Supporting Teachers' Professional Development With Generative AI: The Effects on Higher Order Thinking and Self-Efficacy

Jijian Lu, Ruxin Zheng, Zikun Gong, and Huifen Xu

***Abstract* — Generative AI has emerged as a noteworthy milestone and a consequential advancement in the annals of major disciplines within the domains of human science and technology. This study aims to explore the effects of generative AI-assisted pre-service teaching skills training on pre-service teachers' self-efficacy and higher order thinking. The participants of this study were 215 pre-service mathematics, science, and computer teachers from a university in China. Firstly, a pretest-posttest quasi-experimental design was implemented for an experimental group (teaching skills training by generative AI) and a control group (teaching skills training by traditional methods) by investigating the teacher self-efficacy and higher order thinking of the two groups before and after the experiment. Secondly, a semi-structured interview comprising open-ended questions was administered to 25 pre-service teachers within the experimental group to present their views on generative AI-assisted teaching. The results showed that the scores of pre-service teachers in the experimental group, who used generative AI for teachers' professional development, were considerably higher than those of the control group, both in teacher self-efficacy ($F = 8.589$, $p = 0.0084 < 0.05$) and higher order thinking ($F = 7.217$, $p = 0.008 < 0.05$). It revealed that generative AI can be effective in supporting teachers' professional development. This study produced a practical teachers' professional development method for pre-service teachers with generative AI.**

***Index Terms* — generative AI, pre-service teachers, teaching skills training, teacher self-efficacy, higher order thinking**

I. INTRODUCTION

Currently, artificial intelligence generated content (AIGC) stands out as one of the most prominent and captivating technologies in the contemporary landscape. With some instructions from the users, generative AI can generate content that meets their needs [1-3]. With the development of generative AI, the exploration of its application in the instructional design training of normal college students has gradually attracted attention. ChatGPT is a natural language processing model developed by Open AI [4]. As a prominent representative of generative AI, ChatGPT can generate corresponding text and images according to users' needs, which has shown its absolute advantages in various fields such as real-time question and answer, language translation, and programming code with its powerful interactive capabilities [5]. Since November 2022, with the emergence of ChatGPT and GPT-4, the generative artificial intelligence represented by ChatGPT has become a significant landmark and effective advance in the history of major human science and technology [6].

With the advancement of emerging technologies in the context of the Industry 4.0 revolution, the realm of education has increasingly focused on advancing the reform and development of education. This is achieved by integrating modern information technology to create a more flexible, efficient, and accessible educational system. There is no doubt that AIGC has also evoked polarized reactions in education, having a significant and far-reaching impact on school education in a short period [7]. Education should seize the opportunity to adjust the teaching objectives reasonably, improve the curriculum system and education model, and then promote the quality and equitable advancement of educational institutions in the era of artificial intelligence.

The cultivation of higher order thinking has long been regarded as a crucial factor of academic success, prompting the education field to emphasize its development significantly [8]. Critical thinking, creative thinking, problem-solving, and other higher order thinking skills are essential for students' development [9]. Similarly, there is increasing emphasis in education on evaluating and promoting teachers' higher order thinking [10]. In addition, teachers' self-efficacy also plays an important role in learning and teaching [11, 12]. Teacher self-efficacy is teachers' belief in their ability to perform teaching tasks successfully, which is a critical factor in teachers' teaching motivation and is positively correlated with teachers' teaching quality [13]. Pre-service teachers are both students studying education and future front-line teachers. Therefore, pre-service teachers are essential participants in investigating the development of teacher self-efficacy and higher order thinking because of their special status. In the training process of pre-service teachers, practical teaching skills training is the main pathway for their professional development [14-16]. In an interview with Shaughnessy (2004), Woolfolk pointed out that practical teaching is crucial to developing teacher self-efficacy [17].

In summary, the incorporation of generative AI into the training of pre-service teachers has the potential to positively impact the cultivation of their higher order thinking and teacher self-efficacy. Therefore, this study developed a supportive method of generative AI on the practical training of pre-service teachers' teaching skills and attempted to explore whether the supportive method would have an impact on the pre-service teachers' higher order thinking and self-efficacy through educational experiments to verify the effectiveness of generative AI in assisting the practical training of teaching skills.

This work was supported by the 2021 pilot project of the Ministry of Education's artificial intelligence to promote the construction of teachers' team, Major Humanities and Social Sciences Research Projects in Zhejiang higher education institutions, #2023GH005, and teaching academic and reform project of Hangzhou Normal University. (Corresponding author: Huifen Xu.)

Jijian Lu and Zikun Gong are with the Chinese Education Modernization Research Institute and the Jing Hengyi School of Education, Hangzhou Normal University, 2318 Yuhangtang Road, Hangzhou, Zhejiang, China. (e-mail: lujijian@foxmail.com; zkgong@163.com).

Ruxin Zheng and Huifen Xu are with the Jing Hengyi School of Education, Hangzhou Normal University, 2318 Yuhangtang Road, Hangzhou, Zhejiang, China. (e-mail: zrx_email@163.com; angela422@foxmail.com).

II. PRIOR RESEARCH

A. Generative AI in the Field of Education

A typical application of generative AI in the field of education is the educational application of ChatGPT [18]. With the global popularity of ChatGPT, many researchers have conducted research on the integration of ChatGPT into the educational ecosystem. ChatGPT can not only complete text generation tasks [19] but also solve various problems with its good reasoning ability [20]. Jeon and Lee explored the relationship between ChatGPT and teachers and how ChatGPT was integrated into teachers' teaching through educational experiments. Three teacher roles (coordinating different resources through high-quality teaching decisions, enabling students to become active learners, and raising awareness of AI ethics) and four ChatGPT roles (interlocutor, content provider, teaching assistant, and evaluator) are identified. Based on this, the collaborative model of teachers and ChatGPT is discussed in depth [21]. Humphry and Fuller explored the use of ChatGPT in chemistry education by having ChatGPT complete lab reports. Studies have shown that ChatGPT is weak in chemical analysis, stoichiometry, etc., but is very useful in helping students with calculations and writing [22]. Shoufan, from the perspective of students, used a follow-up survey to find out how students felt about the use of ChatGPT in their learning process. The survey findings indicate a prevailing sentiment among students, asserting that ChatGPT is perceived as highly potent, capable of engendering learning motivation, and providing substantial learning assistance. However, the results also reveal that the accuracy of the answers given by ChatGPT could be higher, which has certain limitations and can not completely replace human learning [23]. Thus, Farrokhnia et al. established a SWOT analysis framework by analyzing extensive literature, comprehensively discussing the advantages and disadvantages of ChatGPT, and discussing its opportunities and threats to education [24].

B. Pre-service Teaching Skills Training For Teachers

In the training process of pre-service teachers, the practical training of teaching skills is an essential and vital content [25]. In 1963, Professor Allen of Stanford University proposed microteaching for the first time. This teaching method is based on psychology, pedagogy, and modern teaching theories. It uses modern teaching methods and information technology to conduct classroom teaching practical training for teachers or normal university students. Microteaching consists of six stages: planning, teaching, observing (criticizing), re-planning, re-teaching, and re-observing [26].

