Using AI Chatbots in Visual Programming: Effect on Programming Self-Efficacy of Upper Primary School Learners

Zipei Zhu[1], Zhaoji Wang[1,*], and Hanhui Bao[2]

[1]Faculty of Education, The University of Hong Kong, Hong Kong SAR, China
[2]Department of Theory and Practice in Teacher Education, The University of Tennessee, Knoxville, USA
Email: zhuzipei@connect.hku.hk (Z. Z); wangzhoj@connect.hku.hk (Z.W); hbao4@vols.utk.edu (H.B.)
*Corresponding author
Manuscript received September 6, 2024; revised September 18, 2024; accepted October 25, 2024; published January 9, 2025

***Abstract*—Programming Self-Efficacy (PSE) is crucial for enhancing learners' skills, cognitive abilities, and career interests. Despite its significance, existing research often overlooks strategies to boost PSE among upper primary school learners. This study evaluates the impact of an innovative e-learning tool, AI Chatbots, on students' PSE. Conducted in a primary school in northern mainland China, the experiment integrated AI Chatbots as an intervention within a visual programming curriculum. It involved 98 fifth-grade students divided into experimental and control groups, both instructed by the same teacher. Interviews were conducted with 6 selected participants from the experimental group after completing all intervention procedures. Findings suggest that, compared to traditional teacher-led instruction, AI Chatbots significantly enhance students' PSE in programming skill, while boost in PSE in programming knowledge remains non-significant. The study also investigates the mechanisms through which AI Chatbots provide students with easily accessible, personalized learning support for self-directed learning and reduce emotional barriers when they seek help.**

***Keywords*—programming self-efficacy, AI-assisted learning, artificial intelligence, visual programming**

I. Introduction

Traditional e-learning frequently fails to accommodate students' individualized learning styles and lacks the specialized functionalities necessary to meet academic requirements [1]. Generative Artificial Intelligence (Generative AI), an emerging e-learning technology within computer science, has demonstrated significant potential when integrated with education, encompassing aspects such as teaching, management, learning, and assessment [2]. AI Chatbots, which leverage Natural Language Processing (NLP) and Natural Language Understanding (NLU) techniques, can offer professional and personalized learning experiences through virtual dialogues for learners [3]. In the wake of considerable advancements and pervasive adoption, AI Chatbots have garnered heightened prominence in the realm of education, facilitating student productivity by providing solutions to queries that emerge throughout the learning experience [4, 5].

Programming, a practical component of the Computer Science (CS) curriculum, is widely used for different stages of K-12 education systems [6–8]. Considering the hierarchical representation of algorithms and the sequential arrays of data procedures, the use of visual displays in programming has increasingly been recognized as an effective tool for illustrating and processing programming languages [9, 10]. Many researchers and educators have expressed a positive attitude towards this method, deeming it suitable for novice students of various ages to learn programming [11]. It is important to consider how to further assist beginners in adapting to visual programming. Additionally, AI Chatbots have been previously applied to teaching various programming languages, including C++ and Python, demonstrating the potential of AI-generated tools to support programming education [12, 13]. This study aims to explore the effectiveness of AI Chatbots in a broader range of programming learning.

In the process of learning programming, self-efficacy is closely related to learning outcomes, influencing learners' attitudes and motivation toward future programming studies [14, 15]. Consciously fostering students' self-efficacy contributes to strategically building academic resilience, facilitating students' personalized development [16]. Most of the prior studies have predominantly focused on the effectiveness of AI Chatbots in enhancing the learning process and improving student academic performance [17, 18]. However, there has been a lack of attention on strengthening the connections between computer science education and desired affective outcomes, particularly at the elementary school level. This study focuses on programming self-efficacy (PSE) as a desired affective outcome and utilizes AI Chatbot as an intervention tool to investigate its effectiveness in cultivating PSE in programming courses. It aims to provide new insights for in-service teachers and policymakers seeking to enhance students' emotional engagement in computer science.

II. Literature Review

A. Programming Self-Efficacy (PSE)

Self-Efficacy (SE), an individual's belief in their ability to handle a given situation, exhibits a robust sense of subjective affect in the face of barriers and challenges [19]. It can facilitate changes in both physical and psychological behaviors within a specific domain [20]. In the educational field, it has been widely accepted that SE has a positive influence on the aspect of academic achievement, including learning attainment, motivation, satisfaction, skill achievement, and so on [21–24].

In the context of CS education, there has been significant exploration and discovery regarding the relationship between SE and programming learning [25, 26]. as Many scholars have advocated for incorporating PSE into educational research. For example, Gordon and O'Rourke indicated that

PSE may impact university students' self-assessment and expectations, which further impact dropout rates [27]. Kanaparan *et al*. confirmed the positive relationship between PSE and perceived enjoyment during programming as established in previous studies and emphasized that programming educators should pay close attention to students' affective demands [28]. Abdunabi *et al*. stated that the increase in PSE can reduce anxiety in the learning process of the program, assisting with course implementation [29]. In light of the findings above, PSE is an indicator of students' interest in the field of programming and is able to prompt the development of students' skills and thinking in the digital era [30].

Furthermore, prior research has suggested that the instructional strategies and tools used in programming teaching influence learners' PSE [31]. Rojas *et al*. found that gamification in programming can improve students' PSE in a quasi-experimental study [32]. Recent advancements in teaching methodologies and technologies have led Wei *et al*. to ascertain that the introduction of partial pair programming significantly enhances PSE [33]. In contrast, Arslan *et al*. evaluated the attitudes and PSE of sophomore students using Arduino as an intervention, revealing that while there were notable improvements in attitude, enhancements in PSE were comparatively less pronounced [34]. These findings highlight the importance of both teaching strategies and assistive tools in effectively influencing students' PSE.

