European Journal of STEM Education,
2026, 11(1), Article 44
eISSN: 2468-4368

Research paper

Artificial Intelligence in STEM Education: Perceived Operational Effectiveness and Critical Ethical Awareness

Ela Suryani [1*], Kartika Yuni Purwanti [1,2], Mochamad Rizqi Adhi Pratama[1,3],
Wulan Aulia Azizah[3]

[1] *Universitas Ngudi Waluyo*, INDONESIA
[2] *Universitas Negeri Yogyakarta,* INDONESIA
[3] *Universitas Negeri Semarang,* INDONESIA

*__Corresponding Author:__ ela.suryani16@gmail.com

This paper is part of the special issue: AI and STEM: Exploring the Impact of Artificial Intelligence on Science, Technology, Engineering, and Mathematics Education

Citation: Suryani, E., Purwanti, K. Y., Pratama, M. R. A., & Azizah, W. A. (2026). Artificial intelligence in STEM education: Perceived operational effectiveness and critical ethical awareness. *European Journal of STEM Education, 11*(1), Article 44. https://doi.org/10.20897/ejsteme/19457

Published: September 28, 2026

ABSTRACT

Artificial intelligence (AI) is increasingly used in education. Yet, evidence remains limited on how educators' perceptions of its practical value relate to their awareness of ethical risks, particularly among educators in Semarang Regency, Indonesia. This quantitative survey examined prior AI use, self-reported confidence in understanding basic AI principles, perceived operational effectiveness, and critical ethical awareness among 218 certified educators in Semarang Regency, Central Java, Indonesia, recruited through purposive sampling. Data were collected using a structured questionnaire and analyzed with descriptive statistics and Pearson's product–moment correlation. Sixty-four percent of respondents reported prior use of AI-based tools, whereas 27% reported confidence in understanding basic AI principles. These separate aggregate indicators show that AI exposure was more prevalent than conceptual confidence in this sample, although they do not establish an individual-level gap. Respondents reported favorable perceptions of AI's operational effectiveness ($M = 4.18$, $SD = 0.63$) and critical ethical awareness above the scale midpoint ($M = 3.42$, $SD = 0.57$). The two constructs were moderately and positively associated, $r(216) = 0.49$, 95% CI [0.38, 0.59], $p < 0.001$, indicating that stronger perceived operational effectiveness co-occurred with greater reported attention to accuracy, privacy, bias, transparency, accountability, and human oversight. These findings indicate that perceived operational effectiveness and critical ethical awareness were positively related dimensions of educators' reported engagement with AI.

Keywords: artificial intelligence, critical ethical awareness, perceived operational effectiveness, STEM education

Digital transformation has reshaped how people learn, work, and communicate. In education, artificial intelligence (AI) has evolved from an experimental technology into an increasingly prominent tool for lesson planning, instructional material development, feedback provision, assessment, and learning analytics (Baker & Siemens, 2014; Holmes et al., 2019, 2021). This transformation is particularly relevant to science, technology, engineering, and mathematics (STEM) education, in which inquiry, modelling, data interpretation, problemsolving, and evidence-based decision-making constitute essential components of the learning process (OECD, 2021; Raza et al., 2025; Suryani et al., 2023; UNESCO, 2017). As international attention to STEM education continues to grow (Suryani et al., 2025), the integration of AI offers new possibilities for supporting learning that is more adaptive and responsive to students' needs (Celik et al., 2022; Li et al., 2020).

Within this context, AI can help teachers identify learning needs, prepare explanations and instructional activities, adjust the difficulty of learning materials, provide timely feedback, and support assessment processes (Zawacki-Richter et al., 2019). Intelligent tutoring systems, generative AI, automated assessment, and learning analytics may also broaden how teachers design problem- and inquiry-based STEM learning experiences (Luckin & Holmes, 2016). Nevertheless, AI does not automatically lead to more meaningful learning. Its educational value continues to depend on teachers' ability to select appropriate tools, verify the scientific accuracy of AI-generated outputs, interpret system recommendations, and situate them within relevant pedagogical contexts (Saavedra & Opfer, 2012; Williamson & Piattoeva, 2019). Teachers should therefore not be regarded as passive users, but as key actors who exercise professional judgement and ensure that technology supports—rather than replaces—the pedagogical relationship between teachers and students (Celik et al., 2022; Miao & Cukurova, 2024).

However, narratives about the transformative potential of AI should be balanced against evidence of its current use among teachers. *Results from the Teaching and Learning International Survey (TALIS 2024)* indicate that, on average across OECD education systems, approximately one-third of lower-secondary teachers reported using AI in their work during the preceding 12 months (OECD, 2025). Among teachers who used AI, 68% employed it to learn about or summarise a topic, while 64% used it to generate lesson plans or instructional activities. Smaller proportions used AI to review data on student participation or performance (25%) or to assess or grade student work (26%) (OECD, 2025). These findings should not be generalized to all STEM educators or to educators outside the study setting. Nevertheless, the pattern suggests that teachers' AI use remains concentrated on planning and information-processing tasks, whereas its use in decisions directly related to student evaluation is more limited. This distinction is particularly important in STEM education because apparently convincing AI-generated outputs may still contain scientific inaccuracies, conceptual oversimplifications, or recommendations that are inconsistent with disciplinary characteristics.

