2026 ASEE ANNUAL CONFERENCE June 21–24, 2026 | Charlotte Convention Center, NC
Host Sponsor UNC Charlotte
Paper ID #52158

Exploring Critical Thinking in Programming Education: A Systematic Literature Review in Computing Education

Ms. Maryam Khalid Multani, University of Florida

Maryam Khalid Multani is currently a second-year PhD student in Engineering Education at the University of Florida.

Ms. Laura Melissa Cruz Castro, University of Florida

Dr. Laura Cruz is an instructional assistant professor in the engineering education department at University of Florida.

Exploring Critical Thinking in Programming Education: A Systematic Literature Review in Computing Education

Abstract

Critical thinking is a fundamental skill in programming education, particularly as automation and generative artificial intelligence (GenAI) increasingly shape how programming is taught and practiced. Programming requires students to analyze ill-structured problems, evaluate multiple solution paths, debug and refine code, and justify design decisions. At the same time, concerns about GenAI have grown, with recent national survey data showing that nearly 90% of faculty in the United States believe these tools may weaken students' critical thinking skills. Despite these concerns, there is limited synthesis of how critical thinking is currently defined, assessed, and supported within programming education research. This study addresses this gap through a systematic review of empirical studies focused on critical thinking in higher education programming contexts.

Following PRISMA guidelines, we reviewed 19 empirical studies to examine the frameworks used to conceptualize critical thinking, the components emphasized, the instruments used for assessment, and the instructional interventions shown to be effective. The findings reveal substantial variation in how critical thinking is defined and measured, with most studies relying on broad educational or cognitive frameworks rather than critical thinking–specific models. However, there was consistent emphasis on analysis and evaluation as central components of critical thinking in programming. Interventions that showed positive outcomes typically involved authentic programming tasks, structured reflection, and timely feedback. Importantly, several studies demonstrated that GenAI tools, when intentionally designed as supports for reflection and evaluation, can contribute positively to students' critical thinking. Together, these results suggest that the impact of GenAI on critical thinking is shaped less by the technology itself and more by the pedagogical choices surrounding its use, pointing to the need for clearer frameworks and theory-informed designs in future programming education research.

Introduction

Critical thinking (CritT) is an essential skill for students and professionals alike [1], [2]. This is especially important in programming, where learners must weigh multiple options and arrive at decisions that are both effective and ethically sound [3]. Programming tasks routinely require students to analyze ill-structured problems, evaluate alternative solutions, debug and refine code, and justify design decisions. Effective programming involves reasoning about trade-offs, anticipating unintended consequences, and applying prior knowledge to new contexts. At the same time, concerns about emerging technologies have heightened the urgency of this skill, with a recent national report indicating that nearly 90% of faculty in the United States believe that generative AI will diminish students' critical thinking skills [4]. Developing students' critical thinking in programming contexts is therefore essential for preparing engineers and computing professionals to address complex, real-world problems in an increasingly AI-mediated landscape. In programming education, efforts to support the development of critical thinking have taken a variety of forms. Researchers have drawn on educational frameworks to inform instructional design, incorporated project-based and experiential learning activities, and emphasized feedback, reflection, and evaluation in programming tasks. Students are often placed in roles that require them to analyze problem requirements, assess alternative solutions, justify design decisions, and revise their work. Together, these approaches reflect sustained interest in fostering critical thinking through authentic programming practices. However, research in programming education has focused more on developing problem-solving and computational thinking skills. Also, there exists an overlap and thus a confounding relationship between computational thinking and critical thinking [3].

This is further compounded by the fear around GenAI tools that excessive dependence on such tools might detrimentally affect students' abilities to think critically and solve problems [4], [5] as they become passive recipients of AI-generated content rather than being agents of their own learning [6], [7]. Thus, a systematic literature review of critical thinking development and assessment in programming contexts is timely and needed.

Literature Review

Although prior systematic reviews have examined computational thinking in programming courses and critical thinking in higher education more broadly [8], [9], and some have examined the effects of Generative Artificial Intelligence (GenAI) on critical thinking skills in higher education [5], there has been no focused synthesis of empirical research that specifically addresses the development and assessment of critical thinking in programming contexts. This gap is increasingly consequential given recent technological shifts, particularly the growing prevalence of GenAI tools, that are reshaping how programming tasks are learned, performed, and evaluated. As programming practice increasingly emphasizes evaluation, judgment, and decision-making, understanding how critical thinking is conceptualized and supported in programming education has become both timely and necessary.

To address this gap, this study presents a systematic literature review of empirical research on critical thinking in programming courses within higher education. The review synthesizes how critical thinking is defined, the frameworks used to guide its development and assessment, the components of critical thinking emphasized, the instruments employed to measure it, and the instructional interventions designed to support its development. By integrating these dimensions, the study aims to provide a clearer picture of how critical thinking is currently addressed in programming education research. Four research questions guide the study: (1) What frameworks have been used or proposed for developing and assessing critical thinking in programming courses? (2) What components of critical thinking are addressed by these frameworks, and to what extent? (3) What instruments have been used to assess critical thinking? (4) What instructional interventions have proven effective in supporting the development of critical thinking skills?