B. The Effect of AI Chatbots on Programming Learning

AI Chatbots accept text-based queries from users and provide answers through messaging techniques, thereby enabling dialogues within relevant contexts under the Chatbot system [35]. Previous studies have demonstrated that utilizing AI Chatbots as both teacher and peer agents enhances student motivation, provides customized educational resources, and offers personalized feedback and guidance within a supportive learning environment [36]. In this case, the advantages of AI Chatbots have been shown in programming learning. Ahmed *et al*. stated that AI Chatbots demonstrate considerable potential in programming education by effectively conveying coherent professional concepts [37]; Savelka *et al*. assessed the AI Chatbots' abilities in evaluating students' academic performance in higher education programming courses [38]; Haindl and Weinberger employed ChatGPT as an assistant tool in facilitating Java programming courses at the undergraduate level, and they found that this AI tool significantly influences the scaffolding processes in the programming learning experience [39]. In other words, this means that AI Chatbots have the potential to assist students and support teachers' programming pedagogical needs to some degree in programming education.

In summary of prior research, while the potential of AI Chatbots is recognized, some gaps are obvious. First, most studies on AI-assisted learning have focused on secondary and post-secondary programming education, leaving the impact on primary education - a foundational stage for building CS skills - largely under-explored when investigating the effect of AI Chatbots on programming learning. Second, while previous research has primarily focused on cognitive learning outcomes facilitated by AI Chatbots, there has been limited investigation into affective learning outcomes, including PSE. Furthermore, even fewer studies delved into the mechanisms by which these outcomes are influenced by AI Chatbots as interventions. To address these research gaps, the following research questions are proposed:

RQ1: Does the adoption of AI Chatbots positively impact programming self-efficacy among upper primary students in visual programming?

RQ2: If so, what mechanisms underlie the influence of AI Chatbots on students' programming self-efficacy?

III. METHODOLOGY

A. Participants

This study was conducted in a primary school in northern mainland China, involving 98 fifth grade students. Participants were selected from two classes, with one class designated as the experimental group (EG) consisting of 48 students and the other as the control group (CG) with 50 students. 77 out of them successfully completed the study, including 36 ($M_{EG} = 11$, $SD_{EG} = 11.33$) assigned to EG and 41 ($M_{CG} = 11$, $SD_{CG} = 11.39$) assigned to CG. Participant details are presented in Table 1. Both groups had approximately one year of similar learning experience in visual programming but no prior exposure to AI Chatbots, ensuring they had a comparable foundation in visual programming. The same computer science teacher instructed both groups to maintain consistency in teaching methods.

Table 1. Demographics of participants

Participant information	Experimental Group (EG)		Control Group (CG)	
	N	%	N	%
All	36	46.75	41	53.25
Gender				
Female	21	58.33	22	53.66
Male	15	41.67	19	46.34
Age				
11	23	63.89	21	51.22
12	8	22.22	15	36.59
Others	5	13.89	5	12.20

B. Procedure

The experiment was carried out from April 1 to May 25, spanning eight weeks and encompassing all required steps (see Fig. 1). In the initial week, the researcher and teacher administered pre-tests and provided instructions on utilizing AI Chatbots. From the second to the seventh week, participants in both the EG and CG were taught the same programming course using Kitten Editor Coding of Codemao. The distinctive feature of the teaching method was the use of AI Chatbots to assist students and teachers with teaching tasks in the EG, whereas the CG adhered to traditional teaching methods without any technological intervention. Specifically, in the EG, the teacher provided students with AI-generated inquiry prompts, allowing them to use AI Chatbots freely during practice sessions to ask questions related to operational steps, knowledge points, and more. Moreover, the researchers and teacher conducted regular weekly reviews of the lessons to discuss aspects that required attention. In the final week, researchers carried out face-to-face interviews with six students, following the completion of the post-tests.

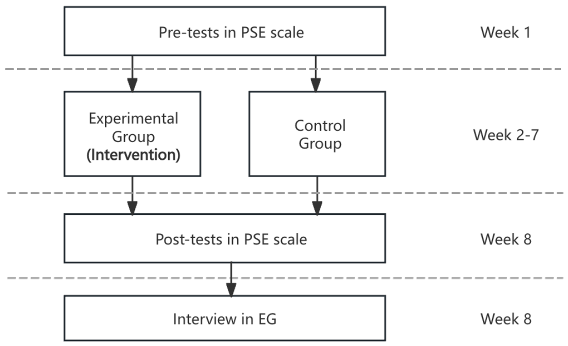

Fig. 1. Flow diagram of the experimental design.

During the 6-week intervention, similar teaching procedures, consisting of 4 teaching steps, were followed in the programming course once a week. Table 2 outlines the visual programming course schedule for the Experimental Group (EG), with each session lasting 40 minutes. In each session, students were assigned specific programming tasks aligned with the programming concepts they had learned. They were required to complete these tasks independently within a set time frame following the given objectives. The tasks were designed to be within the students' Zone of Proximal Development (ZPD), ensuring that most students needed external assistance as scaffolding.

Participants in the EG were allowed to seek help from AI Chatbots when they encountered challenges or confusion. Each student in the EG was provided with a computer, allowing them to either follow researcher-designed prompts or ask their own questions to the Chatbot until they understood the task or knew how to proceed. In contrast, participants in the Control Group (CG) could only seek assistance from the teacher, simulating a traditional classroom environment without AI support. Both EG and CG students were required to design, test, debug, and demonstrate their programming codes.

Table 2. The schedule of programming courses for EG students

Course Session	Time	Details
Review	5 min	The teacher led students to review the core background knowledge.
Project Introduction	10 min	The teacher introduced the background, framework, and course tasks of the programming project to students with prompts for using AI Chatbots.
Programming Exercise	20 min	Students tried to run the program independently by asking for assistance with AI Chatbots.
Presentation & Conclusion	5 min	The teacher summarized the common problems, encouraged students to present their work for class, and asked them to upload the screenshots to the system.

C. *Instruments*

1) *Instructional Tools During Intervention*

As the first generation of AI Chatbots developed by Baidu, Ernie Bot has a strong user base in China [40]. Despite the limitation of regional AI policies [41, 42], it has exhibited fluent response capabilities and excellent information processing skills in answering professional questions across various fields [43, 44]. Hence, Ernie Bot was chosen as the Chatbot tool for the experiments in this study.

In addition, the study adopted Kitten Editor Coding of Codemao as the visual programming editor (see Fig. 2). This editor incorporated game-based learning elements, such as interactive and entertaining animations and modular components, making it well-suited for programming beginners [45].

