Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-024-02751-w OPEN

What factors will affect the effectiveness of using ChatGPT to solve programming problems? A quasi-experimental study

Yuhui Jing[1], Haoming Wang[2], Xiaojiao Chen[1] & Chengliang Wang[2]✉

The emergence of ChatGPT has sparked new expectations for AI-empowered educational transformation. However, it remains unknown which factors affect its effectiveness in empowering learners to solve programming problems. Therefore, this study employed a quasi-experimental research design and used Python graphing in programming education as an example to investigate the factors influencing the effectiveness of learners in applying ChatGPT to problem-solving. Findings: AI literacy significantly influences learners' effectiveness in using ChatGPT to solve problems, with AI awareness and AI usage being key factors. The knowledge base of programming language significantly affects learners' effectiveness in applying ChatGPT to solve programming problems. Learners' cognitive level of ChatGPT significantly influences their effectiveness in applying ChatGPT to problem-solving, while usage intention does not have a significant impact. However, learners' intention to use ChatGPT significantly improves after application. Based on these findings, this study proposes that in the process of empowering education with Artificial Intelligence Generated Content (AIGC) products, the focus on learners should shift from cultivating their AI usage to AI literacy, laying the foundation for empowering learning with AIGC products. It is suggested to shift from mastering specific knowledge to graph-based rules as a method for empowering learning with AIGC products. Additionally, the focus should shift from enhancing learners' intention to use the technology to strengthen their technological awareness, thereby creating practical pathways for empowering learning with AIGC products.

[1] College of Educational Science and Technology, Zhejiang University of Technology, Zhejiang, China. [2] Department of Education Information Technology, Faculty of Education, East China Normal University, Shanghai, China. ✉email: 51264108027@stu.ecnu.edu.cn

Introduction

At the end of 2022, ChatGPT emerged, elevating the application of artificial intelligence (AI) technology in educational reform to a new peak in educational research and practice (Hwang and Chen, 2023; Lodge et al., 2023). ChatGPT, as the current pinnacle of Artificial Intelligence Generated Content (AIGC), even being termed as the 'singularity' in generative AI applications, is bound to compel the establishment of new methods in AIGC-enabled education due to its outstanding or disruptive features (Rahman and Watanobe, 2023). In this context, given ChatGPT's exceptional programming and precise answering capabilities (Luo et al., 2023), some scholars have noted that ChatGPT's functions pose a threat to some programmers (Ouh et al., 2023; Jalil et al., 2023). Others have suggested that with the advent of ChatGPT, the methods of teaching programming need to be rethought (Jacques, 2023). In summary, compared to other subjects, ChatGPT's presence will have an immeasurable impact on programming and its teaching and learning. Thus, exploring the impact of ChatGPT on programming education has become a key research topic that the educational field is focusing on and urgently needs to undertake. However, current research on ChatGPT and programming learning is scarce, with the focus on the impact of ChatGPT on programming learning outcomes, and methods and strategies for programming learning assisted by ChatGPT, including its limitations and risks, mainly in the form of theoretical review articles (Yilmaz and Yilmaz, 2023). There has been no research found that systematically explores the impact of ChatGPT on programming problem-solving from the perspective of influencing factors. To find more effective methods for ChatGPT-assisted programming problem-solving, this study uses a quasi-experimental research method, taking Python drawing in programming education as an example, to explore specific factors affecting learners' programming problem-solving outcomes under the assistance of ChatGPT, and thus form targeted teaching and learning suggestions. Specifically, the study conducted pre- and post-tests from multiple dimensions including AI literacy, foundational programming knowledge, technical cognitive level, and willingness to use, systematically discussing the value and significance of ChatGPT in programming education, and attempting to answer what new teaching logic should be reflected in programming education empowered by new technologies. With empirical exploration as the direction, the study aims to provide evidence reference for building effective teaching methods for ChatGPT-like products from the perspective of programming teaching, to promote high-quality development in AIGC-enabled education.

Literature review

Review of educational applications of artificial intelligence and traditional intelligent conversational Chatbots. Prior to the emergence of ChatGPT, AI had already been widely applied in education, and various scholars have extensively discussed its empowerment in education from the perspectives of theoretical exploration and system development. For instance, in the realm of theoretical exploration, researchers have established ethical principles for AI educational applications (Nguyen et al., 2023), methods and models for AI-empowered teaching (Liu et al., 2021), as well as examined the risks and opportunities AI brings to education (Situmorang et al., 2023; Farrokhnia et al., 2023), among others. At the system development level, a series of AI systems have been developed, such as those guiding novice programmers in learning (Cruz et al., 2017) and assisting teachers in creating programming exercises (Chung et al., 2023).

Intelligent conversational chatbots similar to ChatGPT have also found multiple applications in education (Wu and Yu, 2023; Okonkwo and Ade-Ibijola, 2021; Han et al., 2022), including writing exercises (Kılıçkaya, 2020), mathematics teaching (Lee and Yeo, 2022), and language instruction (Kohnke, 2022), among others. However, previous chatbots before ChatGPT have shown remarkable effectiveness in addressing simple questions, language learning, and text processing (Hiremath et al., 2018; Zhang et al., 2023). Nonetheless, they still have various limitations in understanding contextual situations and providing answers that meet user needs (Hwang and Chang, 2021).

Current analysis of ChatGPT's educational applications. Compared to previous intelligent chatbots, ChatGPT's remarkable text comprehension abilities, long-form text generation capabilities, and programming skills have prompted educational researchers to reconsider its educational applications on top of the existing AI foundations. In the current discourse on ChatGPT's educational applications: in terms of research content, the focus primarily lies on ChatGPT's basic functionalities and internal mechanisms, educational application scenarios and future challenges, educational value and limitations, ethical issues, and potential risks (Zhai, 2022; Cooper, 2023; Kohnke et al., 2023; Humphry and Fuller, 2023; Su and Yang, 2023; Bauer et al., 2023; Tlili et al., 2023). As for the categories of research, these primarily involve ChatGPT's programming and mathematical performance (Surameery and Shakor, 2023; Frieder et al., 2023), as well as its impact on financial education (Dowling and Lucey, 2023), medical education (Kung et al., 2023; Gilson et al., 2022), second-language education (Yan, 2023), engineering education (Qadir, 2022), and programming education (Yilmaz and Yilmaz, 2023). In summary, the aforementioned studies mainly focus on ChatGPT itself, exploring its performance and its impact on teaching and learning across various disciplines.

Current status and review of ChatGPT's application in programming education. Despite the abundant research surrounding the educational applications of ChatGPT, studies focusing on the programming discipline are still extremely scarce, especially high-quality academic journal articles that have undergone peer review. A review of the limited existing research on the combination of ChatGPT and programming reveals that the studies mainly fall into three categories: The first category primarily explores ChatGPT's own programming performance. For instance, Tian et al. (2023) conducted an empirical study on the potential of ChatGPT as an automatic programming assistant, demonstrating its dominant role in handling programming issues while also revealing its limitations in understanding general problems. Surameery and Shakor (2023) discussed the methods and limitations of using ChatGPT to solve programming bugs. Kashefi and Mukerji (2023) explored, for the first time, ChatGPT's capability in programming numerical algorithms. The second category mainly investigates the methods, opportunities, and limitations of learners learning programming with the assistance of ChatGPT, such as the opportunities for scientists and engineers to learn programming using AI tools like ChatGPT (Guo, 2023), and the methods and limitations of ChatGPT in enhancing learners' programming skills (Rahman and Watanobe, 2023). The third category primarily examines the impact of ChatGPT on learners' programming learning performance, with very few related studies such as exploring the impact of programming education using ChatGPT on students' computational thinking abilities, programming self-efficacy, and learning motivation (Yilmaz and Yilmaz, 2023).

Table 1 Participant distribution table.