The use of AI should also not be equated with readiness or conceptual understanding. Teachers may be able to use generative AI interfaces effectively to complete familiar tasks without understanding how training data, model probabilities, algorithmic limitations, and automated recommendation processes shape system outputs. TALIS 2024 further reported that approximately three-quarters of teachers who had not used AI felt that they lacked the knowledge or skills required to teach with it (OECD, 2025). UNESCO's *AI Competency Framework for Teachers* emphasises that teacher readiness is multidimensional, encompassing a human-centred mindset, AI ethics, AI foundations and applications, AI pedagogy, and AI for professional development (Miao & Cukurova, 2024). The AI literacy literature similarly identifies conceptual knowledge, practical use, critical evaluation, and communication with AI as interconnected competencies (Long & Magerko, 2020; Sperling et al., 2024; Wu et al., 2024). Teachers' self-reported confidence in understanding basic AI principles may therefore serve as an indicator of perceived readiness. However, it should not be interpreted as an objective measure of AI knowledge or as evidence of competence in implementing AI in classroom practice.

At the same time, the use of AI in education raises ethical issues that cannot be separated from its operational benefits (Saifudin et al., 2024). The processing of student data, use of generative outputs, and reliance on algorithm-assisted recommendations raise concerns regarding data privacy and security, algorithmic bias, unequal treatment, transparency, information accuracy, accountability, and the potential erosion of teachers' professional judgement (Pedersen et al., 2024; Williamson & Piattoeva, 2019). TALIS 2024 found that approximately four in ten teachers were concerned that AI could amplify bias, reinforce student misconceptions, or threaten data privacy and security. Approximately seven in ten teachers also believed that AI could enable students to present other people's work as their own (OECD, 2025). Such concerns should not automatically be interpreted as resistance to AI. From a human-centred AI perspective, attention to these risks may instead reflect the critical vigilance required to verify AI-generated outputs, protect students, disclose AI use transparently, and maintain meaningful human oversight (Floridi et al., 2018).

Emerging empirical research indicates that teachers view AI as a source of both practical benefits and professional challenges in STEM education. Uğraş et al. (2025), for example, found that teachers perceived ChatGPT as supporting feedback, content personalisation, creativity, planning flexibility, and the relevance of STEM learning. At the same time, teachers identified technical challenges and a need for stronger security-related knowledge. These findings illustrate how perceptions of AI's operational benefits may coexist with attention to the risks and conditions required for its responsible use.

Technical understanding and critical ethical awareness should therefore be examined as related but distinct dimensions of teacher readiness. A better understanding of AI may help teachers recognise limitations, uncertainties, and risks that are less apparent to inexperienced users. Greater AI understanding may consequently not reduce ethical concerns; rather, it may strengthen teachers' capacity to evaluate risks critically. Empirical evidence involving K–12 teachers indicates that AI literacy is positively associated with awareness of AI ethics (Du et al., 2024). In the present study, perceived operational effectiveness of AI refers to the extent to which teachers believe that AI can improve the efficiency and effectiveness of instructional work. Critical ethical awareness of AI refers to teachers' vigilance regarding the risks and responsibilities associated with accuracy, privacy, bias, transparency, accountability, and the need to preserve human judgement. Under these definitions, a higher ethical-awareness score indicates stronger ethical vigilance rather than greater acceptance of AI.

Although research on AI in education is expanding rapidly (Chen et al., 2020; Ertmer & Ottenbreit-Leftwich, 2010; Long & Magerko, 2020; Sperling et al., 2024), three relevant gaps remain. First, numerous studies emphasise the potential of AI or teachers' intentions to use it, but insufficiently explain its actual patterns of use among educators. Second, AI use, readiness, understanding, and acceptance are frequently treated as interchangeable concepts, even though experience with AI and self-reported confidence do not demonstrate conceptual mastery. Third, the operational benefits and ethical implications of AI are often examined separately, leaving limited evidence regarding the relationship between teachers' perceptions of AI's operational effectiveness and their critical ethical awareness. These gaps are particularly consequential in STEM education because inaccurate, biased, or opaque AI-generated outputs may affect scientific reasoning, data interpretation, assessment practices, and students' access to learning opportunities. An analysis that clearly distinguishes experience with AI use, self-reported understanding, perceived operational effectiveness, and critical ethical awareness is therefore warranted.

To address these gaps, this cross-sectional study examines educators' use of AI-based tools, their self-reported confidence in understanding basic AI principles, their perceptions of the operational effectiveness of AI, and their critical ethical awareness concerning the use of AI in education. The study contributes to the literature by shifting the focus from general narratives of technology adoption towards a more specific account of how usage experience, perceived readiness, operational benefits, and ethical vigilance coexist among educators. The investigation is deliberately limited to respondents' self-reported experiences and perceptions; it does not objectively assess AI knowledge, observe classroom practices, measure student learning outcomes, or establish causal relationships.