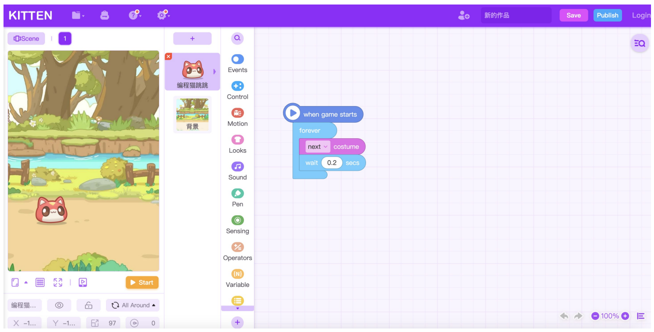

Fig. 2. The interface of programming editor: Kitten editor coding of Codemao.

2) *Programming Self-Efficacy (PSE) scale*

The programming self-efficacy questionnaire in this study was adapted from Wei's Programming Self-Efficacy Scale, which was specifically developed for primary school learners and has been validated in previous research, showing satisfactory reliability and validity [33]. This scale consists of 15 items measured on a 5-point Likert scale, grouping into two sub-dimensions: self-efficacy in programming knowledge and self-efficacy in programming skills respectively. Items 1 to 7 measure learners' self-efficacy in programming knowledge, defined as learners' self-efficacy in programming concepts and knowledge that learners may apply during coding [33]. Items 8 to 15 measure learners' self-efficacy in programming skills, defined as learners' self-efficacy in programming practices related to solving problems with a programming language, such as incrementally and iteratively building solutions, testing and debugging, reusing and remixing code, and applying abstraction and modularization techniques [33]. To ensure the questionnaire provided understandable questions to elementary students, the research team translated and revised the original instrument.

The revised instrument demonstrated satisfactory reliability with Cronbach's alpha value of 0.78 for programming knowledge and 0.88 for programming skills. Additionally, the validity of the instrument was assessed by Confirmatory Factor Analysis (CFA), with all factor loadings ranging from 0.74 to 0.78 for programming knowledge and 0.71 to 0.82 for programming skills. The model fit indices showed that CFI was 0.957 and TLI was 0.936, RMSEA was 0.058, χ^2/df was 1.60 ($p = 0.047$). CFI and TLI values above 0.90 indicate an acceptable fit, while RMSEA below 0.06 indicates a good fit [46]. Therefore, both CFI and TLI indicated a good fit of the model, while a slightly significant p-value can be attributed to the sensitivity of the chi-square test on sample size. Alternatively, the χ^2/df below 3 suggests a good model fit in education research [47].

3) *Semi-structured interview form*

This study also investigated what mechanisms underlie the influence of AI Chatbots on students' programming self-efficacy. To achieve this goal, the researchers adopted the semi-structured interview method. The interview consisted of 10 sample questions about students' experience of completing a programming project. These questions centered on students' perceptions of AI Chatbots, their emotional responses during project execution, the challenges they encountered, and their strategies for addressing programming tasks. The following are some items from this section of questions:

- What difficulties did you encounter during this project? Could you provide an example?
- In what ways did AI Chatbots help you in overcoming these difficulties?
- After using AI Chatbots for some time, how has your confidence or attitude towards problem-solving changed (even after multiple attempts)?

D. Data Analysis

Data from two sources were gathered during the programming project. The objective was to employ a mixed-methods approach to provide a comprehensive exploration of the research questions posited. Quantitative analysis was utilized for RQ1, while RQ2 was addressed using qualitative analysis.

Initially, data collection consisted of gathering information via questionnaires regarding students' self-efficacy from both the experimental and control groups. To answer RQ1, the collected data were analyzed using IBM SPSS. The data were examined to determine whether the two groups demonstrated normal distribution in both pre-tests and post-tests. However, as the data did not conform to a normal distribution, this study adopted non-parametric tests. The Mann-Whitney Test was used to test the differences in performances between EG and CG in the pre-tests, while the Kruskal-Wallis Test was employed to test the differences in performances between pre-tests and post-tests in both groups.

To address RQ2, six participants were purposefully selected from the EG after the intervention, representing slight, intermediate, and large changes in programming self-efficacy, with two students from each level. Guided by the six-phase thematic analysis framework proposed by Braun and Clarke [48], this study analyzed the transcription data. Three coders with academic backgrounds in 'STEM Education' and 'Educational Technologies' were invited to participate in the thematic analysis. After familiarizing coders with the data, the present study conducted inductive coding to identify transcription segments related to 'how programming self-efficacy was developed with the AI Chatbot'. Next, coders refined these codes into themes and sub-themes, which are subsequently defined and named, capturing the essence of the qualitative findings and providing a coherent framework for interpretation

The coding scheme is proposed by the Principal Investigator in this study, and it is reviewed and discussed by all authors. To reduce individual bias, two additional coders independently coded the data. Inter-rate reliability was evaluated by Cohen's Kappa, achieving a coefficient of 0.848 ($p < 0.001$), indicating strong agreement [49]. Additionally, field notes were taken during the interviews to capture the researchers' thoughts and reflections, providing additional context to the interview data. Data triangulation was conducted by cross-validating findings through field notes and member checking, further enhancing the validity of the results.

IV. Result

Results from analysis of quantitative and qualitative data answering the RQ1 and RQ2 respectively.

A. Results from Quantitative Analysis

This quantitative part examines PSE by comparing boosts between pre-tests to post-tests across different treatments addressing RQ1.

The Shapiro-Wilk analysis on pre-tests and post-tests for the two groups indicated that the post-test on PSE in knowledge in EG did not follow a normal distribution (p = 0.04 < 0.05). Therefore, non-parametric analysis was employed to compare the changes in PSE across groups.