	Gender		Major		
	Male	Female	Mechanical engineering	Educational technology	Computer science
N	32	18	13	27	10

In summary, existing research mainly focuses on ChatGPT's programming performance and the undertaking of programming learning by learners assisted by ChatGPT, with the latter aimed at examining the impact of ChatGPT on the effectiveness of learners' programming learning. To my knowledge, there has not yet been a study exploring the specific factors affecting the effectiveness of learners using ChatGPT to solve programming problems, nor how learners undertake programming learning with the assistance of ChatGPT. However, this issue is key to promoting the effective empowerment of programming teaching and learning by ChatGPT. In view of this, our study uses Python drawing as an example to systematically explore the specific factors that affect learners' effectiveness in solving programming problems with ChatGPT.

Research design

Research questions. This study aims to explore the factors that affect the effectiveness of learners' application of ChatGPT in solving programming problems. Previous studies have shown that AI literacy is an important factor in assessing learners' ability to successfully apply AI technologies (Wang et al., 2022; Kong et al., 2022). Additionally, ChatGPT possesses a vast amount of knowledge, and people are highly interested in the emergence of similar ChatGPT products. Will this shift the focus of learners' knowledge acquisition? Moreover, learners' cognitive level and intention to use new technologies are often considered key factors influencing the effectiveness of technology utilization (Chai et al., 2021; Yu et al., 2023; Wang et al., 2023). Based on this, the present study investigates the effects of learners' AI literacy, programming knowledge base, ChatGPT cognitive level, and ChatGPT usage intention. Taking Python drawing in programming education as an example, the study focuses on three research questions:

RQ1: Does learners' AI literacy affect the effectiveness of their application of ChatGPT in solving Python programming problems?

RQ2: Does learners' programming knowledge base affect the effectiveness of their application of ChatGPT in Python programming problem-solving?

RQ3: Does learners' cognitive level of ChatGPT and their intention to use it in teaching affect the effectiveness of their application of ChatGPT in solving Python programming problems? And does their intention to use it in teaching change significantly before and after application?

Research object. The participants of this experiment mainly consisted of 53 undergraduate students from A University in Zhejiang Province. However, due to uncontrollable factors during the experiment, such as software malfunctions and incomplete pre-test and post-test data, 3 participants were excluded, resulting in a final sample size of 50 participants. The basic distribution of the valid participants can be found in Table 1.

Instruments. The materials required for this experiment consist of six components.

Material 1. The basic information survey questionnaire was used to understand the participants' demographics and their familiarity with ChatGPT. This questionnaire primarily collects the following information: participant's gender, major, familiarity with the Python plotting library Matplotlib (Appendix 1), and familiarity with ChatGPT (Appendix 2).

Material 2. AI Literacy Scale was used to assess the participants' AI literacy. This study employed the AI Literacy Scale developed by Wang et al. (2022). The scale is a Likert-type self-report questionnaire consisting of four primary dimensions: AI awareness, AI usage, AI evaluation, and AI ethics. Each primary dimension contains three items. (Appendix 3).

Material 3. ChatGPT Usage Intention Scale was used to assess the participants' attitudes towards using ChatGPT in educational settings. This scale is adapted from the AI Technology Instructional Use Intention Scale developed by Chai et al. (2021). It is also a Likert-type self-report questionnaire consisting of three items. Drawing from various instructional contexts, it investigates learners' intention to accept the application of ChatGPT in education (Appendix 4).

Material 4. Instructional materials are designed to assist learners in understanding the tools used in the experiment. This study focuses on evaluating the effectiveness of learners in completing complex programming tasks with the assistance of ChatGPT. To ensure a smooth instructional process, several scenarios and case studies have been prepared to familiarize learners with the usage of ChatGPT and the basic functionalities of Matplotlib. These materials aim to help participants overcome any obstacles related to tool usage before the formal experiment.

Material 5. Task sheet designed to assist participants in better-completing programming learning tasks. It consists of three sections: prompts for using ChatGPT and the Matplotlib plotting library, task scenario, and criteria for evaluating practical achievements.

Materials 6. The experimental tools required include Python tools and ChatGPT accounts. The Python toolkit includes Python installation packages and Matplotlib. Additionally, to prevent interference among participants and ensure consistent performance and responsiveness when using ChatGPT, this study provided each experimental participant with an independent account for the latest version of ChatGPT from the OpenAI official website. It is important to note that this research chose Python as the programming language for several reasons. Firstly, according to the September 2023 TIOBE index (TIOBE, 2023), Python is currently the most popular programming language. Secondly, Python is relatively easy for beginners to learn. Considering that the participants in the study might include individuals with limited programming experience, Python is more suitable for this research compared to languages like Java, which have higher entry barriers. Additionally, Python is open-source and free, with a large developer community, making it easy to access support and solutions for programming-related issues. Given these factors, Python undoubtedly emerged as the most suitable programming language for this research.

Experimental process. The experimental procedure was conducted in accordance with the general specifications of quasi-experimental research (Ritchie et al., 2013; Babbie, 2016), incorporating design patterns from relevant quasi-experimental empirical studies (Lee et al., 2022; Mo et al., 2022; Shadiev

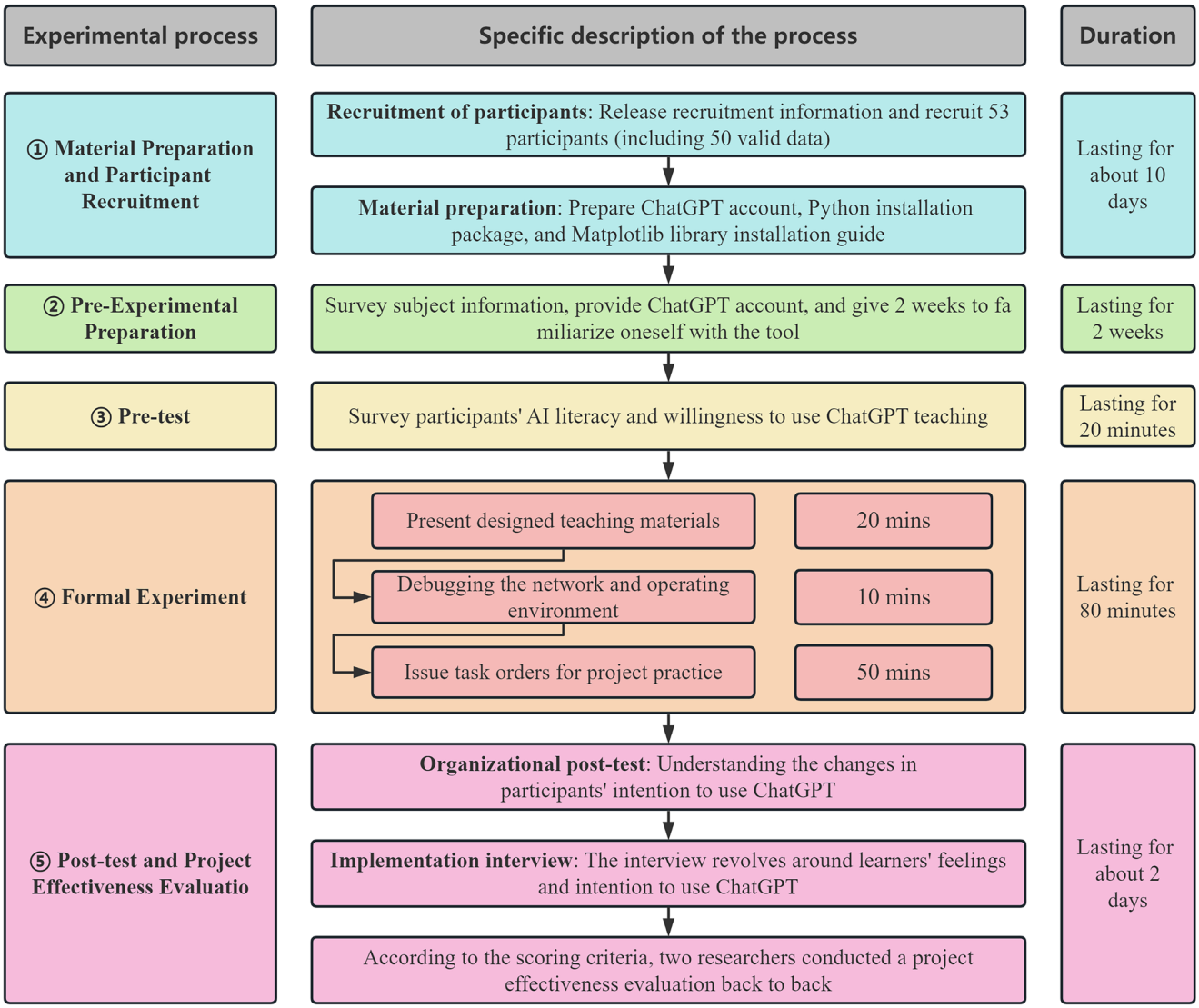

Fig. 1 Experimental design process.

et al., 2021; Hsu et al., 2012; Yang et al., 2021). The experimental process consisted of five stages and the entire experiment lasted approximately four weeks, as depicted in Fig. 1.

