = 4.850$, $SD = 0.560$) and the control group ($M = 4.360$, $SD = 1.150$) in posttest knowledge acquisition ($F(1, 24) = 1.620$, $p = 0.215$); however, the calculated effect size (partial $\eta^2 = 0.063$) exerts a medium effect according to Cohen’s guidelines, suggesting that the intervention may have meaningfully impacted students’ knowledge acquisition despite the non-significant

Table 10 Thematic analysis of students’ experiences with the ILCS

Major theme	Subtheme	Description	Supporting evidence
Effectiveness in solving foundational knowledge problems	Resolution of foundational experimental issues	The system reliably supported students in resolving foundational experimental issues, including waveform anomalies and LED polarity identification	P1: waveform conversion output anomaly; P2: LED polarity question; P3: foundational knowledge problem support
	Support for autonomous knowledge acquisition	The system functioned as an auxiliary tool, facilitating students’ quick access to foundational knowledge during experiments	P3: basic questions solved via PowerPoint, with system providing supplementary guidance
Limitations in handling operational problems	Inability to fully resolve hands-on experimental issues	The system limited capacity to address complex operational problems, particularly circuit operations or image-related tasks	P2: system limited for hands-on guidance; P5: system limited for knowledge questions
	Constraint of single-turn dialogue mode	Current single-turn interactions restrict access to full system functionalities, such as file uploading and summarization	P4: wished for single-turn file uploading and summarization capabilities
Functional enhancement needs	File and image support	Students suggested system support for uploading files and images to summarize experimental materials and improve operational guidance	P4: file uploading and summarization request; P5: image uploading request
	Expectation for improved usability	Students indicated that, while the system helped with foundational knowledge problem-solving, enhancements are needed for practical utility in real-time experiments	P2, P4, and P5: emphasized limitations in hands-on guidance and desire for broader functionalities

Notes. P1–P5 refer to different students. ILCS: intelligent learning companion system, LED: light emitting diode.

p-value. This lack of statistical significance is potentially attributable to the limited sample size ($n = 15$, per group), which reduces the statistical power to detect medium-sized effects.

The performance metrics of the ILCS provide additional context for interpreting these findings. The system achieved a knowledge coverage rate of 96.1%, an answer accuracy of 99.4%, and a personalized feedback rate of 67.2% in the analog circuit experiment scenario. These high levels of content coverage and response accuracy indicate that the system was capable of reliably addressing the students' foundational experimental questions. This system's performance level may help explain the medium effect observed in knowledge acquisition despite the statistical test not achieving statistical significance.

Qualitative insights from the semi-structured interviews provided additional context for this finding. The theme, effectiveness in solving foundational knowledge problems, indicates that the students generally perceived ILCSs as a helpful tool for addressing foundational experimental knowledge and supporting autonomous learning. This qualitative evidence aligns with the quantitative trend, suggesting that, while the statistical test did not achieve significance, the system may contribute positively to knowledge acquisition in practice.

These findings underscore the importance of delineating clear roles prior to learning activities involving human-AI collaboration. In this model, human teachers are responsible for guiding conceptual understanding and providing scaffolding, while students are tasked with carrying out hands-on problem-solving and knowledge application. ILCSs serve as a supplementary resource, offering support for foundational content and immediate queries. Optimizing this differentiated division of labor, such as by integrating system prompts with teacher-led explanations and student practice, can potentially significantly enhance knowledge acquisition effect in future implementations of this system.

6.2 | Learning Effect

Independent-sample *t*-tests were conducted across cognition, skill, and emotion dimensions (Table 7) to examine the impact of instructional approaches on student's learning effect. No statistically significant differences were observed in the cognition dimension ($t = -1.335$, $p = 0.193$, Cohen's $d = -0.480$, medium effect) or the emotion dimension ($t = 0.131$, $p = 0.897$, Cohen's $d = 0.040$, very small effect), suggesting a minimal impact on conceptual understanding and affective engagement. In contrast, a significant improvement was found in the skill dimension for the experimental group supported by the ILCS ($t = -2.709$, $p = 0.011$, Cohen's $d = -0.970$, large effect), indicating that the system

substantially enhanced students' hands-on skill acquisition during analog circuit laboratory instruction. This aligns with recent research finding that ChatGPT significantly enhanced students' research skills by helping them with idea generation, literature review, and data analysis (Li et al., 2025a).