Currently, numerous higher education institutions employ the instructional methodology for teacher training to enhance the professional skills development of prospective educators, thereby augmenting their practical competencies and application proficiency. Many researchers have studied the relationship between teaching skills training and teachers' self-efficacy [27]. Arsal carried out a one-semester educational experiment on 70 pre-service teachers and used the teacher self-efficacy scale to test the pre-service teachers in the experimental group before and after to explore the impact of practical teaching skills training on pre-service teachers' self-efficacy and found it had a positive influence on the cultivation of pre-service teachers' self-efficacy [26]. Aydin.S et al. also conducted a pre- and post-test education experiment on 26 pre-service chemistry teachers and carried out a semi-structured interview with them afterwards. The respondents believed that they had gained a lot of teaching experience in the teaching skills training, especially mastered many teaching strategies, which greatly promoted the development of their teacher self-efficacy [28]. D. Schina et al. integrated educational robots into practical teaching skills training, exploring the impact of integrating them into practical teaching skills training on pre-service teachers' self-efficacy and their views on such practical training methods [29]. The research results show that after the practical training of integrating educational robots into teaching skills, the teacher self-efficacy of pre-service teachers is significantly improved. Generally, there exists a favourable disposition among individuals towards the pragmatic training paradigm, wherein they perceive the training as pioneering, practical, and captivating.

III. RESEARCH QUESTION

Building on the previously mentioned literature review and analysis, this study examines how generative AI, exemplified by ChatGPT, can enhance instructional methodologies used by pre-service teachers. Furthermore, it aims to investigate the impact of using ChatGPT in teaching skills training on pre-service teachers' self-efficacy and higher order thinking. Therefore, the research questions involved in this study are as follows:

- How can ChatGPT assist pre-service teachers in practical teaching skills training?
- Compared with the traditional teaching skills training methods, what are the different effects of ChatGPT assisted training on the pre-service teachers' self-efficacy and higher order thinking?
- What are the opinions of the pre-service teachers in the experimental group about using ChatGPT to assist teaching?

IV. METHODOLOGY

A. Participants

The participants of this study are third-year students majoring in mathematics education, science education, and computer education from a university in Zhejiang Province, China. They come from six different classes, with two classes for each major, totaling approximately 215 students. Their ages range from 20 to 22 years old. According to the curriculum arrangement, they will respectively study practical training courses on teaching skills in mathematics, science, and computer classrooms during their third year. These three courses aim to provide pre-service teachers with practical training on teaching skills through lesson preparation, classroom teaching, and post-class analysis. This research was conducted in these three courses. Before this, all students in these three majors had only studied theoretical courses on teaching methods, namely, Mathematics Education Theory, Science Education Theory, and Computer Education Theory, without receiving practical training on teaching skills.

B. Research Procedures

The specific research ideas are divided into five stages: pre-test, experiment, post-test, interview, and summary.

During the study, 215 students were divided into two groups based on administrative classes, including the experimental group, which used ChatGPT for assisted teaching, and the control group, which did not use ChatGPT for assistance. The experimental group consists of one class from the mathematics education major, one class from the science education major, and one class from the computer education major. Correspondingly, the control group also consists of one class from the mathematics education major, one from the science education major, and one from the computer education major.

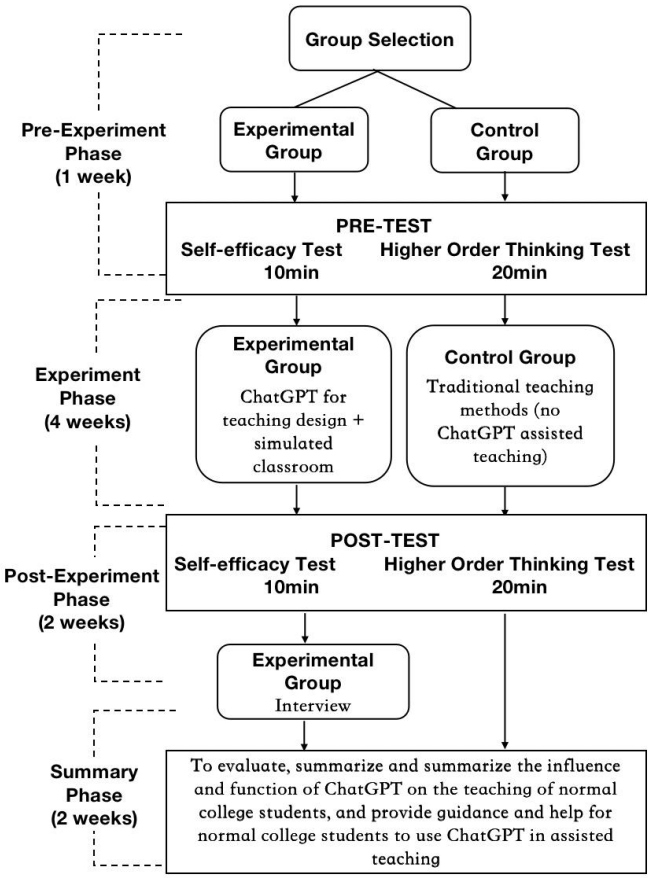

Fig.1. Experimental research idea diagram.

In the pre-experiment phase, the teacher self-efficacy scale and the higher order thinking scale were used to pre-test the two groups of students, respectively, and the data of teacher self-efficacy and higher order thinking of the two groups of students were collected before the experiment.

The experiment lasts for four weeks. During the experiment phase, both the experimental and control groups underwent weekly pre-service teacher training focused on developing pedagogical skills, with each training session lasting approximately 2.5 hours. Traditional teaching skills training can be divided into three stages: preparation, implementation, and evaluation. The training mentors would assign a specific topic to pre-service teachers each week, requiring them to design teaching plans and complete lesson preparations during the preparation stage. Pre-service teachers would conduct microteaching in groups and execute the planned teaching programs during the implementation stage. In the evaluation stage, pre-service teachers would summarize and analyze the teaching training under the guidance of the training mentors. Microteaching is currently widely used as a teaching method in pre-service teacher education. Pre-service teachers would be required to conduct short teaching demonstrations in small classroom environments. The core of microteaching was to have pre-service teachers design and implement lesson plans to cultivate their ability to apply theoretical knowledge and practical teaching skills [30-31]. In the four weeks of the experiment, the topics for teaching skills training for both the experimental and control groups were “Quadratic Equations,” “Rhombus,” “Probability of Events,” and “Sector Graphs.” These four topics cover the three major themes of middle school mathematics: “Numbers and Algebra,” “Geometry and Shapes,” and “Statistics and Probability.”

