ASIA PACIFIC JOURNAL
OF EDUCATORS AND EDUCATION

Volume 41, Issue 1, 2026
https:/ doi.org/10.21315/apjee2026.41.1.15

Research Article:

IEEC: A Framework for AI Literacy in Programming

Ong Yew Chuan[1*], Upasana Gitanjali Singh[2] and Sharifah Sumayyah Engku Alwi[1]

[1]Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Besut Campus, 22200 Besut, Terengganu, Malaysia
[2]College of Law and Management Studies, University of KwaZulu Natal Westville, 36229 KwaZulu Natal, South Africa

*Corresponding author: yewchuan@unisza.edu.my

ABSTRACT

The integration of generative artificial intelligence (GenAI) into higher education presents both opportunities and challenges, particularly in skill-based fields like computer programming. While tools such as ChatGPT can support learning, educators need structured methods to harness their benefits while addressing concerns like over-reliance and ethical misuse. This article introduces and evaluates the IEEC (Introduce, Encourage, Evaluate, Control) framework, developed to guide the thoughtful integration of GenAI tools in programming education. The framework aims to build foundational programming competence alongside essential AI literacy – defined as the ability to use GenAI critically, ethically, and effectively. Implemented in an introductory programming course (N = 60) at Universiti Sultan Zainal Abidin, Malaysia, the study employed a mixed-methods, quasi-experimental pre/post design. Quantitative data included surveys measuring programming self-efficacy, AI attitudes, ethical awareness, and assessment performance related to evaluating GenAI outputs. Qualitative data were drawn from focus group interviews. Findings show statistically significant gains ($p < 0.05$) in programming self-efficacy and students' ability to critically assess AI-generated outputs. Qualitative analysis revealed increased student engagement (Encourage) and deeper reflections on GenAI's capabilities and ethical implications (Introduce/Control). Although students demonstrated greater ethical awareness, challenges in applying ethical principles persisted (Evaluate/Control), and concerns about AI output variability remained. Overall, the IEEC framework supported confidence, critical evaluation skills, and ethical mindfulness alongside technical competence. This study contributes a carefully evaluated pedagogical model for integrating GenAI in programming education and offers a contextually grounded approach to cultivating responsible AI use in technical learning contexts. While the findings are promising, further research across diverse institutions and disciplines is needed to determine the broader applicability of the framework.

Keywords: AI literacy, programming education, generative artificial intelligence, ChatGPT, AI ethics, pedagogical framework

Published: 9 June 2026

To cite this article: Ong, Y. C., Singh, U. G., & Alwi, S. S. E. (2026). IEEC: A framework for AI literacy in programming. *Asia Pacific Journal of Educators and Education, 41*(1), 337–367. https:/ doi.org/10.21315/apjee2026.41.1.15

INTRODUCTION

The rapid proliferation of generative artificial intelligence (GenAI) has marked a pivotal moment in higher education, introducing transformative technologies that are reshaping pedagogical practices and learning environments (Dwivedi et al., 2023). Tools powered by large language models, such as ChatGPT, present a dual paradigm for educators and students. On one hand, they offer unprecedented opportunities for enhancing the educational experience by providing personalised, on-demand learning support, acting as virtual tutors that can explain complex concepts and deliver immediate feedback (Kasneci et al., 2023). This capability can foster a more adaptive, inclusive, and efficient learning ecosystem (Chan & Hu, 2023). On the other hand, the widespread accessibility of these powerful tools raises significant challenges, particularly concerning academic integrity (Cotton et al., 2024), the potential for factual inaccuracies or "hallucinations", and ethical issues related to data privacy and algorithmic bias (Chan & Hu, 2023). This duality compels the academic community to move beyond reactionary responses and proactively develop strategies that harness the benefits of GenAI while mitigating its inherent risks.

THE PROBLEM IN SKILL-BASED EDUCATION

Nowhere are these challenges more pronounced than in skill-based disciplines like computer programming. The capacity of GenAI to generate functional code snippets, debug errors, and explain algorithms offers a powerful scaffold for novice learners grappling with complex syntax and logic (Becker et al., 2023). However, this very utility introduces a significant pedagogical risk: the danger of over-reliance. When students use GenAI as a substitute for genuine cognitive effort – a phenomenon described as "cognitive offloading" – they may successfully complete assignments without internalising the foundational problem-solving skills and logical reasoning that are the bedrock of programming competence (Humble et al., 2023; Li et al., 2025). This surface-level understanding can create a critical knowledge gap, hindering their ability to tackle more advanced challenges and debug their own work effectively. Furthermore, novices often lack the domain expertise to critically evaluate the correctness, efficiency, or security of AI-generated code, leading to the uncritical acceptance of flawed or suboptimal solutions (Akçapınar & Sidan, 2024; Bente et al., 2024). This dynamic threatens to undermine the core objective of programming education: to cultivate independent, critical thinkers who can deconstruct problems and build robust solutions from first principles.

THE GAP IN PEDAGOGICAL APPROACHES

In response to the sudden rise of GenAI, the educational community has been grappling with how to adapt. Recent higher education scholarship has generally moved away from outright bans on GenAI, instead emphasising institutionally guided, transparent, and pedagogically purposeful integration supported by clear policies and course-level guidance (Moorhouse et al., 2023; Luo, 2024; McDonald et al., 2025; An et al., 2025). There

remains a conspicuous absence of structured, evidence-based pedagogical frameworks designed specifically for the thoughtful integration of GenAI into technical curricula. The mere permission to use these tools is insufficient; without a deliberate pedagogical structure, students are left to navigate the complexities of ethical use, critical evaluation, and balanced reliance on their own.

Existing technology integration models often fall short of addressing the unique challenges posed by GenAI's generative capabilities, its potential to automate high-level cognitive tasks, and its inherent ethical complexities (Dwivedi et al., 2023). The unique ability of GenAI to generate novel content creates unprecedented challenges for assessment and academic integrity that older technologies did not pose (Becker et al., 2023). A clear need has therefore emerged for a formal, transferable model that serves as the actionable, empirical bridge between sweeping global policy mandates and concrete student learning. Such a framework is essential to guide instructors in systematically integrating GenAI in a way that supports, rather than supplants, the development of core competencies (Southworth et al., 2023).

THE PROPOSED SOLUTION (THE IEEC FRAMEWORK)

This article introduces and evaluates the Introduce, Encourage, Evaluate, Control (IEEC) framework, a novel pedagogical model designed to address this critical gap. The IEEC framework provides a structured, four-phase approach to integrating GenAI tools into introductory programming education. Its primary objective is twofold: first, to build foundational programming competence by ensuring students master fundamental skills; and second, to simultaneously cultivate essential AI literacy. AI literacy, in this context, is defined as the capacity to critically evaluate, effectively use, and ethically engage with AI technologies (Long & Magerko, 2020). As illustrated in Figure 1, each stage of the IEEC model – Introduce, Encourage, Evaluate, and Control – serves a distinct and purposeful role in guiding students from initial awareness through to independent, ethical application. By moving through the distinct phases of formally *Introducing* the technology and its ethical dimensions, *Encouraging* scaffolded and critical use, *Evaluating* the process of student engagement with AI, and *Controlling* for over-reliance, the IEEC framework offers a robust and actionable alternative to prohibitive or unstructured approaches. It is designed to empower students to become critical consumers and responsible users of AI, transforming it from a potential crutch into a tool for deeper learning and engagement. The detailed implementation of the IEEC framework will be explained in the Methodology section.

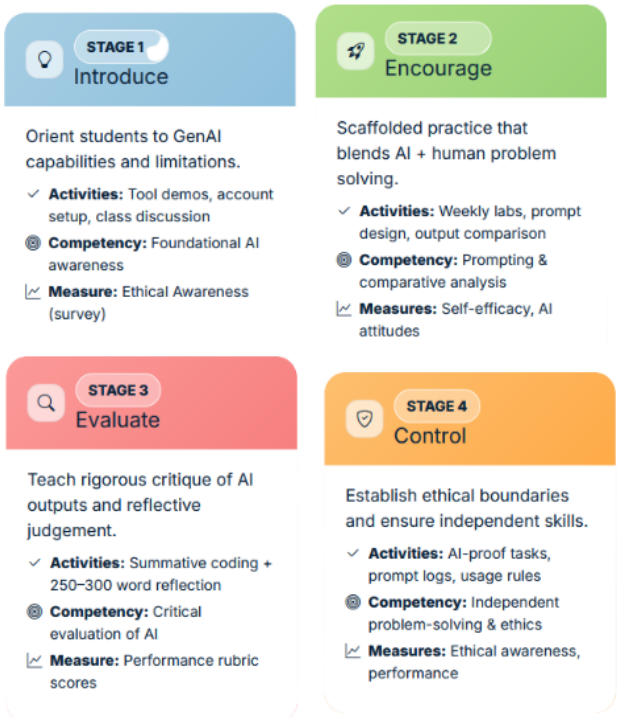

Figure 1. IEEC Framework: A staged approach to integrating GenAI for programming education, aligning activities, competencies and outcome measures.