Qualitative insights from semi-structured interviews complement these quantitative findings. Students generally perceived the ILCS as effective for addressing foundational experimental knowledge and supporting autonomous problem-solving, which with the observed advantage in skill development. The interview themes, such as limitations in handling operational problems and functional enhancement needs, highlight areas where the system's support is constrained, providing a potential explanation for the minimal gains observed in the cognition and emotion dimensions.

From a human-AI collaborative perspective, a key lesson from these findings concerns the design of role boundaries between teachers, students, and ILCS. The lack of a detailed problem classification was a key design flaw within the system. In the experimental setup, students were instructed to first consult the ILCS for all problems without distinguishing between basic issues and more complex tasks. This led to students relying on the system, including those that required emotional support or advanced guidance, which limited the system's effectiveness.

The experiment revealed that a more effective approach would involve classifying problems: assigning basic or technical queries to the ILCS and reserving more complex, conceptual, or emotional tasks for the teacher. This would optimize the learning process and enhance the system's support for students.

6.3 | Flow Experience

Independent-sample *t*-tests were conducted to examine the impact of instructional approaches on students' flow experience across four dimensions: cognitive control, immersion and time transformation, loss of self-consciousness, and autotelic experience (Table 9).

No statistically significant differences were observed in the cognitive control ($t = -1.772$, $p = 0.087$, Cohen's $d = -0.650$) and loss of self-consciousness ($t = -1.458$, $p = 0.156$, Cohen's $d = -0.530$) dimensions. Although not statistically significant, the medium effect sizes for these scores suggest that the ILCS may provide practical benefits in terms of supporting students' cognitive regulation and awareness during experiments.

Significant differences were observed in immersion and time transformation ($t = -4.049$, $p < 0.001$, Cohen's $d = -1.480$) and autotelic experience ($t = -2.190$, $p = 0.037$, Cohen's $d = -0.800$). The *p*-values indicate statistical significance, and their large effect sizes imply that the system's contribution to

increased student engagement and enjoyment is both statistically significant and practically substantial.

Qualitative insights from the semi-structured interviews in this study provide a broader context. Themes, effectiveness in solving foundational knowledge problems and functional enhancement needs, indicate that students valued the system for instant procedural and knowledge support, which helped them maintain focus and flow during the experiments. The theme, limitations in handling operational problems, highlights constraints such as interface restrictions or difficulty in addressing complex hands-on tasks, suggesting areas where the flow could be interrupted. Students also provided suggestions for broader functionalities that could further enhance the students' ability to use and engage with the system.

Optimizing this strategic division of labor by ensuring that the system's assistance complements teacher scaffolding and student practice holds the potential to enhance the flow experience in future implementations, thereby aligning both qualitative insights and quantitative trends.

7 Conclusions

7.1 | Research Findings and Educational Value

This study demonstrates that the proposed ILCS can significantly enhance students' practical skills in analog circuit laboratory instruction; however, no statistically significant differences in cognition and emotion were found. The system's contribution to skill acquisition highlights the importance of refining the division of roles among teachers, students, and the system to facilitate future improvements in learning effect.

First, the division of labor should be optimized. Teachers should continue to focus on higher-order guidance, such as conceptual explanations and emotional support, while students should take more responsibility for carrying out hands-on problem-solving and knowledge application. ILCS should serve as a supplementary resource for addressing foundational knowledge and procedural support. By adjusting the division of labor, the collaborative dynamics among teachers, students, and the system can be enhanced to facilitate a more effective learning experience.

Second, expanding the ILCS's capabilities is essential. Students' feedback indicate that the system should be further developed to handle more complex operational problems, such as image processing and document uploading. These improvements will improve the system's versatility and ensure that it is better equipped to support a broader range of students' needs as they conduct experiments.

Third, given the ILCS's limited ability to provide emotional support, teachers should play a more prominent role in engaging students emotionally. While the system excels at delivering technical support, integrating elements of motivation and emotional scaffolding into the system's design could help students feel more engaged during the learning process, which could lead to more holistic learning experiences.

