ChatGPT in computer programming education: A review of current literature and applications

Maria Ijaz Baig
Faculty of Information Technology, Applied Science Private University, Amman, Jordan

Elaheh Yadegaridehkordi, Ayub Bokani
College of Information and Communications Technology, School of Engineering and Technology, CQUniversity, Sydney, Australia

ChatGPT has gained significant attention in computer programming education due to its advanced capabilities in assisting coding processes and its growing impact on teaching and learning. Despite rapid technological progress and widespread adoption, further research is required to optimise its integration into programming education. This scoping review used the PRISMA–ScR framework to analyse 59 research articles published between 2022 and 2025. The review identified major research areas related to ChatGPT's use in programming education, including its role as a programming assistant, automated assessment and feedback, student and educator perceptions, curriculum design and instructional strategies, learning outcomes and performance, ethical and academic integrity considerations and applications across specific programming domains. It also examined methodological approaches, participant demographics and geographical distribution across the included studies. Findings highlight benefits of integrating ChatGPT, including enhanced student engagement, increased accessibility, support for bridging knowledge gaps and assistance with code optimisation. Meanwhile, challenges include risks of overreliance, reduced critical thinking, accuracy limitations and academic integrity concerns. This review provides practical insights for educators, universities, students and researchers. It emphasises using ChatGPT as a learning assistant, implementing clear policies, tailoring artificial intelligence (AI) tools to diverse student needs and guiding future research on effective and ethical AI-driven programming education.

Implications for practice or policy:

- ChatGPT should support debugging, exploration and collaboration rather than code generation.
- Students must annotate AI outputs, reinforced by oral exams and reflective journals.
- Educators should blend AI feedback with human evaluation through scaffolded, authentic assessments.
- Institutions need clear ethical policies, equitable access and staff training.
- Researchers should use longitudinal, mixed methods studies, while developers design explainable, adaptive and integrity-focused features aligned with course progression.

Keywords: ChatGPT, programming education, systematic review

Introduction

The integration of artificial intelligence (AI) in education offers transformative possibilities (Adıgüzel et al., 2023). Generative AI (GenAI) such as ChatGPT has gained attention for creating human-like text and supporting diverse educational activities (Bewersdorff et al., 2025). ChatGPT, introduced by OpenAI in November 2022 (Bhullar et al., 2024), is a leading GenAI platform. It generates high-quality text, learns continuously and offers a free tier with certain limitations compared to its paid versions (Adeshola & Adepoju, 2023). It supports programming education through coding, explanations and feedback (Yilmaz & Yilmaz, 2023). Programming is challenging in higher education due to abstract concepts and problem-solving demands (Chang et al., 2024). ChatGPT helps bridge this gap by offering real-time support, debugging assistance and code improvements (Rawas, 2024). It also provides contextualised explanations, aiding students in mastering complex programming concepts (Michel-Villarreal et al., 2023). ChatGPT explains code, adds comments for clarity and can translate programs between languages, such as converting Python to Java. It can detect syntax or logical errors, such as missing brackets or incorrect variables and help merge code segments by integrating functions or scripts seamlessly. ChatGPT's instant feedback fosters self-directed learning by helping students quickly identify, correct and refine their programming solutions (Adeshola & Adepoju, 2023; Rahman et al., 2023).

The literature has highlighted ChatGPT's role in enhancing personalised learning, creativity, engagement, interactive coding support and short-term educational outcomes (Baidoo-Anu & Ansah, 2023; Baig & Yadegaridehkordi, 2024; Batista et al., 2024; Sok & Heng, 2023). Chugh et al. (2025) noted that GenAI tools in information and communications technology education have evolved from Codex to more advanced, user-friendly systems such as ChatGPT. Unlike specialised AI coding assistants such as GitHub Copilot or Codex, ChatGPT provides a conversational, multi-turn interface that allows users to generate, debug and explain code using natural language (Nizamudeen et al., 2024). Its accessibility, multilingual flexibility, step-by-step guidance and user-friendly interface make ChatGPT suitable for novice and experienced programmers, driving its adoption in education and professional practice (Akçapınar & Sidan, 2024). Therefore, with growing interest from educational institutions, students and researchers in ChatGPT's potential, a comprehensive examination of the nascent literature on its application in programming education is urgently needed. Reviews have explored ChatGPT's impact on education in general, highlighting benefits, challenges and teaching implications (Baig & Yadegaridehkordi, 2024; Lo, 2023). Some studies have examined ChatGPT's role in programming education, such as Garcia (2025), who conducted a bibliometric analysis to review teaching and learning computer programming using ChatGPT. However, despite growing interest, comprehensive reviews focusing specifically on ChatGPT's role in computer programming education are still lacking. A systematic synthesis of existing studies is therefore needed to consolidate current knowledge and guide future research and pedagogical practice. Meanwhile, researchers have called for comprehensive reviews investigating the educational use of GenAI, particularly studies that explore its potential across diverse disciplines such as computer programming (Bhullar et al., 2024). Thus, this review aimed to examine ChatGPT's role in computer programming education within higher education, focusing on emerging research areas, methodologies, benefits and challenges and offering actionable insights for educators and researchers. The study was guided by the following research questions (RQs):

- RQ1: What are the specific areas of research in relation to the use of ChatGPT in computer programming education?
- RQ2: How do research studies investigate the incorporation of ChatGPT in teaching and learning computer programming?
- RQ3: What are the educational benefits and challenges of using ChatGPT in computer programming education?

This systematic review of research areas, methods, benefits and challenges provides a clear understanding of ChatGPT's role in programming education and guidance for effective integration. Examining prior research areas (RQ1) helps identify the current focus, gaps and emerging themes within this growing field. Investigating research methodologies (RQ2) provides insight into how existing studies have been designed and conducted, which can inform more rigorous and effective future applications of ChatGPT in programming education. Finally, analysing the educational benefits and challenges (RQ3) offers a balanced understanding of ChatGPT's impact, guiding educators, institutions and developers towards responsible and evidence-based implementation. Generally, this study offers valuable insights for educators, researchers, universities and policymakers to effectively implement GenAI technologies in ways that support both student learning and academic integrity.

This review is organised as follows: it begins with the methodology, followed by findings structured by the research questions, then presents a discussion and implications and concludes with limitations.

Research method

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009) enhance systematic review quality by ensuring transparency and completeness through a detailed checklist, helping readers assess validity and reliability. PRISMA Scoping Reviews (PRISMA-ScR; Tricco et al., 2018) expanded the traditional PRISMA framework by allowing broader research questions and more flexible inclusion criteria, unlike standard systematic reviews and meta-analyses. This review followed the PRISMA-ScR guidelines, which provide a structured framework to enhance transparency, consistency and comprehensiveness in scoping review processes. PRISMA-ScR aids educational scoping reviews by ensuring transparency and identifying research gaps (McGowan et al., 2020). Its structured phases of identification, screening, eligibility and inclusion streamline source selection, enhancing clarity and efficiency (Tricco et al., 2018).

Eligibility criteria

Table 1 outlines the inclusion and exclusion criteria for selecting research papers related to ChatGPT in programming. The inclusion process relied on identifying relevant keywords in titles, abstracts or keyword sections. Papers lacking these keywords in these specified sections were excluded to ensure relevance. The review focused on research articles, as they offer original scientific insights and have a significant impact on academic journals. Since GenAI tools like ChatGPT were introduced in November 2022, the review included English-language publications from November 2022 to January 2025. This study explored ChatGPT's role in programming by examining research areas, its integration into education and associated benefits and challenges. Only studies explicitly focused on ChatGPT in programming were included.

Table 1
Inclusion and exclusion criteria

Inclusion	Exclusion
Relevant search terms were identified within the titles, abstracts or keywords sections of the papers.	Research papers lack search terms in the specified sections.
Peer-reviewed journal and conference articles.	Review papers, non-research materials, books, chapters, reports and workshops.
Research papers published in English from November 2022 to January 2025.	Articles published outside this time frame or not in English.
Research articles specifically focused on ChatGPT and programming in higher education contexts.	Research articles are not centred on higher education, ChatGPT and programming.

Search and selection of sources

The search for relevant studies was conducted across Scopus, Web of Science, the ACM Digital Library, the IEEE digital library and the Google Scholar database using a set of keywords combined with Boolean operators: ("ChatGPT") AND ("programming" OR "computer programming" OR "language programming" OR "coding"). The search query was applied to titles, abstracts and keywords to ensure all relevant articles were captured. The database search returned the following records: Scopus (96), IEEE Digital Library (91), ACM Digital Library (78), Web of Science (59) and Google Scholar (72), for a total of 396 records. A total of 291 duplicate records were identified and removed using EndNote. The final screening pool comprised *n* = 105 unique articles. Titles and abstracts were screened to exclude clearly irrelevant records, yielding *n* = 75 articles for full-text review. Each article was read in full and assessed for fit within the higher education context and compliance with all inclusion criteria, resulting in a final set of 59 articles (Figure 1). We independently screened and selected studies, resolving any inclusion or exclusion discrepancies through consensus. After agreeing on the sample pool, we fully reviewed, extracted and synthesised the remaining articles for reporting.