Methods

This study employed a systematic literature review guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework [10]. A structured search string was developed to identify empirical studies examining critical thinking in programming contexts within higher education:

"TITLE-ABS-KEY (('critical thinking') AND (programming OR coding) AND ('higher education' OR college OR university OR undergraduate OR postsecondary))"

Searches were conducted in Scopus, ERIC, and the ACM Digital Library, targeting peer-reviewed journal articles and conference proceedings. The initial search returned 385 records. After filtering for English-language, peer-reviewed publications, 195 records remained. Duplicate removal resulted in 188 unique articles, which were screened at the abstract level by two researchers.

Inclusion criteria (IC) required that studies:

1. Report empirical research, and
2. Focus on, or explicitly aim to develop or assess, critical thinking in a programming course or programming-related learning environment.

Exclusion criteria (EC) eliminated studies that:

1. Were literature reviews or meta-analyses, or
2. Were not implemented in a higher education context.

Discrepancies during abstract screening were resolved through discussion between the two researchers. Following abstract screening, 25 articles were retained for full-text review. Five of these studies were subsequently excluded because their participant populations consisted of school-aged learners rather than postsecondary students. One study was not in programming. This resulted in a final corpus of 19 articles included in the analysis.

The complete study selection process is summarized using a PRISMA flow diagram (see Figure 1).

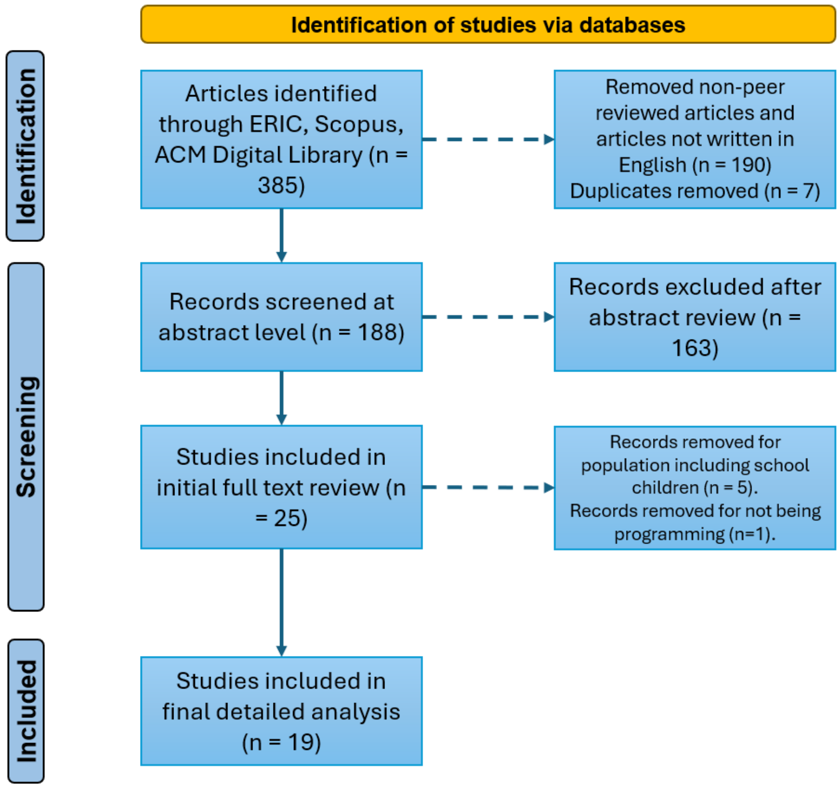

Figure 1 PRISMA flow diagram for systematic literature review

During full-text review, the authors also back-tracked assessment and instructional instruments referenced in the included studies to identify the original critical thinking frameworks or instruments from which adapted measures were derived. Furthermore, the researchers used inductive coding to perform thematic coding analysis [11] on the interventions on critical thinking within the studies to find themes within these interventions.

Results

Of the 19 articles reviewed, nine explicitly proposed or cited a definition of critical thinking, while 10 did not provide a formal definition. Among those without definitions, three framed critical thinking as part of higher-order thinking skills (HOTS), and two treated it as a component of computational thinking, without further elaboration.

RQ1: What frameworks have been used or proposed for developing and assessing critical thinking in programming courses?

Twelve of the nineteen articles employed an explicit framework to guide the development or assessment of critical thinking (see Table 1). The remaining seven articles did not explicitly adopt or propose a framework for developing or assessing critical thinking, despite referencing critical thinking in study goals, outcomes, or discussion sections.

Among the twelve studies that used a framework, several drew models that are not exclusively designed for critical thinking but were nevertheless used to support or operationalize critical-thinking-related constructs. The most frequently used frameworks were revised Bloom's Taxonomy ($n = 3$) and Pintrich's Self-Regulated Learning (SRL) model ($n = 3$). In these studies, Bloom's Taxonomy was used to frame higher-order cognitive processes associated with critical thinking, while SRL frameworks were used to structure planning, monitoring, and reflection processes linked to critical thinking development in programming contexts.

Two studies ($n = 2$) explicitly adopted the Facione Delphi Consensus framework, directly targeting critical thinking skills and subskills. Single studies ($n = 4$) employed the Watson–Glaser Critical Thinking Appraisal, the 6Ts model, Chiruguru's 4Cs of 21st-century skills, and the Paul and Elder framework of critical thinking. Overall, the frameworks used varied in their degree of alignment with critical thinking, ranging from general cognitive or learning frameworks applied to critical thinking outcomes to frameworks explicitly developed for critical thinking assessment.