Nevertheless, a distinction arose: students in the control group received pre-service teacher training through conventional methodologies. In contrast, in addition to traditional methods, the experimental group students were required to use ChatGPT as a supportive tool during the preparation stage. In other words, this study focused on the supporting role of ChatGPT in pre-service teacher training during the preparation stage, and the implementation and evaluation stages of teaching will remain the same for both groups. In the traditional training, pre-service teachers may use regular search engines, teaching reference books, and other tools to prepare for their lessons. In the ChatGPT-assisted training method, pre-service teachers would use ChatGPT in addition to these tools for lesson planning. Specifically, the researchers designed four tasks utilizing ChatGPT as the learning content for the experimental group. The four tasks primarily focused on instructional design, with simulated classroom activities as supplementary exercises. Instructional design involves the planning, designing, hypothesizing, and arranging of classroom instruction, and it

This article has been accepted for publication in IEEE Transactions on Learning Technologies. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TLT.2024.3369690

is an essential component of lesson preparation for teachers. ChatGPT could assist pre-service teachers in generating and optimizing instructional designs based on instructions, thereby improving their lesson preparation efficiency and teaching abilities.

Additionally, simulated classroom activities can help pre-service teachers identify deficiencies in their teaching and make improvements, effectively enhancing their teaching skills. ChatGPT provides a new format for the simulated classroom, where pre-service teachers play the role of the teacher, and ChatGPT assumes the role of the student, engaging in dialogue and simulating classroom activities. It increases pre-service teachers' preconceptions of actual classroom situations and improves their teaching abilities. Therefore, this study selected instructional design and simulated classroom activities as the main content and approach for the experimental group using ChatGPT for instructional assistance.

Meanwhile, one week before the start of the experiment, the researchers introduced the experimental group to the usage of ChatGPT as an instructional aid, as well as the activities related to instructional design and simulated classroom exercises. This ensures that the experimental group can smoothly complete their lesson preparation with the assistance of ChatGPT during the experiment. Throughout a four-week experimental implementation, the investigators engaged with the experimental group on a weekly basis. Assigned tasks were distributed to participants, anticipating that students within the experimental group executed these four prescribed tasks sequentially, adhering to the instructions provided by the researchers.

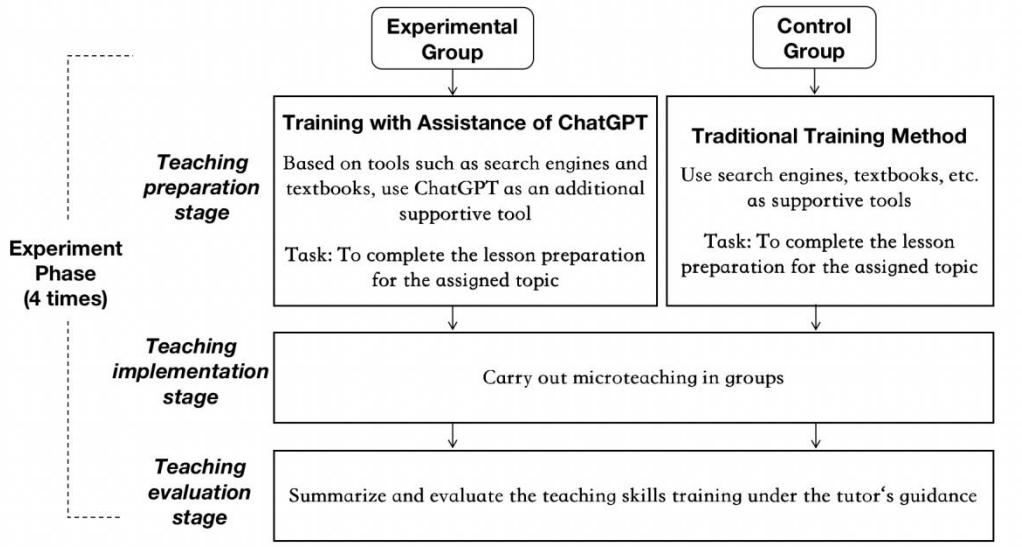

Fig.2. Comparison of training methods in the experiment phase.

Task 1 in the first week was to optimize a given instructional design using ChatGPT. The researchers provided a designated instructional design as a reference. The students in the experimental group needed to supplement, enhance, and improve the deficiencies of this instructional design with the assistance of ChatGPT. Task 2 in the second week was to generate and optimize an instructional design for a specific topic using ChatGPT. The researchers only provided the topic without an instructional design for reference. The students in the experimental group were expected to independently use ChatGPT to generate a complete instructional design, identify its shortcomings, and optimize it with the help of ChatGPT. Task 3 in the third week was to simulate classroom exercises for the interactive section of a designated instructional design using ChatGPT. The researchers provided an instructional design for a specific topic, including interactive sections between teachers and students. The students in the experimental group should use this instructional design as a reference to simulate a classroom exercise, with ChatGPT acting as the student and themselves playing the role of the teacher. Task 4 in the fourth week was to generate and optimize an instructional design using ChatGPT and select suitable sections for simulated classroom exercises. The researchers only provided the topic without any reference content. It means that the students needed to combine Task 2 and Task 3, using ChatGPT to generate and optimize a complete instructional design, and autonomously select appropriate content sections for simulated classroom exercises.

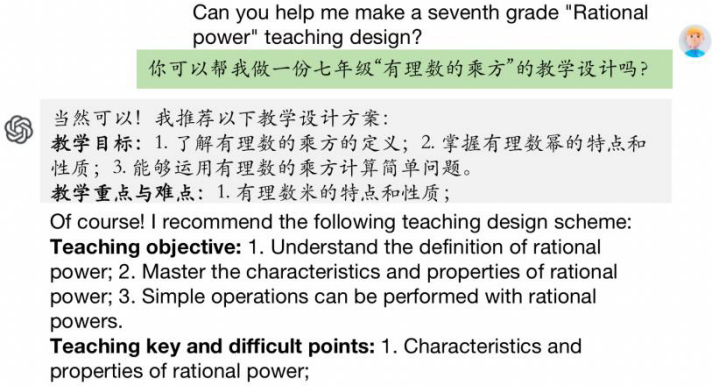

Fig. 3. The example of instructional design using generative AI.

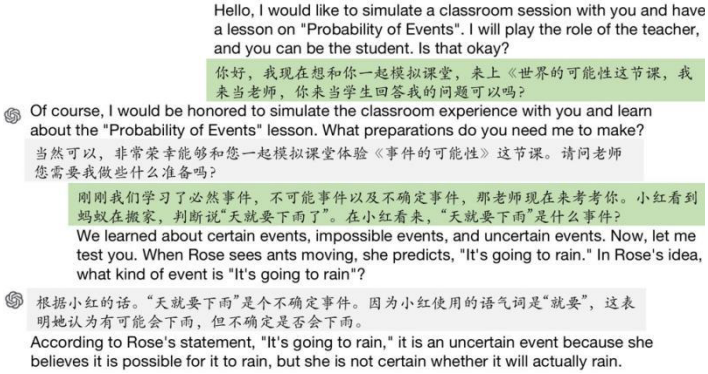

Fig.4. The example of simulated classroom using generative AI.