Step 1: Material preparation and participant recruitment. This stage primarily involves material design, pre-testing, and participant recruitment. Firstly, we designed a programming project called "Stacked Bar Chart Drawing" based on the Python Matplotlib library. Then, five participants were recruited for the pre-test (following the exact procedure of the formal experiment). The pre-test results indicated that the programming project's difficulty level was moderate, and with the assistance of ChatGPT, it took approximately 50 min to complete (with an average score of around 85 among the 5 participants). This suggests that the task was not one that could be completed solely relying on ChatGPT; rather, it was a programming project that required continuous interaction and exploration with ChatGPT. It served as an inquiry-based learning activity. Lastly, formal experimental participants were recruited.

Step 2: Pre-experimental preparation. This stage focuses on conducting a basic information survey of the participants and preparing the necessary tools. Firstly, prior to the start of the experiment, a survey of the participants' basic information is conducted, along with an assessment of their familiarity with ChatGPT. Subsequently, the experimental tools are prepared by distributing a standardized Python installation package (version 3.9.5) and configuring the controlled variable of a unified version of the Matplotlib (version 3.7.0). This ensures the elimination of interference factors caused by differences in tool versions. Following the distribution of the experimental tools, a 2-week period of technical familiarization takes place.

Step 3: Pre-test. Prior to the formal experiment, participants were asked to complete the AI Literacy Survey and the ChatGPT Usage Intention Scale. After the experiment, the reliability and validity of the two scales were verified. The Cronbach's α coefficients for the four dimensions of the AI Literacy Survey were as follows: AI Awareness (0.905), AI Usage Ability (0.933), AI Evaluation Ability (0.861), and AI Ethical (0.863), all exceeding 0.8, indicating good reliability. The overall KMO coefficient was 0.872 (>0.8), indicating good structural validity. Further factor analysis revealed clear factor structures for the four dimensions. The Cronbach's α coefficient of the ChatGPT Instructional Use Intention Scale was 0.863 (>0.8), indicating good reliability.

Task Sheet

[Task Scenario]: The following two-dimensional data represents the weekly library entry and exit records of five dormitories with four residents each. Each row element represents a dormitory (labeled as Dorm 301-305), and the column elements represent the days of the week. For example, the number 3 in the first row and second column indicates that three students from that dormitory entered or exited the library on Tuesday. To visualize the entry and exit records of these five dormitories from Monday to Friday more intuitively, we need to create a visualization. Using ChatGPT, please follow the prompts provided in the task sheet to utilize the Matplotlib plotting library in Python to generate a stacked bar chart.

[[1, 3, 2, 4, 2],
[1, 2, 2, 3, 0],
[0, 1, 3, 0, 0],
[0, 2, 2, 4, 0],
[0, 2, 0, 1, 1]]

On the right is an example:

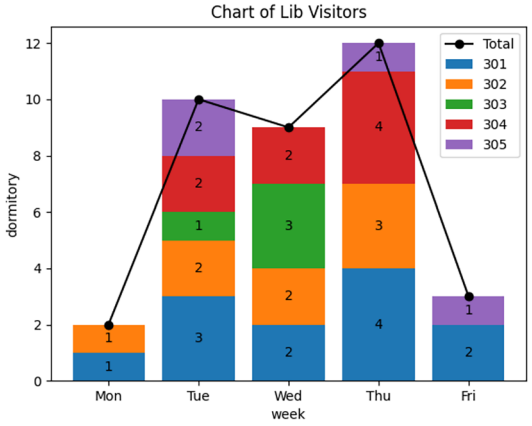

Fig. 2 Task sheet (Task Scenario part).

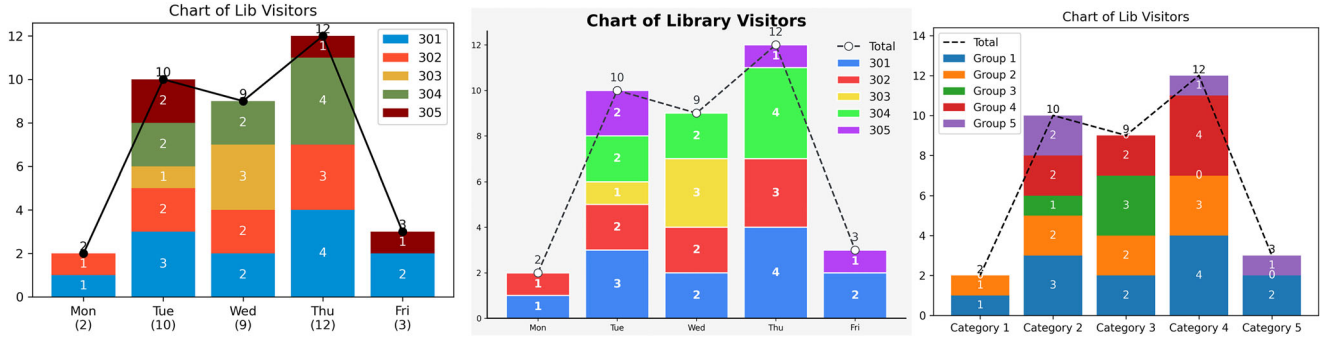

Fig. 3 Examples of project outcomes.

Step 4: Formal experiment. To provide participants with the necessary context for the experiment, they were initially presented with specially designed instructional materials. These materials aimed to familiarize them with the basic usage of ChatGPT and the Matplotlib library. Additionally, efforts were made to ensure a smooth technical setup by debugging the Python runtime environment and network, thus minimizing potential technical issues that could impact the experiment. As part of the briefing, participants were introduced to the programming task scenario and project requirements, and they were provided with a task sheet that included the task scenario as shown in Fig. 2. This figure depicted a two-dimensional representation of the weekly library entry and exit records of five dormitories, each housing four residents. Each row element represented a dormitory, labeled as Dorm 301–305, while the column elements represented the days of the week. For example, the number 3 in the first row and second column indicated that three students from that dormitory entered or exited the library on Tuesday. These visual representations were crucial in helping participants understand the practical task at hand—to utilize ChatGPT and the provided data along with the Matplotlib library to draw a stacked bar chart that visualizes the entry and exit records of these five dormitories from Monday to Friday more intuitively. With this contextual information, participants were officially instructed to commence the practical phase, with a time limit of 50 min to complete the task.

Step 5: Post-test and project effectiveness evaluation. After completing the programming project task, participants submitted their practical outcomes (stacked bar charts) through an online platform (a sample of project outcomes is shown in Fig. 3). They also filled out the ChatGPT Instructional Use Intention Scale again. Following the scale completion, six students were randomly selected for semi-structured interviews regarding their experiences using ChatGPT and their instructional use intentions. After the interviews, the project entered the effectiveness evaluation phase. Research personnel assessed 50 project outcomes based on evaluation criteria. The criteria included coordinates, legends, colors, overall coherence, numerical indicators, titles, and stacking patterns, comprising seven observation points. Each observation point consisted of 2–4 rating criteria, resulting in a total of 20 rating criteria. Each rating criterion had a maximum score of 5 points, amounting to a total of 100 points. To ensure objective and fair scoring, two researchers independently rated the project outcomes and the average score was taken as the final score. The assessment revealed consistent scores for 46 project outcomes, with minor errors below 5 points observed in 4 scores. According to the formula for inter-rater reliability calculation, the coefficient R was 0.958, indicating good reliability of the scoring results (Gaur and Kumar, 2018).