Table 1 Summary of the components of critical thinking addressed by the studies by using each framework

Framework	Components of Critical Thinking addressed
Bloom's Taxonomy [12]	**Analysis**; **Evaluation**; Creation
Pintrich's Self-regulated Learning [13], [14]	Applying prior knowledge to new situations; **Analysis; Evaluation**
Watson-Glaser Appraisal [15]	Recognition of Assumptions; **Evaluation of Arguments**; and Drawing logical Conclusions
6Ts model [16]	Logical thinking; Systems thinking; Creative thinking; Sustainable thinking; Strategic thinking; Logical thinking
Chiruguru's 4Cs [17]	Reasoning; **Analysis; Evaluation**; Synthesis; Interpretation; Problem solving
Paul & Elder (PA-GPT) [18]	**Evaluating** other's work; Articulating understanding
Facione Delphi Consensus framework on critical thinking skills and subskills [19]	Interpretation; **Analysis; Evaluation**; Inference; Self-regulation

RQ2: What components of critical thinking are addressed by these studies, and to what extent?

Across frameworks, analysis and evaluation were the most frequently addressed components of critical thinking. Analysis was explicitly referenced in four frameworks (Bloom's Taxonomy; Pintrich's SRL model; Chiruguru's 4Cs; and the Facione Delphi Consensus framework), while evaluation appeared in six frameworks (Bloom's Taxonomy; Pintrich's SRL model; Watson–Glaser Appraisal; Chiruguru's 4Cs; Paul & Elder; and the Facione Delphi Consensus framework). This recurrence indicates a consistent overlap across frameworks in the emphasis placed on analysis and evaluation as central components of critical thinking in computing and programming contexts.

Other critical thinking components, including interpretation, inference, reasoning, problem-solving, and self-regulation, were addressed less consistently and appeared in fewer frameworks. Frameworks such as the Facione Delphi Consensus and Chiruguru's 4Cs incorporated a broader range of components, whereas others focused on more narrowly defined skills or higher-order cognitive processes.

With respect to the extent of articulation, three articles referenced critical thinking as part of higher-order thinking skills [20], [21], [22], one as part of problem solving [23], and one as a component of computational thinking [24] , without specifying which critical thinking components were targeted or how they were enacted.

RQ3: What instruments have been used to assess critical thinking?

The most commonly used instrument was the 4C1P Awareness questionnaire [25], followed by the Motivated Strategies for Learning Questionnaire (MSLQ) [14] (see Table 2). In many studies, the instruments as well as components of critical thinking skills assessed by the instruments were not specified. The reviewers backtracked the instruments referred to extract this information.

Table *2* Study, Critical Thinking Frameworks, CritT Components covered, Instruments, CritT Components in Instrument, Instructional Interventions, and CritT Evidence

Study	Framework used	Components of Critical thinking covered	Instrument used	Components of Critical thinking covered by instrument	Intervention type	Evidence for CritT
[26]	Watson-Glaser Appraisal	Inference, recognition of assumptions, deduction, interpretation, evaluation of arguments	Watson-Glaser Appraisal	Inference; recognition of assumptions; deduction; interpretation; evaluation of arguments	Type of programming course (procedure vs. object oriented)	No evidence of improved critical thinking skills between the two groups.
[22]	**Bloom's taxonomy for HOTS included in GenAI-Ped framework**	**Applying; Analysis; Evaluation**	**Custom Critical Thinking Rubric**	**Error Identification, Debugging Strategies, Optimizing Code, Evaluating Solutions, Use of Logical**	**GenAI integration into the course activities using a Self-regulated learning model [27]**	**Gains in critical thinking components, particularly "Conditional Thinking," which rose from 1.680 to 2.120, and "Use of Logical Constructs" rising from 1.560 to 1.920**

				Constructs, Conditional Thinking		
[16]	6Ts model	Logical thinking, systems thinking, creative thinking, sustainable thinking, strategic thinking, logical thinking.	Self-reflection questionnaire	Logical thinking, systems thinking, creative thinking, sustainable thinking, strategic thinking, logical thinking.	Developing materials for the course centered around the 6Ts model. Including, experiential learning, self-reflection, and peer learning.	No evidence of Critical Thinking improvement, yet there was evidence of a stronger understanding of the 6Ts (from −1 to 0 and 0 to 1). The only exception was systems thinking where 5 out of 8 transitions from 1 to 0.
[28]	**Facione Delphi Consensus framework on critical thinking skills and subskills**	**Interpretation; Analysis; Evaluation; Inference; Self-regulation**	**Custom survey based on Critical thinking skills but focusing more on critical thinking practices**	**Gathering & Examining; Information for Significance; Analyzing Information; Evaluating Information; Justification & Rationale; Reflecting; Self-direction**	**Using different GenAI tools as pair programmers and peer reviewers.**	**A self-reported critical thinking questionnaire had average positive changes between the pre and post report.**
[29]	Facione Delphi Consensus framework on critical thinking skills and subskills	Interpretation; **Analysis; Evaluation**; Inference; Self-regulation	Custom survey based on Critical thinking skills but focusing more on critical thinking practices	Gathering & Examining; Information for Significance; Analyzing Information; Evaluating Information; Justification & Rationale; Reflecting; Self-direction	Use of different GenAI tools as a pair programmer and peer reviewer	Learner opinions and practices changed after engaging with the Generative AI lab assessment but were not significant.
[23]	**Not specified**	Not specified, but exploring problem solving strategies said to enhance critical thinking	problem-solving ability and critical thinking tendency Questionnaire developed by Lai and Hwang [30]	Not specified.	The study compared a GPT-supported intervention (GPT-CMDP) with a conventional approach (C-CMDP). The GPT-CMDP intervention guided students through six steps, including	Statistically significant differences for critical thinking awareness and problem-solving skills between GPT-CMDP (M=4.28, SD=0.59) and C-CMDP groups (M=3.95, SD=0.63).