In summary, the students in the control group would conduct lesson preparation with the assistance of conventional search engines and reference books. The students in the experimental group would additionally use ChatGPT to assist them in the manner of completing the four tasks during the teaching preparation stage. The teaching implementation and evaluation stages remained the same for both the control and experimental groups in terms of content and methods of teacher training.

The post-experiment phase can be divided into three stages: post-test, interview, and summary. In the post-test stage, the teacher self-efficacy scale and the higher order thinking scale were used again to conduct the post-test on the two groups of students, respectively. The teacher self-efficacy and higher-order thinking data of the two groups of students after the experiment were collected and compared with the pre-test data, and the changes were analyzed.

In the interview stage, 25 students from the experimental group were selected and interviewed about their experiences, feelings, and problems that they encountered in the process of using ChatGPT for assisted teaching so as to deeply understand their views on ChatGPT assisted teaching and the role and impact of ChatGPT on pre-service teachers' teaching skills.

In the summary stage, the help and influence of ChatGPT on teaching were analyzed based on the coding of pre- and post-test data and interview content. Finally, the supportive role of ChatGPT in the practical teaching of pre-service teachers was summarized, and guidance was provided for the reasonable application of ChatGPT to improve the teaching skills of pre-service teachers.

C. Data Collection

1) **Higher order thinking**

This study makes an overall study of higher order thinking and related research [22-34] and organizes the relevant higher order thinking scale based on the Bloom educational goal classification framework. According to the definition of higher order thinking mentioned before, this study divides higher order thinking into four dimensions: critical thinking, creative thinking, meta-cognitive thinking, and problem-solving ability. A corresponding scale was developed. Subsequently, through discussions and revisions with numerous experts, the researchers added a dimension of computational thinking to higher order thinking, eventually resulting in the "Higher Order Thinking Scale".

According to the definition of higher order thinking mentioned above, the scale divides higher order thinking into the following five dimensions: critical thinking, creative thinking, meta-cognitive thinking, problem-solving ability, and computational thinking. The first part of the scale is used to investigate students' basic information, mainly involving the content of students' demographic variables and the content closely related to the development of higher order thinking. The second part of the scale is 66 questions, given in tables. Questions 1-16 mainly evaluate the development level of students' critical thinking. The questions involve critical thinking aspects that students may encounter while learning mathematics, specifically abilities such as questioning, deconstruction, and dialectics. Questions 17-25 mainly evaluate the development level of students' creative thinking, involving the ability of students to expand, construct, and diverge in the process of learning. Questions 26-34 mainly evaluate students' meta-cognitive thinking development level, involving students' ability to plan, adjust, and monitor the learning process. Questions 35-44 are used to evaluate students' problem-solving ability. Questions 45-66 mainly assess the development level of students' computational thinking, involving students' abilities such as transferring, questioning, and optimization in the learning process.

The scale was tested and yielded an overall Cronbach's Alpha coefficient of 0.931 and a KMO value of 0.871. Additionally, the principal component analysis extract five common factors and the cumulative variance explanation rate of the model is 72.998%, which indicates that this measurement scale has good structural validity, indicating that the division and expression of higher order thinking have a certain scientific nature. The reliability of each dimension of the measurement scale is higher than 0.7, indicating that the measurement scale has good reliability and the construction of higher order thinking has good credibility. Moreover, Amos22.0 was used to conduct confirmatory factor analysis on the higher order thinking scale, and all fitting indexes reached or approached 0.9 (χ^2 / df = 2.208, RMSEA = 0.024, NFI = 0.958, CFI = 0.929, IFI =0.934, TLI=0.939). It shows that the scale has good construction validity. Through validity and reliability tests, the scale can be used as a measurement tool for further research.

2) **Teacher self-efficacy**

The teacher self-efficacy scale was first created by the author in a research project for online teachers in Nevada [35]. Generative AI, a relatively new information technology product, requires teachers to possess a certain level of technological proficiency to use it for instructional purposes. It challenges teachers to integrate information technology with teaching and use it to facilitate knowledge dissemination, which aligns with the principles of TPACK. Therefore, This paper selected the adapted version of Archambault and Crippen [36]. The scale uses TPACK as a framework to assess teachers' self-efficacy in teaching and includes 24 questions assessed using a 5-point Likert scale (1 = very inconsistent, 5 = very consistent). The 24 questions are composed of seven areas: educational knowledge, technical knowledge, content knowledge, technical content knowledge, teaching content knowledge, technical teaching knowledge, and technical teaching content knowledge. There are 3~4 questions in each area. The scale assessed teachers' self-efficacy by allowing them to rate their abilities in seven domains. The scale had been tested, yielding an overall Cronbach's Alpha coefficient of 0.872 and a KMO value of 0.804. The tool had demonstrated sufficient reliability and validity to be used for

further research.

3) **Interview data**

At the interview stage, 25 students from the experimental group were selected for semi-structured interviews about their feelings and evaluation of ChatGPT in the process of assisted teaching and their thoughts on further promotion of ChatGPT. The interview questions were as follows:

a) How do you feel when using ChatGPT? For example, you can talk about the effectiveness of the feedback given by ChatGPT, the problems encountered, and so on.
b) Do you feel that using ChatGPT has impacted you?
c) Do you find ChatGPT helpful in your teaching? If so, which of its features (on which issues) are more helpful?
d) As a teacher, where else do you think ChatGPT can be used to assist in the teaching process?
e) Do you intend to employ ChatGPT for teaching assistance in future educational endeavors?
f) If ChatGPT is integrated into a broader range of pre-service teacher teaching, what effects do you think it will produce, or what problems do you think it will face?

D. Data analysis

1) **Quantitative analysis**

SPSS 23 was used to conduct a one-way analysis of variance for quantitative data. This study employed ANOVA with repeated measures to examine the interactive effects of time (pre-test, post-test) and group (experimental group, control group) on the teacher self-efficacy and higher order thinking development of 215 pre-service teachers. The objective was to scrutinize the augmentative function of generative AI in the instructional skills development of pre-service educators.

2) **Qualitative analysis**

The qualitative data of the experimental group was analyzed using NVivo 12. First, the data obtained from the interviews were transferred to the program. Next, the method of content analysis was used to merge similar codes into the same sub-theme. This process was carried out by two researchers who coded the data in a back-to-back manner. Finally, the qualitative results were presented in a visual report.