To make sure the participants had similar baselines in PSE, the Mann-Whitney Test was adopted to compare the baselines of PSE scores across EG and CG (see Table 3) prior to intervention. The results indicate that there is no significant difference in students' PSE before the intervention, with PSE in programming knowledge ($M_{EG} = 4.65$, $SD_{EG} = 1.06$, $M_{CG} = 4.91$, $SD_{CG} = 1.18$, $p = 0.329 > 0.05$) and programming skills ($M_{EG} = 4.16$, $SD_{EG} = 1.28$, $M_{CG} = 4.27$, $SD_{CG} = 1.44$, $p = 0.653 > 0.05$).

Table 3. Results of Mann-Whitney test for PSE before intervention

Programming Self-Efficacy	Experimental Group (EG)		Control Group (CG)		*p*-value between Groups
	Mean	SD	Mean	SD	
Programming Knowledge	4.65	1.06	4.91	1.18	0.329
Programming Skills	4.16	1.28	4.27	1.44	0.653

To compare the PSE changes across groups (see Fig. 3), the Kruskal-Wallis Test was employed to analyze changes in two dimensions of PSE during pre-tests and post-tests.

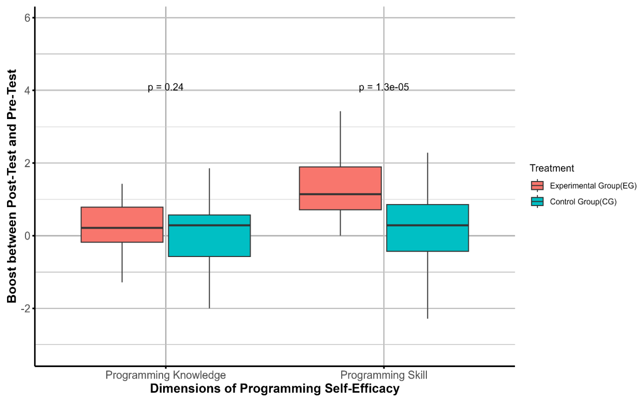

Fig. 3. Boosts of PSE in between Experimental Group (EG) and Control Group (CG).

As shown in Table 4, there is a significant difference in the boost of PSE in programming skills ($p < 0.001$) among EG and CG, with a stronger boost in EG ($n = 36$, $M_{EG} = 1.55$, $SD_{EG} = 1.34$) compare to CG ($n = 41$, $M_{CG} = 0.19$, $SD_{CG} = 1.22$), while the difference in boost of PSE in programming knowledge is not significant ($p = 0.237$). Based on these results, it can be concluded that using AI Chatbots in programming classes positively impacts students' PSE in programming skills compared to employing traditional teacher-led instruction. It indicates that there is no statistically significant difference in the improvement of PSE in programming knowledge between the EG and the CG.

Table 4. Results of the Kruskal-Wallis for changes in PSE during intervention

Programming Self-Efficacy	Experimental Group (EG)		Control Group (CG)		p-value
	Mean	SD	Mean	SD	
Programming Knowledge	0.45	1.11	-0.05	1.21	0.237
Programming Skill	1.55	1.34	0.19	1.22	<0.001

B. Interviews with Students

Based on the interview content, most of the interviewees showed a positive attitude and high willingness to use AI Chatbots during programming learning. Thematic analysis on transcription identified two primary themes and four sub-themes that illustrate the mechanisms of how AI Chatbots influence PSE, as detailed in Table 5. The 'N of Responses' refers to the number of respondents who mentioned codes during the interview.

Table 5. Coding scheme: How AI Chatbot influence learners' programming self-efficacy

Main Theme	Sub-Theme	Definition	N of Responses
Support on Content Knowledge and Skills	Enhancing Accessibility of Effective Learning Support	AI Chatbots provide learners with immediate, personalized assistance, making students accessible to guidance, feedback, and resources that are essential for programming.	6
	Facilitating Access to Extended Knowledge	AI Chatbots provide extended information and resources beyond the programming task itself, helping learners deal with programming tasks divergently.	2
Support on Meta-cognitive Awareness and Emotional Well-being	Promoting Self-Directed Learning	AI Chatbots boost students the flexibility and autonomy with controlling their learning progress, allowing them to manage their learning independently.	4
	Reducing Social and Emotional Barriers	AI Chatbots create a supportive, non-judgmental environment where learners feel comfortable repeating questions or making mistakes.	3

1) Support on content knowledge and skills

This theme highlights how AI Chatbots enhance students' programming self-efficacy by supporting the acquisition of content knowledge and skills. Specifically, AI Chatbots primarily achieve this by improving access to effective learning resources and extending knowledge beyond what is immediately available.

Several students expressed that AI Chatbots offer accessible and convenient resources for programming tailored to their individual needs. For instance, S4 noted, "If you can ask more clearly, AI Chatbots will give you the answer directly, and then you can easily find the module code. " S6 affirmed, "AI Chatbots always give me ideas and suggestions to think about like telling me which step to start with." S3 shared a specific experience, "When I can't find where a modular code operation is, I can ask AI Chatbots, and it gives a clear answer like where this code should go." S2 also added, "The teacher's explanations are sometimes quick and easy to forget, but I can revisit the answers from AI Chatbots." S1 shared a similar perspective. Additionally, S4 highlighted another advantage: "Everyone has different questions, so asking AI Chatbots saves me time compared to asking the teacher." This suggests that AI Chatbots provide an efficient, time-saving solution that tailors individual needs and overcomes the limitations of relying on a single teacher.

Another important aspect of this theme is the AI Chatbots' ability to facilitate access to extended knowledge. This includes offering broader insights and alternative suggestions that guide students through more complex tasks, such as different programming methods or new coding ideas. S4 noted, "AI Chatbots provide more information like programming tips beyond just the codes, which makes me more confident in programming." S3 elaborated further, stating, "With AI Chatbots, I feel I understand these codes and programming better. It explains the whole process in detail, unlike the teacher who just gives a simple explanation."

2) Support on meta-cognitive awareness and emotional well-being

In terms of promoting meta-cognitive awareness and emotional well-being, two key codes emerged under this theme: encouraging self-directed learning and reducing social and emotional barriers.