Results

In this study, the effectiveness of learners in applying ChatGPT to solve programming problems was assessed based on their project outcomes' scores (S). The study also incorporates other related concepts, including AI literacy and programming knowledge base.

Table 2 Analysis of the impact of AI literacy on the effectiveness of using ChatGPT to solve programming problems.

Group	N	Score (S)	SD	t-value	df	p
Low AI Literacy Group	24	66.458	14.005	−5.027	48	0.000
High AI Literacy Group	26	85.577	12.890			

AI literacy refers to the learners' capability to comprehend AI, utilize AI for task completion, assess the data and information provided by AI, and understand the ethical and moral norms related to AI usage (Wang et al., 2022; Chai et al., 2020). It is inherently a non-operational concept. In this study, the AI literacy scale developed by Wang et al. (2022) was employed to measure AI literacy across four dimensions: AI awareness, AI usage, AI evaluation, and AI ethics.

Furthermore, when referring to the programming knowledge base in this study, it primarily pertains to learners' proficiency in Python and their knowledge of Matplotlib. Matplotlib, which encompasses data visualization in Python, falls under the category of fine-grained knowledge within the Python programming language.

Next, we will analyze the influencing factors on learners' effectiveness in applying ChatGPT to solve programming problems from three major dimensions.

The influence of AI literacy on the effectiveness of using ChatGPT to solve programming problems

The impact of AI literacy on the effectiveness of using ChatGPT to solve programming problems. This study drew inspiration from the research design of Yang et al. (2021) and adopted the scale-based experimental data collection approach described by Yilmaz and Yilmaz (2023) for the competency section of this study. Initially, participants were categorized into high AI literacy and low AI literacy groups based on the results of the AI literacy survey. Subsequently, an independent samples t-test was employed to analyze the programming problem-solving effectiveness (S) among participants with different levels of AI literacy. The results are shown in Table 2. The study found a significant correlation between the level of AI literacy and the effectiveness of using ChatGPT to solve programming problems ($t = -5.027$, $p = 0.000 < 0.05$), with participants in the high AI literacy group ($S = 85.577$) achieving significantly higher scores than those in the low AI literacy group ($S = 66.458$).

Since AI literacy is a relatively novel concept that has not yet been widely applied in quasi-experimental and experimental studies, we can compare the results of this study with research that explores computational thinking and Artificial Intelligence capability (Yilmaz and Yilmaz, 2023; Wang et al., 2023). For example, in the study by Wang et al. (2023), it was found that there is a positive correlation between Artificial Intelligence capability and learner performance. Learners with higher Artificial Intelligence capability tend to achieve better results in tasks, which aligns with the findings of this study.

Impact of AI literacy's different dimensions on the effectiveness of using ChatGPT to solve programming problems. Based on the measurement results of AI awareness, AI usage, AI evaluation, and AI ethics, participants were categorized into high and low AI literacy groups (Yang et al., 2021). Independent samples t-tests were conducted to examine the score (S) of solving programming problems for different groups within each of the four dimensions.

Table 3 Analysis of the impact of AI literacy's different dimensions on the effectiveness of using ChatGPT to solve programming problems.

Specific literacy	Group	N	Score (S)	SD	t-value	df	p
AI Awareness	Low	25	66.800	14.905	−5.066	48	0.000
	High	25	86.000	11.704			
AI Usage	Low	24	69.063	13.689	−3.332	48	0.002
	High	26	83.173	16.040			
AI Evaluation	Low	25	73.000	17.245	−1.481	48	0.145
	High	25	79.800	15.154			
AI Ethics	Low	25	75.800	17.526	−0.256	48	0.799
	High	25	77.000	15.596			

Specific results are presented in Table 3. From Table 3, it can be observed that AI awareness and AI usage significantly affect the effectiveness of learners in using ChatGPT to solve programming problems, while AI evaluation and AI ethics standards did not have a significant impact.

AI awareness literacy primarily refers to the ability to identify and understand AI before or during the use of AI-related applications, which is considered a cognitive process (Calvani et al., 2009; Hallaq, 2016). From this perspective, AI awareness focuses on the perceptiveness of AI before its use (Wang et al., 2022). In this study, AI awareness refers to the materials presented to the participants before the project implementation to understand the assistive role of ChatGPT in learning. During this process, learners with higher AI awareness are more likely to effectively apply their objective understanding of ChatGPT's capabilities to solve programming problems, thereby achieving better completion of programming tasks within the given timeframe.

AI usage literacy refers to the capacity to use AI to accomplish tasks, which emphasizes the operational aspect of AI literacy. It includes proficiency in AI applications and related products, familiarity with the operating standards of AI tools, and the ability to integrate different types of AI applications and tools to solve problems. AI usage refers to the capacity to use AI to accomplish tasks. It includes proficiency in AI applications and related products, familiarity with the operating standards of AI tools, and the ability to integrate different types of AI applications and tools to solve specific problems. The study found that AI usage significantly impacts the effectiveness of participants in using ChatGPT to solve programming problems.

AI evaluation literacy primarily refers to the user's ability to compare similar products and select the optimal solution. AI ethics literacy refers to the user's concern for ethical issues that AI may raise, such as privacy breaches (Wang et al., 2022). In this study, there are possible reasons why these two dimensions did not have a significant impact on learners' effectiveness in solving programming problems: firstly, the study specified ChatGPT as the tool for completing the project tasks, eliminating the need for learners to make independent choices; secondly, during the project implementation, a relatively stable and secure ChatGPT environment was provided to learners, aiming to minimize the occurrence of ethical issues.

The impact of programming knowledge base on the effectiveness of using ChatGPT to solve programming problems

Impact of Python knowledge base on the effectiveness of using ChatGPT to solve programming problems. In this study, participants came from different major backgrounds, and due to variations in the curriculum plans across these majors, some students had completed a 3-credit Python course (Group A, Major: Educational Technology), some had completed a

Table 4 Analysis of the impact of Python knowledge base on the effectiveness of learners in applying ChatGPT to solve programming problems.

Group	Python knowledge base	N	Score (S)	SD	F	p
Group A	3 Credits Python Fundamentals	27	79.907	15.008	7.039	0.002
Group B	4 Credits Python Fundamentals	10	83.750	15.954		
Group C	No System Python Base	13	63.462	13.012		

4-credit Python course (Group B, Major: Computer Science), while others had not taken any Python course (Group C, Major: Mechanical Engineering). This resulted in differences in the participants' programming knowledge base. The relationship between the Python knowledge base and the effectiveness of using ChatGPT to solve programming problems was analyzed using a one-way analysis of variance (ANOVA) method, as shown in Table 4. The results indicated that the Python knowledge base significantly impacted the effectiveness ($F = 7.039$, $p = 0.002 < 0.05$). Further post-hoc multiple comparison analyses revealed significant differences in the programming problem-solving score between Group A and Group C ($p = 0.002 < 0.05$) and between Group B and Group C ($p = 0.002 < 0.05$). However, there was no significant difference between Group A and Group B ($p = 0.484 > 0.05$).

Based on the analysis above: Firstly, the Python knowledge base is an important factor that affects the effectiveness of learners in applying ChatGPT to solve programming problems. Learners with a solid foundation in Python have a significant advantage in the effectiveness of using ChatGPT to solve programming problems. Secondly, learners who possess a solid foundation in Python (Group A and B) show no significant differences in the effectiveness of solving programming problems using ChatGPT, despite variations in their knowledge levels.

By examining the question records of 50 participants stored in their ChatGPT accounts, a significant correlation was found between the above research findings and the characteristics of the ChatGPT tool. Specifically, using ChatGPT requires learners to accurately express their problem requirements. In the context of this study, solving programming problems necessitates learners to have a certain foundation in programming in order to accurately break down programming tasks and express the corresponding problem statements. Participants who lack a knowledge base in Python find it challenging to transform programming tasks into a series of questions suitable for natural language processing using ChatGPT. On the other hand, participants with a solid foundation in Python can provide ChatGPT with clear and precise instructions based on their own Python knowledge, thus enabling them to complete programming tasks with the assistance of ChatGPT. Despite the variation in the level of the Python knowledge base, the final effectiveness of solving programming problems did not show significant differences because ChatGPT addressed the specific programming details.