V. RESULTS

A. Quantitative analysis

This study conducted a repeated-measures analysis of variance to examine whether there were significant differences in teacher self-efficacy and higher order thinking average scores between the experimental and control groups in the pre-test and post-test. The results revealed that after a four-week experiment, pre-service teachers' self-efficacy in both groups had improved considerably ($p<0.05$). In terms of higher order thinking, pre-service teachers in the experimental group have significantly enhanced their higher order thinking [$F(1,108)=21.144$, $p=0.000<0.05$], while there is no significant improvement in the control group [$F(1,105)=2.988$, $p=0.085>0.05$]. Comparing the two groups, it was found that there were significant differences in teacher self-efficacy [$F(1,213)=5.282$, $p=0.023<0.05$] and higher order thinking [$F(1,213)=4.592$, $p=0.033<0.05$] among the two groups of pre-service teachers, under the interaction effect of time and group. This suggests that with the assistance of ChatGPT, pre-service teachers demonstrated higher levels of teacher self-efficacy and higher order thinking compared to those who received traditional teaching skill training.

TABLE I
SELF-EFFICACY AND HIGHER ORDER THINKING SCORE (M±SD) OF PRE-SERVICE TEACHERS

	Experimental Group		Control Group		F	p	η^2
	Pre-Test	Post-Test	Pre-Test	Post-Test			
self-efficacy	3.87±0.40	3.97±0.43	3.81±0.41	3.73±0.46	5.282*	0.023	0.024
higher order thinking	3.64±0.25	3.80±0.27	3.67±0.19	3.73±0.27	4.592*	0.033	0.021
critical thinking	3.72±0.35	3.79±0.32	3.68±0.28	3.86±0.40	3.061	0.082	0.014
creative thinking	3.55±0.43	3.67±0.49	3.58±0.41	3.80±0.53	1.189	0.277	0.006
meta-cognitive thinking	3.80±0.41	3.91±0.48	3.90±0.45	3.83±0.38	4.054*	0.045	0.019
problem-solving ability	3.45±0.33	3.65±0.32	3.47±0.25	3.44±0.35	8.704*	0.004	0.039
computational thinking	3.84±0.36	4.04±0.50	3.90±0.30	4.00±0.46	4.913*	0.028	0.023

Note. *$p < .05$

Specifically, this study also conducted a repeated-measures analysis of variance on the five sub-dimensions of higher order thinking. The results revealed that after a four-week teaching experiment, under the interaction effect of time and group, there were significant differences among pre-service teachers in three dimensions: meta-cognitive thinking [$F(1,213)=4.054$, $p=0.045<0.05$], problem-solving ability [$F(1,213)=8.704$, $p=0.004<0.05$], and computational thinking [$F(1,213)=4.913$, $p=0.028<0.05$]. However, there were no significant differences in the dimensions of critical thinking [$F(1,213)=3.061$, $p=0.082>0.05$] and creative thinking [$F(1,213)=1.189$, $p=0.277>0.05$].

B. Qualitative analysis

The qualitative results of this study showed that generative AI has both advantages and disadvantages when used to assist in the practical training of teaching skills. When asked, "Will you continue to use ChatGPT and other generative AI for assisted teaching in the future?" 20 of the 25 interviewees said that they would continue to use ChatGPT and other generative AI, while five indicated that they might not use them to

support teaching in the future. Therefore, the qualitative data obtained from the interviews in the experimental group were content analyzed, focusing on pre-service teachers' views on using ChatGPT for teaching assistance. Two researchers independently coded the interview content following the "back-to-back" principle, with a coding consistency of 0.876. Subsequently, the two researchers analyzed and discussed the coding results, reaching a consensus and generating Figure 5. From the figure, using ChatGPT for teaching assistance brought various experiences for the pre-service teachers, including collecting information, providing content references, inconvenience, and so on. Providing content references, simulating students' responses, low accuracy, and unexpected responses were high-frequency keywords. Therefore, the following sections will provide a discussion of these four aspects.

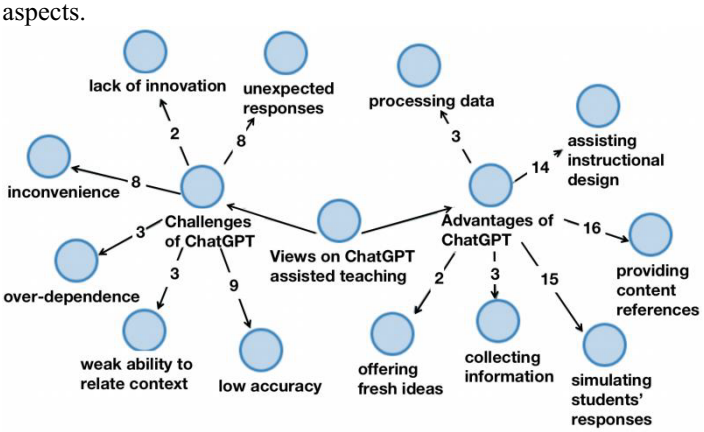

Fig.5. Code of pre-service teachers' views on ChatGPT and other generative AI assisted teaching.

1) **Give content reference**

Of the 25 pre-service teachers interviewed, 16 mentioned ChatGPT and other generative AI as an aid to teaching by providing content references.

"I intend to persist in utilizing ChatGPT due to its proficiency in addressing issues. Entrusting ChatGPT with problem-solving yields a structured output as a foundational framework, offering me the flexibility to make modifications and providing a solid reference point." PT-2

"As a new teacher, you have to write a lot of lesson plans, materials, etc. ChatGPT is very helpful for us to develop a first draft, some basic framework, and then you can revise from the basis." PT-4

2) **Simulating students' response**

Of the 25 pre-service teachers interviewed, 15 agreed that using generative AI to simulate real teacher-student interaction by presupposing students' possible responses was a very effective way to assist.

"Simulating teacher-student interaction is very good. On the one hand, it can give students' answers that I can't imagine. On the other hand, ChatGPT may give wrong answers, and students will certainly make many mistakes in the actual teaching process, so I think it is more effective to simulate a real classroom in this way." PT-11

"I will use ChatGPT to simulate the dialogue between teachers and students because we cannot simulate good responses from students only from our own perspective as teachers, and it can provide a lot of reference in simulating students' perspectives." PT-22

3) **Low accuracy**

During the interview, nine interviewees mentioned the low accuracy of the answers given by ChatGPT and other generative AI.

"ChatGPT has a low accuracy rate in its responses. For example, if you ask whether differentiability implies continuity, it might answer that it does not necessarily imply continuity. It makes obvious mistakes like this as well." PT-7

"ChatGPT's responses are not very satisfactory. I have to ask detailed questions, and it doesn't seem as intelligent. Additionally, the answers it provides are not necessarily correct. This is particularly evident in programming code generation, where some of the provided code proves to be non-functional." PT-15

4) **Unexpected responses**

Eight respondents pointed out that when using ChatGPT for instructional design assistance, ChatGPT often failed to provide satisfactory answers and generated responses that did not meet their expectations, sometimes even giving irrelevant answers.