					structured content, micro-level syntax breakdown, interactive exploration with ChatGPT, immediate feedback, reflection, and self-paced learning. This design aimed to scaffold learning while promoting active engagement and metacognitive practice.	
[21]	**Revised Bloom's taxonomy for HOTS**	**Higher order nature of CritT addressed, "HOT is essential for learners to make connections between concepts, solve complex problems, and apply knowledge in unfamiliar contexts "**	**4C1P Awareness questionnaire (4C1PA) [25]**	**Not specified.**	**Feedback mechanisms in VR learning environment**	**Critical thinking findings not reported separately but as part of HOTS. Pre and post tests revealed significant improvement in ability to think critically.**
[18]	**PA-GPT adheres to Paul and Elder framewor k**	**Evaluating other's work; Articulating understanding**	**4C1P Awareness questionnaire (4C1PA) [25]**	**Not specified.**	**GenAI tool as a peer reviewer and engaging with student questions**	**Significant differences between experimental and control group in critical thinking (F = 37.00, p < 0.001).**
[31]	**HOTS, Revised Bloom's taxonomy.**	**Analysis; Evaluation**	**Chinese translation of a HOTS questionnaire (4C1P Awareness questionnaire (4C1PA) [25])**	**Not specified.**	**Computational Thinking-based diagnostic reflective report system**	**After controlling for pre-test scores, significant differences were found between the experimental and control groups in critical thinking (F = 15.54, p < 0.001)**
[20]	**Revised Bloom's Taxonomy**	**Analysis; Evaluation**	**4C1P Awareness**	**Not specified.**	**Two-stage feedback processes incorporating**	**There were significant differences between the groups in critical**

			questionnaire (4C1PA) [25]		student reflection and revision	thinking (t = 5.45, p < 0.001)
[32]	Pintrich's SRL model	Applying prior knowledge to new situations; Analysis; Evaluation	Motivated Strategies for Learning Questionnaire (MSLQ) [14]	Applying prior knowledge to new situations, analysis and evaluation.	Exposition to DSLab Tool, which was a tool designed by the authors to teach distributed systems.	60% of students applied prior knowledge to new situations. 64% of students were able to analyze and evaluate information in a thoughtful manner.
[33]	Pintrich's SRL model	Not specified in paper but from the model: Applying prior knowledge to new situations, analyzing and evaluating information thoughtfully	Motivated Strategies for Learning Questionnaire (MSLQ) [14]	Applying prior knowledge to new situations, analyzing and evaluating information thoughtfully	Control group: traditional classroom tutoring review activities, experimental group: personalized intervention review activities.	Experimental group has significantly higher abilities [ANCOVA critical thinking (F = 11.93,p < .05)]. (ANOVA) showed activity × critical thinking (F = 9.53, p = .000)
[24]	Not specified but Critical thinking considered a component of Computational thinking	Not specified.	Pre and post questionnaires of computational thinking, semi-structured interview	Not specified.	The Experimental Group (EG) was exposed to Progressive flowcharts based on interpolation, extrapolation, and reinterpretation model.	Pre-test had no significant differences between groups, EG obtained significantly higher scores than CG in Critical thinking EG (M=4.06, SD = 0.54), CG (M = 3.58, SD = 0.53)
[34]	Not specified	Not specified.	Industry experts assessed CritT. Instrument not specified	Identify functional and non-functional requirements. Analyze the given problem; Evaluate options, considering data and pros/cons; Address industry-related complex issues; Propose innovative	The intervention was composed of three stages. Stage 1 (Bridging the gap); Stage 2 (Experiential Learning); Stage 3 (Industry collaboration).	Enhancement in all components of (Industry related) critical thinking skills.

				solutions to current issues		
[17]	**Chiruguru 's 4Cs**	**Not specified**	**Observation report, self and peer assessment**	**Not specified**	**Digital Making Activities**	**Critical thinking mean difference increased by 11 points between assessments According to observation report: "participants were able to solve problems and think critically"**
[35]	**Not specified.**	**Analysis; Evaluation**	**Critical thinking questionnaire (Assessing multidimensional students' perceptions of twenty-first-century learning practices) [36]**	**Analysis; Reasoning; Evaluation**	**Type of learning approach, Conventional Scrum Project (C-SP) for control and Peer Assessment Scrum Project (PA-SP) for experimental**	**Critical thinking scores were higher n PA-SP group. Significant difference between the two groups on the post-questionnaire scores (F = 15.90, p < 0.001).**
[37]	**Not specified**	Not specified.	Anonymous survey	Not specified	Case study on pill-sized capsule robot formulated as a course module	Critical thinking (average 3.46). However, there were no pre-post or control groups.
[38]	**Not specified.**	Not specified.	Survey used to measure critical thinking skills.	Not specified.	Students were exposed to a visual programming Tool GameSalad for Digital Media Course	Not mentioned.
[39]	**CritT as a strategy of SRL. Not clearly specified but possibly Pintrich's SRL model**	**Not specified.**	**Eye tracking movements measured: Total fixation count (TFC), Total Fixation Duration (TFD); Scanpath analysis.**	**Not specified**	**Eye-tracking and retrospective think-aloud data for program comprehension.**	**Statistically significant differences between LG and HG for critical thinking strategy CT2 for Program Code 1 (U = 240.000, |Z| = 3.714, p < 0.001) and for Program Code 2 critical thinking strategies CT1 to CT3 (U = 225.000–946.000, |Z| = 2.702–6.590, p < 0.01)**