Students emphasized the value of AI Chatbots in fostering Self-Directed Learning (SDL) as they provide the freedom to develop programming skills independently. This allowed students to make informed decisions about their learning progress, including what to learn and how to approach it. S1 expressed a similar view: "It helps me watch and do things at the same time. For example, after looking at the answers from AI, I can find what I need to do and follow each step one by one." This aligns with the principles of SDL, where learners actively diagnose their own learning needs and take charge of their progress [50]. S2 reflected, "AI Chatbots let me explore more. I might analyze things more and think harder. Compared to the teacher's help, this feels different. I feel more capable of completing the programming myself, which gives me a strong sense of achievement." S5 echoed this sentiment noting, "If I use AI Chatbots, I can ask more questions on my own," reinforcing how the technology supports learners' autonomy in seeking further understanding. In particular, S6 saw AI Chatbots as a personal assistant stating, "After AI Chatbots give me answers, I feel like I did it independently, without relying on anyone else, which boosts my confidence." This growing sense of independence is crucial for developing SDL as AI Chatbots help build confidence in the ability to tackle and complete tasks without relying on human assistance.

The second aspect identified was the ability of AI Chatbots to reduce social and emotional barriers. Students described feeling less pressure and more at ease when interacting with AI because the technology provides a non-judgmental environment for learning. For example, S1 commented, "AI Chatbots can explain a problem patiently over and over again. The robot won't get annoyed." S2 expressed a similar view: "Compared to teachers, I feel that AI will not have emotional issues. At least I won't get scolded." S4 added an insightful observation: "AI Chatbots give me a more private space to ask questions, especially since I'm shy. Sometimes I'm afraid my classmates will laugh at me for asking silly questions, so it helps protect my self-esteem." This suggests that students value privacy and boundaries in their learning experiences, which can influence their sense of self-efficacy.

V. Discussion

To summarize, we found the following findings:

Finding 1: The use of AI Chatbots significantly enhanced learners' PSE in programming skills compared to traditional teacher-led instruction.

Finding 2: In contrast, AI Chatbots did not significantly improve learners' PSE in programming knowledge relative to traditional teacher-led instruction.

Finding 3: AI Chatbots contribute to learners' PSE by providing easily accessible, personalized learning support for self-directed learning and by reducing emotional barriers to seeking help.

Findings 1&2 are interesting since they indicate the different effects of AI Chatbots on two sub-dimensions of PSE. The effect of AI Chatbots on learners' PSE in programming skills is significant compared to the CG, which is consistent with prior research by Yilmaz *et al.* conducted among undergraduate students [51]. However, the boosts of PSE in programming knowledge do not show a significant difference between the groups.

To explain these findings, we combined our qualitative data with conclusions from prior studies. Firstly, in programming courses typically structured around project-based topics, students are often assigned specific programming tasks to complete independently. The qualitative results indicate that AI Chatbots primarily provide direct instructions on programming tasks when requested by students, rather than introducing underlying concepts or theoretical knowledge, which implies that the AI Chatbots tend to explain concepts or terminologies only when participants directly ask for clarifications. The response paradigm of AI Chatbots may explain the disparity in its effect on PSE in knowledge and skills.

Secondly, based on the theoretical framework of Programming Self-Efficacy proposed by Kong [52] and Bloom's Taxonomy [53], programming skill can be understood as "knowing how to do," which corresponds to "procedural knowledge," while programming knowledge refers to "knowing what it is," aligning with "declarative knowledge." Procedural knowledge, such as using loops, debugging code, or employing specific functions, involves hands-on practice and iteration. In our study, AI Chatbots provided personalized, repetitive interactions, helping learners enhance their procedural knowledge and thereby increasing their PSE in programming skills, consistent with the findings of the previous study [54]. Conversely, declarative knowledge often requires scaffolded learning, where learners gradually build on prior understanding to form a complete conceptual framework. This form of learning requires more nuanced and adaptive teaching techniques, which traditional human teachers are generally better equipped to provide compared to AI Chatbots [55]. Therefore, AI Chatbots showed no significant effect on PSE in programming knowledge.

To address the disparity between PSE skill and knowledge, a specially trained AI Chatbot for programming learning is necessary, which may provide memory adaptive functions for scaffolded learning to students to support their PSE in knowledge.

Finding 3 reveals the underlying mechanism of how AI Chatbots influence learners' PSE.

Firstly, AI Chatbots help alleviate emotional barriers such as social anxiety that students may experience when seeking help. Some interviewees reported feeling ashamed or anxious about being criticized by teachers or ridiculed by peers when asking for help, which negatively impacted their PSE. Bandura posits that self-efficacy is influenced by personal experiences, vicarious experiences, verbal persuasion, and physiological states, when students fear criticism or judgment from teachers or peers [56]. Prior research also revealed that primary school students' social anxiety when seeking help from others is particularly salient in contemporary classroom instruction [57]. Consistent with Stowell *et al.*, students often favor anonymous idea exchange, particularly when experiencing feelings of embarrassment and anxiety in public settings [58]. By providing anonymous and non-judgmental feedback, AI Chatbots create a psychologically safe environment for students, which enhances their self-efficacy [59].

Secondly, AI Chatbots provide students with greater flexibility in regulating their learning process. Students can ask the AI to re-explain concepts, clarify specific lines of code, or explore alternative solutions at their own pace, which is difficult for human instructors to offer due to time and resource constraints. According to the Self-Determination Theory [60], when learners feel autonomous, competent, and supported, their self-efficacy increases. By allowing students to control their learning pace, AI Chatbots directly contribute to fostering these key elements, thereby promoting greater programming self-efficacy.

Additionally, AI Chatbots not only provide answers but also suggest extended content and alternative solutions. This exposure encourages students to think divergently, exploring multiple approaches to a problem. As a result, students' programming self-efficacy is enhanced because they gain a sense of mastery by understanding and applying diverse problem-solving methods. As suggested by Bandura's Social

Cognitive Theory, mastery experience is the most powerful source of self-efficacy, it derives from learners' feeling of mastery of multiple solutions and it strengthens learners' PSE because it demonstrates that learners can finish the programming task [61–64].