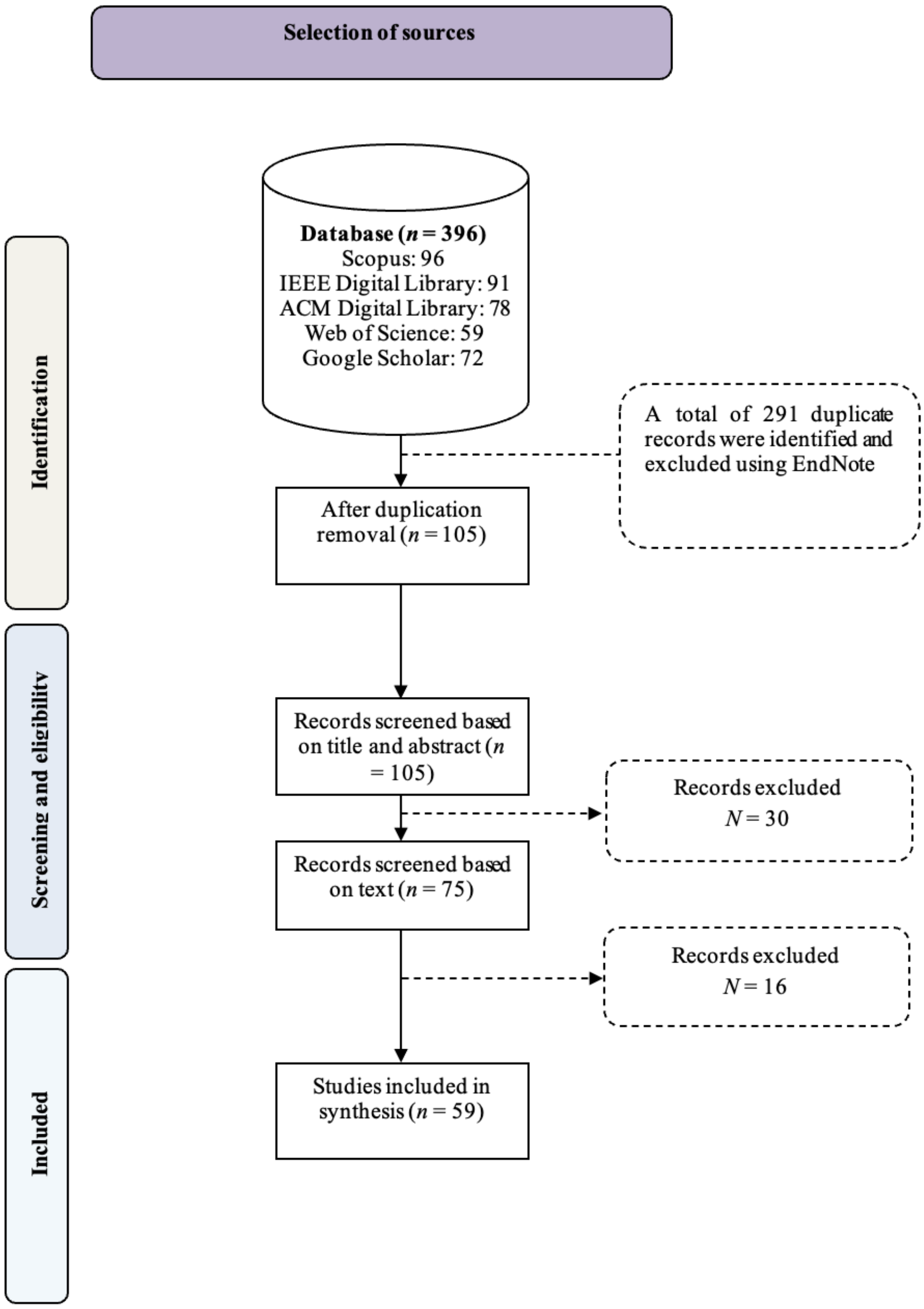

Figure 1. Selection of sources

Findings

RQ1: What are the specific areas of research in relation to the use of ChatGPT in computer programming education?

This study analysed 59 selected studies, categorising them into seven groups, as shown in Table 2. The research areas summarised in Table 2 were derived inductively from the main objectives and focus of the reviewed studies. Each study was assigned to a single primary category based on its main focus to maintain clarity and avoid overlap.

Table 2
Areas of research in relation to the use of ChatGPT in computer programming education

Areas of research groups	Definition	No. & percentage	References
Role of ChatGPT as a programming assistant	The use of ChatGPT to support programming learning and practice by generating, explaining, debugging and optimising code while providing interactive feedback.	13 (22.03%)	(Berrezueta-Guzman & Krusche, 2023; Groothuijsen et al., 2024; Hasrod et al., 2024; Kuramitsu et al., 2023; McDaniel & Zibran, 2024; Moon et al., 2023; Nizamudeen et al., 2024; Ouh et al., 2023; Rees & Gibert, 2024; Sun et al., 2024; Urhan et al., 2024; Yang et al., 2024; Yunianto et al., 2024)
Automated assessment and feedback in programming education	The use of ChatGPT to evaluate code quality and provide real-time, personalised automated feedback to students in programming assessment.	10 (16.94%)	(Callejo et al., 2024; Estévez-Ayres et al., 2024; Grandel et al., 2024; Kimmel et al., 2024; Kizilcec et al., 2024; McGowan & Zibran, 2024; Piccolo et al., 2023; Savelka et al., 2023; Smolansky et al., 2023; Willert & Würz, 2025)
Student and educator behaviour and perceptions	How students and instructors use, experience and evaluate ChatGPT in programming education, including motivation and engagement, technology acceptance and adoption of drivers or barriers.	10 (16.94%)	(Batac et al., 2024; Boguslawski et al., 2024; Boudia, 2024; García-Alonso et al., 2024; Hernandez et al., 2023; Husain, 2024; Padilla et al., 2023; Prather et al., 2024; Țală et al., 2024; Zhang et al., 2024)
Programming curriculum design and instructional strategies	How ChatGPT is embedded into the design and delivery of courses, generating tasks and assessments.	9 (15.25%)	(Al Hajj & Sah, 2023; Baláž et al., 2024; Dickey et al., 2024; French et al., 2023; Lau & Guo, 2023; Liu et al., 2024; Maher et al., 2023; Qureshi, 2023; Speth et al., 2023)
Student learning outcomes and performance	The impact of ChatGPT use on students' learning achievements, programming proficiency and critical thinking skills.	7 (11.86%)	(Ellis et al., 2024; Ma et al., 2024; Park & Kim, 2025; Popovici, 2024; Styve et al., 2024; Wu et al., 2025; Yilmaz & Yilmaz, 2023)
Ethical and academic integrity concerns	Ethical issues and academic integrity risks, including plagiarism, cheating detection, misinformation and the responsible use of ChatGPT in teaching and learning contexts.	6 (10.16%)	(Akçapınar & Sidan, 2024; Al Ahmed & Sharo, , 2023; Amaro et al., 2024; Denzler et al., 2024; Laker & Sena, 2023; Malinka et al., 2023)
ChatGPT application in specific programming domains	Use of ChatGPT for specific programming education courses or tasks.	4 (6.77%)	(Abdulla et al., 2024; García-Alonso et al., 2024; Johnson et al., 2024; Ma'ruf et al., 2023; Valverde-Rebaza et al., 2024)

Role of ChatGPT as a programming assistant

A total of 13 studies (22.03%) examined ChatGPT as a programming assistant. Ouh et al. (2023) evaluated its ability to generate Java code, highlighting its syntax accuracy. Sun et al. (2024) found that ChatGPT-assisted programming improves debugging efficiency and accuracy, highlighting its role as an interactive programming assistant for troubleshooting and code refinement. Nizamudeen et al. (2024) explored student perceptions of ChatGPT as a programming assistant, emphasising its usefulness in explanations, debugging and code suggestions. Berrezueta-Guzman and Krusche (2023) highlighted ChatGPT's ability to generate correct code and explanations. Groothuijsen et al. (2024) examined ChatGPT's impact on programming education, highlighting its role in speeding up debugging and code comprehension. Hasrod et al. (2024) explored its assistance for non-programmers in GUI development, showing how it lowers entry barriers and enables users to create functional applications with AI guidance. Moon et al. (2023) compared ChatGPT with human mentors, highlighting its quick, context-aware suggestions and immediate feedback. Yunianto et al. (2024) examined its role in enhancing computational thinking in programming and debugging using GeoGebra, showing its effectiveness in guiding command-based tasks. McDaniel and Zibran (2024) assessed ChatGPT's impact on source code improvement. Rees and Gibert (2024) explored its role in explaining Python concepts to Spanish-speaking students, reporting improved comprehension through detailed responses.

Yang et al. (2024) introduced PyTutor, a ChatGPT-powered tutoring system providing structured hints and code explanations for beginner Python programmers. Kuramitsu et al. (2023) developed KOGI, a ChatGPT-integrated support system for programming education in Jupyter notebooks. Urhan and Kocadere (2024) found that ChatGPT supports computational thinking in Python problem-solving but requires structured tasks and carefully designed prompts for effective use.

Automated assessment and feedback in programming education

A total of 10 studies (16.94%) examined ChatGPT's use for automated assessment and feedback in programming education. Willert and Würz (2025) assessed ChatGPT's role in quality assurance for formative assessments, identifying response patterns. McGowan et al. (2024) introduced a ChatGPT-based feedback system for university programming assignments. Piccolo et al. (2023) demonstrated its effectiveness in real-time bioinformatics feedback, achieving high success rates. Similarly, Callejo et al. (2024) examined AI-assisted feedback in big data courses, supporting students with diverse backgrounds. Kimmel et al. (2024) found that ChatGPT improved students' understanding of programming mistakes. Savelka et al. (2023) analysed GPT's performance in Python course assessments. Estévez-Ayres et al. (2024) and Grandel et al. (2024) highlighted ChatGPT's role in complex programming exercises, improving grading efficiency in computer science courses. Similarly, Smolansky et al. (2023) and Kizilcec et al. (2024) examined AI's impact on coding assessments and critical thinking.

Student and educator behaviour and perceptions

A total of 10 studies (16.94%) examined student and educator behaviour and perceptions. Boguslawski et al. (2024) highlighted ChatGPT's role in social learning analytics and student engagement. García-Alonso et al. (2024) used the technology acceptance model (Davis, 1989) to show that prior experience and training shape student attitudes. Similarly, Zhang et al. (2024) analysed student perceptions of ChatGPT-generated feedback in a Java course, noting mixed preferences on tone and specificity. Husain (2024) explored instructors' views, balancing instructional benefits with dependency concerns. Prather et al. (2024) found that AI accelerates learning but fosters overconfidence in novice programmers. Hernandez et al. (2023) and Batac et al. (2024) linked ChatGPT adoption to academic stress and self-efficacy. Likewise, Boudia (2024) emphasised psychological and behavioural factors, including academic pressure. Padilla et al. (2023) identified ChatGPT's efficiency benefits, while Țală et al. (2024) found positive attitudes towards GenAI correlate with perceived creativity enhancement.