RESEARCH OBJECTIVES

This study aims to rigorously evaluate the pedagogical effectiveness of the IEEC framework within an authentic higher education context. The research was guided by the following objectives:

1. To implement the IEEC framework in an introductory computer programming course and assess its impact on students' programming self-efficacy (RO1).
2. To measure changes in students' attitudes toward GenAI, including their perceptions of its usefulness and their intention for future use (RO2).
3. To evaluate the framework's effectiveness in enhancing students' critical evaluation skills and their ethical awareness regarding GenAI use (RO3).
4. To qualitatively explore the student learning experience within the IEEC framework, gathering insights into how each component influenced their engagement, confidence, and critical thinking (RO4).

LITERATURE REVIEW

The rapid proliferation of GenAI is transforming higher education in profound and urgent ways, compelling educators to rethink traditional pedagogical practices. As these tools become increasingly embedded in academic and professional activities, scholars have explored their potential benefits alongside significant challenges. The literature reflects a clear duality: while GenAI offers opportunities for enhancing learning and teaching efficiency, it also raises pressing concerns about academic integrity, ethical responsibility, and the preservation of core disciplinary competencies. This section synthesises relevant research in four domains central to this study:

1. The role of GenAI in higher education.
2. Its applications and challenges in computer science and programming education.
3. The imperative of fostering AI literacy and critical evaluation skills.
4. The need for pedagogical frameworks tailored to GenAI integration, culminating in the identification of a gap addressed by the IEEC framework.

GENERATIVE AI IN HIGHER EDUCATION

The integration of Generative AI (GenAI) in higher education has emerged as a multifaceted phenomenon defined by a fundamental tension between pedagogical efficiency and cognitive preservation. On one hand, the literature highlights unprecedented opportunities for enhancing the educational experience through personalised, on-demand support (Kasneci et al., 2023; Dwivedi et al., 2023). By functioning as virtual tutors capable of explaining complex concepts and providing immediate feedback, GenAI tools facilitate a more inclusive and adaptive learning ecosystem (Chan & Hu, 2023; Garcia, 2025). For educators, the automation of routine content creation (e.g., quizzes and lesson plans) theoretically redirects instructional time toward higher-order teaching and mentoring, potentially improving broader learning outcomes (Kasneci et al., 2023; Humble et al., 2023; Dwivedi et al., 2023).

However, this efficiency is sharply counterbalanced by the risk of "cognitive offloading," where the tool's utility as a scaffold transforms into a substitute for genuine intellectual effort (Humble et al., 2023; Li et al., 2025). The ability of GenAI to produce high-quality academic artifacts with minimal human intervention creates an immediate crisis for academic integrity and assessment design, as traditional methods of identifying plagiarism and dishonest practices become increasingly obsolete (Cotton et al., 2024; Shardlow & Latham, 2023). This risk is not merely behavioural but developmental; excessive reliance on AI-generated solutions may weaken the very problem-solving and critical thinking skills that higher education seeks to cultivate (Humble et al., 2023; Garcia, 2025; Li et al., 2025).

Compounding these pedagogical risks is the inherent unreliability and ethical opacity of the technology itself. The literature underscores a persistent "trust gap" caused by factual inaccuracies or "hallucinations," biased algorithmic perspectives, and a lack of transparency in model design (Kasneci et al., 2023; Dwivedi et al., 2023; Chan & Hu, 2023). These technical limitations, alongside urgent concerns regarding data privacy and algorithmic bias, necessitate a high degree of critical evaluation from users – a skill many students have yet to master (Southworth et al., 2023; Chan & Hu, 2023).

In light of these contradictions, the academic consensus has shifted away from prohibitive bans toward a mandate for structured, responsible integration (Humble et al., 2023; Dwivedi et al., 2023). The literature suggests that the challenges of GenAI cannot be resolved through policy alone but require targeted training for both staff and students to ensure that AI serves as a supplement to, rather than a replacement for, the active learning process.

THE ROLE OF AI IN COMPUTER SCIENCE AND PROGRAMMING EDUCATION

In the domain of computer science (CS) and programming, GenAI functions as a critical pedagogical scaffold, lowering the historically high barriers to entry related to syntax and initial conceptual understanding (Becker et al., 2023). Tools such as ChatGPT and GitHub Copilot offer a real-time feedback loop for debugging, code explanation, and snippet generation, which enables novice programmers to maintain the momentum necessary to navigate complex logical problems (Bente et al., 2024; Garcia, 2025). This immediate support translates into measurable outcomes, with empirical evidence suggesting that AI-assisted learners often achieve significantly higher task completion rates and exam performance compared to those working without such tools (Akçapınar & Sidan, 2024; Sivasakthi & Meenakshi, 2025).

Beyond mere performance gains, the automation of routine coding tasks has the potential to reconfigure student cognition. By delegating syntax and boilerplate generation to AI, students can theoretically reserve cognitive capacity for higher-order algorithmic design and creative experimentation (Yilmaz & Karaoglan Yilmaz, 2023). This shift is further reinforced by the psychological impact of the technology; when students perceive GenAI as a non-judgemental "learning companion," the emotional frustration typical of early programming – often a primary cause of attrition – is significantly reduced, thereby fostering greater self-efficacy and engagement (Boguslawski et al., 2025).

However, the literature identifies a significant risk where this "functional success" may mask a lack of conceptual depth. A heavy reliance on AI for solution generation can lead to "cognitive outsourcing," a state where students produce functional code without internalising the logic or principles that underpin the solution (Becker et al., 2023; Humble et al., 2023). This is particularly problematic in programming education, where the iterative struggle of debugging and refinement is considered an essential driver of deep learning (Li

et al., 2025). Without these struggles, students may develop a surface-level understanding that fails when they are faced with more advanced, unassisted challenges.
Furthermore, this "illusory competence" is exacerbated by the novice's inability to critically judge the quality of AI-generated work. Because GenAI can produce code that is functional but insecure, inefficient, or logically flawed, novice programmers often lack the domain expertise to detect these errors, leading to the uncritical acceptance of suboptimal solutions (Bente et al., 2024; Akçapınar & Sidan, 2024). As the distinction between legitimate pedagogical assistance and academic dishonesty becomes increasingly opaque, the risk of plagiarism rises, necessitating a shift in how educators' scaffold and monitor the active learning process (Deriba et al., 2024; Garcia, 2025).

FOSTERING AI LITERACY AND CRITICAL EVALUATION SKILLS

As established in the preceding sections, the "illusory competence" created by uncritical AI assistance necessitates a shift in programming education toward a global mandate of AI literacy (Southworth et al., 2023). This capability is defined not merely by technical usage, but by a composite of critical evaluation, effective application, and ethical engagement (Long & Magerko, 2020). To mitigate the cognitive risks of GenAI, the literature suggests a necessary epistemic shift in the student's role: moving from a passive recipient of code to a "critical consumer" who actively interrogates and validates algorithmic outputs (Bente et al., 2024; Kohen-Vacs et al., 2025). This transition is particularly vital in programming, where the AI's propensity for logical reasoning failures and the fabrication of methods requires a "human-in-the-loop" approach to maintain technical integrity (Bente et al., 2024).

The challenge for educators lies in operationalising this literacy through concrete pedagogical actions. Current scholarship suggests that critical evaluation skills are best cultivated through deliberate exposure to AI fallibility, such as assignments that require students to detect errors and compare contrasting AI solutions (Bente et al., 2024; Garcia, 2025). These evidence-based strategies are directly integrated into the "**E**valuate" stage of the proposed IEEC framework, which draws theoretical support from Long and Magerko's (2020) emphasis on providing explicit instruction regarding both the affordances and limitations of AI tools. By forcing students to justify code modifications and reflect on ethical implications, this stage serves as a pedagogical counterweight to "blind trust," ensuring human logic supersedes automated suggestions (Kohen-Vacs et al., 2025; Garcia, 2025).