Fourth, the ILCS's ability to reduce teacher's instructional load should be further improved. The experimental group, which used the system, required significantly fewer teacher interventions than the control group (11 versus 21), suggesting that the system can effectively alleviate the teacher's instructional burden. Future iterations of the system should further increase its capacity to address routine queries. This would allow teachers to focus on more complex, higher-order teaching tasks, freeing up their time to provide more personalized and detailed support to their students and fostering a highly immersive learning environment.

7.2 | Role Optimization in Human-AI Collaborative Teaching

In this study, situated in the context of complex engineering education, we propose a human-AI collaborative teaching model designed to optimize the unique strengths of the teacher, student, and the ILCS through a clear division of labor. The teacher's role within this model is redefined as a facilitator who focuses on higher-order pedagogical tasks. The student is an active explorer who takes responsibility for their own learning autonomy, while the ILCS serves as a collaborative tutor that provides students with instant, personalized technical support. [Table 11](#) details the specific roles, core responsibilities, and collaborative relationships of each party in this synergistic model.

This study reveals the profound impact of human-AI collaborative teaching on traditional instructional roles in the context of complex engineering education. The proposed ILCS is not merely a technical supplement to engineering education; rather, it necessitates a fundamental restructuring of the duties and relationships among key participants in human-AI collaborative teaching: teacher, student, and the system.

The findings indicate that the intelligent learning companion is highly effective in reducing students' low-level cognitive load and enhancing their performance on skill-based tasks. The system's performance presents a unique opportunity to optimize the role teachers play in education. Teachers can be freed from having to conduct repetitive knowledge transfers and basic skill coaching to focus on areas of greater human and professional value. Specifically, teachers within this model remains responsible for building a systematic

Table 11 Roles and responsibilities of participants in human–AI collaborative teaching

Participant	Core responsibility	Collaborative relationship
Teacher	Providing a systematic knowledge framework, guiding higher-order pedagogical tasks, evaluating complex tasks, managing the overall teaching process, and offering emotional support	Collaborating with the intelligent learning companion system to focus on tasks requiring human intelligence and emotional interaction, such as conceptual explanations and in-depth guidance
Student	Acting as an active explorer, participating in experiments, exploring knowledge application, and engaging in self-regulation and reflection	Autonomously completing practical tasks with the immediate support of the intelligent learning companion and seeking assistance from the teacher for higher-level problems
Intelligent learning companion	Providing instant, personalized support and feedback, answering lower-level knowledge questions, reducing cognitive load, and supporting skill acquisition	Creating a synergistic effect with the teacher by taking on some pedagogical tasks, allowing the teacher to focus on higher-level educational activities

knowledge framework and for filling the gaps of intelligent systems in terms of integrating and connecting higher-level knowledge. In complex engineering education, this role involves seamlessly integrating foundational theory with practical applications and guiding students so that they may grasp the deeper logic behind the principles they learned. Moreover, teacher’s distinctive value lies in their ability to guide students in critical thinking and creative problem-solving, they design and facilitate open-ended discussions that encourage students to explore multiple solutions rather than relying solely on a single answer provided by an intelligent agent. As education concerns not only knowledge transfer but also emotion connection and value formation, teachers should provide students with emotional support, which may involve asking students to face setbacks, motivating them intrinsically, and guiding them on professional ethics and engineering responsibilities.

The optimization of the student’s role in this framework is the key to the success of human–AI collaboration. Students should be cultivated as active self-regulators who are adept at leveraging intelligent tools for autonomous learning while strategically engaging with teachers. Students should first seek immediate assistance from an intelligent learning companion for specific factual or procedural questions. This helps students resolve problems quickly and fosters their capacity for independent learning and effective tool utilization. When the intelligent learning companion is unable to solve a problem or students require a deeper conceptual understanding of the topic, students should proactively turn to their teachers. This “system first, then teacher” model ensures that the teacher’s guidance is reserved for the most valuable and intellectually demanding interactions.