Programming curriculum design and instructional strategies

A total of nine studies (15.25%) examined ChatGPT's role in programming education, focusing on curriculum design and instructional strategies. Dickey et al. (2024) introduced the AI-Lab framework to integrate ChatGPT while maintaining foundational skills. Similarly, Speth et al. (2023) found AI accelerates exercise creation but requires manual revisions. French et al. (2023) highlighted AI's role in personalised learning, aiding problem-solving and code refinement in game programming. Liu et al. (2024) presented an AI-based tutoring system at Harvard, providing real-time guidance. Likewise, Al Hajj and Sah (2023) found that AI significantly improves performance in PHP programming. Baláž et al. (2024) showed AI users complete tasks nearly three times faster. Maher et al. (2023) stressed balancing AI support with skill development. Lau and Guo (2023) identified a divide in instructor perspectives on AI integration. Qureshi (2023) found ChatGPT improved coding performance in data structures but noted inconsistencies in AI-generated solutions.

Student learning outcomes and performance with ChatGPT

A total of seven studies (11.86%) explored the impact of ChatGPT on student learning outcomes and programming performance. These studies highlighted ChatGPT's prospective to develop computational thinking, problem-solving abilities and enhance academic performance. Yilmaz and Yilmaz (2023) found that students using ChatGPT showed greater improvements in computational thinking and motivation than a control group. Popovici (2024) reported enhanced coding performance in functional programming tasks. Ma et al. (2024) and Park and Kim (2025) emphasised ChatGPT's role in personalised learning and real-time support, particularly in large-scale tasks like big data analytics. Wu et al. (2025) highlighted AI’s contribution to higher-order thinking and active learning. Ellis et al. (2024) compared student assignments with AI-generated code, assessing ChatGPT’s effectiveness in high-level programming. Likewise, Styve et al. (2024) underscored its role in fostering critical thinking, essential for programming proficiency.

Ethical and academic integrity concerns

A total of six studies (10.16%) explored ethical and academic integrity concerns related to the misuse of ChatGPT in academic contexts. Akçapınar and Sidan (2024) investigated students’ susceptibility to incorrect AI-generated information. Similarly, Denzler et al. (2024) found increased *style anomalies* in programming assignments, indicating AI-generated solutions and raising academic dishonesty concerns. Laker and Sena (2023) examined ChatGPT in business analytics education, highlighting integrity challenges alongside learning benefits. Malinka et al. (2023) and Al Ahmed and Sharo (2023) identified plagiarism risks and AI’s potential to aid problem-solving. Amaro et al. (2024) analysed AI-generated misinformation’s impact on user trust, stressing the need for critical engagement with AI tools.

ChatGPT application in specific programming domains

A total of four studies (6.77%) specifically examined the application of ChatGPT in particular programming domains or for specific tasks. Johnson et al. (2024) found that ChatGPT-assisted programming enhanced agriculture students' self-efficacy and microcontroller coding performance. Valverde-Rebaza et al. (2024) highlighted the benefits of integrating ChatGPT with visualisation tools for data analytics learning among non-computational students. Abdulla et al. (2024) demonstrated performance gains in visual programming. Ma’ruf et al. (2023) examined task collaboration via ChatGPT, reporting high usability ratings and widespread academic acceptance.

In summary, research on ChatGPT in programming education has focused on its role as a programming assistant and its use in automated assessment and feedback, improving efficiency and learning outcomes. Studies have also examined student and educator perceptions, noting positive attitudes alongside concerns. Ethical issues and domain-specific applications remain under-explored.

RQ2: How do research studies investigate the incorporation of ChatGPT in teaching and learning computer programming?

To address this research question, the extracted articles were categorised and analysed based on key aspects, including research method, research approach, sample population or user group and country of research (Table 3).

Table 3
Research methods and contexts in ChatGPT for computer programming research

Research method	Research approach	User group	Country	References
Quantitative 24 (40.67%)	Experimental research (*N* = 8)	Students	Germany	(Berrezueta-Guzman & Krusche, 2023)
		Students	Turkey	(Yilmaz & Yilmaz, 2023)
		Students	Taiwan	(Wu et al., 2025)
		Students	Not specified	(Laker & Sena, 2023)
		Educational researchers	Not specified	(Savelka et al., 2023)
		Students	Spain	(Estévez-Ayres et al., 2024)
		Educational researchers	United States of America	(Piccolo et al. 2023)
		Students	Not specified	(Denzler et al., 2024)
	Survey research (*N* = 7)	Students	Indonesia	(Ma'ruf et al., 2023)
		Students	United States	(Dickey et al., 2024)
		Students	The Philippines	(Batac et al., 2024)
		Students	Romania	(Țala et al., 2024)
		Students	The Philippines	(Padilla et al., 2023)
		Students	Spain	(García-Alonso et al., 2024)
		Students	Not specified	(Hernandez et al., 2023)
	Quasi-experimental design (one-group pretest–post-test) (*N* = 9)	Students	Turkey	(Akçapınar & Sidan, 2024)
		Students	Saudi Arabia	(Qureshi, 2023)
		Students	Not specified	(Johnson et al., 2024)
		Educational researchers	Germany	(Willert & Würz, 2025)
		Students	United States of America	(Grandel et al., 2024)
		Students	Not specified	(Yang et al., 2024)
		Students	Slovakia	(Baláž et al., 2024)
		Students	Cyprus	(Al Hajj & Sah, 2023)
		Students	Not specified	(Callejo et al., 2024)
Mixed methods 22 (37.28%)	Survey + interview (*N* = 7)	Students and educators	Spain	(Rees & Gibert, 2024)
		Students and educators	Germany	(Boguslawski et al., 2024)
		Students	The Netherlands	(Nizamudeen et al., 2024)
		Students	South Korea	(Park & Kim, 2025)
		Students and educators	Algeria	(Boudia, 2024)
		Students and educator	The Netherlands	(Groothuijsen et al., 2024)

Research method	Research approach	User group	Country	References
		Students	Not specified	(McGowan et al., 2024)
	Survey + open-ended responses (*N* = 6)	Students	United States of America	(Kimmel et al., 2024)
		Students	Japan	(Ma et al., 2024)
		Students and educators	Australia, Cyprus, United States of America	(Kizilcec et al., 2024)
		Students and educators	Australia, United States of America	(Smolansky et al., 2023)
		Students	United States of America	(Zhang et al., 2024)
		Students	Not specified	(Styve et al., 2024)
	AI-generated responses + discussion (*N* = 2)	--	Czechia	(Malinka et al., 2023)
		50 essay papers from university students	Not specified	(Al Ahmed & Sharo, 2023)
	Quasi-experimental pre and post-test + interview (*N* = 1)	Students	United States of America	(Maher et al., 2023)
	Quasi-experimental + video coding (*N* = 1)	Students	China	(Sun et al., 2024)
	Observation + usage data analysis (*N* = 1)	Students	Japan	(Kuramitsu et al., 2023)
	Experimental + survey (*N* = 1)	Education researchers and students	Romania	(Popovici, 2024)
	Experiment + user perceptions (*N* = 1)	Students	Italy	(Amaro et al., 2024)
	Interviews + experimental (*N* = 1)	Students	Oman	(Abdulla et al., 2024)
	Comparative analysis + System development + Experimental study (*N* = 1)	Students	Korea	(Moon et al., 2023)
Qualitative 13 (22.03%)	Case study (*N* = 4)	Students and professionals	Mexico	(Valverde-Rebaza et al., 2024)
		Students	Germany	(Speth et al., 2023)
		Students	United States of America	(Liu et al., 2024)
		Students	United Kingdom	(French et al., 2023)
	Observation and interaction analysis (*N* = 2)	Students	Turkey	(Urhan et al., 2024)
		Students and visiting scholars	Austria	(Yunianto et al., 2024)

Research method	Research approach	User group	Country	References
	Content analysis (*N* = 2)	Educational researchers	United States of America	(McDaniel & Zibran, 2024)
		Students	Not specified	(Ouh et al.., 2023)
	Comparative analysis (*N* = 1)	Students	Not specified	(Ellis et al., 2024)
	Semi-structured interviews (*N* = 1)	Educators	Not specified	(Lau & Guo, 2023)
	Observations and semi-structured interviews (*N* = 1)	Students	United States of America	(Prather et al., 2024)
	In-depth interviews (*N* = 1)	Educators	Jordan	(Husain, 2024)
	ChatGPT-assisted software development (*N* = 1)	Education researchers	Not specified	(Hasrod et al., 2024)

Note. The numbers in the first column represent the number of studies employing each research method, and the percentages in parentheses indicate the proportion of studies in that category relative to the total number of studies reviewed.

Quantitative research in ChatGPT-assisted programming education

A total of 24 quantitative studies (40.67%) examined ChatGPT's integration into programming education using experimental, survey and quasi-experimental methods.