The effectiveness of such explicit literacy interventions is increasingly supported by empirical evidence. For instance, Chuan et al. (2025) demonstrate that when AI literacy is treated as a structured pedagogical component, students show measurable improvements in complex tasks like object-oriented design while simultaneously developing more robust reflective habits. These findings suggest that AI literacy should be treated as an integrated technical competency rather than a secondary "soft skill." However, while the literature confirms the importance of evaluation, there remains a lack of structured models that

sequence these skills to prevent early over-reliance – a gap addressed by the staged design of the IEEC framework.

PEDAGOGICAL FRAMEWORK FOR TECHNOLOGY INTEGRATION

The challenge of integrating new technologies into education is not new, and established pedagogical frameworks such as TPACK and SAMR have long been used to guide educational technology adoption. TPACK highlights the importance of aligning technological knowledge with pedagogy and disciplinary content, while SAMR conceptualises technology use in terms of substitution, augmentation, modification, and redefinition (Mishra & Koehler, 2006; Koehler & Mishra, 2009; Hasan & Kabilan, 2024). These models remain influential because they help educators think systematically about how technologies can support teaching and learning.

However, the literature increasingly suggests that GenAI exposes the limitations of these legacy models. Unlike many earlier educational technologies, GenAI does not simply mediate learning activities; it can generate original text, code, explanations, and solutions, thereby reshaping how cognitive work is distributed between students and machines (Kasneci et al., 2023; Miao & Holmes, 2023). This shifts the pedagogical question from "how do we integrate a tool into teaching?" to "how do we structure human-AI collaboration in ways that preserve learning, agency, and integrity?" General technology-integration models do not fully address this challenge because they were not designed for tools that can automate complex cognitive tasks, produce uncertain outputs, and introduce new ethical and authorship dilemmas (Hamilton et al., 2016; Dwivedi et al., 2023).

More recent scholarship has begun to propose frameworks more responsive to these conditions. In programming education, Bente et al. (2024) outline a staged conceptual approach in which AI use is initially restricted to support foundational learning, then made explicit to cultivate critical awareness, and later used more flexibly as students develop the capacity for responsible engagement. This work is important because it recognises that GenAI integration should be developmental and pedagogically sequenced rather than immediate and unrestricted. Yet such emerging models are still relatively limited in number, and many remain conceptual rather than empirically tested in authentic classroom contexts. As a result, there is still a shortage of structured, evidence-based pedagogical frameworks that help educators systematically introduce, scaffold, evaluate, and regulate GenAI use within real courses.

It is within this gap that the IEEC framework is positioned. While it draws on the broader literature on AI literacy and staged AI integration, it advances existing models in several important ways. First, unlike TPACK and SAMR, IEEC is not a generic framework for technology integration; it is specifically designed for GenAI contexts where the tool can generate disciplinary outputs and influence student cognition directly. Second, unlike broad AI-literacy frameworks, IEEC translates these competencies into a concrete course-level pedagogical sequence. The Introduce phase builds foundational awareness of AI

tools, affordances, limitations, and expectations. The Encourage phase creates guided opportunities for students to use AI productively in learning activities. The Evaluate phase develops critical judgement by requiring students to interrogate and compare AI-generated outputs. The Control phase addresses over-reliance, ethical boundaries, and academic integrity through explicit regulation and pedagogical design. Third, unlike many recent conceptual discussions, IEEC is intended as a structured and empirically examinable model for authentic higher education practice.

Taken together, the literature indicates that higher education needs more than policy statements or broad competency frameworks. It needs pedagogical models that can operationalise responsible AI integration in discipline-specific ways. In introductory programming, this means designing learning environments in which AI supports rather than supplants the development of foundational knowledge, critical evaluation, and ethical awareness. The IEEC framework responds to this need by providing a staged and pedagogically actionable model for GenAI integration that builds on existing scholarship while addressing its current limitations.

METHODOLOGY

To capture both the quantitative outcomes and the qualitative nuances of the IEEC framework intervention, a mixed-methods approach was employed. This chapter details the research design, the context and participants, the specifics of the IEEC intervention, and the methods used for data collection and analysis

RESEARCH DESIGN

The study utilised a mixed-methods, quasi-experimental pre-test/post-test design. This approach was selected as the most appropriate for achieving the research objectives for several reasons. The quasi-experimental structure, a common and practical approach in educational settings where random assignment to a true control group is not feasible (Kazemitabaar et al., 2023), allowed for the investigation of the framework's impact within an authentic classroom environment. The pre-test/post-test component enabled the quantitative measurement of changes in key student variables, such as programming self-efficacy and ethical awareness, over the duration of the 14-week intervention.

By incorporating a mixed-methods design, the study aimed to produce a more holistic and robust set of findings. While the quantitative data provides evidence of what changes occurred, the qualitative data, gathered through focus group interviews, offers crucial insights into how and why these changes transpired (Boguslawski et al., 2025). This integration of quantitative and qualitative approaches provides a deeper, more nuanced understanding of the student experience, aligning with contemporary trends in educational research that seek to combine empirical outcomes with rich, contextualised insights (Sun et al., 2024).

PARTICIPANTS AND SETTING

The research was conducted at the Universiti Sultan Zainal Abidin (UniSZA) in Malaysia, within the CSF12003 Problem Solving and Computer Programming course. This is a compulsory introductory programming course (equivalent to CS1) for all undergraduate computer science students. The study population consisted of 60 undergraduate students enrolled in the course. The cohort was comprised of 43% male and 57% female students. This setting provided an ideal context for the study, as the participants were novice programmers for whom the foundational skills of programming and the novel introduction of AI tools were both highly relevant.

The course structure consisted of a one-hour lecture and a two-hour guided lab session each week over a 14-week semester. This regular, structured contact time provided a consistent environment for the phased implementation of the IEEC framework. All students enrolled in the course during the study period were invited to participate. Participation was voluntary, and informed consent was obtained from all students who agreed to contribute their data to the research.

ETHICAL CONSIDERATIONS

This study adhered to strict ethical protocols for pedagogical research. The project was supported by a university Scholarship of Teaching and Learning (SoTL) grant, which required a formal review of the research design. The complete research protocol encompassing participant recruitment, informed consent procedures, and data management plans underwent a rigorous peer review by two senior academics with expertise in educational research and was approved prior to the study's commencement. All participants provided informed consent after being explicitly assured that their participation was voluntary, would not affect their grades or course standing, and that they retained the right to withdraw at any time without consequence.

Participant confidentiality was protected through robust data anonymisation and security measures. All personally identifiable information was removed from research data and replaced with unique participant codes. All digital materials, including anonymised survey data, focus group transcripts, and copies of student prompt logs, were stored in a secure, institutionally managed Google Workspace (enterprise) environment. This system employs advanced security protocols, including end-to-end encryption and restricted access controls monitored by audit trails. To mitigate privacy risks associated with third-party platforms, prompt logs were immediately anonymised upon export and stored exclusively on the secure university server. Access to all identifiable and anonymised data was strictly limited to the authors only.

THE IEEC FRAMEWORK AND INTERVENTION

The core of the study was the implementation of the IEEC framework as the primary pedagogical intervention. The framework is designed to systematically and ethically integrate GenAI tools into the curriculum, aligning with the four key aspects of AI literacy: knowing and understanding AI, using and applying AI, evaluating and creating with AI, and understanding its ethical implications (Long & Magerko, 2020). Each component of the framework was introduced sequentially and reinforced throughout the 14-week semester.

For clarity, each phase of the intervention is described below in terms of its pedagogical purpose, the learning activities through which it was implemented, and the intended student outcomes. This parallel structure reflects the staged design of the framework and clarifies how each phase contributed to the development of programming competence, critical evaluation, and ethical awareness.

Introduce (I)

The initial phase of the framework focused on formally introducing students to the concept of AI-assisted programming. During the first class, a dedicated lecture segment was used to expose students to the available GenAI tools, including ChatGPT, Google Gemini, and GitHub Copilot. This introduction was not merely a technical demonstration; it was a balanced presentation of both the benefits (e.g., accelerated debugging, code explanation) and the significant limitations (e.g., potential for inaccuracy, inherent biases) of using AI in programming tasks. This aligns with recommendations from the literature that students must first understand the strengths and weaknesses of AI tools before they can use them effectively (Bente et al., 2024). To ensure all students had access, they were required to create a ChatGPT account if they did not already have one.