The impact of Matplotlib knowledge base on the effectiveness of using ChatGPT to solve programming problems. To confirm the above viewpoint, this study further used the level of knowledge in the Matplotlib plotting library as an independent variable (measured items can be seen in Appendix 1) to categorize the participants into three groups: no base in Matplotlib knowledge (0–2 correct answers), weak base in Matplotlib knowledge (3-4 correct answers), and good base in Matplotlib knowledge (5 correct answers). One-way analysis of variance (ANOVA) was used to test the differences in practical achievement scores corresponding to different levels of the Matplotlib knowledge base (see Table 5). It was found that the level of Matplotlib knowledge base did not have a significant impact on the effectiveness of learners in using ChatGPT to solve programming problems ($F = 1.283$, $p = 0.287 > 0.05$).

Table 5 Analysis of the impact of Matplotlib base on the effectiveness of learners in using ChatGPT to solve programming problems.

Matplotlib knowledge base	N	Score (S)	SD	F	p
Not in contact with Matplotlib	25	72.700	17.011	1.283	0.287
Weak Matplotlib knowledge base	21	80.119	14.566		
Good Matplotlib knowledge base	4	80.000	14.142		

Based on the above results, the involvement of ChatGPT in helping learners solve domain-specific problems, combined with a systematic foundation in the subject, highlights the importance of mastering scientific knowledge framework, basic expression standards (Python language specifications), and the ability to decompose complex tasks and transform them into problem-solving strategies. This importance surpasses the mere accumulation of specialized knowledge.

It should be noted that this study contributes some novel findings in this regard. Previous research has not focused on the impact of learners' programming knowledge base on the effectiveness of applying ChatGPT (or other AI technologies) to solve programming problems. Many studies have stopped at discovering that ChatGPT (or other AI technologies) can improve the efficiency and quality of programming problem-solving (Yilmaz and Yilmaz, 2023; Katchapakirin et al., 2022). Based on the findings of this study, we can further analyze and point out that for learners with no programming background, the assistance provided by ChatGPT is limited. However, for learners with some programming knowledge, ChatGPT can significantly enhance their upper limit of programming ability, enabling them to better and faster complete relatively complex programming tasks. For learners with a higher programming knowledge base, the assistance effect of ChatGPT is not as pronounced as for those with some programming knowledge. Although our study focused on Python as a tool, we believe that this conclusion will still hold true for other programming languages, and we will explore this further in future research.

This original conclusion from our study actually illuminates a new logic for learning in the era of education empowered by ChatGPT. Knowledge acquisition and application become inseparable, especially in programming learning, where greater emphasis is placed on problem decomposition. Learners need to focus more on problem-solving thinking and logic, rather than just coding ability and efficiency.

The impact of learners' cognitive level and usage intention towards ChatGPT on the effectiveness of solving programming problems

The impact of learners' cognitive level of ChatGPT on the effectiveness of solving programming problems. Building upon the

Table 6 Analysis of the impact of learners' cognitive level of ChatGPT on the effectiveness of using it to solve programming problems.

Group	N	Score (S)	SD	t	df	p
Low cognitive level group	23	70.543	14.789	−2.441	48	0.018
High cognitive level group	27	81.389	16.355			

Table 7 Analysis of the impact of learners' intention towards ChatGPT on the effectiveness of using it to solve programming problems.

Group	N	Score (S)	SD	t-value	df	p
Low Intention Group	25	77.800	13.117	0.599	42.192	0.552
High Intention Group	25	75.000	19.365			

previous research, this study further explores the impact of learners' cognitive level and usage intention towards ChatGPT on their effectiveness in using ChatGPT to solve programming problems, as well as the changes in usage intention before and after usage.

Learners' cognitive level of ChatGPT mainly refers to their understanding of ChatGPT prior to usage. In this study, a cognitive level scale was developed based on the technical awareness scale by Venkatesh (2000) to measure the participants' cognitive level of ChatGPT (see Appendix 2). Based on the measurement results, the 50 participants were divided into a low cognitive-level group and a high cognitive-level group (Yang et al., 2021). An independent samples t-test was conducted to examine the learning effectiveness between the two groups (see Table 6), revealing a significant impact of participants' cognitive level of ChatGPT on the effectiveness of using it to solve programming problems ($t = -2.441$, $p = 0.018$).

The higher the learners' cognitive level of ChatGPT, the clearer their understanding of its principles and functions. On one hand, they can effectively delegate the parts of programming tasks that are suitable for ChatGPT to it for completion. On the other hand, they can easily provide ChatGPT with instructions that are easy to comprehend. As a result, they can achieve better, faster, and even surpass the target in completing programming projects. In some cases, individual learners enhanced the esthetics of stacked bar charts by utilizing ChatGPT, building upon the requirements of their programming projects.

The impact of learners' usage intention towards ChatGPT on the effectiveness of solving programming problems. Learners' usage intention towards ChatGPT refers to their attitude towards using ChatGPT in teaching and learning contexts. The scale used to measure learners' teaching usage intention towards ChatGPT in this study was adapted from the AI Usage Intention Scale developed by Chai et al. (2021) (see Appendix 4). Using this scale, the usage intention towards ChatGPT of the 50 participants was measured, and based on the measurement results, they were divided into low intention group and high intention group (Yang et al., 2021). Subsequently, an independent samples t-test was conducted to examine the practical achievement scores between the two groups (see Table 7). From Table 7, it can be observed that the teaching usage intention towards ChatGPT did not have a significant impact on the participants' effectiveness in using ChatGPT to solve programming problems ($t = 0.599$, $p = 0.552 > 0.5$). This result contradicts the findings of international studies that suggest a significant impact of usage intention towards technology on learners' learning performance (Sung et al., 2017).

In order to further explore and explain the reasons behind the above results, this study conducted semi-structured interviews with six participants. Two main reasons were identified: First, high-achieving students tended to be resistant to using ChatGPT to complete programming projects. This resistance stemmed from their high confidence in problem-solving and the belief that the intervention of ChatGPT would diminish their learning advantage compared to other classmates. Second, participants who had not developed a comprehensive understanding of ChatGPT were influenced by their subjective norms when it came to their intention to use it. Some participants believed that ChatGPT was capable of anything, thus enhancing their intention to use it. These findings were attributed to the participants' limited knowledge of ChatGPT and the hype surrounding its disruptive and outstanding performance compared to previous technologies. Therefore, this serves as a reminder that when faced with new technologies, it is important to avoid both conservative attitudes and blind optimism. Developing a rational and objective understanding of the technology is key to effectively utilizing it and maximizing its educational benefits.

The usage intention towards ChatGPT of participants before and after project implementation. This study further collected participants' usage intention towards ChatGPT after completing programming projects using it. Then, a paired samples t-test was conducted to analyze the usage intention towards ChatGPT of participants before and after project implementation (see Table 8). The results revealed that the usage intention towards ChatGPT increased significantly after the completion of programming projects, with a mean score of 8.420 (compared to the pre-test mean score of 7.180), indicating a significant improvement in usage intention towards ChatGPT ($p = 0.017 < 0.05$). This suggests that the application of ChatGPT not only helps participants develop a comprehensive understanding of ChatGPT but also enhances their intention to use this new technology in learning. In other words, learners' perception of AI technologies such as ChatGPT influences their intention to use them. The conclusion on this point is consistent with previous research. For example, Huang et al. (2023) found that personalized video recommendations based on AI can increase learners' motivation to learn. Studies by Fryer et al. (2019) and Huang & Qiao (2022) also indicate that using chatbots in language learning can enhance learner engagement and motivation. Therefore, it can be inferred that when learners become aware of a technology with significant potential benefits for learning, they tend to have higher satisfaction and, consequently, an increased willingness to use it.

In this study, it was found that the intention to use ChatGPT among participants did not significantly affect the effectiveness of solving programming problems (detailed analysis has been conducted, but further validation is needed in other disciplines). However, previous research has demonstrated that AI, represented by ChatGPT, can indeed enhance the efficiency of teaching and learning (Zheng et al., 2021), and they have the potential to empower the intelligent transformation of education from multiple perspectives. With the exponential development of information technology, the integration of emerging technologies with AI as the core of educational reforms has become an unstoppable trend (Popenici and Kerr, 2017). Therefore, guiding and promoting learners to enhance their intention to use ChatGPT in teaching and learning is not only a requirement

Table 8 Analysis of learners' usage intention before and after project implementation.