- Experimental studies (*N* = 8) compared groups using ChatGPT with those that did not, assessing programming performance and learning outcomes. Berrezueta-Guzman and Krusche (2023) evaluated 65 German university students' performance on 22 exercises via the Artemis platform. Yilmaz and Yilmaz (2023) conducted a pretest-post-test study with 45 Bartin University students, measuring computational thinking, self-efficacy and motivation using analysis of variance. Wu et al. (2025) performed a randomised controlled trial in Taiwan with 61 students. Laker and Sena (2023) analysed coursework performance in the United States of America using analysis of variance across six semesters. Savelka et al. (2023) compared AI outputs on multiple-choice and coding tasks. Estévez-Ayres et al. (2024) assessed AI feedback versus instructor grading for 104 Spanish students. Piccolo et al. (2023) tested ChatGPT and Bard on 184 coding exercises at Brigham Young University. Denzler et al. (2024) examined style anomalies and similarity results with 50 students per lab across six labs.
- Survey-based studies (*N* = 7) explored student perceptions and behaviours towards ChatGPT. Ma'ruf et al. (2023) surveyed 102 Indonesian students. Dickey et al. (2024) collected anonymous responses from 41 students. Batac et al. (2024) surveyed 131 students in the Philippines with partial least squares structural equation modelling. Ţala et al. (2024) analysed 364 Romanian students' data via IBM SPSS. Padilla et al. (2023) surveyed 298 computing students in the Philippines. García-Alonso et al. (2024) surveyed 216 Spanish students and Hernandez et al. (2023) analysed 299 responses using exploratory data analysis and partial least squares structural equation modelling.
- Quasi-experimental studies (*N* = 9), mainly one-group pretest–post-test designs, evaluated ChatGPT's impact on learning by comparing performance before and after use. Akçapınar and Sidan (2024) analysed 45 first-year students' scores using paired *t*tests. Qureshi (2023) tracked programming performance in Saudi Arabia through descriptive and inferential statistics. Johnson et al. (2024) assessed Arduino programming outcomes, while Willert and Würz (2025) compared AI-human classification accuracy in German courses. Grandel et al. (2024) used rubric-based evaluations with 26 United States of America students. Yang et al. (2024) compared 35 PyTutor users with 36 controls. Baláž et al. (2024) examined linked list tasks with or without ChatGPT. Al Hajj and Sah (2023) analysed quiz scores of 50 Cyprus undergraduates. Callejo et al. (2024) studied 31 students across entry-level and engineering groups using pretests and in-class exercises.

Mixed methods research in ChatGPT-assisted programming education

Mixed methods research accounted for 22 (37.28 %) and combines quantitative and qualitative approaches to examine ChatGPT in programming education. Studies integrated surveys, interviews, quasi-experiments, observations, usage data analysis, experimental designs and AI-generated response evaluations.

- Survey and interview combinations were common (*N* = 7). Rees and Gibert (2024) analysed 33 Spanish students. Boguslawski et al. (2024) interviewed six educators and 44 students. Nizamudeen et al. (2024) surveyed 39 students and interviewed five, analysing familiarity with GenAI using descriptive statistics. Park and Kim (2025) collected 343 student surveys. Boudia (2024) surveyed 57 Algerian students with 20 Likert-scale questions, recorded ChatGPT interactions and conducted semi-structured interviews, analysing responses across Arabic, French and English. Groothuijsen et al. (2024) studied 29 students and one educator, applying statistical and qualitative coding analysis. McGowan et al. (2024) compared ChatGPT- and human-generated feedback for 108 postgraduate students.
- A survey with open-ended responses is another mixed method, combining quantitative and qualitative insights, accounting for a total of six studies. Kimmel et al. (2024) analysed submission logs and surveys from 52 United States of America students. Ma et al. (2024) surveyed 26 Japanese undergraduates, using Likert-scale analysis and thematic coding. Kizilcec et al. (2024) collected data from students and educators across three universities, using statistics, thematic coding and GPT-4.

Smolansky et al. (2023) examined data from 389 students and 36 educators in Australia and the United States of America, comparing responses via descriptive statistics and thematic analysis. Zhang et al. (2024) assessed 22 assignments via Artemis and Styve et al. (2024) analysed pretest and post-test surveys from 90 students, focusing on critical thinking practices.

- Maher et al. (2023) used a quasi-experimental pretest and post-test and interviews with three students in the United States of America, analysing task time, ChatGPT interactions, quiz results and learning themes. Sun et al. (2024) conducted a controlled experiment with 30 Chinese students comparing unprompted versus prompt-based learning. Kuramitsu et al. (2023) observed AI use among 127 Japanese students, tracking errors and usage. Popovici (2024) surveyed 181 Romanian students and researchers on AI responses. Malinka et al. (2023) analysed Czech student discussions. Al Ahmed and Sharo (2023) compared 50 essays. Amaro et al. (2024) studied 62 Italian students' perceptions. Abdulla et al. (2024) combined interviews and experiments with 26 Omani students, and Moon et al. (2023) integrated analysis, system development and experiments for Korean students.

Qualitative research in ChatGPT-assisted programming education

Qualitative research accounted for 13 (22.03%) of the reviewed studies, employing comparative analysis, case studies, interviews, content analysis, observations and software development. These studies offer in-depth insights into student and educator experiences with AI-assisted programming education.

- Four case studies explored AI's role in programming education within real-world settings, using small sample sizes. For instance, Valverde-Rebaza et al. (2024) studied 59 participants at Tecnológico de Monterrey, Mexico. Speth et al. (2023) examined nine students in beginner-to-intermediate programming courses in Germany. Liu et al. (2024) analysed AI's impact on coding style among 70 United States of America students. French et al. (2023) investigated AI interactions in a games programming course at London University, focusing on student projects.
- Two studies using observation and interaction analysis examined student collaboration with ChatGPT in problem-solving. For example, Urhan et al. (2024) analysed video recordings of three students' pair programming sessions in Turkey, assessing their practices and ChatGPT's response accuracy. Yunianto et al. (2024) analysed screen recordings and conversations in Austria, focusing on prompts, final codes and iteration counts.
- Two studies used content analysis to systematically evaluate ChatGPT's feedback. McDaniel and Zibran (2024) analysed ChatGPT's responses to C++ code from graduate-level courses at Idaho State University, assessing strengths, weaknesses and recommendations through evaluation and subjective agreement ratings. Similarly, Ouh et al. (2023) examined ChatGPT's solutions for undergraduate exercises, evaluating accuracy, readability and efficiency.
- Lau and Guo (2023) conducted semi-structured interviews with 20 university educators from nine countries, analysing AI's role in introductory programming courses using thematic analysis. Ellis et al. (2024) used a comparative method to examine 36 student and ChatGPT-generated code submissions, evaluating quality, structure and adherence to requirements. Prather et al. (2024) observed 19 students, combining observations with semi-structured interviews in the United States of America. Husain (2024) conducted in-depth interviews with 12 programming instructors at Al al-Bayt University, Jordan, using thematic analysis. Hasrod et al. (2024) examined ChatGPT-assisted software development. Educational researchers interacted with ChatGPT to troubleshoot and refine GUI code, analysing its effectiveness in coding and debugging tasks.

In summary, research to date on ChatGPT in programming education is predominantly quantitative, emphasising measurable outcomes and experimental validation. Mixed methods studies combine statistical findings with qualitative insights, while qualitative research offers contextual perspectives on experiences and classroom dynamics. Experimental and quasi-experimental designs are most common, alongside surveys of student and educator perceptions. Studies are globally distributed, mainly from Europe and Asia, with additional contributions from North America and Australia, reflecting widespread interest in AI integration.

RQ3: What are the educational benefits and challenges of using ChatGPT in computer programming education?

While ChatGPT has emerged as a valuable tool for programming students and educators, offering numerous benefits, its integration also raises concerns. The following discussion explores both the advantages and challenges of using ChatGPT in this discipline.

Benefits

Adaptive learning in computer programming education

One key advantage of ChatGPT in programming education is its ability to provide personalised learning (Park & Kim, 2025). Unlike traditional classrooms, where instant help may be limited, ChatGPT delivers immediate explanations, examples and step-by-step guidance tailored to individual needs. This flexibility enables self-paced learning, reinforcing programming concepts (Yang et al., 2024). Additionally, ChatGPT adapts it responses based on inquiry complexity, benefiting both beginners and advanced learners (Berrezueta-Guzman & Krusche, 2023).

Automated assessment and feedback

Traditional programming courses require extensive manual grading, which is time-consuming for educators. ChatGPT can evaluate code, detect syntax errors and suggest improvements in real time (McGowan et al., 2024). This reduces instructor workload while providing students with instant feedback, enabling them to correct mistakes and refine their code without delay. Immediate feedback is crucial for learning, helping students identify errors early and understand the logic behind corrections (Smolansky et al., 2023).

Enhanced student engagement and motivation

Incorporating ChatGPT into programming education enhances student engagement and motivation (Boguslawski et al., 2024). AI-driven learning makes problem-solving more interactive and less daunting. Instead of struggling alone, students can engage in real-time conversations with ChatGPT, simulating a mentor–student dynamic (Boudia, 2024). This fosters curiosity, encourages experimentation with coding techniques and deepens computational thinking (Padilla et al., 2023). Additionally, ChatGPT introduces real-world applications, helping students recognise the practical relevance of programming.

Code optimisation

Beyond assisting with coding, ChatGPT can optimise code quality and encourage best programming practices. It suggests improvements in code efficiency, readability and adherence to educational standards (McDaniel & Zibran, 2024). By guiding students towards writing structured and maintainable code, ChatGPT helps them develop habits that align with professional software development practices (Hasrod et al., 2024). This feature is particularly useful for students who struggle with writing clean and optimised code, as they can learn from AI-generated suggestions and refine their programming style.

Bridging knowledge gaps and enhancing accessibility

Another key benefit of ChatGPT is bridging knowledge gaps among students with diverse backgrounds. It simplifies complex programming concepts and offers explanations in multiple ways (Park & Kim, 2025). Students can request practical examples or step-by-step guidance to matches their learning preferences (Boudia, 2024). ChatGPT also supports students with disabilities or learning difficulties, providing adaptive, interactive learning methods (Țală et al., 2024).

Challenges

Overreliance and reduced critical thinking

Despite its numerous benefits, integrating ChatGPT into programming education comes with significant challenges. One of the most pressing concerns is the risk of student overreliance on AI-generated solutions (Akçapınar & Sidan, 2024) When students depend too heavily on ChatGPT for coding assistance, they may not develop essential problem-solving skills. While AI can provide immediate answers, it does not replace the critical thinking required to debug, optimise and fully understand code (Styve et al., 2024).