Encourage (E)

Following the introduction, the second phase was designed to motivate and guide students in using AI tools for learning. This was primarily achieved through weekly guided lab exercises that integrated GenAI use in a structured manner. For example, in an exercise on flowcharts and pseudocode, students were not only asked to solve the problem themselves but were also tasked with writing a prompt to have ChatGPT produce a solution. They were then required to critically compare the AI-generated response with their own, identifying strengths and weaknesses in both. This activity was designed to develop their prompt engineering skills and critical analysis of AI outputs. Another exercise, focused on Java string manipulation, intentionally used a task where ChatGPT is known to produce inconsistent and incorrect results. This was a deliberate pedagogical choice to directly illustrate the tool's limitations and reinforce the necessity for students to verify all AI-generated outputs, a critical component of AI literacy. These carefully scaffolded tasks

encouraged students to view AI not as an infallible oracle but as a powerful, yet imperfect, collaborator.

Evaluate (E)

The third component of the framework focused on teaching and assessing students' ability to critically evaluate AI-generated content. The use of AI tools was formally incorporated into summative assessments, shifting the focus from simply producing a correct answer to demonstrating a critical and reflective use of the tool. In a key lab report, where students developed a Fibonacci-based GUI application, the assessment brief explicitly required them to use an LLM. Students had to submit a 250–300 words reflective paragraph detailing how the LLM assisted them, what guidance they sought, how they adapted the AI's suggestions, and how this led to an improved code submission. Student reflections were evaluated using a clear rubric with criteria including Clarity, Conciseness, Relevance, Comprehensiveness, and the use of specific Examples. This practice aligns with the growing consensus that assessment in the age of AI must evolve to measure not just the final product, but also the process, including the student's ability to critically engage with AI tools (Shardlow & Latham, 2023).

Control (C)

The final component of the framework was to manage and regulate the use of AI to prevent over-reliance and ensure academic integrity. The framework employed a dual strategy of establishing clear guidelines and designing AI-proof assessment components. To ensure students developed foundational GUI-building skills without AI assistance, one part of the assessment required them to submit a screencast demonstrating their use of the "Drag and Drop" feature in the NetBeans IDE. This task is not currently performable by even the most advanced, publicly available generative AI tools, thus serving as an AI-proof component. For the parts of the assessment where AI use was encouraged, strict guidelines were provided. A prominent notice in the assessment brief instructed students: "Do Not Copy-Paste Directly" and emphasized that AI-generated code should serve as guidance that "must be understood, interpreted, and adapted." Crucially, to enforce this and provide an evidence trail, students were required to maintain a history of all prompts used in their interaction with AI tools, which could be requested by the marker at any time. This measure aligns with recommendations for fostering ethical AI use and providing a mechanism for verifying the authenticity of student work (Shardlow & Latham, 2023).

DATA COLLECTION AND INSTRUMENTS

A mixed-methods data collection strategy was employed to gather a comprehensive dataset before and after the 14-week intervention.

Quantitative Instruments

Three quantitative instruments were administered in a pre-test/post-test format.

Programming self-efficacy scale

To measure students' confidence in their programming abilities, a scale adapted from the well-established Computer Programming Self-Efficacy Scale by Ramalingam and Wiedenbeck (1998) was used. This instrument is widely cited and validated in the literature for its effectiveness in capturing novice programmers' self-perceptions (Yilmaz & Karaoglan Yilmaz, 2023). The scale consists of items rated on a Likert scale, asking students to assess their confidence in performing specific programming tasks, from simple syntax-based operations to more complex problem-solving.

AI attitude and perceptions survey

A survey was developed to measure students' attitudes toward GenAI, drawing heavily on the constructs of the Technology Acceptance Model (TAM) (Venkatesh & Davis, 2000). As demonstrated in studies on student acceptance of new technologies (Sun et al., 2024), TAM provides a robust theoretical foundation. The survey included subscales measuring:

1. Perceived Usefulness: Students' beliefs about how GenAI can enhance their learning efficiency and programming performance.
2. Perceived Ease of Use: Students' perceptions of the clarity and simplicity of using GenAI tools.
3. Intention to Use: Students' willingness to continue using GenAI in their future academic and professional work.
4. Ethical Awareness: A custom subscale was developed to assess students' ability to recognise ethical issues such as over-reliance on technology, the potential for misinformation, and concerns related to academic integrity.

The initial items underwent content validation by two senior academics with expertise in educational technology and programming pedagogy to ensure their validity and clarity. The experts independently rated the relevance and clarity of each item on a 4-point scale and provided qualitative feedback. Items with low ratings were revised or removed based on their feedback, and the process was repeated until a high degree of content validity was established.

Performance assessment on evaluating GenAI outputs

As part of their final summative assessment, students were required to complete the task described in the "Evaluate" phase, which involved correcting flawed AI-generated code and providing a written reflection on the process. This task was designed to directly

measure their ability to critically assess AI outputs. The written reflection was scored by the course instructors using a detailed rubric with criteria for clarity, conciseness, relevance, and comprehensiveness, providing a quantitative score of their critical evaluation skills.

Qualitative Instruments

Focus Group Interviews: Following the post-test, semi-structured focus group interviews were conducted. A purposive sample of eight students who had completed the course was selected to provide rich, in-depth data. The choice of focus groups over individual interviews was deliberate, as this method is particularly effective for exploring shared experiences and allowing participants to build on each other's comments, leading to a more dynamic and comprehensive dataset (Boguslawski et al., 2025). The semi-structured interview protocol included open-ended questions designed to elicit detailed narratives about their experiences with each component of the IEEC framework. Sample questions included:

1. "Can you describe your initial reaction to the *Introduction* of GenAI tools in this course?"
2. "In what ways did the weekly lab exercises (*Encourage*) change how you approached problem-solving?"
3. "Tell us about the process of writing the reflective analysis for your lab report (*Evaluate*). What did you learn from that?"
4. "What were your thoughts on the guidelines for using AI, such as the requirement to log prompts (*Control*)?"

DATA ANALYSIS

The quantitative and qualitative data were analysed separately before being integrated to provide a comprehensive interpretation of the findings.

Quantitative Instruments

Prior to conducting the paired-samples t-tests, the data was screened to ensure that the necessary assumptions were met. The differences between pre- and post-test scores were checked for outliers, and the normality of these differences was assessed through visual inspection of Q-Q plots. The results indicated no violations of assumptions, supporting the use of the paired-samples *t*-test.

Paired-samples *t*-tests were then used to compare mean scores from the pre- and post-test surveys, an appropriate approach for a single-group pre/post design. Statistical significance was evaluated at the $p < 0.05$ level, and Cohen's d was calculated to provide a standardised measure of effect size, offering insight into the magnitude of observed changes beyond *p*-values alone (Akçapınar & Sidan, 2024). Performance assessment scores were summarised

using descriptive statistics (means and standard deviations) to characterise students' ability to critically evaluate AI outputs.

Qualitative Instruments

The audio recordings from the focus group interviews were transcribed verbatim. The transcripts were then analysed using the six-phase thematic analysis process described by Braun and Clarke (2006), a method widely recognised for its rigour and flexibility in qualitative research (Boguslawski et al., 2025). This systematic process involved:

1. Initial familiarisation with the data by reading and re-reading the transcripts.
2. Generating initial codes from the raw text.
3. Searching for potential themes by organising the codes into broader patterns.
4. Reviewing and refining these themes against the dataset.
5. Defining and naming the final themes.
6. Producing the final analysis.

This structured approach ensured a comprehensive and data-driven interpretation of the student experience.

To ensure the reliability and trustworthiness of the analysis, a formal intercoder agreement process was implemented. Two of the authors independently coded a representative 25% sample of the verbatim transcripts using an initial codebook developed from the research objectives. Intercoder reliability was then calculated, yielding a strong initial agreement indicated by a Cohen's Kappa coefficient of $\kappa = 0.84$, which is considered 'almost perfect' agreement. All discrepancies in coding were discussed and resolved. This process was not merely for consensus-building; it served to refine the codebook itself by clarifying ambiguous code definitions and merging conceptually overlapping codes. Once this revised and more robust codebook was finalised, it was applied consistently by the lead researcher to the entire dataset. This systematic, collaborative process significantly enhanced the interpretive validity of the resulting themes.

RESULTS

The implementation of the IEEC framework was evaluated using a mixed-methods approach, integrating quantitative survey and assessment data with qualitative focus group insights. This section presents the results in alignment with the four research objectives, offering both statistical evidence of change and thematic interpretations from student narratives. Table 1 summarises the pre- and post-test survey results for self-efficacy and AI-related perceptions, demonstrating statistically significant improvements across all measured constructs following the implementation of the IEEC framework. Figure 2

visually illustrates these changes, highlighting consistent gains across programming self-efficacy, attitudes toward GenAI, and ethical awareness.