Future ILCS should evolve beyond having simple Q&A capabilities to become pedagogically wise collaborative partners that are better suited to assist students in complex educational settings. This requires such companions to integrate principles from intelligent tutoring systems, develop stronger cognitive diagnostic abilities, and employ more sophisticated, heuristic-based questioning. By using a phased, progressive

approach to guide students’ thinking instead of providing them with direct answers, these intelligent learning companions can facilitate students’ knowledge internalization and development of higher-order thinking. System design should also incorporate affective computing and adaptive interaction mechanisms. This would enable future companions to recognize and respond to students’ emotional states, thereby providing students with more holistic support. Ultimately, the core value of ILCS lies in their immediacy and personalization. These companions should play a consistently auxiliary role and should be focused on providing real-time technical support and feedback to effectively reduce students’ low-level cognitive load during experiments rather than replacing the teacher in any capacity.

8 Limitations

Despite this study’s contributions, it has three limitations that provide valuable directions for future research.

First, the study’s relatively small sample size may limit the generalizability of its findings. The study’s experiment was conducted within a single, intact class of the electric circuit experiments course with a class size ($n = 30$, with 15 students per group) determined by the curriculum. This constraint poses a challenge to the external validity of our results. Future studies should aim to expand the sample size to enhance the representativeness of the findings. While this study compared the effects of traditional teacher-led instruction with those of instruction guided by the ILCS, it did not account for the potential interaction effects between different instructional approaches. Future research could adopt a multigroup experimental design to explore the combined impact of teacher instruction and intelligent learning companions.

Second, this study primarily focused on knowledge acquisition, learning effect, and flow experience dimensions. It did not incorporate other potentially influential factors that may have confounded the results, such as students’ prior knowledge and learning

motivation. To better understand the multifaceted nature of AI-supported learning, future research should incorporate a broader range of variables into the experimental design.

Third, the ILCS has technical limitations, particularly its ability to address complex experimental problems and its limited capacity for emotional support. The system is currently unable to understand and respond to students' emotional states, which is a crucial aspect of human teaching. Future enhancements to the system's architecture to include a more comprehensive knowledge base and improved reasoning capabilities are necessary for the system to overcome these challenges and boost its performance. Moreover, the relatively short duration of the experiment, a constraint imposed by the course schedule, may have led to the study's inability to fully capture the long-term effects of the ILCS. Subsequent studies should consider extending the intervention period to assess its sustained impact over time.

Acknowledgments This work was supported in part by the Chinese Academy of Engineering Science and Technology Strategy Consulting Project (Grant No. 2025-XZ-136), in part by the Digital Intelligence Beijing University of Posts and Telecommunications Integrated Innovation Program (Grant No. RCXM-2025-047), in part by the BUPT Innovation and Entrepreneurship Support Program (Grant No. 2025-YC-T044), and in part by the 2024 Beijing Higher Education Undergraduate Teaching Reform and Innovation Project entitled "Exploration and Practice in Building Innovative Teaching Teams Empowered by Digital Intelligence Under the Background of Three-Education Collaboration."

Author Contributions Xinyi Tian was responsible for conceptualizing the research, developing the ILCS, and drafting the manuscript. Jianwei Li supervised the study, guided the research methodology, and provided critical revisions to the manuscript. Yanli Ji contributed to data collection, experimental instruction, and statistical analysis. All authors read and approved the final manuscript.

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

Ethics Statements The authors declare that their Institutional Ethics Committee confirmed that no ethical review was required for this study. Written informed consent for participation was not required because all participants' data was anonymized before the statistical analyses were conducted.

Data Availability Statements The authors confirm that all data generated or analyzed during this study are included in this published article.