If students simply copy and paste ChatGPT-generated code without analysing it, their ability to think through programming logic may weaken over time. This raises concerns about whether AI is fostering genuine learning or merely enabling shortcuts.

Accuracy and reliability concerns

ChatGPT's accuracy and reliability remain a challenge. While AI has advanced, it can generate incorrect or suboptimal code, misleading students (Yilmaz & Yilmaz, 2023). Unlike human instructors, it lacks contextual understanding and real-world experience, making complex debugging difficult (Ellis et al., 2024). Students who are not yet proficient in programming may find it difficult to recognise when ChatGPT provides incorrect information, which can negatively impact their learning process.

Limited context awareness

A related issue is ChatGPT's limited context awareness. While it can process queries effectively, it does not always understand vague or poorly structured questions, leading to responses that may not fully address a student's needs. Programming problems often require a deep understanding of a student's intent, which ChatGPT may not always capture accurately (Ma et al., 2024). Instructors usually tailor explanations to a student's thinking, but ChatGPT can produce generic or misaligned responses due to its limited dynamic interaction (García-Alonso et al., 2024).

Ethical and academic integrity concerns

Ethical and academic integrity concerns also present significant challenges in AI-assisted programming education (Al Ahmed & Sharo, 2023). With ChatGPT readily available, there is a higher risk of plagiarism and academic dishonesty. Students may use AI to generate code for assignments without putting in the necessary effort to understand the logic behind it (Laker & Sena, 2023). This undermines the purpose of programming education, which is to develop computational thinking and problem-solving skills (Al Ahmed & Sharo, 2023). Educational institutions must implement guidelines to ensure that AI is used responsibly, balancing its benefits while discouraging dishonest practices (Amaro et al., 2024).

Challenges in debugging code

ChatGPT often identifies coding errors but lacks detailed explanations or alternative solutions (Abdulla et al., 2024). Effective debugging requires understanding logic, yet ChatGPT provides single fixes without alternatives. This may hinder students' debugging skills, which are crucial for real-world programming (McDaniel & Zibran, 2024), as understanding errors is key to learning.

Integration challenges with curriculum and instruction

Integrating ChatGPT into programming education presents challenges in curriculum design and instructional strategies. Many programming courses follow structured curricula that emphasise fundamental programming concepts, problem-solving techniques and software development methodologies (Dickey et al., 2024). Educators should integrate ChatGPT strategically to complement teaching, prevent learning gaps and may require training for effective AI use (Maher et al., 2023).

In summary, ChatGPT enhances programming education through personalised learning, automated feedback, engagement and accessibility. However, effective integration requires ethical guidance and educator support. Challenges include student overreliance, reduced critical thinking, accuracy limitations, academic integrity concerns, limited debugging explanations and difficulties aligning AI use with existing curricula.

Discussion and implications

This review examined ChatGPT's use in computer programming education and discusses key implications for different stakeholders (students, educators, researchers, institutions and developers). ChatGPT functions as an effective coding assistant, enabling students to debug errors, generate code more efficiently and grasp complex programming concepts. However, consistent with prior studies, concerns persist about overreliance on AI-generated code and weakened conceptual understanding that may undermine long-term learning and professional competence (Silva et al., 2024). Educators are therefore encouraged to integrate ChatGPT into programming units or laboratory sessions that emphasise collaborative problem-solving and experimentation rather than basic code generation. Students should explain and annotate AI-generated code, while complementary assessments like oral exams reinforce critical thinking and mastery (Al Ahmed & Sharo, 2023; Urhan & Kocadere, 2024). Institutions need clear policies to ensure ChatGPT is used ethically as a supportive, not replacement learning tool. Ongoing professional development and best-practice sharing can help educators design AI-supported activities, like guided debugging, pair programming and AI-assisted code reviews, to boost student engagement and skills.

In response to ChatGPT's growing role in automated programming feedback, educators and institutions are encouraged to adopt formative, scaffolded assessments over purely summative ones. Authentic assessment can then be employed through reflection activities where students explain their reasoning and debugging strategies. Institutions should adopt existing frameworks (e.g., the AI Assessment Scale by Furze et al., 2024) or develop appropriate new ones to guide educators in redesigning assessments for the GenAI era and in systematically evaluating student work that incorporates AI tools. Institutions should provide professional development that enables educators to responsibly integrate AI into teaching and assessment. This includes training in prompt engineering, AI critique and the redesign of assessments to ensure transparency and integrity. The research community can advance this field by creating frameworks to assess AI-generated feedback's validity and pedagogy and by comparing student performance in hybrid versus traditional feedback models. Drawing on other stakeholder guidance (e.g., research and institutional), developers can create explainable feedback features and interactive prompts that align AI tools with learning objectives and student-centred practice (Grandel et al., 2024).

In the area of student and educator behaviour and perceptions, weekly learning journals give students structured opportunities to reflect on AI-assisted coding. This reflection helps them assess their choices, balance reliance on ChatGPT and strengthen self-regulated learning skills (Hernández et al., 2023; Silva et al., 2024). Building on these reflections, AI literacy workshops enable instructors to guide students in using AI feedback responsibly and applying suggestions thoughtfully (García-Alonso et al., 2024; Zhang et al., 2024). Insights from behavioural research reveal patterns in motivation, engagement and technology acceptance, providing researchers with evidence to design strategies that reduce overdependence on AI and enhance learning outcomes (Batac et al., 2024; Boguslawski et al., 2024). Institutions can reinforce these efforts by establishing clear AI policies and structured programs that promote equitable access (Prather et al., 2024). Developers can add interactive prompts, reflective scaffolds and alerts to promote ChatGPT as a supportive not substitutive coding tool (Țală et al., 2024).

In the area of programming curriculum design and instructional strategies, integrating AI tools allows students to engage in exploratory coding challenges and iterative problem-solving exercises. Maintaining structured reflections and mini code define sessions helps learners apply AI-generated suggestions while developing critical reasoning skills (Baláž et al., 2024). When instructors design AI-supported exercises, they can promote deeper understanding of programming concepts and creativity (Dickey et al., 2024). Insights from curriculum-focused findings can provide researchers with evidence to evaluate the effectiveness of AI integration and course design optimisation (Liu et al., 2024). Institutions can strengthen curriculum innovation by providing resources, frameworks and policies that support AI-enhanced instruction (French et al., 2023). Finally, developers can collaborate with educators to co-design AI tools that align with course progression, ensuring the technology enhances rather than replaces pedagogical design (Al Hajj & Sah, 2023).

In the area of student learning outcomes and performance, AI-assisted exercises allow students to explore new programming tasks and test alternative coding approaches. Structured practice cycles that combine AI-supported activities with independent assignments help learners build confidence and fluency in applying programming concepts (Yilmaz & Yilmaz, 2023). When instructors design activities with incremental challenges and provide opportunities for peer review, students can better internalise course material and track their progress (Popovici, 2024; Wu et al., 2025). These study findings offer researchers evidence on the measurable impact of AI on learning efficiency and performance consistency across programming courses (Ellis et al., 2024). Institutions can enhance outcomes by monitoring student analytics, identifying knowledge gaps and providing targeted support where needed (Park et al., 2025). Developers can create AI tools that guide learners through stepwise coding exercises and provide adaptive prompts to reinforce understanding (Silva et al., 2024).

In the area of ethical and academic integrity concerns, ensuring that students submit AI-assisted code with accurate citations and annotations promotes accountability and transparency (Denzler et al., 2024; Akçapınar & Sidan, 2024). Educators can integrate ethics modules in programming courses, clearly outlining acceptable AI practices and consequences for misuse (Al Ahmed & Sharo, 2023; Amaro et al., 2024). Research evaluating AI-detection systems and integrity interventions provides evidence for identifying unauthorised AI use, code plagiarism and inconsistencies in programming submissions (Malinka et al., 2023). Institutions can support ethical practices by establishing formal policies, implementing monitoring systems and ensuring adherence to academic standards (Denzler et al., 2024). Developers can promote integrity by adding traceability, role-based access and prompt-justification features to support responsible AI-assisted learning (Denzler et al., 2024).

Considering the findings related to how research studies investigate ChatGPT in programming education (RQ2), future research should move beyond methodological fragmentation towards greater coherence and standardisation. The literature already to date has employed diverse approaches ranging from experiments and surveys to design-based case studies. But most studies have relied on small-scale quantitative or single-institution designs. Future studies should use longitudinal, cross-institutional designs to assess long-term learning, behavioural and pedagogical impacts. Mixed methods approach using surveys, interviews, system logs and performance analytics can offer a holistic view of how ChatGPT supports teaching and learning. Researchers should also develop standardised instruments and evaluation frameworks to ensure consistency and comparability across studies. Comparative studies that examine learner characteristics (e.g., novices vs advanced students, native vs non-native speakers), AI access (free vs paid), level of study (undergraduate vs postgraduate) and mode of study (face-to-face vs online) can provide new insights and inform developers and institutions to tailor ChatGPT and support to diverse student needs (Moon et al., 2023). Expanding research into under-represented regions such as Africa, the Middle East and the Americas can reveal cross-cultural adoption patterns and challenges. Comparative studies could be conducted to explore regional differences in access, perception and pedagogical integration of ChatGPT. The findings from such studies could then inform developers and policymakers to customize AI systems and educational resources that better align with local needs languages and cultural contexts.

The findings related to RQ3 on benefits can help educators strategically use ChatGPT to enhance learning and support student understanding. At the same time, recognising the challenges enables institutions to design targeted interventions, training and policies to ensure responsible use (Adeshola et al., 2023). For researchers, this synthesis guides future studies on long-term learning outcomes and best practices for AI integration in programming education (Adel et al., 2024).

The key implications for research, practice and policy are summarised in Figure 2.