Table 1. Pre-test vs. post-test survey results on self-efficacy and AI perceptions

Construct/Subscale	Pre-test mean (SD)	Post-test mean (SD)	*t*-value	*p*-value	Cohen's *d*
Programming self-efficacy	3.12 (0.85)	3.85 (0.76)	5.42	< .001	0.88
Simple tasks	3.55 (0.91)	4.15 (0.68)	4.98	< .001	0.75
Complex tasks	2.69 (0.88)	3.55 (0.81)	5.81	< .001	0.99
AI attitude & perceptions					
Perceived usefulness	3.45 (1.02)	4.21 (0.79)	4.88	< .001	0.83
Intention to use	3.61 (1.10)	4.35 (0.75)	4.65	< .001	0.79
Ethical awareness	3.05 (0.95)	3.98 (0.82)	6.03	< .001	1.01

Note: * $N \approx 60$. Scores are on a 5-point Likert scale. *p*-values < .05 are statistically significant.

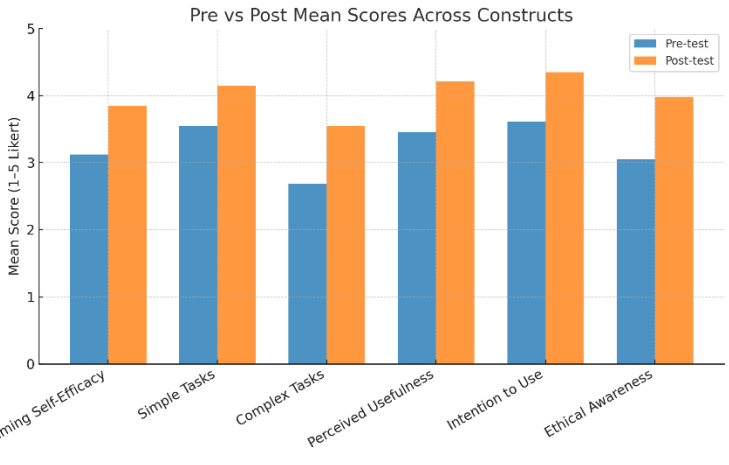

Figure 2. Comparison of pre- and post-test mean scores (± SD) across constructs measured on a 5-point Likert scale

OBJECTIVE 1: IMPACT ON PROGRAMMING SELF-EFFICACY

Quantitative Findings

Pre- and post-intervention surveys indicated a statistically significant improvement in students' programming self-efficacy. Overall self-efficacy scores increased from $M = 3.12$, SD = 0.85 to $M = 3.85$, SD = 0.76, $t(59) = 5.42$, $p < .001$, $d = 0.88$, reflecting a large effect size. Improvements were consistent across both simple programming tasks ($p < .001$, $d = 0.75$) and complex programming tasks ($p < .001$, $d = 0.99$). These results indicate a

positive association between the framework and skill growth across the full spectrum of task difficulty.

Qualitative Insights

The *Encourage* phase, characterised by scaffolded, AI-assisted lab activities, was repeatedly cited as central to the increase in confidence. Students described GenAI as a non-judgemental, persistent "partner" that reduced frustration and encouraged experimentation with more complex challenges:

> Before, getting stuck on a bug was so frustrating… With ChatGPT, it felt like I had a partner… It made me want to try more complex things and not give up so easily.

This narrative suggests that the AI functions as an affective buffer. By absorbing the perceived judgement of a syntax error or a logical failure, the AI prevents the novice learner from entering a demoralising avoidance loop. However, this shift reveals a potential tension of agency: while the student feels more capable, their confidence is initially tethered to the presence of the tool.

The Encourage phase therefore acts as a 'psychological bridge' – a temporary support that lowers the emotional cost of failure – but one that relies on the subsequent Control phase to ensure that students do not mistake facilitated success for unassisted mastery. Because this persistence is rooted in such a structured, low-stakes environment, the framework successfully reduced the emotional barriers often associated with debugging. This aligns with broader evidence that AI-supported learning environments can strengthen intrinsic motivation by promoting a sense of control over the learning process, provided that the support is eventually faded to allow for independent competence.

OBJECTIVE 2: CHANGES IN ATTITUDES TOWARD GENAI

Quantitative Findings

Survey data revealed significant positive shifts in students' perceptions and intentions regarding GenAI:

1. "Perceived usefulness" increased from $M = 3.45$, SD = 1.02 to $M = 4.21$, SD = 0.79, $p < .001$, $d = 0.83$.
2. "Intention to use" rose from $M = 3.61$, SD = 1.10 to $M = 4.35$, SD = 0.75, $p < .001$, $d = 0.79$.

These effect sizes suggest a substantial enhancement in openness to continued and responsible engagement with GenAI tools.

Qualitative Insights

Students attributed these positive shifts to two primary design features:

1. The *Introduce (I)* phase's clear, balanced overview of GenAI's strengths and limitations, which dispelled misconceptions and built trust.
2. The *Encourage (E)* phase's opportunities for low-risk experimentation, allowing students to explore functionality without fear of academic penalty.

As one student observed, "It didn't just give me the answer; it helped me understand the error." Such experiences reinforced that GenAI could serve as a cognitive partner rather than a shortcut, supporting learning through guided interaction rather than passive consumption of answers.

OBJECTIVE 3: ENHANCEMENT OF CRITICAL EVALUATION SKILLS AND ETHICAL AWARENESS

Quantitative Findings

The largest effect size observed across all constructs was for Ethical Awareness, which increased from $M = 3.05$, SD = 0.95 to $M = 3.98$, SD = 0.82, $p < .001$, $d = 1.01$*. This marked gain suggests the potential effectiveness of the framework in cultivating an awareness of the ethical complexities surrounding AI-assisted programming.

The summative "Evaluate" phase task – debugging flawed AI-generated Java code and producing a reflective commentary – yielded a mean score of 81.5% (SD = 9.2%), indicating strong attained competence in both technical and evaluative dimensions. Although this measure lacked a pre-test comparison, performance levels suggest that post-intervention students possessed the intended analytical capabilities.

Qualitative Insights

Students described a clear shift from "blind trust" to critical partnership with AI tools. Early tendencies to accept AI outputs uncritically were disrupted by deliberate pedagogical interventions, notably exercises in which AI-generated code contained embedded flaws.

One student reflected, "At first, I assumed the code from ChatGPT was perfect… That Java string exercise where it was just wrong was a real wake-up call. Now, I double-check everything."

The requirement to submit reflective analyses reinforced this critical stance, compelling students to justify code modifications and articulate decision-making processes. This process not only strengthened technical accuracy but also nurtured a habit of methodical

scepticism – a transferable skill relevant to both AI-mediated and traditional coding practices.

OBJECTIVE 4: EXPLORATION OF THE STUDENT LEARNING EXPERIENCE

Thematic analysis of the focus group interviews revealed three overarching themes that closely aligned with the phases of the IEEC framework.

Increased Confidence and Engagement (The Encourage Phase)

Reflecting the goals of the Encourage phase, students described how the AI-assisted lab activities transformed moments of frustration into opportunities for problem-solving, shifting their approach from avoidance to persistence. The non-judgmental nature of GenAI played a crucial role in lowering the barrier to attempting challenging problems. Without fear of criticism, students felt more at ease experimenting with code and seeking solutions iteratively. This supportive environment encouraged them to persist in the face of errors and emboldened them to attempt more ambitious projects. The combination of immediate feedback and the safe, structured learning space cultivated a sense of agency, enabling learners to see errors not as failures but as stepping stones toward mastery.

Developing a Critical Eye for AI Outputs (The Evaluate Phase)

The Evaluate phase triggered a fundamental realignment of epistemic authority in the classroom. By forcing students to justify AI-driven modifications, the framework moved the source of truth from the AI (the oracle) back to the student (the auditor). This transition is critical because it highlights a threshold of effort: students only began to think critically once they encountered disruptive fallibility (the flawed string exercise). This suggests that AI literacy is not a passive skill but a performative disposition that must be triggered by deliberate pedagogical friction.

Navigating Ethical Boundaries (The Introduce and Control Phases)

The Introduce and Control phases together established a strong ethical foundation for AI use. The initial ethics lecture and explicit guidelines clarified academic integrity expectations from the outset, providing students with a clear framework for responsible practice. Control mechanisms, such as prompt logging, were initially perceived as restrictive; however, as the course progressed, many students recognised these tools as valuable for fostering intentional and accountable AI use. These structures prompted learners to think critically about the prompts they crafted and the purposes behind their requests. As one student reflected, "The first lecture on ethics was really important. We knew from day one what the rules were. The prompt log was a bit annoying at first, but it made me think twice before just copy-pasting something. It made me more intentional about what I was asking and why."