RQ4: What instructional interventions have proven effective in supporting the development of critical thinking skills?

To address Research Question 4, we examined the extent to which the reviewed instructional interventions demonstrated measurable effects on students' critical thinking or related subcomponents. Of the 19 studies included in the review, 13 reported positive effects on critical thinking skills or specific dimensions such as analysis, evaluation, or inference, while one additional study identified improvements in critical thinking awareness rather than skill development; the remaining five studies reported no significant effects on critical thinking or critical awareness (see Table 2). The following themes were identified.

Technology-Supported Interventions

Eight of the thirteen studies reporting positive effects incorporated technology-enhanced learning environments, including GenAI tools [18], [22], [23], VR environments [20], and custom-built instructional systems [10], [26]. GenAI was used most effectively when positioned as a peer reviewer, pair programmer, or reflective scaffold, rather than a content generator [17], [28], [29] .

Reflection and Feedback Cycles

Eleven interventions incorporated structured reflection or enhanced feedback mechanisms, including self-reflection questionnaires [16], two-stage feedback with revision[20], GenAI-mediated reflection [21], [29], and diagnostic reflective reporting systems [31]. These studies consistently showed gains in critical thinking or its subcomponents, suggesting that opportunities to reflect, revise, and justify reasoning are central mechanisms.

Self-Regulated Learning and Metacognitive Scaffolding

Interventions grounded in SRL frameworks showed strong evidence of effectiveness. Studies using Pintrich's SRL model [32], [33], and GPT-supported metacognitive designs [23] reported statistically significant improvements in critical thinking awareness, analysis, and evaluation. The GPT-CMDP intervention [23] demonstrated significant gains over a conventional approach by combining structured content, micro-steps, immediate feedback, reflection, and self-paced learning.

Human and Social Enhancement

A smaller subset of interventions emphasized human-centered enhancements, such as industry collaboration, peer review, and expert feedback [18], [34], [35]. These approaches supported deeper evaluative and justificatory thinking, particularly complex, ill-structured problems.

Innovative Learning Activities

Two studies demonstrated gains by restructuring course activities themselves, including progressive flowcharts [24] and digital making activities [17]. These approaches emphasized problem interpretation, solution evaluation, and iterative refinement.

Discussion

This review examined how critical thinking is conceptualized, operationalized, assessed, and supported in programming education research. Across the four research questions, the findings point to both convergence around certain practices and fragmentation in construct definition and measurement, particularly with respect to how critical thinking is distinguished from adjacent constructs.

RQ1: What frameworks have been used or proposed for developing and assessing critical thinking in programming courses?

Only a small subset of the frameworks used in the reviewed studies were originally developed as critical thinking–specific frameworks. Rather than proposing or using frameworks specific to critical thinking, most studies instead adopted broader educational or cognitive frameworks such as revised Bloom's Taxonomy, Pintrich's Self-Regulated Learning model, or Chiruguru's 4Cs of 21st-century skills; and positioned critical thinking as a subcomponent of higher-order thinking, self-regulation, or broader computational thinking skills. This pattern suggests that, in programming education research, critical thinking is often treated as an embedded or secondary construct rather than as a primary object of instructional design or assessment. While these frameworks offer useful lenses, their adoption may also contribute to conceptual ambiguity when critical thinking is not explicitly delineated from related constructs.

With adjacent frameworks being used more rather than critical thinking frameworks, this could be due to lack of clarity among instructors and researchers as to what critical thinking constitutes in programming contexts. Thus, it becomes important to study whether new critical thinking frameworks specifically for programming education should be developed.

RQ2: What components of critical thinking are addressed by these studies, and to what extent?

It is also evidenced from Table 2 that **evaluation** and **analysis** were the most commonly appearing components of critical thinking within the frameworks used. This indicates that these components could be fundamental in a critical thinking framework developed specifically for programming education.

Across the studies and frameworks examined, there was substantial convergence around cognitive components such as analysis, evaluation, reasoning, and the application of prior knowledge. However, the extent to which these components were systematically addressed or operationalized within individual studies was often unclear. Many articles referenced a framework at a high level without detailing how specific critical thinking components were enacted in instructional activities or measured in student work. As a result, it is difficult to determine whether observed gains reflect comprehensive development of critical thinking or improvement in a narrower subset of skills.

This finding points to a need for more explicit alignment between frameworks, instructional design, and targeted critical thinking components in future research.

RQ3: What instruments have been used to assess critical thinking?