Group	*N*	Intention score (*I*)	SD	*t*-value	df	*p*
Pre-test of intention to use	50	7.180	2.238	−2.471	49	0.017
Post-test of intention to use	50	8.420	2.658			

for AI-enabled educational transformation but also a crucial aspect of adapting to and driving the digital transformation of education. It highlights the importance of learners' digital literacy skills.

Discussion and implications

Based on the aforementioned research findings, during the educational digital transformation empowered by AIGC products, teaching should focus on fostering three aspects of learners' abilities:

From AI capability to AI literacy: establishing the foundation for empowering education with AIGC. With the iterative upgrades of the ChatGPT core model and the emergence of many similar products, there will be more applications and products integrating technological advantages, forming a diverse range of multimodal AIGC technology tools. In this context, learners who only possess passive AI usage and lack the proactive skills to identify and select the most appropriate AI tools and adhere to AI ethics may struggle to timely and effectively access and utilize advanced AIGC technologies, impeding their ability to engage in safe and effective learning. Therefore, in addition to enhancing learners' AI usage, it is essential to help them develop AI awareness, AI evaluation, and AI ethics, thus fostering AI literacy. This requirement not only establishes the base for learners to efficiently leverage AI, represented by ChatGPT, but also enables them to adapt to the core AI-driven educational transformation and effectively apply AI to enhance their problem-solving effectiveness. It should be noted that participants were provided with specific AI tools and were required to complete programming tasks within specific time and space constraints. This setup may have limited the participants' ability to fully utilize their AI evaluation and exercise judgment regarding AI ethics. However, as an abundance of related tools emerges in the future, issues such as data privacy and network security will become increasingly pressing. Therefore, AI literacy is crucial for ensuring high-quality teaching and learning, as it equips individuals with the necessary knowledge and skills to navigate these challenges effectively.

From specific knowledge to graph-based rules: empowering education with AIGC through grafting methods. Based on the findings of this study (systematic knowledge base significantly affects the effectiveness of programming problem-solving, but the knowledge base of Matplotlib does not have a significant impact when having a systematic knowledge base), it can be concluded that:

Firstly, mastering the subject knowledge graph, basic conventions, and rule-based instructions (in programming learning) becomes more important. From the perspective of programming learning, ChatGPT has powerful code-writing capabilities. As long as users can select a specific programming language type and provide relevant execution rules and instructions based on language conventions (such as invoking the Matplotlib library), they can obtain programming results that lead to the achievement of task goals. In this process, although the specific details of the programming process are taken care of by ChatGPT, learners who lack systematic knowledge of the relevant programming language, basic conventions, and instruction rules will have difficulty effectively utilizing ChatGPT and truly transforming it into their "external brain." Therefore, mastering the subject knowledge graph, basic conventions, and rule-based instructions is more important compared to acquiring a greater amount of specific specialized knowledge.

Secondly, possessing the ability to decompose complex tasks and ask questions that lead to the achievement of task goals becomes more important. The ability to decompose complex learning tasks (such as programming graphics in this study) and transform the decomposed series of tasks into problem expressions that can be unambiguously understood in natural language is a prerequisite for ChatGPT to operate effectively and provide efficient solutions. Therefore, in the context of empowering AIGC products, it becomes important for learners to systematically master the subject knowledge graph, basic conventions, and relevant knowledge invocation instructions, as well as possess the ability to decompose complex tasks and transform task-related problems. This enables the development of learning approaches that integrate knowledge and skills from ChatGPT (forming complex learners with a composite brain), which is more important compared to solely acquiring more programming knowledge and single-complex problem-solving abilities. In conclusion, this study further addresses the question of whether we still need to learn knowledge or focus on which specific knowledge in the context of empowering AIGC products.

From intention to technological cognition: creating drivers to empower education with AIGC. Based on the findings of this study, it is evident that learners' level of cognitive understanding of ChatGPT significantly affects their effectiveness in solving programming problems. Moreover, as their cognitive understanding of ChatGPT improves, their intention to use it also increases significantly. Therefore, for the effective application and promotion of AIGC products, it is important to enhance learners' technological cognition. This can be achieved by providing learners with application guides for technical tools and assisting them in experiencing the application of such tools. By helping learners establish a rational and accurate understanding of AIGC products with astonishing intelligence, their intention to use them can be enhanced. This, on the one hand, contributes to transforming learners into proactive and efficient composite learners who adapt to the intelligent era. On the other hand, it empowers AIGC products to drive discipline-based learning effectively, enabling learners to fully utilize and unleash the educational value of AIGC products. Through complementary human-machine integration, high-quality teaching and learning can be truly achieved.

Conclusion and prospect

Conclusion. AIGC products, represented by ChatGPT, are revolutionizing people's existing perceptions of AI-enabled education with their unprecedented outstanding performance. As a new driving force for AI-enabled education, the current education community lacks sufficient empirical evidence to determine its value in educational applications. It remains unclear which factors influence the effectiveness of solving programming problems under the support of ChatGPT and requires further investigation. In this study, we adopted a quasi-experimental research method, taking Python

plotting in programming education as an example, to explore the factors influencing the effectiveness of learners in solving problems with ChatGPT, including AI literacy, knowledge base, and usage intention towards ChatGPT. Our findings are as follows:

Firstly, AI literacy significantly affects the effectiveness of learners in problem-solving with ChatGPT. Specifically, AI awareness and AI usage significantly influence the effectiveness of learners in using ChatGPT to solve problems, while AI evaluation and AI ethics do not have a significant impact on learners' effectiveness in problem-solving with ChatGPT in the context of this study.

Secondly, The presence or absence of a Python knowledge base significantly influences the effectiveness of learners in using ChatGPT to solve programming problems. However, for learners who have all studied Python courses, different levels of Python knowledge base do not significantly affect the effectiveness of using ChatGPT to solve problems. This suggests that in small-scale development projects similar to this experiment, ChatGPT can greatly enhance the practical skills of learners who already possess a certain level of programming knowledge.

Thirdly, The usage intention towards ChatGPT does not significantly affect the effectiveness of learners in using ChatGPT to solve problems. However, after completing relevant project tasks using ChatGPT, learners' usage intention towards ChatGPT significantly increases compared to before.

These findings collectively contribute to our understanding of the factors that influence the effectiveness of learners in utilizing ChatGPT for problem-solving, shedding light on the potential of AIGC products in educational applications.

Limitations and future research. The study employed a quasi-experimental research method to explore the factors influencing learners' programming problem-solving effectiveness with the assistance of ChatGPT. Using "Python programming problem-solving" as an example, this study helps clarify the specific factors affecting the effectiveness of learners' problem-solving with ChatGPT and other AIGC products. It provides a theoretical reference for effectively empowering subject teaching through the construction of AIGC products and offers practical support for AIGC products' educational applications.

It should be noted that, in practical terms, the value of this research is not limited to guiding Matplotlib plot instruction itself. Any programming-based plotting, such as pyecharts for project development and business analysis, Seaborn for statistical modeling, or ggplot2 for R language, can benefit from the patterns identified in this study. Furthermore, this research provides some insights into the broader context of programming learning.

However, as the first empirical study applying AIGC to programming education, this research undoubtedly has some limitations. Firstly, the research was conducted in a contextualized and project-based manner. While the ChatGPT accounts provided recorded the interactions between learners and ChatGPT, this study did not find a suitable method to analyze this poorly structured data. Therefore, it could not elucidate or deconstruct the internal activities of learners when using ChatGPT for project work. The specific role of ChatGPT in this practical process remains unclear. In view of this, future analysis can incorporate human-computer interaction data as a crucial entry point to understanding the inherent mechanisms of ChatGPT in empowering teaching. Furthermore, the study lacked a focus on the efficiency of learners in using ChatGPT. It should be noted that "Prompt Engineering" is a highly meaningful learning material and should have been incorporated into the experiment design as a prerequisite for learners to efficiently utilize ChatGPT. Future research will delve into whether "Prompt Engineering" can assist learners in completing programming tasks more efficiently. Lastly, this study was constrained by the discipline and types of programming languages. Subsequent research is needed to expand beyond the Python language foundation in programming education to encompass a broader range of programming languages (such as Java, MATLAB, and C++) and different disciplines. Additionally, there is a need to increase both the number and diversity of participants, as well as consider variations in experimental task types and durations. These steps are essential to further enrich and unveil the conclusions drawn from this study.