Despite this strengthened ethical awareness, students acknowledged ongoing difficulty in consistently applying these principles, particularly in light of the variability in AI output quality. This suggests that while the framework successfully promoted ethical mindfulness, sustained reinforcement and real-world contextualisation are necessary for these practices to become fully internalised.

SUMMARY OF KEY RESULTS

The findings suggest that the IEEC framework was effective in achieving its intended outcomes through a phase-linked combination of pedagogical strategies:

1. "Introduce" phase activities appeared to effectively establish conceptual and ethical foundations.
2. "Encourage" phase tasks leveraged AI's strengths to build competence and confidence.
3. "Evaluate" phase challenges were associated with the development of critical engagement with AI outputs.
4. "Control" phase mechanisms seemed to provide structural reinforcement for responsible, mindful use.

Statistical evidence confirmed significant improvements in self-efficacy, AI attitudes, and ethical awareness, while thematic analysis illuminated the mechanisms driving these changes. Together, these findings suggest that a structured, deliberate integration of GenAI into programming education can simultaneously enhance technical competence, foster evaluative judgement, and promote ethical mindfulness.

DISCUSSION

The rapid proliferation of GenAI in higher education presents both opportunities and challenges, underscoring the need for structured, evidence-based pedagogical frameworks. This study evaluated the IEEC framework in an introductory programming course, with the aim of supporting technical skill development while cultivating metacognitive and ethical competencies – key components of AI literacy. The results indicate that IEEC can effectively nurture student confidence, critical engagement with AI, and ethical awareness. This section synthesises these findings, considers their pedagogical and theoretical implications, and outlines limitations and future research directions.

ACHIEVEMENT OF RESEARCH OBJECTIVES

The study successfully addressed all four research objectives. First, the integration of IEEC led to statistically significant gains in programming self-efficacy (d = 0.88). While maturation over a 14-week semester naturally builds confidence, the disproportionate surge

in efficacy for complex tasks (d = 0.99) compared to simple tasks (d = 0.75) suggests that the framework provided a specific 'affective buffer' against the plateaus typical of novice instruction. Second, students' attitudes toward GenAI shifted positively, with significant improvements in perceived usefulness and intention for future use. Third, the framework enhanced critical evaluation and ethical awareness; the latter showed the largest measured effect size (d = 1.01), while students demonstrated strong debugging performance and reflective capacity. Fourth, thematic analysis confirmed that each IEEC phase contributed distinct learning benefits, ranging from increased persistence during the Encourage phase to deeper critical engagement in the Evaluate phase. Collectively, these findings support the view of IEEC as an adaptable model for integrating GenAI into technical curricula, though its applicability across diverse cultural and disciplinary settings requires further replication.

Having established that the framework addressed the study objectives, the discussion now turns from reporting what was achieved to interpreting what these outcomes mean pedagogically and theoretically. The following section therefore focuses on how the quantitative and qualitative findings together explain the developmental value of the IEEC framework.

INTERPRETATION OF KEY FINDINGS

The mixed-methods approach enriched the interpretation of results, enabling quantitative shifts to be contextualised through qualitative narratives. Three overarching themes emerged: the development of student confidence through scaffolded support, the adoption of a critical evaluative stance toward AI, and the emergence of sustained ethical mindfulness.

Taken together, these themes suggest that students' relationship with GenAI shifted over the course of the intervention: from seeing AI primarily as a convenient source of help to engaging with it more reflectively as a learning support, an object of critique, and a tool requiring ethical judgement. This deeper progression is important because it shows that the value of GenAI in programming education lies not simply in access to automated assistance, but in how that assistance is pedagogically structured.

Building Confidence Through Scaffolded AI Support

The significant increase in programming self-efficacy (p < .001, d = 0.88) was illuminated by student reflections on the Encourage phase. GenAI was perceived not simply as a utility, but as a supportive, non-judgemental partner that reframed frustration into learning opportunities. As one participant remarked, "it helped me understand the error. This made me want to try more complex things and not give up so easily." This aligns with evidence that AI can reduce the stress associated with learning to code by providing immediate, personalised feedback (Boguslawski et al., 2025). The IEEC's structured lab sessions created a psychologically safe environment for experimentation, lowering affective barriers and promoting intrinsic motivation (Yilmaz & Karaoglan Yilmaz, 2023). This design

fostered active engagement and a sense of agency, key drivers of skill acquisition in novice programmers.

More specifically, the theme suggests that increased confidence was not simply a result of students completing tasks more easily. Rather, the Encourage phase appears to have functioned as a form of pedagogical scaffolding in which AI support reduced the emotional burden of debugging while still keeping students actively involved in problem-solving. In this respect, the findings resonate with literature describing AI as a supportive learning companion when embedded within structured educational conditions rather than used as an unrestricted substitute for effort (Boguslawski et al., 2025; Garcia, 2025). The qualitative accounts therefore indicate that confidence developed through supported participation in programming tasks, not through the removal of difficulty altogether.

Nevertheless, it is critical to distinguish between "Independent Self-Efficacy" and "AI-Mediated Confidence." There is a risk that the recorded $d = 0.88$ gain represents a "phantom competence," where students feel capable only when the tool is available. Yet, the qualitative focus on "understanding the error" rather than just "getting the answer" suggests that the Encourage phase successfully converted AI output into a cognitive scaffold, though future research must test if this confidence survives when the tool is removed.

Cultivating Critical Evaluation Skills

A central aim of IEEC was to move beyond uncritical use of AI toward an informed, evaluative relationship. The strong average performance (81.5%) on the debugging-and-reflection summative task suggests this aim was met. While such a score could be viewed as a byproduct of general coding familiarity gained over the semester, we argue it represents a specialised "adversarial literacy" developed specifically through the framework. Traditional programming assessments primarily reward functional output; however, the Evaluate phase required students to overcome automation bias – the psychological tendency to trust machine-generated solutions. The high attainment suggests that deliberate exposure to "flawed" AI outputs forced a cognitive shift that natural maturation alone likely does not produce. Consequently, students transitioned from 'blind trust' to a "critical partnership," a shift catalysed by the realisation that AI outputs are provisional and require human auditing. This strategy effectively countered the pervasive tendency for students to accept AI suggestions without scrutiny (Akçapınar & Sidan, 2024).

The Evaluate phase – especially the requirement for written reflection on AI-assisted code – compelled students to articulate their reasoning and justify modifications, embedding metacognitive processes into their workflow. This reinforces calls for AI-era assessment models that measure critical engagement and process, not just final products (Shardlow & Latham, 2023).

From the perspective of digital pedagogy, this finding is important because it shows that GenAI was positioned not as an invisible shortcut, but as an object of critique within the learning process. The reflective requirement made students' judgement visible and

encouraged them to monitor, question, and regulate their own interactions with AI-generated outputs. This theme can be interpreted as an epistemic shift in how students positioned themselves in relation to AI-generated code. Rather than treating AI output as authoritative, students increasingly approached it as provisional material that needed to be checked, tested, and refined. This interpretation connects strongly with Long and Magerko's (2020) conception of AI literacy as involving critical evaluation, and with programming education literature arguing that students must learn to question, validate, and improve AI-generated solutions rather than merely consume them (Bente et al., 2024; Kohen-Vacs et al., 2025). Importantly, students' accounts also suggest that this evaluative stance extended beyond AI-generated code to their own debugging practices, indicating that the framework may have supported the development of transferable habits of methodical scepticism.

Advancing Ethical Awareness

The most substantial quantitative gain was in Ethical Awareness ($p < .001$, $d = 1.01$), indicating the effectiveness of the Introduce and Control phases in raising students' consciousness of AI's ethical dimensions. Students described the initial ethics lecture as "eye-opening" and viewed the Control phase – particularly the prompt log – as a mechanism for fostering intentional, reflective AI use. These qualitative findings suggest that students were not only able to identify ethical issues in principle, but were beginning to reflect more carefully on how, when, and why they used GenAI in their own programming practice. This is important because it shows that ethical awareness in the IEEC framework was developing through concrete learning activities, rather than through abstract discussion alone.