"We use ChatGPT for instructional design, but I think it can only be used as a supporting tool. For example, in the introductory phase of the teaching process in instructional design, the ideas provided by ChatGPT are not particularly innovative, so we still need to design some more innovative situations that better fit the characteristics of the current era based on its foundation." PT-22

"Especially the content of the teaching process, for example, when I did the instructional design of "Possibility of events" last time, ChatGPT could not recognize that the balls of the same color were the same on the problem of grasping small balls, and it took many times to make the correct answer, but if I wrote it myself, it would soon be written out." PT-4

VI. Conclusion

This study combined AIGC with teaching to explore how generative AI can assist pre-service teachers in the practical teaching skills training and then produced a practical training method for pre-service teachers with the assistance of ChatGPT as an example of generative AI. This study also explores whether generative AI-assisted teaching training has different effects on pre-service teachers' higher order thinking and self-efficacy compared with traditional training methods. In addition, through interviews, this study deeply explored the views of pre-service teachers using generative AI to assist teaching and their attitudes toward integrating generative AI into teaching.

For research question 1, in terms of how generative AI assisted pre-service teachers' practical teaching training, this study took instructional design as the main line and integrated ChatGPT into the instructional design of pre-service teachers. First, let pre-service teachers use ChatGPT to generate and optimize instructional design. Second, simulated classroom exercises are carried out for teacher-student interaction in instructional design. This is consistent with the findings of Jeon et al., in which teachers were free to decide how to use ChatGPT as a teaching aid, and all teachers used ChatGPT as an interlocutor to communicate in real time, simulating a real classroom [21].

For research question 2, after a four-week educational experiment, this study showed that the experimental group, which used ChatGPT to assist teacher skills training, scored significantly higher on higher order thinking and teacher

self-efficacy than the control group, which used traditional teaching practice methods. Specifically, this study divides higher order thinking into five sub-dimensions: critical thinking, creative thinking, meta-cognitive thinking, problem-solving ability and computational thinking. The research results show that after the experiment, pre-service teachers in the experimental group are considerably higher than those in the control group in three sub-dimensions of meta-cognitive thinking, problem-solving ability, and computational thinking, while in two sub-dimensions of critical thinking, and creative thinking, there is no significant difference between the two groups.

For research question 3, this study obtained the views of pre-service teachers on using generative AI in teaching through interviews. During the interview, among the 25 pre-service teachers who were interviewed in the experimental group, 20 teachers indicated that they would continue to use ChatGPT and other generative AI in their future teaching, while 5 teachers indicated that they might not use them in the future. Pre-service teachers with an optimistic attitude towards generative AI generally believe that it can greatly help in instructional design, such as textbook analysis and classroom introduction design. Pre-service teachers with negative attitudes towards generative AI generally believe that the feedback given by generative AI often does not align with their expectations. Additionally, ordinary search engines, such as Baidu, can also meet their needs, so there is no need to use generative AI. Meanwhile, the interviewed pre-service teachers gave polarized feedback on using AIGC in simulated classrooms. Some believed that AIGC could well simulate the generation of students and give teachers unexpected answers and feedback from students' perspectives, thus improving teachers' capability to control the classroom. Others argued that AIGC could not really think like a student, and its communication could not simulate a real classroom.

VII. Discussion

The researchers discussed the reasons for the above conclusions.

First, the findings are supported by many previous studies, which show that modern technology-assisted teaching improves students' academic performance and their higher order thinking skills much more than traditional classrooms [37]. For this study, the modern technology is generative AI, which can provide rich and immediate learning feedback to pre-service teachers by impersonating real students in a simulated classroom. Many scholars have conducted similar research regarding this functionality. Schneider et al. found that real-time learning feedback can improve learners' self-awareness and self-assessment [38]. Chat bot software with real-time interactive performance can give students relatively immediate and detailed feedback, thereby stimulating students' cognition and positively influencing students' learning performance [39]. Jamani and Angeli found that using robot technology in college courses can promote pre-service teachers' self-efficacy [40]. These are all consistent with the results of this study.

Second, the sub-dimensions of higher order thinking are discussed. In terms of creative thinking, the results are supported by Chao et al. [41]. They categorized the richness of media functions into four dimensions: availability of immediate feedback, use of multiple cues, linguistic diversity, and personal attention, and then investigated the potential impact of these functions on learners' creativity. The results showed that among the four dimensions, only the availability of immediate feedback does not have a remarkable correlation with the cognitive performance of creativity. Similarly, as a natural language processing model, the most important functions of generative AI are real-time interaction and timely feedback. Therefore, the findings of this study, which shows that using generative AI to assist teaching skills training cannot improve the creative thinking of pre-service teachers, are consistent with the conclusion of Chao et al. However, Asiri et al. found that real-time interactive feedback can improve students' critical thinking, which is contrary to the conclusion of critical thinking in this study [42]. To explain this, the characteristics of generative AI are analyzed in more depth. Generative AI is based on pre-trained models, which rely on existing data and patterns. It may lack creative problem-solving solutions and critical analysis abilities. Moreover, the semantic understanding capability of generative AI is limited, rendering it challenging to comprehend and interpret intricate semantics or implicit information accurately. Therefore, the results of this study show that generative AI does not contribute to the development of pre-service teachers' critical thinking and creative thinking.

In terms of meta-cognitive thinking, generative AI can assist pre-service teachers in reflecting on their teaching processes and strategies by providing immediate feedback and assessment. This prompts pre-service teachers to clarify their teaching objectives and progress and make timely adjustments to their strategies, thereby enhancing their meta-cognitive thinking. In terms of problem-solving ability, generative AI can help pre-service teachers solve teaching-related problems by providing explanations, answers, and guidance. Through this process, pre-service teachers can acquire problem-solving methods and approaches, gradually enhancing their problem-solving ability. In terms of computational thinking, AIGC may involve complex calculations or logical reasoning. Therefore, the interaction between pre-service teachers and generative AI can help cultivate and improve their computational thinking, enabling them to decompose problems, analyze relationships, and make reasoned inferences.

Third, to explain the interview results, the researchers analyzed the correlation between the attitudes of the pre-service teachers towards generative AI and their performance in completing the four tasks during the experimental phase. It is observed that pre-service teachers with a positive attitude toward generative AI tend to engage in "probing" interactions with it, which leads to more extensive and comprehensive question-answer sessions. Their questioning skills were more advanced, with well-structured and explicit expressions. This enables generative AI to deliver precise and accurate responses consistently. On the other hand, pre-service teachers with a negative attitude towards generative AI typically exhibited lower levels of questioning skills during their interactions. Their questions lacked logical coherence and often contained incomplete sentences or grammatical errors, resulting in irrelevant answers. Furthermore, their questions were usually too broad and failed to highlight key points, resulting in insufficiently targeted answers. However, it is worth noting that even though five pre-service teachers expressed the opinion that generative AI was "inconvenient," three of them still acknowledged its effectiveness in the simulated classroom. According to this,

the researchers consider that the conclusion of "inconvenience" is limited to situations where generative AI is used for instructional design with low-level questioning methods. Therefore, the researchers believe that during the teaching skills training, the training mentors can allow pre-service teachers to use generative AI as instructional assistance by providing them with comprehensive and detailed guidance. This would enable pre-service teachers to engage in smooth and effective interaction with generative AI, thereby improving the quality of training and enhancing their teacher self-efficacy and higher order thinking.