The assessment of critical thinking relied primarily on previously developed and validated instruments, including the 4C1P Awareness questionnaire, the Motivated Strategies for Learning Questionnaire, and critical thinking or problem-solving questionnaires adapted from prior work [30], [36]**.** While the use of established instruments supports reliability and comparability, many of these tools assess constructs that overlap with critical thinking, such as self-regulation, higher-order thinking, or problem solving, rather than critical thinking in isolation. Although some studies developed custom instruments or rubrics tailored to programming contexts, these were less common. Taken together, these findings highlight ongoing challenges related to construct validity and raise questions about how well commonly used instruments capture the multidimensional nature of critical thinking in programming tasks. These patterns suggest a need for assessment approaches that more directly and transparently align with explicitly defined critical thinking constructs in programming contexts.

RQ4: What instructional interventions have proven effective in supporting the development of critical thinking skills?

The instructional interventions associated with improvements in critical thinking consistently emphasized active engagement in authentic or quasi-authentic tasks, coupled with timely feedback. Effective interventions often placed students in roles analogous to workplace contexts such as peer reviewers, pair programmers, or evaluators of AI-generated code; thus, requiring them to analyze solutions, justify decisions, and revise their work. These approaches align closely with core critical thinking processes while also intersecting with self-regulation, computational thinking, and problem-solving skills.

A recent national report suggests that 90% of faculty in the United States believe that generative AI will diminish students' critical thinking skills [5]. This concern is understandable, particularly in programming contexts where automation can easily replace effortful reasoning. However, our review tells a more nuanced story. We identified four interventions in programming education in which generative AI tools, when intentionally paired with reflection and feedback grounded in educational theory, were associated with positive impacts on students' critical thinking skills. These findings suggest that the risks often attributed to generative AI are not inherent to the tools themselves but are strongly shaped by how they are designed and used in instructional settings.

More broadly, despite ongoing concerns about the negative effects of generative AI on students' thinking, this review shows that many researchers are actively experimenting with ways to integrate these tools to support, rather than undermine, critical thinking. This pattern highlights a sense of resilience within the programming education community and a willingness to engage with emerging technologies, even those widely viewed as problematic. Instead of rejecting generative AI outright, educators are exploring how it can become a productive part of the learning process. These efforts frame generative AI as an opportunity that can either hinder or enhance critical thinking, depending on the presence of scaffolding, reflection, and feedback. For the broader computing education community, these findings point to the importance of moving beyond simple narratives of harm and toward theory-informed approaches that intentionally use generative AI to support the development of critical thinking skills.

Conclusion and Future Directions

As programming education continues to evolve in response to automation and generative AI, critical thinking is becoming increasingly central to what students must learn beyond syntax and tools. This review examined how critical thinking is conceptualized, assessed, and supported in programming education research and found both convergence and fragmentation. While many studies drew on broad educational frameworks rather than critical thinking–specific models, there was clear agreement around core components such as analysis and evaluation. At the same time, critical thinking was often treated as an embedded or secondary construct, with limited clarity in how it was operationalized or distinguished from related constructs such as problem solving or self-regulation. Assessment practices reflected similar challenges, relying largely on instruments that capture overlapping constructs rather than critical thinking in programming tasks directly. Together, these findings point to a need for greater conceptual precision and alignment among frameworks, instructional design, and assessment.

Importantly, instructional interventions associated with positive outcomes consistently emphasized active engagement in authentic tasks and timely feedback. Despite widespread concerns that generative AI may erode students' critical thinking, this review identified multiple studies in which AI tools, when intentionally coupled with reflection and feedback grounded in educational theory, supported critical thinking development. These findings suggest that in computing education the impact of generative AI on critical thinking is not inherent to the technology but depends on how it is pedagogically framed. Future work should prioritize the development of programming-specific critical thinking frameworks and assessments, as well as theory-informed designs that leverage generative AI as a tool for thinking rather than a substitute for it.

Acknowledgements

The authors would like to thank Ms. Olivia Suban for serving as one of the reviewers during abstract review and initial full text review stages.