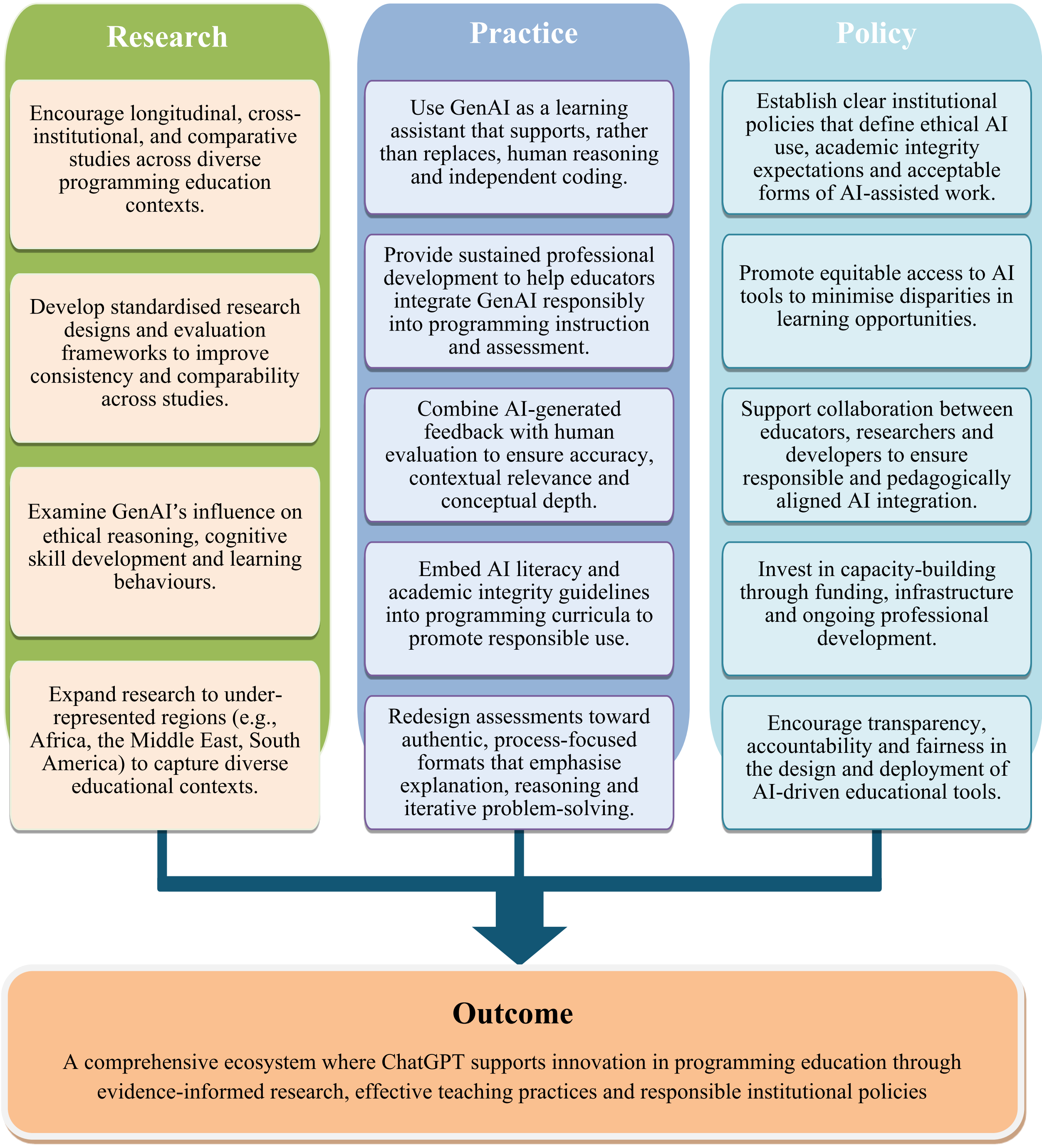

Figure 2. Key implications across research, practice and policy domains

Conclusions and limitations

This scoping review examined the role of ChatGPT in computer programming education using the PRISMA-ScR framework (Tricco et al., 2018). It mapped key research areas, explored how ChatGPT has been integrated into programming teaching and learning and identified the main benefits and challenges reported in the literature. Fifty-nine studies published between 2022 and early 2025 were retrieved from Scopus, Web of Science, the ACM Digital Library, the IEEE Digital Library and Google Scholar. The synthesis, guided by three RQs, offers insights for educators seeking to incorporate AI tools into curricula, for policymakers developing ethical-use guidelines and for developers aiming to improve AI systems for diverse learner needs. The review establishes a foundation for further investigation into effective and ethical AI-driven learning. Several limitations can be acknowledged. English-language studies indexed in major databases were included, excluding books, technical reports and workshop papers. Many studies employed small samples, particularly in qualitative or experimental designs, limiting generalisability; these findings were interpreted cautiously. Methodological diversity also reduced comparability across studies. Finally, the exclusive focus on ChatGPT offers depth but constrains generalisation to other generative AI tools. Future work should examine a broader range of systems to capture varied capabilities and educational impacts.

Author contributions

Maria Ijaz Baig: Conceptualisation, Investigation, Writing – original draft, Writing – review and editing; **Elaheh Yadegaridehkordi:** Data curation, Investigation, Formal analysis, Writing – review and editing; **Ayub Bokani:** Writing – review and editing.

References

Abdulla, S., Ismail, S., Fawzy, Y., & Elhag, A. (2024). Using ChatGPT in teaching computer programming and studying its impact on students performance. *Electronic Journal of e-Learning, 22*(6), 66–81. https://doi.org/10.34190/ejel.22.6.3380

Adel, A., Ahsan, A., & Davison, C. (2024). ChatGPT promises and challenges in education: Computational and ethical perspectives. *Education Sciences, 14*(8), Article 814. https://doi.org/10.3390/educsci14080814

Adeshola, I., & Adepoju, A. P. (2023). The opportunities and challenges of ChatGPT in education. *Interactive Learning Environments, 32*(10), 6159–6172. https://doi.org/10.1080/10494820.2023.2253858

Adıgüzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology, 15*(3), Article ep429. https://doi.org/10.30935/cedtech/13152

Akçapınar, G., & Sidan, E. (2024). AI chatbots in programming education: Guiding success or encouraging plagiarism. *Discover Artificial Intelligence*, *4*, Article 87. https://doi.org/10.1007/s44163-024-00203-7

Al Hajj, J. A., & Sah, M. (2023). Assessing the impact of ChatGPT in a PHP programming course. In G. Cankaya (Chair), *Proceeedings of the 7th International Symposium on Innovative Approaches in Smart Technologies* (pp. 25–44). IEEE. https://doi.org/10.1109/ISAS60782.2023.10391549

Al Ahmed, Y. A., & Sharo, A. (2023). On the education effect of ChatGPT: Is AI ChatGPT to dominate education career profession? In J. Lloret (Chair), *Proceedings of the2023 International Conference on Intelligent Computing, Communication, Networking and Services* (pp. 79–84). IEEE. https://doi.org/10.1109/ICCNS58795.2023.10192993

Amaro, I., Barra, P., Greca, A. D., Francese, R., & Tucci, C. (2024). Believe in artificial intelligence? A user study on the ChatGPT's fake information impact. *IEEE Transactions on Computational Social Systems, 11*(4), 5168–5177. https://doi.org/10.1109/TCSS.2023.3291539

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI, 7*(1), 52–62. https://doi.org/10.61969/jai.1337500

Baig, M. I., & Yadegaridehkordi, E. (2024). ChatGPT in higher education: A systematic literature review and research challenges. *International Journal of Educational Research, 127*, Article 102411. https://doi.org/10.1016/j.ijer.2024.102411

Baláž, N., Porubän, J., Horváth, M., & Kormaník, T. (2024). Using ChatGPT during implementation of programs in education. In A. Santos (Ed.), *Proceedings of the 5th International Computer Programming Education Conference (*pp. 18:1–18:9*).* OASIcs. https://doi.org/10.4230/OASIcs.ICPEC.2024.18

Batac, C. A., Baroja, M. J., Caballero, D. J. D., Coloma, L. G., Tan, L. M., & Ebardo, R. (2024). Do human beliefs and traits influence the adoption of ChatGPT among programming students? In T. G. T.

Nindhia (Chair), *Proceedings of the 10th International Conference on Computing and Artificial Intelligence* (pp. 339–344). ACM. https://doi.org/10.1145/3669754.3669806

Batista, J., Mesquita, A., & Carnaz, G. (2024). Generative AI and higher education: Trends, challenges, and future directions from a systematic literature review. *Information, 15*(11), Article 676. https://doi.org/10.3390/info15110676

Berrezueta-Guzman, J., & Krusche, S. (2023). Recommendations to create programming exercises to overcome ChatGPT. In H. Washizaki (Chair), *Proceedings of the 35th International Conference on Software Engineering Education and Training* (pp. 147–151). IEEE. https://doi.org/10.61969/jai.1337500

Bewersdorff, A., Hartmann, C., Hornberger, M., Seßler, K., Bannert, M., Kasneci, E., Zhai, X., & Nerdel, C. (2025). Taking the next step with generative artificial intelligence: The transformative role of multimodal large language models in science education. *Learning and Individual Differences, 118*, Article 102601. https://doi.org/10.1016/j.lindif.2024.102601

Bhullar, P. S., Joshi, M., & Chugh, R. (2024). ChatGPT in higher education: A synthesis of the literature and a future research agenda. *Education and Information Technologies, 29*(16), 21501–21522. https://doi.org/10.1007/s10639-024-12723-x

Boguslawski, S., Deer, R., & Dawson, M. G. (2024). Programming education and learner motivation in the age of generative AI: Student and educator perspectives. *Information and Learning Sciences*, *126* (1/2), 91–109. https://doi.org/10.1108/ILS-10-2023-0163

Boudia, C. (2024). Early insights into SLA with ChatGPT: Navigating CS teachers and student perspectives in an opinion-based exploration. *Edelweiss Applied Science and Technology, 8*(5), 648–661. https://doi.org/10.55214/25768484.v8i5.1729