Despite the findings of the current research on generative AI assisting teaching, there are three limitations in this study that suggest the need for further empirical research. First, the small sample size in our study may lead to underrepresentation. We exclusively recruited pre-service teachers from a well-developed city in China, resulting in a modest number of participants. In this respect, the future research should broaden the scope of the study and include a larger sample size to enhance the persuasiveness of the findings. Second, due to time constraints, the experiment was only conducted over a four-week period, which is not sufficient to provide a comprehensive understanding of long-term effects of using generative AI for pre-service teachers. It is possible that different results may have been obtained if the duration of the experiment was extended. Third, only ChatGPT was discussed in this study as an example of generative AI, without considering other models. It is recommended that future research could investigate the use of various generative AI models as research tools. Conducting comparative studies among multiple models could also offer valuable insights into their effectiveness for pre-service teachers.

REFERENCES

[1] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. MatenaA, et al., "Exploring the limits of transfer learning with a unified text-to-text transformer," J. Mach.Learn.Res., vol. 21, no. 1, pp. 140, Jan. 2020, doi:10.48550/arXiv.1910.10683.

[2] H. Y. Du, Z. H. Li, D. Niyato, J. W. Kang, Z. H. Xiong, X. M. Shen, et al., "Enabling AI-generated content (AIGC) services in wireless edge networks," *arXiv preprint*, 2023, doi:10.48550/arXiv.2301.03220.

[3] Y. Ming, N. Hu, C. Fan, F. Feng, J. Zhou and H. Yu, "Visuals to Text: A Comprehensive Review on Automatic Image Captioning," in *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 8, pp. 1339-1365, August 2022, doi: 10.1109/JAS.2022.105734.

[4] J. Qadir, A. E. Taha, K. L. Yau, J. Ponciano, S. Hussain, et al., "Leveraging the Force of Formative Assessment & Feedback for Effective Engineering Education," in *2020 ASEE Annual Conference & Exposition*, Online Event, Jun. 2020, pp.21-24, doi:10.18260/1-2--34923.

[5] *Gpt-4 technical report*, 2023, [online]. Available: https://cdn.openai.com/papers/gpt-4.pdf.

[6] J. Qadir, "Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for Education," *2023 IEEE Global Engineering Education Conference (EDUCON),* Kuwait, Kuwait, 2023, pp. 1-9, doi:10.1109/EDUCON54358.2023.10125121.

[7] M. M. Rahman and Y. Watanobe, "ChatGPT for Education and Research: Opportunities, Threats, and Strategies," *Applied Sciences*, vol. 13, no. 9, pp. 1-21, 2023, doi:10.20944/preprints202303.0473.v1.

[8] J.Lee, H. Choi, "What affects learner's higher-order thinking in technology-enhanced learning environments? The effects of learner factors," Computers & Education, vol. 115, pp. 143-152, Dec. 2017, doi: 10.1016/j.compedu.2017.06.015.

[9] M. Barak, D. C. Ben and U. Zoller, "Purposely Teaching for the Promotion of Higher-order Thinking Skills: A Case of Critical Thinking," *Research in ence Education*, vol.37, no.4, pp.353-369, 2007, doi: 10.1007/s11165-006-9029-2.

[10] E. Driana, S. Pascasarjana, U. Muhammadiyah, et al., "Teachers' Understanding and Practices in Assessing Higher Order Thinking Skills at Primary Schools," *Acitya*, vol. 1, no. 2, pp. 110-118, 2019, doi: 10.30650/ajte.v1i2.233.

[11] M. T. Moran and A. W. HOY, "The differential antecedents of self-efficacy beliefs of novice and experienced teachers," *Teaching and Teacher Education*, vol. 23, no.6, pp. 944-956, Aug. 2007, doi: 10.1016/j.tate.2006.05.003.

[12] B. Albert, "Self-efficacy: toward a unifying theory of behavioral change," *Advances in Behaviour Research and Therapy*, vol. 1, no. 4, pp. 139-161, 1978, doi:10.1016/0146-6402(78)90002-4.

[13] R. Lazarides and U. Schiefele, "The relative strength of relations between different facets of teacher motivation and core dimensions of teaching quality in mathematics - A multilevel analysis," *Learning and Instruction*, vol. 76, pp. 1-16, Dec. 2021, doi:10.1016/j.learninstruc.2021.101489.

[14] J. B. Kupper, "The Microteaching Experience: Student Perspectives," *Education*, vol. 121, no. 4, pp. 830-835, Jun. 2001.

[15] M. L. Fernandez and M. A. Robinson, "Prospective Teachers' Perspectives on Microteaching Lesson Study," *Education*, vol. 127, no. 2, pp. 203-215, Dec. 2006.

[16] E. Kpanja, "A Study of the Effects of Video Tape Recording in Microteaching Training," *British Journal of Educational Technology*, vol. 32, no. 4, pp. 483-486, Dec. 2002, doi:10.1111/1467-8535.00215.

[17] M. F. Shaughnessy, "An Interview With Anita Woolfolk: The Educational Psychology of Teacher Efficacy," *Educational Psychology Review,* vol. 16, no. 2, pp. 153-176, Jun. 2004, doi:10.1023/B:EDPR.0000026711.15152.1f.

[18] T. Wu et al., "A Brief Overview of ChatGPT: The History, Status Quo and Potential Future Development," in *IEEE/CAA Journal of Automatica Sinica*, vol. 10, no. 5, pp. 1122-1136, May 2023, doi: 10.1109/JAS.2023.123618.

[19] P. Liang, R. Bommassani, T. Lee, D. Tsipras, D. Soylu, M. Yasunaga, et al., "Holistic evaluation of language models," *Computer Science*, vol. 1525, pp. 140-146, Nov. 2022, doi:10.48550/arXiv.2211.09110.

[20] C. W. Qin, A. Zhang, Z. S. Zhang, J. A. Chen, M. Yasunaga and D. Y. Yang, "Is ChatGPT a general-purpose natural language processing task solver?", *ArXiv*, Feb. 2023, doi:10.48550/arXiv.2302.06476.

[21] J. Jeon and S. Lee, "Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT," *Education and Information Technologies*, pp. 1-20, May 2023, doi:10.1007/s10639-023-11834-1.

[22] T. Humphry and A. L. Fuller, "Potential ChatGPT Use in Undergraduate Chemistry Laboratories," *Journal of Chemical Education*, vol. 100, no. 4, pp. 1434-1436, Mar. 2023, doi:10.1021/acs.jchemed.3c00006.

[23] A. Shoufan, "Exploring Students' Perceptions of ChatGPT: Thematic Analysis and Follow-Up Survey," in *IEEE Access*, vol. 11, pp. 38805-38818, 2023, doi: 10.1109/ACCESS.2023.3268224.

[24] M. Farrokhnia, S. K. Banihashem, O. Noroozi, et al., "A SWOT analysis of ChatGPT: Implications for educational practice and research," *Innovations in Education and Teaching International*, pp. 1-15, Mar. 2023, doi:10.1080/14703297.2023.2195846.