References

[1] S. ŽivkoviĿ, “A Model of Critical Thinking as an Important Attribute for Success in the 21st Century,” *Procedia - Social and Behavioral Sciences*, vol. 232, pp. 102–108, Oct. 2016, doi: 10.1016/j.sbspro.2016.10.034.
[2] T. J. Brudvig, A. Dirkes, P. Dutta, and K. Rane, “Critical Thinking Skills in Health Care Professional Students: A Systematic Review,” *Journal of Physical Therapy Education*, vol. 27, no. 3, p. 12, Fall 2013.
[3] S. Maesschalck, “Critical Thinking: The Code to Crack Computer Science Education,” *JITE:IIP*, vol. 23, p. 013, 2024, doi: 10.28945/5387.
[4] E. Townsend and staff, “Elon/AAC&U national survey: 95% of college faculty fear student overreliance on AI,” Today at Elon. Accessed: Jan. 21, 2026. [Online]. Available: https://www.elon.edu/u/news/2026/01/21/elon-aacu-national-survey-95-of-college-faculty-fear-student-overreliance-on-ai/
[5] P. P. Premkumar, M. R. K. N. Yatigammana, and S. Kannangara, “Impact of Generative AI on Critical Thinking Skills in Undergraduates: A Systematic Review,” *sljo-j-jdrra*, vol. 2, no. 1, pp. 199–215, Dec. 2024, doi: 10.4038/jdrra.v2i1.55.
[6] A. Ayari and L. Ouerfelli, “Generative Ai in the Classroom: Balancing Innovation, Fear, and Necessity,” in *2025 IEEE Global Engineering Education Conference (EDUCON)*, Apr. 2025, pp. 1–6. doi: 10.1109/EDUCON62633.2025.11016392.
[7] T. Baeten, “The Influence of Artificial Intelligence Adoption on Educator Perceptions”.
[8] N. Rothinam, R. Vengrasalam, S. Naidu, S. Nachiappan, and S. Jabamoney, “Systematic literature review on critical thinking in higher education,” *2576-8484*, vol. 9, no. 5, pp. 2046–2063, May 2025, doi: 10.55214/25768484.v9i5.7377.
[9] F. J. Agbo, S. S. Oyelere, J. Suhonen, and S. Adewumi, “A Systematic Review of Computational Thinking Approach for Programming Education in Higher Education Institutions,” in *Proceedings of the 19th Koli Calling International Conference on Computing Education Research*, in Koli Calling ’19. New York, NY, USA: Association for Computing Machinery, Nov. 2019, pp. 1–10. doi: 10.1145/3364510.3364521.
[10] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and the P. Group*, “Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement,” *Annals of Internal Medicine*, Aug. 2009, Accessed: Jan. 13, 2025. [Online]. Available: https://www.acpjournals.org/doi/10.7326/0003-4819-151-4-200908180-00135
[11] J. W. Creswell and J. D. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications, 2017.
[12] B. S. Bloom, *A taxonomy for learning, teaching, and assessing: A revision of Bloom’s taxonomy of educational objectives*. Longman, 2010.
[13] P. R. Pintrich, “A Conceptual Framework for Assessing Motivation and Self-Regulated Learning in College Students,” *Educational Psychology Review*, vol. 16, no. 4, pp. 385–407, Dec. 2004, doi: 10.1007/s10648-004-0006-x.
[14] P. R. Pintrich and E. V. De Groot, “Motivational and self-regulated learning components of classroom academic performance,” *Journal of Educational Psychology*, vol. 82, no. 1, pp. 33–40, 1990, doi: 10.1037/0022-0663.82.1.33.
[15] G. Watson, *Watson-Glaser critical thinking appraisal*, vol. 3. Psychological Corporation San Antonio, TX, 1980.
[16] B. Kalluri, P. Prasad, P. Sharma, and D. Chippa, “Developing Future Computational Thinking in Foundational CS Education: A Case Study From a Liberal Education University

in India," *IEEE Transactions on Education*, vol. 67, no. 6, pp. 944–953, 2024, doi: 10.1109/TE.2024.3394060.
[17] D. Konstantinou, A. Parmaxi, and P. Zaphiris, "Evaluating the Development of Soft Skills Through the Integration of Digital Making Activities in Undergraduate Computing Courses," presented at the Lecture Notes in Computer Science, 2023, pp. 601–618. doi: 10.1007/978-3-031-34411-4_41.
[18] W. Ting-Ting, L. Hsin-Yu, C. Pei-Hua, L. Chia-Ju, and H. Yueh-Min, "Integrating Peer Assessment Cycle into ChatGPT for STEM Education: A Randomised Controlled Trial on Knowledge, Skills, and Attitudes Enhancement," *Journal of Computer Assisted Learning*, vol. 41, no. 1, 2025, doi: 10.1111/jcal.13085.
[19] P. A. Facione, "Using the Holistic Critical Thinking Scoring Rubric to Train the Discovery of Evidence of Critical Thinking".
[20] H. Yueh-Min, W. Wei-Sheng, L. Hsin-Yu, L. Chia-Ju, and W. Ting-Ting, "Empowering Virtual Reality with Feedback and Reflection in Hands-On Learning: Effect of Learning Engagement and Higher-Order Thinking," *Journal of Computer Assisted Learning*, vol. 40, no. 4, pp. 1413–1427, 2024, doi: 10.1111/jcal.12959.
[21] W.-S. Wang, M. Pedaste, and Y.-M. Huang, "Enhancing Higher-Order Thinking and Hands-On Performance in VR Learning Environments Through CTML-Based Feedback Design," presented at the Lecture Notes in Computer Science, 2026, pp. 373–381. doi: 10.1007/978-3-031-98185-2_39.
[22] J. Nathaniel, S. S. Oyelere, J. Suhonen, and M. Tedre, "Investigating the impact of generative AI integration on the sustenance of higher-order thinking skills and understanding of programming logic," *Computers and Education: Artificial Intelligence*, vol. 9, 2025, doi: 10.1016/j.caeai.2025.100460.
[23] Y.-F. Lee, G.-J. Hwang, and L.-C. Cheng, "Impacts of a ChatGPT-supported concept mapping approach on students' database programming achievement and their problem-solving and critical thinking awareness," *Interactive Learning Environments*, 2025, doi: 10.1080/10494820.2025.2523395.
[24] J.-H. Zhang, B. Meng, L.-C. Zou, Y. Zhu, and G.-J. Hwang, "Progressive flowchart development scaffolding to improve university students' computational thinking and programming self-efficacy," *Interactive Learning Environments*, vol. 31, no. 6, pp. 3792–3809, 2023, doi: 10.1080/10494820.2021.1943687.
[25] G.-J. Hwang, C.-L. Lai, J.-C. Liang, H.-C. Chu, and C.-C. Tsai, "A long-term experiment to investigate the relationships between high school students' perceptions of mobile learning and peer interaction and higher-order thinking tendencies," *Education Tech Research Dev*, vol. 66, no. 1, pp. 75–93, Feb. 2018, doi: 10.1007/s11423-017-9540-3.
[26] N. S. Grant, "A study on critical thinking, cognitive learning style, and gender in various information science programming classes," in *Proceedings of the 4th Conference on Information Technology Curriculum*, in CITC4 '03. New York, NY, USA: Association for Computing Machinery, 2003, pp. 96–99. doi: 10.1145/947121.947142.
[27] P. Prasad and A. Sane, "A Self-Regulated Learning Framework using Generative AI and its Application in CS Educational Intervention Design," in *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*, in SIGCSE 2024. New York, NY, USA: Association for Computing Machinery, Mar. 2024, pp. 1070–1076. doi: 10.1145/3626252.3630828.