References

- Abdullah, A., Adil, M., Rosenbaum, L., Clemmons, M., Shah, M., Abrahamson, D., & Neff, M. (2017). Pedagogical agents to support embodied, discovery-based learning. In: *Proceedings of the 17th International Conference on Intelligent Virtual Agents*. Stockholm: Springer, 1–14.
- Al-Abri, A. (2025). Exploring ChatGPT as a virtual tutor: A multi-dimensional analysis of large language models in academic support. *Education and Information Technologies*, 30(12), 17447–17482.
- Bas, C., & Kiraz, A. (2025). Primary school teachers' needs for AI-supported STEM education. *Sustainability*, 17(15), 7044.
- Baylor, A. L., & Kim, S. (2009). Designing nonverbal communication for pedagogical agents: When less is more. *Computers in Human Behavior*, 25(2), 450–457.
- Beard, K. S. (2015). Theoretically speaking: An interview with Mihaly Csikszentmihalyi on flow theory development and its usefulness in addressing contemporary challenges in education. *Educational Psychology Review*, 27(2), 353–364.
- Chen, Y. X. (2022). *Research on the learning effect of college students based on the blended teaching model—Taking Nanjing University of Posts and Telecommunications as an example*. Thesis for the Master's Degree. Nanjing: Nanjing University of Posts and Telecommunications. (in Chinese).
- Chiu, T. K. F., & Rospigliosi, P. (2025). Encouraging human–AI collaboration in interactive learning environments. *Interactive Learning Environments*, 33(2), 921–924.
- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. 2nd ed. New York: Routledge.
- Cohn, C., Snyder, C., Fonteles, J. H., T. S., A., Montenegro, J., & Biswas, G. (2025). A multimodal approach to support teacher, researcher and AI collaboration in STEM + C learning environments. *British Journal of Educational Technology*, 56(2), 595–620.
- Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety: Experience flow in work and play*. San Francisco: Jossey-Bass.
- Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York: Harper & Row.
- Dai, L., Jung, M. M., Postma, M., & Louwse, M. M. (2022). A systematic review of pedagogical agent research: Similarities, differences and unexplored aspects. *Computers & Education*, 190, 104607.
- Ding, L., Li, T., Jiang, S. Y., & Gapud, A. (2023). Students' perceptions of using ChatGPT in a physics class as a virtual tutor. *International Journal of Educational Technology in Higher Education*, 20(1), 63.
- El Shazly, R. (2021). Effects of artificial intelligence on English speaking anxiety and speaking performance: A case study. *Expert Systems*, 38(3), e12667.
- Fan, Y. Z., Tang, L. Z., Le, H. X., Shen, K. J., Tan, S. F., Zhao, Y. Y., Shen, Y., Li, X. Y., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*, 56(2), 489–530.