Data availability

The datasets analyzed during the current study are available in the Dataverse repository: https://doi.org/10.7910/DVN/2ERVNE.

Received: 24 August 2023; Accepted: 25 January 2024;
Published online: 26 February 2024

References

Babbie ER (2016) Practice of social research. Cengage Learning

Bauer E, Greisel M, Kuznetsov I, Berndt M, Kollar I, Dresel, M., ... & Fischer, F (2023) Using natural language processing to support peer-feedback in the age of artificial intelligence: a cross-disciplinary framework and a research agenda. Br J Educ Technol. https://doi.org/10.1111/bjet.13336

Calvani A, Fini A, Ranieri M (2009) Assessing digital competence in secondary education. Issues, models and instruments. Issues in information and media literacy. Educ Pract Pedagog 2:153–172

Chai CS, Lin PY, Jong MSY, Dai Y, Chiu TKF, Qin J (2021) Perceptions of and behavioral intentions towards learning artificial intelligence in primary school students. Educ Technol Soc 24(3):89–101. https://www.jstor.org/stable/27032858

Chai CS, Wang X, Xu C (2020) An extended theory of planned behavior for the modelling of Chinese secondary school students' intention to learn artificial intelligence. Mathematics 8(11):2089. https://doi.org/10.3390/math8112089

Chung CY, Hsiao IH, Lin YL (2023) AI-assisted programming question generation: Constructing semantic networks of programming knowledge by local knowledge graph and abstract syntax tree. J Res Technol Educ 55(1):94–110. https://doi.org/10.1080/15391523.2022.2123872

Cooper G (2023) Examining science education in ChatGPT: an exploratory study of generative artificial intelligence. J Sci Educ Technol, 1–9. https://doi.org/10.1007/s10956-023-10039-y

Cruz G, Jones J, Morrow M, Gonzalez A, Gooch B (2017, July) An AI system for coaching novice programmers. In: Learning and collaboration technologies. Technology in Education: 4th International Conference, LCT 2017, Held as Part of HCI International 2017. pp. 12–21

Dowling M, Lucey B (2023) ChatGPT for (finance) research: the Bananarama conjecture. Financ Res Lett 53:103662. https://doi.org/10.1016/j.frl.2023.103662

Farrokhnia M, Banihashem SK, Noroozi O, Wals A (2023) A SWOT analysis of ChatGPT: Implications for educational practice and research. Innov Educ Teach Int 1–15. https://doi.org/10.1080/14703297.2023.2195846

Frieder S, Pinchetti L, Griffiths RR, Salvatori T, Lukasiewicz T, Petersen PC, ... Berner J (2023) Mathematical capabilities of ChatGPT. https://doi.org/10.48550/arXiv.2301.13867

Fryer LK, Nakao K, Thompson A (2019) Chatbot learning partners: Connecting learning experiences, interest and competence. Comput Hum Behav 93:279–289. https://doi.org/10.1016/j.chb.2018.12.023

Gaur A, Kumar M (2018) A systematic approach to conducting review studies: an assessment of content analysis in 25 years of IB research. J World Bus 53(2):280–289. https://doi.org/10.1016/j.jwb.2017.11.003

Gilson A, Safranek C, Huang T, Socrates V, Chi L, Taylor R A, Chartash D (2022) How well does ChatGPT do when taking the medical licensing exams? The implications of large language models for medical education and knowledge assessment. medRxiv https://doi.org/10.1101/2022.12.23.22283901

Guo PJ (2023) Six opportunities for scientists and engineers to learn programming using AI tools such as ChatGPT. Comput Sci Eng 25(3):73–78. https://doi.org/10.1109/MCSE.2023.3308476

Hallaq T (2016) Evaluating online media literacy in higher education: validity and reliability of the digital online media literacy assessment (DOMLA). J Media Lit Educ 8(1):62–84

Han JW, Park J, Lee H (2022) Analysis of the effect of an artificial intelligence chatbot educational program on non-face-to-face classes: a quasi-