However, the qualitative data revealed a persistent gap between ethical understanding and application. Students acknowledged challenges in consistently enacting ethical practices, citing the temptation to copy-paste and difficulties navigating ambiguous situations. This reflects the broader educational challenge of bridging the knowing–doing gap, particularly in contexts involving complex, dynamic technologies. This finding also connects with existing literature which notes that students may recognise concerns such as plagiarism, over-reliance, and inappropriate delegation of thinking, yet still struggle to apply ethical principles consistently in authentic learning contexts (Cotton et al., 2024; Dwivedi et al., 2023; Southworth et al., 2023). The findings reaffirm the need for sustained human oversight – the "human-in-the-loop" principle (Dwivedi et al., 2023) – and ongoing reinforcement of ethical habits.

This also suggests that AI literacy should be understood not only as conceptual awareness, but as a situated socio-ethical practice. In other words, students may recognise ethical principles in theory while still needing structured opportunities to enact them consistently in authentic learning situations. The prompt log, explicit rules, and controlled assessment design therefore functioned not merely as compliance mechanisms, but as pedagogical support for developing responsible habits of human-AI collaboration. The qualitative

evidence thus strengthens the interpretation that the Introduce and Control phases did more than impose rules: they helped make ethical decision-making visible and habitual within students' day-to-day engagement with AI-assisted programming. This interpretation also supports literature emphasising process, transparency, and accountability as central to responsible AI-integrated assessment and learning design (Shardlow & Latham, 2023).

Finally, we must acknowledge the pedagogical overhead of the IEEC framework. Its success is intrinsically linked to the instructor's role in curating flawed exercises and auditing logs. This indicates that IEEC is not a 'plug-and-play' technological fix, but a structured teaching philosophy. Its efficacy likely depends on an instructor's willingness to transition from a provider of knowledge to a facilitator of critical human-AI collaboration.

IMPLICATIONS OF THE IEEC FRAMEWORK

The findings have significant implications for both educational practice and theory. Building on the preceding interpretation of the results, this subsection shifts the discussion from what the findings reveal to what they imply for future pedagogical design and theoretical understanding. The implications are considered in two related ways: first, in terms of practical guidance for programming educators, and second, in terms of the framework's contribution to wider debates on AI-integrated learning.

Practical Pedagogical Implications

For programming educators seeking to integrate GenAI responsibly, this study presents IEEC as a structured, empirically-supported framework that offers an alternative to both prohibition and laissez-faire adoption:

1. Adopt a Phased Approach – Introducing GenAI should begin with formal orientation to its capabilities, limitations, and ethical implications. This sets a foundation for responsible use and prevents early over-reliance.
2. Integrate GenAI into Learning Activities – AI use should be embedded within intentional tasks, as seen in the Encourage and Evaluate phases, rather than permitted ad hoc. Purposeful integration fosters learning and critical thinking rather than superficial task completion.
3. Assess Process as Well as Product – Incorporating reflections, AI use justifications, and prompt logs shifts the assessment focus toward reasoning, decision-making, and ethical engagement—skills central to AI literacy (Shardlow & Latham, 2023).
4. Implement Control Measures as Scaffolds – Rather than restricting innovation, clear guidelines and AI-proof assessment elements function as guardrails, helping students develop responsible habits.

Theoretical Contributions

The IEEC framework advances pedagogical theory by offering a more explicit account of how generative AI can be integrated into teaching in ways that support learning, reflection, and responsibility simultaneously. While broad models such as SAMR and TPACK remain useful for understanding technology integration in general, they do not sufficiently account for the distinctive characteristics of GenAI, particularly its capacity to generate disciplinary outputs, shape student reasoning, and introduce new ethical and authorship dilemmas (Mishra & Koehler, 2006; Koehler & Mishra, 2009; Blundell et al., 2022; Dwivedi et al., 2023). IEEC contributes greater specificity by framing AI integration not as simple tool adoption, but as a staged pedagogical process that must be intentionally designed and regulated.

First, the findings contribute to learning theory by showing that GenAI can operate productively as a scaffold for novice learners when embedded within structured pedagogical conditions. The significant gains in programming self-efficacy, supported by student accounts of reduced frustration and greater persistence, indicate that AI-assisted learning was most effective when students remained actively engaged in problem-solving rather than passively accepting AI-generated solutions. This suggests that the educational value of GenAI lies not in automation itself, but in its capacity to extend learner performance within carefully designed boundaries that preserve agency and participation. In this sense, IEEC offers an empirically grounded model of scaffolded human-AI learning in programming education.

Second, the framework contributes to digital pedagogy by demonstrating that productive AI integration requires the design of tasks that make student reasoning visible. The Evaluate phase did not reward students simply for obtaining functional code; instead, it required them to interrogate flawed AI outputs, justify revisions, and reflect on the role AI played in their decision-making. The observed shift from "blind trust" to "critical partnership" indicates that the framework fostered metacognitive regulation and reflective judgement. This is theoretically important because it positions AI not merely as a support technology, but as a pedagogical object through which learners can develop habits of critique, verification, and reflective engagement.

Third, the study extends AI literacy theory by showing that AI literacy is best understood as an integrated capability involving technical confidence, evaluative judgement, and ethical self-regulation. The framework operationalised these dimensions through its sequential phases: Introduce established awareness of affordances, limitations, and ethical implications; Encourage enabled guided use; Evaluate developed critique; and Control reinforced responsible boundaries and accountability. Importantly, the findings also show that ethical awareness alone is insufficient. Students' continued difficulty in applying ethical principles consistently suggests that AI literacy must be treated not as a static body of knowledge, but as a practice developed through repeated reflection, contextual judgement, and sustained human oversight.

Taken together, these findings suggest that IEEC advances existing pedagogical models in three ways. It provides a GenAI-specific framework where general technology models remain too broad; it translates broad AI-literacy principles into concrete course-level pedagogical actions; and it offers empirical evidence from an authentic programming context showing how structured human-AI collaboration can foster confidence, critical evaluation, and ethical mindfulness at the same time. The contribution of IEEC therefore lies not only in proposing a staged model, but in demonstrating how such a model can connect theories of learning, digital pedagogy, and AI literacy in actual higher education practice.

LIMITATIONS

The findings of this study should be considered in light of two primary limitations: the absence of a control group and the single-institution context. The quasi-experimental design prevents the definitive attribution of observed effects solely to the IEEC intervention, as natural skill development and increased familiarity with course content over the semester may have also contributed to the observed improvements in student outcomes. Accordingly, the quantitative findings should be interpreted as showing an association with the IEEC intervention rather than definitive causal evidence of its independent effect. Future studies employing randomised controlled designs are needed to establish causal claims. Furthermore, the study was conducted in a single Malaysian institution. This context-specific setting means the results may not be directly generalisable to other cultural or educational environments, highlighting the need for cross-cultural replication studies to validate the framework's broader applicability.

FUTURE RESEARCH DIRECTIONS

Building on these findings, future research could:

1. Test IEEC in Controlled Designs – Employ experimental methods with control groups to isolate the framework's effects from other variables.
2. Expand Across Contexts – Replicate the study in varied cultural, disciplinary (e.g., humanities, engineering), and institutional settings to test the framework's generalisability and explore contextual influences on its efficacy.
3. Extend Longitudinally – Track students over multiple semesters or into professional contexts to conduct longitudinal analysis. This would be crucial for gauging the retention of technical and critical evaluation skills and, most importantly, for investigating whether increased ethical awareness translates into sustained ethical practice when students are no longer in the structured course environment.
4. Adapt to Other Disciplines – Modify and evaluate IEEC for fields with distinct AI-related challenges, such as humanities scholarship, journalism, or engineering design.

5. Investigate Ethical Application – Examine interventions that close the knowing–doing gap, potentially integrating scenario-based training or real-world ethical dilemmas.

In summary, this study demonstrates that the IEEC framework can effectively integrate GenAI into programming education, fostering technical proficiency, critical evaluation skills, and ethical awareness. By offering both structure and flexibility, IEEC addresses the dual imperatives of empowering students to leverage AI constructively while guiding them toward responsible and reflective use. The model provides a timely and practical contribution to the evolving landscape of AI-enabled education; however, its relevance across other disciplines and institutions should be established through further testing in varied educational settings.

CONCLUSION

The rapid and pervasive integration of GenAI into the academic landscape has created a significant pedagogical challenge for higher education, particularly in skill-intensive disciplines like computer programming. Educators are confronted with the dual task of leveraging powerful AI tools like ChatGPT and Google Gemini to enhance learning while simultaneously mitigating the inherent risks of over-reliance, cognitive offloading, and ethical misuse. This study was motivated by the conspicuous absence of structured, evidence-based frameworks to navigate this complex new terrain. Consequently, the primary purpose of this research was to introduce, implement, and rigorously evaluate the IEEC pedagogical framework, a model designed to thoughtfully integrate GenAI into an introductory programming course to build foundational technical competence alongside essential AI literacy.