- Fang, H. G., Wang, S. C., Xue, S. S., & Wang, X. L. (2020). Research on human–computer cooperative teaching supported by artificial intelligence robot assistant. In: Pinkwart, N., & Liu, S., eds. *Artificial intelligence supported educational technologies*. Cham: Springer, 45–58.
- Filippi, S., & Motyl, B. (2024). Large language models (LLMs) in engineering education: A systematic review and suggestions for practical adoption. *Information*, 15(6), 345.
- García-Méndez, S., de Arriba-Pérez, F., & Somoza-López, M. C. (2025). A review on the use of large language models as virtual tutors. *Science & Education*, 34(2), 877–892.
- Gürdür Broo, D., Kaynak, O., & Sait, S. M. (2022). Rethinking engineering education at the age of Industry 5.0. *Journal of Industrial Information Integration*, 25, 100311.
- Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning? *Educational Research Review*, 6(1), 27–54.
- Heutte, J., Fenouillet, F., Martin-Krumm, C., Gute, G., Raes, A., Gute, D., Bachelet, R., & Csikszentmihalyi, M. (2021). Optimal experience in adult learning: Conception and validation of the flow in education scale (EduFlow-2). *Frontiers in Psychology*, 12, 828027.
- Holstein, K., & Alevén, V. (2022). Designing for human–AI complementarity in K-12 education. *AI Magazine*, 43(2), 239–248.
- Jang, H. (2016). Identifying 21st century STEM competencies using workplace data. *Journal of Science Education and Technology*, 25(2), 284–301.
- Jeon, J. (2023). Chatbot-assisted dynamic assessment (CA-DA) for L2 vocabulary learning and diagnosis. *Computer Assisted Language Learning*, 36(7), 1338–1364.
- Ji, H., Han, I., & Ko, Y. (2022). A systematic review of conversational AI in language education: Focusing on the collaboration with human teachers. *Journal of Research on Technology in Education*, 55(1), 48–63.
- Kaye, L. K. (2012). *Motivations, experiences and outcomes of playing videogames*. Dissertation for the Doctoral Degree. Preston: University of Central Lancashire.
- Kim, J. (2024a). Leading teachers' perspective on teacher–AI collaboration in education. *Education and Information Technologies*, 29(7), 8693–8724.
- Kim, J. (2024b). Types of teacher–AI collaboration in K-12 classroom instruction: Chinese teachers' perspective. *Education and Information Technologies*, 29(10), 17433–17465.
- Kim, Y., & Baylor, A. L. (2006a). A social–cognitive framework for pedagogical agents as learning companions. *Educational Technology Research and Development*, 54(6), 569–596.
- Kim, Y., & Baylor, A. L. (2006b). Pedagogical agents as learning companions: The role of agent competency and type of interaction. *Educational Technology Research and Development*, 54(3), 223–243.
- Kim, Y., & Baylor, A. L. (2016). Research-based design of pedagogical agent roles: A review, progress, and recommendations. *International Journal of Artificial Intelligence in Education*, 26(1), 160–169.
- Kong, X. M., Fang, H. G., Chen, W. L., Xiao, J. J., & Zhang, M. H. (2025). Examining human–AI collaboration in hybrid intelligence learning environments: Insight from the synergy degree model. *Humanities and Social Sciences Communications*, 12(1), 821.
- Li, Y., Sadiq, G., Qambar, G., & Zheng, P. Y. (2025a). The impact of students' use of ChatGPT on their research skills: The mediating effects of autonomous motivation, engagement, and self-directed learning. *Education and Information Technologies*, 30, 4185–4216.
- Li, H., Wang, Z., Ding, L., Zhang, J. W., & Wang, G. H. (2025b). The facts about the effects of pedagogical agents on learners' cognitive load: A meta-analysis based on 24 studies. *Frontiers in Psychology*, 16, 1635465.
- Lin, L. J., Ginns, P., Wang, T. H., & Zhang, P. L. (2020). Using a pedagogical agent to deliver conversational style instruction: What benefits can you obtain? *Computers & Education*, 143, 103658.
- Martin, S. A., Azevedo, R., Taub, M., Mudrick, N. V., Millar, G. C., & Grafsgaard, J. F. (2016). Are there benefits of using multiple pedagogical agents to support and foster self-regulated learning in an intelligent tutoring system? In: *Proceedings of the 13th International Conference on Intelligent Tutoring Systems*. Zagreb: Springer, 273–279.
- Oker, A., Pecune, F., & Declercq, C. (2020). Virtual tutor and pupil interaction: A study of empathic feedback as extrinsic motivation for learning. *Education and Information Technologies*, 25(5), 3643–3658.
- Pham, T., Nguyen, B., Ha, S., & Nguyen, N. T. (2023). Digital transformation in engineering education: Exploring the potential of AI-assisted learning. *Australasian Journal of Educational Technology*, 39(5), 1–19.
- Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. New York: Cambridge University Press.
- Schroeder, N. L. (2017). The influence of a pedagogical agent on learners' cognitive load. *Educational Technology & Society*, 20(4), 138–147.
- Sikström, P., Valentini, C., Sivunen, A., & Kärkkäinen, T. (2022). How pedagogical agents communicate with students: A two-phase systematic review. *Computers & Education*, 188, 104564.
- Sikström, P., Valentini, C., Sivunen, A., & Kärkkäinen, T. (2024). Pedagogical agents communicating and scaffolding students' learning: High school teachers' and students' perspectives. *Computers & Education*, 222, 105140.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285.
- Tokatlidis, C., Tselegkaridis, S., Rapti, S., Sapounidis, T., & Papakostas, D. (2024). Hands-on and virtual laboratories in electronic circuits learning—knowledge and skills acquisition. *Information*, 15(11), 672.
- Ubani, S., & Nielsen, R. (2022). Review of collaborative intelligent tutoring systems (CITS) 2009–2021. In: *Proceedings of the 2022 the 11th International Conference on Educational and Information Technology*. Chengdu: IEEE, 67–75.
- Yang, D., Wu, X. P., Liu, J. L., & Zhou, J. C. (2023).

CiteSpace-based global science, technology, engineering, and mathematics education knowledge mapping analysis. *Frontiers in Psychology*, 13, 1094959.

Yung, H. L., & Paas, F. (2015). Effects of cueing by a pedagogical agent in an instructional animation: A cognitive load approach. *Educational Technology & Society*, 18(3), 153–160.