VI. Conclusion

This study demonstrates the effectiveness of AI Chatbots in enhancing Programming Self-Efficacy (PSE) among upper primary school students. Specifically, it highlights the different effects of AI Chatbots on two sub-dimensions of PSE. Furthermore, we investigated the mechanism underlying how AI Chatbots influence PSE through a combination of empirical data and theoretical perspectives. These findings have implications for policymakers and teachers. To be specific, integrating AI-assisted learning into programming curricula may yield significant improvements in students' PSE.

Additionally, our findings provide insights into the comparative strengths and limitations of AI Chatbots versus human teachers when teaching programming in primary schools. Apart from the desirable outcomes of using AI Chatbots as an educational intervention, there are several concerns regarding the adoption of AI tools in teaching. Firstly, excessive reliance on AI tools may hinder students' ability to develop effective information retrieval skills. Compared to "Googling it" or "looking it up in worksheets," asking AI directly acts as a shortcut to obtaining "truth" for primary school students. In the long term, this approach may lead to an over-reliance on AI and hinder students' critical information retrieval skills.

Secondly, adopting AI tools without proper supervision can lead to issues such as plagiarism, where students may directly copy and paste answers without fully understanding the content [65]. To address these concerns, further efforts should be directed toward developing educational AI Chatbots specifically designed to promote learning while discouraging academic dishonesty and fostering essential research skills.

Several limitations of this study should be acknowledged. First, although the intervention was initially planned for a full 12-week semester, unforeseen circumstances, such as cancellations due to hardware issues or class time being reallocated to other subjects, led to some sessions not being delivered as planned. Additionally, the study was conducted in a single school in China, which may introduce biases and limit the generalizability of the findings. Furthermore, potential covariates such as age, gender, parental educational level, and socioeconomic status were not fully controlled, which may have impacted the accuracy of the results [66]. Future research should explore the effects of participants' demographic information and extend the intervention duration to improve the robustness of findings.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Zipei Zhu conducted the research, data analysis and wrote the paper. Zhaoji Wang conducted the data analysis. Hanhui Bao did the proofreading. All authors had approved the final version.

Declarations

A. Ethical Approval and Consent to Participate

This research project has received IRB approval from Research Ethics Committee of Faculty of Education, The University of Hong Kong.

B. Consent for Publication

Participation in the survey implied consent.

Acknowledgment

We would like to express our sincere gratitude to Neal Alexander Monroe for his valuable advice on our manuscripts.

References

[1] A. Z. Al Rawashdeh, E. Y. Mohammed, A. R. Al Arab, M. Alara, and B. Al-Rawashdeh, "Advantages and disadvantages of using e-learning in university education: Analyzing students' perspectives: EJEL," *Electronic Journal of e-Learning,* vol. 19, no. 3, pp. 107–117, 2021.

[2] Z. Hu and H. Wang, "Relevant concepts, key technologies and typical application scenarios of AI + education," *Journal of Physics: Conference Series,* vol. 1861, no. 1, 2021. doi: 10.1088/1742-6596/1861/1/012059

[3] M. Aleedy, E. Atwell, and S. Meshoul, "Using AI Chatbots in education: Recent advances challenges and use case," in *Artificial Intelligence and Sustainable Computing*, Singapore, M. Pandit, M. K. Gaur, P. S. Rana, and A. Tiwari, Eds., Springer Nature Singapore, 2022, pp. 661–675.

[4] L. Liu, R. Subbareddy, and C. G. Raghavendra, "AI intelligence Chatbot to improve students learning in the higher education platform," *Journal of Interconnection Networks,* vol. 22, no. 2, 2143032, 2022. doi: 10.1142/S0219265921430325

[5] T. A. Baha, M. E. Hajji, Y. Es-Saady, and H. Fadili, "The impact of educational chatbot on student learning experience," *Education and Information Technologies,* pp. 1–24, 2023. doi: 10.1007/s10639-023-12166-w

[6] C. Angeli and K. Georgiou, "Investigating the effects of gender and scaffolding in developing preschool children's computational thinking during problem-solving with Bee-Bots," *Frontiers in Education (Lausanne),* vol. 7, 2023. doi: 10.3389/feduc.2022.757627

[7] L. El-Hamamsy *et al.*, "A computer science and robotics integration model for primary school: evaluation of a large-scale in-service K-4 teacher-training program," *Education and Information Technologies,* vol. 26, no. 3, pp. 2445–2475, 2021. doi: 10.1007/s10639-020-10355-5

[8] E. A. Kassa and E. A. Mekonnen, "Computational thinking in the Ethiopian secondary school ICT curriculum," *Computer Science Education,* vol. 32, no. 4, pp. 502–531, 2022. doi: 10.1080/08993408.2022.2095594

[9] A. Prasad, K. Chaudhary, and B. Sharma, "Programming skills: Visualization, interaction, home language and problem solving,", *Education and Information Technologies,* vol. 27, no. 3, pp. 3197–3223, 2022. doi: 10.1007/s10639-021-10692-z

[10] M. Mladenović, Ž. Žanko, and M. Aglić Čuvić, "The impact of using program visualization techniques on learning basic programming concepts at the K–12 level," *Computer Applications in Engineering Education,* vol. 29, no. 1, pp. 145–159, 2021.

[11] Kanika, S. Chakraverty, and P. Chakraborty, "Tools and techniques for teaching computer programming: A review," *Journal of Educational Technology Systems,* vol. 49, no. 2, pp. 170–198, 2020. doi: 10.1177/0047239520926971

[12] J. Nayak, T. Keane, T. Linden, and A. Molnar, "Teaching high school students artificial intelligence by programming Chatbots," in *Teaching Coding in K-12 Schools: Research and Application*, T. Keane and A. E. Fluck Eds., Cham: Springer International Publishing, 2023, pp. 263–276.