Callejo, P., Alario-Hoyos, C., & Delgado-Kloos, C. (2024). Evaluating ChatGPT impact on the programming learning outcomes of students in a big data course. *International Journal of Engineering Education*, *40*(4), 863–872. https://www.ijee.ie/contents/c400424.html

Chang, L. C., Lin, H. R., & Lin, J. W. (2024). Learning motivation, outcomes, and anxiety in programming courses—A computational thinking–centered method. *Education and Information Technologies, 29*(1), 545–569. https://doi.org/10.1007/s10639-023-12313-3

Chugh, R., Turnbull, D., Morshed, A., Sabrina, F., Azad, S., Rahman, M. M., Kaisar, S., & Subramani, S. (2025). The promise and pitfalls: A literature review of generative artificial intelligence as a learning assistant in ICT education. *Computer Applications in Engineering Education, 33*(2), Article e70002. https://doi.org/10.1002/cae.70002

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology, *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Denzler, B., Vahid, F., Pang, A., & Salloum, M. (2024). Style anomalies can suggest cheating in CS1 programs. In L. Barker (Chair), *Proceedings of the 2024 on Innovation and Technology in Computer Science Education* (Vol. 1, pp. 381–387). ACM. https://doi.org/10.1145/3649217.3653626

Dickey, E., Bejarano, A., & Garg, C. (2024). AI-Lab: A framework for introducing generative artificial intelligence tools in computer programming courses. *SN Computer Science, 5*, Article 720. https://doi.org/10.1007/s42979-024-03074-y

Ellis, M. E., Casey, K. M., & Hill, G. (2024). ChatGPT and Python programming homework. *Decision Sciences Journal of Innovative Education, 22*(2), 74–87. https://doi.org/10.1111/dsji.12306

Estévez-Ayres, I., Callejo, P., Hombrados-Herrera, M. Á., Alario-Hoyos, C., & Delgado-Kloos, C. (2024). Evaluation of LLM tools for feedback generation in a course on concurrent programming. *International Journal of Artificial Intelligence in Education*, *35*(2), 774–790. https://doi.org/10.1007/s40593-024-00406-0

French, F., Levi, D., Maczo, C., Simonaityte, A., Triantafyllidis, S., & Varda, G. (2023). Creative use of OpenAI in education: Case studies from game development. *Multimodal Technologies and Interaction, 7*(8), Article 81. https://doi.org/10.3390/mti7080081

Furze, L., Perkins, M., Roe, J., & MacVaugh, J. (2024). The AI assessment scale (AIAS) in action: A pilot implementation of GenAI-supported assessment. *Australasian Journal of Educational Technology, 40*(4), 38–55. https://doi.org/10.14742/ajet.9434

Garcia, M. B. (2025). Teaching and learning computer programming using ChatGPT: A rapid review of literature amid the rise of generative AI technologies. *Education and Information Technologies*, *30*(12), 16721–16745. https://doi.org/10.1007/s10639-025-13452-5

García-Alonso, E. M., León-Mejía, A. C., Sánchez-Cabrero, R., & Guzmán-Ordaz, R. (2024). Training and technology acceptance of ChatGPT in university students of social sciences: A netcoincidental analysis. *Behavioral Sciences, 14*(7), 612–632. https://doi.org/10.3390/bs14070612

Grandel, S., Schmidt, D. C., & Leach, K. (2024). Applying large language models to enhance the assessment of parallel functional programming assignments. In L. O'Conner (Ed.), *Proceedings of the 1st International Workshop on Large Language Models for Code* (pp. 102–110). ACM. https://doi.org/10.1145/3643795.3648375

Groothuijsen, S., van den Beemt, A., Remmers, J. C., & van Meeuwen, L. W. (2024). AI chatbots in programming education: Students' use in a scientific computing course and consequences for learning. *Computers and Education: Artificial Intelligence, 7*, Article 100290. https://doi.org/10.1016/j.caeai.2024.100290

Hasrod, T., Nuapia, Y. B., & Tutu, H. (2024). ChatGPT helped me build a chemistry app, and here's how you can make one also. *Journal of Chemical Education*, *101*(2), 653–660. https://doi.org/10.1021/acs.jchemed.3c01170

Hernandez, A. A., Padilla, J. R. C., & Montefalcon, M. D. L. (2023). Information seeking behavior in ChatGPT: The case of programming students from a developing economy. In R. Adnan (Chair), *Proceedings of the 13th International Conference on System Engineering and Technology* (pp. 72–77). IEEE. https://doi.org/10.1109/ICSET59111.2023.10295122

Husain, A. (2024). Potentials of ChatGPT in computer programming: Insights from programming instructors. *Journal of Information Technology Education: Research, 23, 1–25*. https://doi.org/10.28945/5240

Johnson, D. M., Doss, W., & Estepp, C. M. (2024). Agriculture students' use of generative artificial intelligence for microcontroller programming. *Natural Sciences Education, 53*(2), Article e20155. https://doi.org/10.1002/nse2.20155

Kimmel, B., Geisert, A.L., Yaro, L., Gipson, B., Hotchkiss, R.T., Osae-Asante, S.K., Vaught, H., Wininger, G. & Yamaguchi, C., (2024). Enhancing programming error messages in real time with generative AI. In F. F. Mueller, P. Kyburz, J. R. Williamson & C. Sas (Eds.), *Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp. 1-7). ACM. https://doi.org/10.48550/arXiv.2402.08072

Kizilcec, R. F., Huber, E., Papanastasiou, E. C., Cram, A., Makridis, C. A., Smolansky, A., Zeivots, S., & Raduescu, C. (2024). Perceived impact of generative AI on assessments: Comparing educator and student perspectives in Australia, Cyprus, and the United States. *Computers and Education: Artificial Intelligence, 7*, Article 100269. https://doi.org/10.1016/j.caeai.2024.100269

Kuramitsu, K., Obara, Y., Sato, M., & Obara, M. (2023). KOGI: A seamless integration of ChatGPT into Jupyter environments for programming education. In M. Q. Feldman (Chair), *Proceedings of the ACM SIGPLAN International Symposium on SPLASH-E* (*pp. 50–59*). ACM. https://doi.org/10.1145/3622780.3623648

Laker, L. F., & Sena, M. (2023). Accuracy and detection of student use of ChatGPT in business analytics courses. *Issues in Information Systems, 24*(2), 153–163. https://doi.org/10.48009/2_iis_2023_113

Lau, S., & Guo, P. (2023). From "Ban it till we understand it" to "Resistance is futile": How university programming instructors plan to adapt as more students use AI code generation and explanation tools such as ChatGPT and GitHub Copilot. In K. Fisler, P. Denny, D. Franklin, & M. Hamilton (Eds.), *Proceedings of the ACM Conference on International Computing Education Research* (*Vol. 1,* pp. 106–121). ACM. https://doi.org/10.1145/3568813.3600138

Liu, R., Zenke, C., Liu, C., Holmes, A., Thornton, P., & Malan, D. J. (2024). Teaching CS50 with AI: leveraging generative artificial intelligence in computer science education. In B. Stephenson (Chair), . *Proceedings of the 55th ACM Technical Symposium On Computer Science Education* (Vol. 1, pp. 750–756). ACM. https://doi.org/10.1145/3626252.3630938

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13*(4), 410–424. https://doi.org/10.3390/educsci13040410

Ma, B., Chen, L., & Konomi, S. I. (2024). Enhancing programming education with ChatGPT: A case study on student perceptions and interactions in a Python course. In J. Filipe, A. Ghosh, & L. Zhou (Eds.), *Proceedings of the International Conference on Artificial Intelligence in Education* (pp. 113–126). Springer. https://doi.org/10.48550/arXiv.2403.15472

Maher, M. L., Tadimalla, S. Y., & Dhamani, D. (2023). An exploratory study on the impact of AI tools on the student experience in programming courses: An intersectional analysis approach. In T. Hammond (Chair), *Proceedings of the IEEE Frontiers in Education Conference* (pp. 1–5). IEEE. https://doi.org/10.1109/FIE58773.2023.10343037

Malinka, K., Peresíni, M., Firc, A., Hujnák, O., & Janus, F. (2023). On the educational impact of ChatGPT: Is artificial intelligence ready to obtain a university degree? In M. J. Laakso (Chair), *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education* (Vol. 1, pp. 47–53). ACM. https://doi.org/10.1145/3587102.3588827

Ma'ruf, H., Aditya, B. R., Hernawati, E., Gunawan, T., & Wijayanto, P. W. (2023). Usability testing of ChatGPT website as a medium for task collaboration using the System Usability Scale Method (SUS). In C. K. Dewa (Chair), *Proceedings of the 8th International Conference on Information Technology and Digital Applications* (pp. 86–90). IEEE. https://doi.org/10.1109/ICITDA60835.2023.10427187

McDaniel, S., & Zibran, M. F. (2024). Improving source code with assistance from AI — A pilot case study with ChatGPT. In A. Gokhale (Chair), *Proceedings of the 7th International Conference on Information and Computer Technologies* (pp. 332–337). IEEE. https://doi.org/10.1109/ICICT62343.2024.00060

McGowan, A., Anderson, N., & Smith, C. (2024, May). The use of ChatGPT to generate summative feedback in programming assessments that is consistent, prompt, without bias and scalable. In K. Erentürk (Chair), *Proceedings of the Cognitive Models and Artificial Intelligence Conference* (pp. 39–43). ACM. https://doi.org/10.1145/3660853.3660863

McGowan, J., Straus, S., Moher, D., Langlois, E. V., O'Brien, K. K., Horsley, T., Aldcroft, A., Zarin, W., Garritty, C. M., Hempel, S., Lillie, E., Tunçalp, Ö., & Tricco, A. C. (2020). Reporting scoping reviews—PRISMA ScR extension. *Journal of Clinical Epidemiology, 123,* 177–179. https://doi.org/10.1016/j.jclinepi.2020.03.016

Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Education Sciences, 13*(9), Article 856. https://doi.org/10.3390/educsci13090856

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group, T. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151*(4), 264–269. https://doi.org/10.1371/journal.pmed.1000097

Moon, J., Yang, R., Cha, S., & Kim, S. B. (2023). chatGPT vs Mentor: Programming language learning assistance system for beginners. In W. I. S. W Din (Chair), *Proceedings of the 8th International Conference On Software Engineering and Computer Systems* (pp. 106–110). IEEE. https://doi.org/10.1109/ICSECS58457.2023.10256295

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence, 2*, Article 100041. https://doi.org/10.1016/j.caeai.2021.100041

Nizamudeen, F., Gatti, L., Bouali, N., & Ahmed, F. (2024). Investigating the impact of code generation tools (ChatGPT & GitHub CoPilot) on programming education. In O. Poquet, A. Ortega-Arranz, O. Viberg, I.-A. Chounta, B. McLaren, & J. Jovanovic (Eds.), *Proceedings of the 16th International Conference on Computer Supported Education* (Vol. 2, pp. 221–229). Science and Technology Publications. https://doi.org/10.5220/0012628000003693

Ouh, E. L., Gan, B. K. S., Jin Shim, K., & Wlodkowski, S. (2023). ChatGPT, can you generate solutions for my coding exercises? An evaluation on its effectiveness in an undergraduate Java programming course. In M. J . Laakso (Chair), *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education* (Vol. 1, pp. 54–60). ACM.