experimental study. BMC Med Educ 22(1):830. https://doi.org/10.1186/s12909-022-03898-3
Hiremath G, Hajare A, Bhosale P, Nanaware R, Wagh KS (2018) Chatbot for education system. Int J Adv Res Idea Innov Technol 4(3):37–43
Hsu CK, Hwang GJ, Chuang CW, Chang CK (2012) Effects on learners' performance of using selected and open network resources in a problem-based learning activity. Br J Educ Technol 43(4):606–623. https://doi.org/10.1111/j.1467-8535.2011.01235.x
Huang AY, Lu OH, Yang SJ (2023) Effects of artificial Intelligence-Enabled personalized recommendations on learners' learning engagement, motivation, and outcomes in a flipped classroom. Comput Educ 194:104684. https://doi.org/10.1016/j.compedu.2022.104684
Huang X, Qiao C (2022) Enhancing computational thinking skills through artificial intelligence education at a STEAM high school. Sci Educ 1–21. https://doi.org/10.1007/s11191-022-00392-6
Humphry T, Fuller AL (2023) Potential ChatGPT use in undergraduate chemistry laboratories. J Chem Educ 100(4):1434–1436. https://doi.org/10.1021/acs.jchemed.3c00006
Hwang GJ, Chen NS (2023) Editorial position paper: exploring the potential of generative artificial intelligence in education: applications, challenges, and future research directions. Educ Technol Soc 26(2). https://doi.org/10.30191/ETS.202304_26(2).0014
Hwang GJ, Chang CY (2021) A review of opportunities and challenges of chatbots in education. Interact Learn Environ 1–14. https://doi.org/10.1080/10494820.2021.1952615
Jacques L (2023) Teaching CS-101 at the dawn of ChatGPT. ACM Inroads 14(2):40–46. https://doi.org/10.1145/3595634
Jalil S, Rafi S, LaToza TD, Moran K, Lam W (2023, April) Chatgpt and software testing education: Promises & perils. In 2023 IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW) pp. 4130–4137
Kashefi A, Mukerji T (2023) Chatgpt for programming numerical methods. J Mach Learn Model Comput 4(2):1–74. https://doi.org/10.1615/JMachLearnModelComput.2023048492
Katchapakirin K, Anutariya C, Supnithi T (2022) ScratchThAI: A conversation-based learning support framework for computational thinking development. Educ Inf Technol 27(6):8533–8560. https://doi.org/10.1007/s10639-021-10870-z
Kılıçkaya F (2020) Using a Chatbot, Replika, to practice writing through conversations in L2 English: a case study. In: Kruk M, Peterson M (eds.) Advances in educational technologies and instructional design. IGI Global. pp. 221–238. https://doi.org/10.4018/978-1-7998-2591-3.ch011
Kohnke L (2022) A pedagogical Chatbot: a supplemental language learning tool. RELC J 00336882211067054. https://doi.org/10.1177/00336882211067054
Kohnke L, Moorhouse BL, Zou D (2023) ChatGPT for language teaching and learning. RELC J 00336882231162868. https://doi.org/10.1177/0033688223116286
Kong SC, Cheung WMY, Tsang O (2022) Evaluating an artificial intelligence literacy programme for empowering and developing concepts, literacy and ethical awareness in senior secondary students. Educ Inf Technol 1–22. https://doi.org/10.1007/s10639-022-11408-7
Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, Tseng V (2023) Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. PLoS Digit health 2(2):e0000198. https://doi.org/10.1371/journal.pdig.0000198
Lee D, Yeo S (2022) Developing an AI-based chatbot for practicing responsive teaching in mathematics. Comput Educ 191:104646. https://doi.org/10.1016/j.compedu.2022.104646
Lee YF, Hwang GJ, Chen PY (2022) Impacts of an AI-based chatbot on college students' after-class review, academic performance, self-efficacy, learning attitude, and motivation. Educ Technol Res Dev 70(5):1843–1865. https://doi.org/10.1007/s11423-022-10142-8
Liu J, Wang L, Zhou H (2021) The application of human–computer interaction technology fused with artificial intelligence in sports moving target detection education for college athlete. Front Psychol 12:677590. https://doi.org/10.3389/fpsyg.2021.677590
Lodge JM, Thompson K, Corrin L (2023) Mapping out a research agenda for generative artificial intelligence in tertiary education. Australas J Educ Technol 39(1):1–8. https://doi.org/10.14742/ajet.8695
Luo W, He H, Liu J, Berson IR, Berson MJ, Zhou Y, Li H (2023) Aladdin's Genie or Pandora's box for early childhood education? Experts chat on the roles, challenges, and developments of ChatGPT. Early Educ Dev 1–18. https://doi.org/10.1080/10409289.2023.2214181
Mo CY, Wang CL, Dai J, Jin PQ (2022) Video playback speed influence on learning effect from the perspective of personalized adaptive learning: a study based on cognitive load theory. Front Psychol 13:839982. https://doi.org/10.3389/fpsyg.2022.839982
Nguyen A, Ngo HN, Hong Y et al. (2023) Ethical principles for artificial intelligence in education. Educ Inf Technol 28:4221–4241. https://doi.org/10.1007/s10639-022-11316-w
Okonkwo CW, Ade-Ibijola A (2021) Chatbots applications in education: a systematic review. Comput Educ Artif Intell 2:100033. https://doi.org/10.1016/j.caeai.2021.100033
Ouh EL, Gan BKS, Shim KJ, Wlodkowski S (2023) ChatGPT, can you generate solutions for my coding exercises? An evaluation on its effectiveness in an undergraduate Java programming course. https://doi.org/10.48550/arXiv.2305.13680
Popenici SAD, Kerr S (2017) Exploring the impact of artificial intelligence on teaching and learning in higher education. Res Pract Technol Enhanced Learn 12(1). https://doi.org/10.1186/s41039-017-0062-8
Qadir J (2022) Engineering education in the era of ChatGPT: promise and pitfalls of generative AI for education. TechRxiv https://doi.org/10.36227/techrxiv.21789434.v1
Rahman MM, Watanobe Y (2023) ChatGPT for education and research: Opportunities, threats, and strategies. Appl Sci 13(9):5783. https://doi.org/10.3390/app13095783
Ritchie J, Lewis J, Nicholls CM, Ormston R (2013) Qualitative research practice: a guide for social science students and researchers. SAGE
Shadiev R, Wang X, Huang YM (2021) Cross-cultural learning in virtual reality environment: facilitating cross-cultural understanding, trait emotional intelligence, and sense of presence. Educ Technol Res Dev 69(5):2917–2936. https://doi.org/10.1007/s11423-021-10044-1
Situmorang DDB, Salim RMA, Ifdil I, Liza LO, Rusandi MA, Hayati IR, ... Fitriani A (2023) The current existence of ChatGPT in education: a double-edged sword?. J Public Health, fdad082. https://doi.org/10.1093/pubmed/fdad082
Su J, Yang W (2023) Unlocking the power of ChatGPT: a framework for applying generative AI in education. ECNU Rev Educ, 20965311231168423. https://doi.org/10.1177/20965311231168423
Sung HY, Hwang GJ, Lin CJ, Hong TW (2017) Experiencing the Analects of Confucius: an experiential game-based learning approach to promoting students' motivation and conception of learning. Comput Educ 110:143–153. https://doi.org/10.1016/j.compedu.2017.03.014
Surameery NMS, Shakor MY (2023) Use ChatGPT to solve programming bugs. Int J Inf Technol Comput Eng (IJITC) 3(01):17–22. https://doi.org/10.55529/ijitc.31.17.22
Tian H, Lu W, Li TO, Tang X, Cheung SC, Klein J, Bissyandé TF (2023) Is ChatGPT the ultimate programming assistant—how far is it?. *arXiv preprint* https://doi.org/10.48550/arXiv.2304.11938
TIOBE (2023) TIOBE Index | TIOBE - The Software Quality Company. Tiobe.com. https://www.tiobe.com/tiobe-index/
Tlili A, Shehata B, Adarkwah MA et al. (2023) What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learn Environ 10:15. https://doi.org/10.1186/s40561-023-00237-x
Venkatesh V (2000) Determinants of perceived ease of use: Integrating control, intrinsic motivation, and emotion into the technology acceptance model. Inf Syst Res 4:342–365. https://doi.org/10.1287/isre.11.4.342.11872
Wang B, Rau PLP, Yuan T (2022) Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale. Behav Inf Technol 1–14. https://doi.org/10.1080/0144929X.2022.2072768
Wang CL, Dai J, Zhu KK, Yu T, Gu XQ (2023) Understanding the continuance intention of college students toward new E-learning spaces based on an integrated model of the TAM and TTF. Int J Human-Comput Interact 1–14, https://doi.org/10.1080/10447318.2023.2291609
Wang S, Sun Z, Chen Y (2023) Effects of higher education institutes' artificial intelligence capability on students' self-efficacy, creativity and learning performance. Educ Inf Technol 28:4919–4939. https://doi.org/10.1007/s10639-022-11338-4
Wu R, Yu Z (2023) Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. Br J Educ Technol. https://doi.org/10.1111/bjet.13334
Yan D (2023) Impact of ChatGPT on learners in a L2 writing practicum: an exploratory investigation. Educ Inform Technol, 1–25. https://doi.org/10.1007/s10639-023-11742-4
Yang G, Chen YT, Zheng XL, Hwang GJ (2021) From experiencing to expressing: a virtual reality approach to facilitating pupils' descriptive paper writing performance and learning behavior engagement. Br J Educ Technol 52(2):807–823. https://doi.org/10.1111/bjet.13056
Yilmaz R, Yilmaz FGK (2023) The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. Comput Educ Artif Intellig 100147. https://doi.org/10.1016/j.caeai.2023.100147
Yu T, Dai J, Wang CL (2023) Adoption of blended learning: Chinese university students' perspectives. Humanit Soc Sci Commun 10:390. https://doi.org/10.1057/s41599-023-01904-7
Zhai X (2022) ChatGPT user experience: implications for education. SSRN Scholarly Paper. https://doi.org/10.2139/ssrn.4312418
Zhang S, Shan C, Lee JSY, Che S, Kim JH (2023) Effect of chatbot-assisted language learning: a meta-analysis. Educ Inf Technol 1–21. https://doi.org/10.1007/s10639-023-11805-6

Zheng L, Niu J, Zhong L, Gyasi JF (2021) The effectiveness of artificial intelligence on learning achievement and learning perception: a meta-analysis. Interactive Learn Environ 1–15. https://doi.org/10.1080/10494820.2021.2015693

Acknowledgements
This research was supported by the 2022 National Social Science Foundation Education Youth Project "Research on the Strategy of Creating Learning Space Value and Empowering Classroom Teaching under the background of 'Double Reduction'" (Grant No. CCA220319).

Author contributions
Conceptualization: HW, CW; methodology, XC, CW; software, HW, CW; writing—original draft preparation, YJ, CW; writing—review and editing, YJ, XC, CW; supervision, YJ, CW; project administration, CW; funding acquisition, YJ. All authors read and approved the final manuscript. All authors have read and approved the re-submission of the manuscript.

Competing interests
The authors declare no competing interests.

Ethical approval
The study obtained approval from the Research Ethical Board at Zhejiang University of Technology. Although no specific approval number was assigned, the research was conducted in accordance with the ethical guidelines set forth by the university. Simultaneously, in studies involving human participants, all procedures adhered to the ethical standards of the institutional and/or national research committee, as well as the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed consent
The researcher sought and gained the consent of the participants to take part in the study. Out of the 53 sampled participants, all 53 accepted and voluntarily participated in the study after the researcher assured them of anonymity and that their responses were solely for academic purposes.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-024-02751-w.

Correspondence and requests for materials should be addressed to Chengliang Wang.