[25] Z. Y. Avci and R. Osman, "Incorporation of a software engineering evaluation method in an educational software course in Turkey," *Educational Studies*, vol. 47, no. 40, pp. 383-402, Dec. 2019, doi:10.1080/03055698.2019.1702509.

[26] Z. Arsal, "Microteaching and pre-service teachers' sense of self-efficacy in teaching," *European Journal of Teacher Education*, vol. 37, no. 4, pp. 453-464, Oct. 2014, doi:10.1080/02619768.2014.912627.

[27] A. G. Mergler and D. Tangen, "Using microteaching to enhance teacher efficacy in pre-service teachers," T*eaching Education*, vol. 21, no. 2, pp. 199-210, May 2010, doi:10.1080/10476210902998466.

[28] S. A. GüNBATAR , B. Demirdogen and A. T. ÇELIKKIRAN, "Are they efficacious? Exploring pre-service teachers' teaching efficacy beliefs during practicum," *The Asia-Pacific Education Researcher*, vol. 21, no. 1, pp. 203-2013, Jan. 2012.

[29] D. Schina, C. V. Bautista, A. B. Riera, M. Usart and V. E. González, "An associational study: preschool teachers' acceptance and self-efficacy towards Educational Robotics in a pre-service teacher training program," *Int J Educ Technol High Educ*, vol. 18, no. 1, pp. 28, May 2021, doi:10.1186/s41239-021-00264-z.

[30] C. Yan, C. He, "Pair microteaching:an unrealistic pedagogy in preservice methodology courses?" *Journal of Education for Teaching*, vol. 43, no. 2, pp. 206-218, Feb. 2017, doi: 10.1080/02607476.2017.1286783.

[31] Z. Arsal, "Microteaching and pre-service teachers' sense of self-efficacy in teaching", *European Journal of Teacher Education*, vol. 37, no. 4, pp. 453-464, May 2014, doi: 10.1080/02619768.2014.912627.

[32] M. Priyaadharshini and B. V. Sundaram, "Evaluation of higher-order thinking skills using learning style in an undergraduate engineering in flipped classroom," *Computer Applications in Engineering Education*, vol. 26, no. 6, pp. 2237-2254, Jul. 2018, doi:10.1002/cae.22035.

[33] D. Fung , "Promoting critical thinking through effective group work: A teaching intervention for Hong Kong primary school students," *International Journal of Educational Research*, vol. 66, pp. 45-62, 2014, doi:10.1016/j.ijer.2014.02.002.

[34] G. K. Montgomery, "Effects of performance evaluation and anxiety on cardiac response in anticipation of difficult problem solving,"

Psychophysiology, vol. 14, no. 3, pp. 251-257, May 1977, doi:10.1111/j.1469-8986.1977.tb01170.x.

[35] L. M. Archambault and K. J. Crippen, *The Preparation and Perspective of Online K-12 Teachers in Nevada*. Publ. 2006.

[36] L. M. Archambault and K. J. Crippen, "Examining TPACK Among K-12 Online Distance Educators in the United States," *Contemporary Issues in Technology and Teacher Education*, vol. 9, no. 1, pp. 71-88, Mar. 2009.

[37] D. Liu and H. Zhang, "Improving Students' Higher Order Thinking Skills and Achievement Using WeChat based Flipped Classroom in Higher Education," *Education and Information Technologies*, vol. 27, no. 5, pp. 7281-7302, Feb. 2022, doi:10.1007/s10639-022-10922-y.

[38] J. Schneider, D. Börner, P. van Rosmalen and M. Specht, "Can You Help Me with My Pitch? Studying a Tool for Real-Time Automated Feedback," in *IEEE Transactions on Learning Technologies*, vol. 9, no. 4, pp. 318-327, 1 Oct.-Dec. 2016, doi: 10.1109/TLT.2016.2627043.

[39] M. Fidan and N. Gencel, " Supporting the Instructional Videos With Chatbot and Peer Feedback Mechanisms in Online Learning: The Effects on Learning Performance and Intrinsic Motivation," *Journal of Educational Computing Research*, vol. 60, no. 7, pp. 1716-1741, Feb. 2022, doi:.org/10.1177/07356331221077901.

[40] K. J. Jamani and C. Angeli, "Effect of Robotics on Elementary Preservice Teachers' Self-Efficacy, Science Learning, and Computational Thinking," *Journal of Science Education and Technology*, vol. 26, no. 2, pp. 175-192, Apr. 2017, doi:10.1007/s10956-016-9663-z.

[41] S. H. Chao, JIANG J. Jiang, C. H. Hsu, Y. T. Chiang, E. Ng and W. T. Fang, "Technology-Enhanced Learning for Graduate Students: Exploring the Correlation of Media Richness and Creativity of Computer-Mediated Communication and Face-to-Face Communication," *Applied Sciences*, vol. 10, no. 5, pp. 1602, Feb. 2020, doi:10.3390/app10051602.

[42] Y. A. Asiri, D. E. Millard and M. J. Weal, "Assessing the Impact of Engagement and Real-Time Feedback in a Mobile Behavior Change Intervention for Supporting Critical Thinking in Engineering Research Projects," in IEEE Transactions on Learning Technologies, vol. 14, no. 4, pp. 445-459, 1 Aug. 2021, doi: 10.1109/TLT.2021.3104817.

Jijian Lu received the Ph.D. degree in educational science from the Beijing Normal University, China, in 2019. He is currently an Associate Professor with the Jing Hengyi School of Education, Hangzhou Normal University, Hangzhou, Zhejiang, China. He is also a Research Member in the Chinese Education Modernization Research Institute, Hangzhou Normal University, Hangzhou, Zhejiang, China. His main research interests include centered around studying how computers affect educational behavior in various areas, such as teacher education, curriculum and teaching practice, etc.

Ruxin Zheng is currently a master student majoring in mathematics teaching at Jing Hengyi School of Education, Hangzhou Normal University, Hangzhou, Zhejiang, China. Her main research interests centered around studying how computers affect educational behavior in various areas, such as teacher education, curriculum and teaching practice, etc.

Zikun Gong received the Ph.D. degree in educational science from the Southwest University, China, in 2006. He is currently an Professor with the Jing Hengyi School of Education, Hangzhou Normal University, Hangzhou, Zhejiang, China. He is also a Research Member in the Chinese Education Modernization Research Institute, Hangzhou Normal University, Hangzhou, Zhejiang, China. His main research interests include centered around studying how computers affect educational psychology and behavior in various areas, such as teacher education, curriculum and teaching practice, etc.

Huifen Xu is currently a doctoral student majoring in curriculum and teaching theory at Jing Hengyi School of Education, Hangzhou Normal University, Hangzhou, Zhejiang, China. Her main research interests centered around studying how computers affect psychology and behavior in various areas, such as teacher education, curriculum and teaching practice, etc.