Padilla, J. R. C., Montefalcon, M. D. L., & Hernandez, A. A. (2023). Language AI in programming: A case study of ChatGPT in higher education using natural language processing. In R. Adnan (Chair), *Proceedings of the 11th Conference on Systems, Process & Control* (pp. 276–281). IEEE. https://doi.org/10.1109/ICSPC59664.2023.10420194

Park, A., & Kim, T. (2025). Code suggestions and explanations in programming learning: Use of ChatGPT and performance. *The International Journal of Management Education, 23*(2), Article 101119. https://doi.org/10.1016/j.ijme.2024.101119

Piccolo, S. R., Denny, P., Luxton-Reilly, A., Payne, S. H., & Ridge, P. G. (2023). Evaluating a large language model's ability to solve programming exercises from an introductory bioinformatics course. *PLOS Computational Biology, 19*(9), Article e1011511. https://doi.org/10.1371/journal.pcbi.1011511

Popovici, M. D. (2024). ChatGPT in the classroom: Exploring its potential and limitations in a functional programming course. *International Journal of Human–Computer Interaction, 40*(22), 7743–7754. https://doi.org/10.1080/10447318.2023.2269006

Prather, J., Reeves, B. N., Leinonen, J., MacNeil, S., Randrianasolo, A. S., Becker, B. A., Kimmel, B., Wright, J., & Briggs, B. (2024). The widening gap: The benefits and harms of generative AI for novice programmers. In P. Denny, L. Porter, M. Hamilton, & B. Morrison (Eds.), *Proceedings of the 2024 ACM Conference on International Computing Education Research* (Vol., 1, pp. 469–486). ACM. https://doi.org/10.48550/arXiv.2405.17739

Qureshi, B. (2023). ChatGPT in computer science curriculum assessment: An analysis of its successes and shortcomings. In M. D. Ventura (Chair), *Proceedings of the 2023 9th International Conference on e-Society, e-Learning and e-Technologies* (pp. 7–13). ACM. https://doi.org/10.1145/3613944.3613946

Rahman, M. S., Sabbir, M. M., Zhang, J., Moral, I. H., & Hossain, G. M. S. (2023). Examining students' intention to use ChatGPT: Does trust matter? *Australasian Journal of Educational Technology, 39*(6), 51–71. https://doi.org/10.14742/ajet.8956

Rawas, S. (2024). ChatGPT: Empowering lifelong learning in the digital age of higher education. *Education and Information Technologies, 29*(6), 6895–6908. https://doi.org/10.1007/s10639-023-12114-8

Rees, G. P., & Gibert, I. (2024). A textbook or ChatGPT: Which helps novice programmers most with unknown terms? In K. Š. Despot, A. Ostroški Anić, & I. Brač (Eds.), *Proceedings of the XXI EURALEX International Congress* (pp. 265–275). European Association for Lexicography. http://euralex.org/wp-content/themes/euralex/proceedings/Euralex%202024/EURALEX2024_Pr_p265-275_Rees-Gibert.pdf.pdf

Savelka, J., Agarwal, A., An, M., Bogart, C., & Sakr, M. (2023). Thrilled by your progress! Large language models (GPT-4) no longer struggle to pass assessments in higher education programming courses. In K. Fisler, P. Denny, D. Franklin, & M. Hamilton (Eds.), *Proceedings of the 2023 ACM Conference on International Computing Education Research* (Vol. 1. pp. 78–92). ACM. https://doi.org/10.1145/3568813.3600142

Silva, C. A. G. D., Ramos, F. N., De Moraes, R. V., & Santos, E. L. D. (2024). ChatGPT: Challenges and benefits in software programming for higher education. *Sustainability, 16*(3), Article 1245. https://doi.org/10.3390/su16031245

Smolansky, A., Cram, A., Raduescu, C., Zeivots, S., Huber, E., & Kizilcec, R. F. (2023). Educator and student perspectives on the impact of generative AI on assessments in higher education. In D. Spikol (Chair), *Proceedings of the Tenth ACM Conference on Learning@ Scale* (pp. 378–382). ACM. https://doi.org/10.1145/3573051.3596191

Sok, S., & Heng, K. (2023). ChatGPT for education and research: A review of benefits and risks. *Cambodian Journal of Educational Research, 3*(1), 110–121. https://doi.org/10.2139/ssrn.4378735

Speth, S., Meißner, N., & Becker, S. (2023). Investigating the use of AI-generated exercises for beginner and intermediate programming courses: A ChatGPT case study. In H. Washizaki (Chair), *Proceedings of the 35th International Conference on Software Engineering Education and Training* (pp. 142–146). IEEE. https://doi.org/10.1109/CSEET58097.2023.00030

Styve, A., Virkki, O. T., & Naeem, U. (2024). Developing critical thinking practices interwoven with generative AI usage in an introductory programming course. In C. Douligeris (Chair), *Proceedings of the 2024 IEEE Global Engineering Education Conference* (pp. 01–08). IEEE. https://doi.org/10.1109/EDUCON60312.2024.10578746

Sun, D., Boudouaia, A., Yang, J., & Xu, J. (2024). Investigating students' programming behaviors, interaction qualities and perceptions through prompt-based learning in ChatGPT. *Humanities and Social Sciences Communications, 11*, Article 1447. https://doi.org/10.1057/s41599-024-03991-6

Ţală, M. L., Müller, C. N., Năstase, I. A., & Gheorghe, G. (2024). Exploring university students' perceptions of generative artificial intelligence in education. *Amfiteatru Economic Journal, 26*(65), 71–88. https://doi.org/10.24818/EA/2024/65/71
Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., Weeks, L., Hempel, S., Ekl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garrity, C., … Strauss, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine, 169*(7), 467–473. https://doi.org/10.7326/m18-0850
Urhan, S., & Kocadere, S. A. (2024). Problem-solving through pair-programming: The mediational role of ChatGPT. In *Proceedings of the 5th International Conference in Electronic Engineering, Information Technology & Education* (pp. 1–10). IEEE. https://doi.org/10.1109/EEITE61750.2024.10654434
Valverde-Rebaza, J., González, A., Navarro-Hinojosa, O., & Noguez, J. (2024). Advanced large language models and visualization tools for data analytics learning. *Frontiers in Education*, *9*, Article 1418006. https://doi.org/10.3389/feduc.2024.1418006
Willert, N., & Würz, P. K. (2025). Assisting quality assurance of examination tasks: Using a GPT model and Bayesian testing for formative assessment. *Computers and Education: Artificial Intelligence, 8*, Article 100343. https://doi.org/10.1016/j.caeai.2024.100343
Wu, T. T., Lee, H. Y., Chen, P. H., Lin, C. J., & Huang, Y. M. (2025). Integrating peer assessment cycle into ChatGPT for STEM education: A randomised controlled trial on knowledge, skills, and attitudes enhancement. *Journal of Computer Assisted Learning, 41*(1), Article e13085. https://doi.org/10.1111/jcal.13085
Yang, A. C., Lin, J. Y., Lin, C. Y., & Ogata, H. (2024). Enhancing Python learning with PyTutor: Efficacy of a ChatGPT-based intelligent tutoring system in programming education. *Computers and Education: Artificial Intelligence, 7*, Article 100309. https://doi.org/10.1016/j.caeai.2024.100309
Yilmaz, R., & Yilmaz, F. G. K. (2023). 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, 4*, Article 100147. https://doi.org/10.1016/j.caeai.2023.100147
Yunianto, W., Lavicza, Z., Kastner-Hauler, O., & Houghton, T. (2024). Investigating the use of ChatGPT to solve a GeoGebra-based mathematics + computational thinking task in a geometry topic. *Journal on Mathematics Education*, *15*(3), 1027–1052. https://doi.org/10.22342/jme.v15i3.pp1027-1052
Zhang, Z., Dong, Z., Shi, Y., Price, T., Matsuda, N., & Xu, D. (2024). Students' perceptions and preferences of generative artificial intelligence feedback for programming. *Proceedings of the AAAI Conference on Artificial Intelligence*, *38*(21), 23250–23258. https://doi.org/10.1609/aaai.v38i21.30372

Corresponding author: Elaheh Yadegaridehkordi**,** yellahe@gmail.com

Please cite as: Baig, M. I., Yadegaridehkordi, E., & Bokani, A. (2026). ChatGPT in computer programming education: A review of current literature and applications. *Australasian Journal of Educational Technology*, *42*(2), 138–160. https://doi.org/10.14742/ajet.10588