Our mixed-methods evaluation indicates that the IEEC framework can provide a potentially effective pathway for achieving this dual objective. The quantitative findings revealed statistically significant improvements in students' programming self-efficacy, particularly in tackling complex tasks, and a marked increase in their ethical awareness regarding GenAI use – the latter showing the largest effect size of all measured constructs. Qualitatively, our findings illuminated the mechanisms behind these results. The structured *Encourage* phase transformed student frustration into engaged problem-solving by positioning GenAI as a supportive "partner," thereby boosting confidence. The *Evaluate* phase successfully cultivated critical thinking by compelling students to move beyond uncritical acceptance of AI outputs to a more mature "critical partnership," a shift catalysed by exercises involving deliberately flawed AI-generated code. Furthermore, the foundational *Introduce* and *Control* phases established clear expectations for responsible and ethical engagement from the outset.

The key contribution of this study is the empirical evaluation of the IEEC framework, presenting it as a practical and contextually validated model for the responsible integration of GenAI in an introductory programming course. It offers a clear alternative to the

ineffective extremes of outright prohibition and unstructured, laissez-faire adoption. By operationalising the principles of AI literacy into a clear sequence of pedagogical actions, the framework guides educators in designing learning experiences that foster not just task completion, but the crucial metacognitive and ethical skills required to use AI effectively and wisely. While our findings confirm the framework's success in enhancing students' critical awareness, they also highlight the persistent gap between knowing and doing, reinforcing the need for continuous, scaffolded practice in applying ethical principles.

Looking forward, the collaboration between human intellect and artificial intelligence is poised to become a defining feature of the modern professional and academic world. The central question for educators is no longer if GenAI will be used, but how we can prepare students to navigate this future responsibly. This study provides compelling evidence that intentional, structured pedagogical strategies are indispensable. Frameworks like IEEC offer a vital blueprint for transforming GenAI from a potential crutch into a powerful tool for deeper learning, ensuring we cultivate a new generation of critical thinkers who are not only proficient with technology but are also prepared to be its ethical and mindful masters.

ACKNOWLEDGEMENTS

The authors wish to express their sincere appreciation to Universiti Sultan Zainal Abidin (UniSZA) for its generous support of this research through the grant (UniSZA/2024/SoTL/14/RK088). This support played a pivotal role in enabling the comprehensive design, rigorous implementation, and thorough analysis undertaken in this study.

REFERENCES

An, Y., Yu, J. H., & James, S. (2025). Investigating the higher education institutions' guidelines and policies regarding the use of generative AI in teaching, learning, research, and administration. *International Journal of Educational Technology in Higher Education, 22*, 10. https://doi.org/10.1186/s41239-025-00507-3

Akçapınar, G., & Sidan, E. (2024). AI chatbots in programming education: Guiding success or encouraging plagiarism. *Discover Artificial Intelligence, 4*(1), 87. https://doi.org/10.1007/s44163-024-00203-7

Becker, B. A., Denny, P., Finnie-Ansley, J., Luxton-Reilly, A., Prather, J., & Santos, E. A. (2023). Programming is hard-or at least it used to be: Educational opportunities and challenges of ai code generation [Paper presentation]. *Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1*, Toronto, Canada, 15–18 March, pp. 500–506. https://doi.org/10.1145/3545945.3569759

Bente, S., Randall, N., & Wäckerle, D. (2024). A conceptual framework to transform coding education in times of generative AI. In A. Schmolitzky, & S. Klikovits (Eds.), *SEUH 2024, Lecture Notes in Informatics (LNI)* (pp. 93–104). Gesellschaft für Informatik.

Boguslawski, S., Deer, R., & Dawson, M. G. (2025). 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

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education, 20*(1), 43. https://doi.org/10.1186/s41239-023-00411-8

Chuan, O. Y., Singh, U., Alwi, S. S. E., Usop, N. S., Muda, R., & Bongsu, R. H. R. (2025, September). Integrating AI literacy in programming education: Enhancing object-oriented design and problem-solving skills. In *2025 IEEE 14th International Conference on Engineering Education (ICEED)* (pp. 298-303). IEEE. https://doi.org/10.1109/ICEED66794.2025.11400487

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Deriba, F., Sanusi, I. T., Campbell, O., & Oyelere, S. S. (2024, July 10). Computer programming education in the age of generative AI: Insights from empirical research. *SSRN*. https://doi.org/10.2139/ssrn.4891302

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., … Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management, 71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

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

Hamilton, E. R., Rosenberg, J. M., & Akcaoglu, M. (2016). The substitution augmentation modification redefinition (SAMR) model: A critical review and suggestions for its use. *TechTrends, 60*(5), 433–441. https://doi.org/10.1007/s11528-016-0091-y

Hasan, M. M., & Kabilan, M. K. (2024). Practices and effectiveness of online teaching of English in Bangladeshi universities: Implications for a revised TPACK. *Asia Pacific Journal of Educators and Education, 39*(1), 259–295. https://doi.org/10.21315/apjee2024.39.1.11

Humble, N., Boustedt, J., Holmgren, H., Milutinovic, G., Seipel, S., & Östberg, A. -S. (2023). Cheaters or AI-enhanced learners: Consequences of ChatGPT for programming education. *Electronic Journal of E-Learning, 21*(5), 16–29. https://doi.org/10.34190/ejel.21.5.3154

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kazemitabaar, M., Chow, J., Ma, C. K. T., Ericson, B. J., Weintrop, D., & Grossman, T. (2023). Studying the effect of AI code generators on supporting novice learners in introductory programming [Paper presentation]. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, Hamburg, Germany, 23–28 April, pp. 1–23. https://doi.org/10.1145/3544548.3580919

Koehler, M. J., & Mishra, P. (2009). What is technological pedagogical content knowledge? *Contemporary Issues in Technology and Teacher Education, 9*(1), 60–70.

Kohen-Vacs, D., Usher, M., & Jansen, M. (2025). Integrating generative AI into programming education: Student perceptions and the challenge of correcting AI errors. *International Journal of Artificial Intelligence in Education, 35*, 3166–3184. https://doi.org/10.1007/s40593-025-00496-4

Li, S., Liu, J., & Dong, Q. (2025). Generative artificial intelligence-supported programming education: Effects on learning performance, self-efficacy and processes. *Australasian Journal of Educational Technology, 41*(3), 1–25. https://doi.org/10.14742/ajet.9932

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations [Paper presentation]. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3313831.3376727

Luo, J. (2024). A critical review of GenAI policies in higher education assessment: A call to reconsider the "originality" of students' work. *Assessment & Evaluation in Higher Education, 49*(5), 651–664. https://doi.org/10.1080/02602938.2024.2309963

McDonald, N., Johri, A., Ali, A., & Hingle Collier, A. (2025). Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines. *Computers in Human Behavior: Artificial Humans, 3*, 100121. https://doi.org/10.1016/j.chbah.2025.100121

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record, 108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Moorhouse, B. L., Yeo, M. A., & Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers & Education Open, 5*, 100151. https://doi.org/10.1016/j.caeo.2023.100151

Ramalingam, V., & Wiedenbeck, S. (1998). Development and validation of scores on a computer programming self-efficacy scale and group analyses of novice programmer self-efficacy. *Journal of Educational Computing Research, 19*(4), 367–381.

Shardlow, M., & Latham, A. (2023). *ChatGPT in computing education: A policy whitepaper*. Council of Professors and Heads of Computing.

Sivasakthi, M., & Meenakshi, A. (2025). Generative AI in programming education: Evaluating ChatGPT's effect on computational thinking. *SN Computer Science*, *6*(5), 541. https://doi.org/10.1007/s42979-025-04051-9

Southworth, J., Migliaccio, K., Glover, J., Glover, J., Reed, D., McCarty, C., Brendemuhl, J., & Thomas, A. (2023). Developing a model for AI across the curriculum: Transforming the higher education landscape via innovation in AI literacy. *Computers and Education: Artificial Intelligence, 4*, 100127. https://doi.org/10.1016/j.caeai.2023.100127

Sun, D., Boudouaia, A., Zhu, C., & Li, Y. (2024). Would ChatGPT-facilitated programming mode impact college students' programming behaviors, performances, and perceptions? An empirical study. *International Journal of Educational Technology in Higher Education, 21*(1), 14. https://doi.org/10.1186/s41239-024-00446-5

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Yilmaz, R., & Karaoglan Yilmaz, F. G. (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*, 100147. https://doi.org/10.1016/j.caeai.2023.100147