[13] V. L. M. Nguyen *et al.*, "Codebot—A Vietnamese Chatbot system for answering C++ and python-related questions," *Journal of Science,* vol. 18, no. 9, pp. 1711–1723, 2021.
[14] A. Kovari and J. Katona, "Effect of software development course on programming self-efficacy," *Education and Information Technologies,* vol. 28, no. 9, pp. 10937–10963, 2023. doi: 10.1007/s10639-023-11617-8
[15] H. Özyurt and Ö. Özyurt, "Analyzing the effects of adapted flipped classroom approach on computer programming success, attitude toward programming, and programming self-efficacy," *Computer Applications in Engineering Education,* vol. 26, no. 6, pp. 2036–2046, 2018.
[16] T.-T. Wu, L. M. Silitonga, B. Dharmawan, and A. T. Murti, "Empowering students to thrive: The role of CT and self-efficacy in building academic resilience," *Journal of Educational Computing Research,* vol. 62, no. 3, pp. 816–845, 2024. doi: 10.1177/07356331231225468
[17] J. Rajala, J. Hukkanen, M. Hartikainen, and P. Niemelä, "\" call me Kiran\"—ChatGPT as a tutoring Chatbot in a computer science course," in *Proc. the 26th International Academic Mindtrek Conference*, 2023, pp. 83–94.
[18] S. I. Malik, M. W. Ashfaque, R. Mathew, J. Jabbar, R. S. Al-Nuaimi, and A. Alsideiri, "Fostering the learning process in a programming course with a Chatbot," *International Journal of Online Pedagogy and Course Design (IJOPCD),* vol. 12, no. 1, pp. 1–17, 2022.
[19] A. Bandura and S. Wessels, *Self-Efficacy,* 1994.
[20] S. Schweder, "The role of control strategies, self-efficacy, and learning behavior in self-directed learning," *International journal of school & educational psychology,* vol. 7, no. 1, pp. 29–41, 2019. doi: 10.1080/21683603.2018.1459991
[21] T. Sun and C. Wang, "College students' writing self-efficacy and writing self-regulated learning strategies in learning English as a foreign language," *System,* vol. 90, 102221, 2020. doi: 10.1016/j.system.2020.102221
[22] D. Hampton *et al.*, "Self-efficacy and satisfaction with teaching in online courses," *Nurse Educator,* vol. 45, no. 6, pp. 302–306, 2020. doi: 10.1097/NNE.0000000000000805
[23] K. M. Mitchell, S. Zumbrunn, D. N. Berry, and L. Demczuk, "Writing self-efficacy in postsecondary students: A scoping review," *Educational psychology review,* vol. 35, no. 3, p. 82, 2023. doi: 10.1007/s10648-023-09798-2
[24] M. Trautner and M. Schwinger, "Integrating the concepts self-efficacy and motivation regulation: How do self-efficacy beliefs for motivation regulation influence self-regulatory success?" *Learning and Individual Differences,* vol. 80, 101890, 2020. doi: 10.1016/j.lindif.2020.101890
[25] L. Malmi, J. Sheard, P. Kinnunen, Simon, and J. Sinclair, *Theories and Models of Emotions, Attitudes, and Self-Efficacy in the Context of Programming Education*, New York, NY, USA, ACM, 2020, pp. 36–47. doi: 10.1145/3372782.3406279
[26] P. Steinhorst, A. Petersen, and J. Vahrenhold, "Revisiting self-efficacy in introductory programming," presented at the Proceedings of the 2020 ACM Conference on International Computing Education Research, Virtual Event, New Zealand, 2020.
[27] J. Gorson and E. O'Rourke, "Why do CS1 students think they're bad at programming?: Investigating self-efficacy and self-assessments at three universities," New York, NY, USA, ACM, 2020, pp. 170–181. doi: 10.1145/3372782.3406273
[28] G. Kanaparan, R. Cullen, and D. Mason, "Effect of self-efficacy and emotional engagement on introductory programming students," *Australasian Journal of Information Systems,* vol. 23, 2019. doi: 10.3127/ajis.v23i0.1825
[29] R. Abdunabi, I. Hbaci, and H.-Y. Ku, "Towards enhancing programming self-efficacy perceptions among undergraduate information systems students," *Journal of Information Technology Education: Research,* vol. 18, 2019.
[30] T. Kantathanawat, S. Chinchua, and S. Tuntiwongwanich, "Programming self-efficacy in higher education research: A systematic literature review," *Journal of Higher Education Theory and Practice,* vol. 23, no. 19, pp. 19–28, 2023. doi: 10.33423/jhetp.v23i19.6673
[31] K. Davidson, L. Larzon, and K. Ljunggren, "Self-efficacy in programming among STS students," *Retrieved August,* vol. 12, p. 2013, 2010.
[32] M. E. Ortiz Rojas, K. Chiluiza, and M. Valcke, "Gamification in computer programming: Effects on learning, engagement, self-efficacy and intrinsic motivation," in *Proc. 11th European Conference on Game-Based Learning (ECGBL)*, Acad Conferences LTD, 2017, pp. 507–514.
[33] X. Wei, L. Lin, N. Meng, W. Tan, S.-C. Kong, and Kinshuk, "The effectiveness of partial pair programming on elementary school students' computational thinking skills and self-efficacy," *Computers & Education,* vol. 160, 104023, 2021. doi: 10.1016/j.compedu.2020.104023
[34] K. Arslan and Z. Tanel, "Analyzing the effects of Arduino applications on students' opinions, attitude and self-efficacy in programming class," *Education and Information Technologies,* vol. 26, no. 1, pp. 1143–1163, 2021. doi: 10.1007/s10639-020-10290-5
[35] S. A. Thorat and V. Jadhav, "A review on implementation issues of rule-based chatbot systems," in *Proc. the International Conference on Innovative Computing & Communications (ICICC)*, 2020.
[36] M. A. Kuhail, N. Alturki, S. Alramlawi, and K. Alhejori, "Interacting with educational chatbots: A systematic review," *Education and Information Technologies,* vol. 28, no. 1, pp. 973–1018, 2023. doi: 10.1007/s10639-022-11177-3
[37] Z. Ahmed, S. S. Shanto, and A. I. Jony, "Potentiality of generative AI tools in higher education: Evaluating ChatGPT's viability as a teaching assistant for introductory programming courses," *STEM Education,* vol. 4, no. 3, pp. 165–182, 2024.
[38] J. Savelka, A. Agarwal, C. Bogart, Y. Song, and M. Sakr, "Can Generative Pre-trained Transformers (GPT) pass assessments in higher education programming courses?" presented at the 2023 Conference on Innovation and Technology in Computer Science Education V. 1, Finland, 2023.
[39] P. Haindl and G. Weinberger, "Students' experiences of using ChatGPT in an undergraduate programming course," *IEEE Access,* vol. 12, pp. 43519–43529, 2024. doi: 10.1109/ACCESS.2024.3380909
[40] J. Rudolph, S. Tan, and S. Tan, "War of the chatbots: Bard, bing chat, ChatGPT, ernie and beyond. The new AI gold rush and its impact on higher education," *Journal of Applied Learning and Teaching,* vol. 6, no. 1, pp. 100–120, 2023.
[41] M. Sheehan, "China's AI regulations and how they get made," *Horizons: Journal of International Relations and Sustainable Development,* no. 24, pp. 108–125, 2023.
[42] M. Romero and A. Urmeneta, *Citizenship, Censorship, and Democracy in the Age of Artificial Intelligence*, Switzerland: Palgrave Macmillan, 2024.
[43] Z. Mi and K. Li, "A comparative analysis of different large language models in evaluating student-generated questions," in *Proc. 2024 13th International Conference on Educational and Information Technology (ICEIT)*, 2024, pp. 24–29. doi: 10.1109/ICEIT61397.2024.10540914
[44] S. Guo and Q. Wang, "Application of knowledge distillation based on transfer learning of ERNIE model in intelligent dialogue intention recognition," *Sensors*, vol. 22, no. 3. doi: 10.3390/s22031270
[45] X. Peng, *The Study of Digital Gamebased Learning on Motivating Chinese Primary Students to Study Mathematics*, 2019.
[46] Y. Xia and Y. Yang, "RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods," *Behavior research methods,* vol. 51, no. 1, pp. 409–428, 2019. doi: 10.3758/s13428-018-1055-2
[47] B. M. Byrne, *Structural Equation Modeling with LISREL, PRELIS, and SIMPLIS: Basic Concepts, Applications, and Programming*, Psychology Press, 2013.
[48] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Research in Psychology,* vol. 3, no. 2, pp. 77–101, 2006.
[49] M. L. McHugh, "Interrater reliability: The kappa statistic," *Biochemia Medica,* vol. 22, no. 3, pp. 276–282, 2012.
[50] B. Bhandari, D. Chopra, and K. Singh, "Self-directed learning: Assessment of students' abilities and their perspective," *Advances in Physiology Education,* vol. 44, no. 3, pp. 383–386, 2020. doi: 10.1152/ADVAN.00010.2020
[51] R. Yilmaz and F. G. K. Yilmaz, "The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation," *Computers and Education: Artificial Intelligence,* vol. 4, 100147, 2023. doi: 10.1016/j.caeai.2023.100147
[52] S.-C. Kong, "Development and validation of a programming self-efficacy scale for senior primary school learners," in *Proc. the International Conference on Computational Thinking Education*, The Education University of Hong Kong, 2017, pp. 97–102.
[53] L. W. Anderson and D. R. Krathwohl, *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives: Complete Edition*, Addison Wesley Longman, Inc., 2001.