[28] A. Styve, O. T. Virkki, and U. Naeem, “Developing Critical Thinking Practices Interwoven with Generative AI Usage in an Introductory Programming Course,” presented at the IEEE Global Engineering Education Conference, EDUCON, 2024. doi: 10.1109/EDUCON60312.2024.10578746.
[29] U. Naeem, A. Styve, and O. T. Virkki, “Stimulating Critical Thinking in a Web Programming Module with Generative AI Tools,” presented at the IEEE Global Engineering Education Conference, EDUCON, 2025. doi: 10.1109/EDUCON62633.2025.11016298.
[30] C.-L. Lai and G.-J. Hwang, “Effects of mobile learning time on students’ conception of collaboration, communication, complex problem–solving, meta–cognitive awareness and creativity,” *International Journal of Mobile Learning and Organisation*, Jan. 2015, Accessed: Jan. 21, 2026. [Online]. Available: https://www.inderscienceonline.com/doi/10.1504/IJMLO.2014.067029
[31] T.-T. Wu, H.-Y. Lee, P.-H. Chen, W.-S. Wang, and Y.-M. Huang, “Integrating Computational Thinking Diagnostic Mechanism and Reflective Learning: An Innovative Approach to Enhance Learning Outcomes in Introductory Programming,” *Journal of Computer Assisted Learning*, vol. 41, no. 5, 2025, doi: 10.1111/jcal.70121.
[32] J. M. Marquès Puig, T. Daradoumis, M. Arguedas, and L. Calvet Liñan, “Using a Distributed Systems Laboratory to Facilitate Students’ Cognitive, Metacognitive and Critical Thinking Strategy Use,” *Journal of Computer Assisted Learning*, vol. 38, no. 1, pp. 209–222, 2022, doi: 10.1111/jcal.12605.
[33] A. Y. Q. Huang *et al.*, “Personalized Intervention based on the Early Prediction of At-risk Students to Improve Their Learning Performance,” *Educational Technology and Society*, vol. 26, no. 4, pp. 69–89, 2023, doi: 10.30191/ETS.202310_26(4).0005.
[34] A. S. Nayak, N. D. Hiremath, F. M. Umadevi, P. B. Patil, and S. Chickerur, “Fostering Higher Cognitive Skills Through Industry Collaboration: A Microcontroller-Based Approach,” *Journal of Engineering Education Transformations*, vol. 37, no. Special Issue 2, pp. 688–694, 2024, doi: 10.16920/jeet/2024/v37is2/24107.
[35] C. Shao-Chen and C. Wongwatkit, “Effects of a Peer Assessment-Based Scrum Project Learning System on Computer Programming’s Learning Motivation, Collaboration, Communication, Critical Thinking, and Cognitive Load,” *Education and Information Technologies*, vol. 29, no. 6, pp. 7105–7128, 2024, doi: 10.1007/s10639-023-12084-x.
[36] C. S. Chai, F. Deng, P.-S. Tsai, J. H. L. Koh, and C.-C. Tsai, “Assessing multidimensional students’ perceptions of twenty-first-century learning practices,” *Asia Pacific Educ. Rev.*, vol. 16, no. 3, pp. 389–398, Sep. 2015, doi: 10.1007/s12564-015-9379-4.
[37] Y. Guo, S. Zhang, A. Ritter, and H. Man, “A Case Study on a Capsule Robot in the Gastrointestinal Tract to Teach Robot Programming and Navigation,” *IEEE Transactions on Education*, vol. 57, no. 2, pp. 112–121, May 2014, doi: 10.1109/TE.2013.2281025.
[38] S. Dekhane, X. Xu, and M. Y. Tsoi, “Mobile app development to increase student engagement and problem solving skills,” *Journal of Information Systems Education*, vol. 24, no. 4, pp. 299–308, 2013.
[39] G. Cheng, D. Zou, H. Xie, and F. L. Wang, “Exploring differences in self-regulated learning strategy use between high- and low-performing students in introductory programming: An analysis of eye-tracking and retrospective think-aloud data from program comprehension,” *Computers and Education*, vol. 208, 2024, doi: 10.1016/j.compedu.2023.104948.