[54] A. Bahrini *et al.*, "ChatGPT: Applications, opportunities, and threats," in *Proc. 2023 Systems and Information Engineering Design Symposium (SIEDS)*, IEEE, 2023, pp. 274–279.
[55] H. Rasku-Puttonen, A. Eteläpelto, M. Arvaja, and P. Häkkinen, "Is successful scaffolding an illusion?–Shifting patterns of responsibility and control in teacher-student interaction during a long-term learning project," *Instructional Science,* vol. 31, pp. 377–393, 2003.
[56] S. M. Abdullah, "Social cognitive theory: A Bandura thought review published in 1982–2012," *Psikodimensia: Kajian Ilmiah Psikologi,* vol. 18, no. 1, pp. 85–100, 2019.
[57] M. G. Nakabugo, C. Opolot-Okurut, C. M. Ssebbunga, J. S. Maani, and A. Byamugisha, "Large class teaching in resource-constrained contexts: Lessons from reflective research in Ugandan primary schools," *Journal of International Cooperation in Education,* vol. 11, no. 3, pp. 85–102, 2008.
[58] J. R. Stowell, T. Oldham, and D. Bennett, "Using student response systems ("Clickers") to Combat conformity and shyness," *Teaching of Psychology,* vol. 37, no. 2, pp. 135–140, 2010. doi: 10.1080/00986281003626631
[59] Q. Xia, T. K. Chiu, C. S. Chai, and K. Xie, "The mediating effects of needs satisfaction on the relationships between prior knowledge and self-regulated learning through artificial intelligence chatbot," *British Journal of Educational Technology,* vol. 54, no. 4, pp. 967–986, 2023.
[60] R. M. Ryan and E. L. Deci, "Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being," *American psychologist,* vol. 55, no. 1, p. 68, 2000.
[61] C. Y. Chang, G. J. Hwang, and M. L. Gau, "Promoting students' learning achievement and self-efficacy: A mobile chatbot approach for nursing training," *British Journal of Educational Technology,* vol. 53, no. 1, pp. 171–188, 2022. doi: 10.1111/bjet.13158
[62] A. Bandura, "Self-efficacy: Toward a unifying theory of behavioral change," *Psychological Review,* vol. 84, no. 2, p. 191, 1977.
[63] A. Bandura, "Self-efficacy mechanism in human agency," *American Psychologist,* vol. 37, no. 2, p. 122, 1982.
[64] A. Bandura, "Organisational applications of social cognitive theory," *Australian Journal of Management,* vol. 13, no. 2, pp. 275–302, 1988.
[65] D. R. Cotton, P. A. Cotton, and J. R. Shipway, "Chatting and cheating: Ensuring academic integrity in the era of ChatGPT," *Innovations in Education and Teaching International,* vol. 61, no. 2, pp. 228–239, 2024.
[66] J. A. Kitchen, C. Chen, G. Sonnert, and P. Sadler, "The impact of participating in college-run STEM clubs and programs on students' STEM career aspirations," *Teachers College Record,* vol. 124, no. 2, pp. 117–142, 2022. doi: 10.1177/01614681221086445