Unlike studies that treat AI use, readiness, acceptance, and ethical concerns as interchangeable indicators, this study distinguishes teachers' prior AI use, self-reported confidence in basic AI principles, perceived operational effectiveness, and critical ethical awareness. Its specific contribution is to examine whether teachers who perceive AI as useful for instructional work also report stronger ethical vigilance regarding privacy, bias, transparency, accuracy, accountability, and human oversight. By examining this relationship among certified STEM educators in Semarang Regency, Indonesia, the study provides context-specific evidence that operational value and ethical vigilance may coexist as related dimensions of responsible AI engagement rather than as competing orientations. Based on this scope, the study addresses the following research questions:

1. What proportion of educators report having used AI-based tools and feeling confident in their understanding of basic AI principles?
2. What is the relationship between educators' perceived operational effectiveness of AI and their critical ethical awareness of AI?

METHOD

Research design

This study employed a cross-sectional correlational survey design. The survey design was considered suitable because it facilitates the systematic accumulation of data from a defined population to characterize their perceptions, attitudes, and behaviors concerning a specific phenomenon (Creswell, 2014). This design allowed the researchers to describe educators' reported AI use, their self-reported confidence in understanding basic AI principles, their perceived operational effectiveness of AI, and their critical ethical awareness, as well as to examine the association between the two perceptual constructs.

Participants

This study involved 218 certified professional educators who were actively teaching at least one science, technology, engineering, or mathematics (STEM) subject at the primary or secondary education level in Semarang Regency, Central Java, Indonesia. In this study, certified professional educators refer to teachers who held a professional educator registration number issued by the Indonesian Ministry of Education and Culture and were employed at a participating institution at the time of data collection. Eligible participants were teachers who taught science, mathematics, technology, or engineering, had access to the online survey, and provided informed consent. Teachers who were not actively teaching a STEM subject, as well as duplicate and incomplete responses, were excluded from the analysis.

Research instrument

EFA was conducted on 12 five-point Likert-scale items using responses from 130 educators as the instrument-validation sample. Instrument validation was conducted using an independent sample of 130 educators. None of the participants in the validation sample were included in the main analytic sample of 218 respondents. The validation sample was recruited using the same eligibility criteria, namely professionally certified educators actively teaching a STEM subject at the primary or secondary level in Semarang Regency, Central Java, during 2025. The data were suitable for factor analysis, as indicated by a KMO value of 0.894. Bartlett's test of sphericity was also statistically significant, $\chi^2(66) = 424.74$, $p < 0.001$, indicating that the correlation matrix was suitable for factor analysis. Factors were extracted using principal axis factoring with direct oblimin rotation because the two constructs were theoretically expected to be correlated. The eigenvalue-greater-than-one criterion, visual inspection of the scree plot, and parallel analysis supported retaining two factors. Together, the two factors accounted for 42.85% of the total variance after extraction.

The perceived operational effectiveness factor comprised six items with pattern coefficients ranging from 0.617 to 0.745, while critical ethical awareness comprised six items with coefficients ranging from 0.559 to 0.686. Item communalities ranged from 0.312 to 0.555. No cross-loadings of 0.30 or higher were identified; therefore, all items were retained. The interfactor correlation of .31 indicated that the constructs were positively related but empirically distinguishable. Both factors demonstrated adequate-to-good internal consistency, with Cronbach's alpha coefficients of 0.84 and 0.79, respectively. Overall, the findings provided preliminary support for the instrument's two-factor structure and reliability. The correlation of 0.31 represents the factor correlation obtained from the validation sample, rather than the correlation between composite scores in the main survey. The complete wording of the 12 items, response coding, communalities, pattern matrix, interfactor correlation, and instrument-validation results are provided in Supplementary Material AI-STEM Instrument and EFA.

Data collection procedure

A total of 245 formal invitations containing unique access tokens were sent to eligible educators within the networks, resulting in 218 verified completions (89% response rate). Despite this controlled distribution, the study remains subject to self-selection bias. Educators who already maintain an active interest in or preliminary engagement with digital technologies may have been more intrinsically motivated to complete the survey, which could inflate baseline scores for prior AI usage and positive perceptions of effectiveness.

Data analysis

Data were analyzed using descriptive statistics and Pearson's product-moment correlation in SPSS version 26.0. Composite scores were calculated as the mean of the six items representing each construct; higher scores indicated greater perceived operational effectiveness or critical ethical awareness. Incomplete and duplicate responses were excluded before analysis. Internal consistency was evaluated using Cronbach's alpha. Statistical significance was set at $p < 0.05$, and 95% confidence intervals were reported. Following data screening, all 218 respondents had complete responses to the 12 scale items and therefore had complete composite scores for both constructs; no item-level imputation was performed. Pearson assumptions were examined before the correlation analysis. The scatterplot indicated an approximately linear positive relationship, and the quadratic term did not improve model fit ($\Delta R^2 = 0.0054$, F change$(1, 215) = 1.544$, $p = 0.215$). Standardized residuals showed no marked curvature and no absolute value exceeded 3. No influential cases were identified: the maximum Cook's distance was 0.0143, below the 4/N criterion of 0.0183, and the largest absolute studentized deleted residual was 1.652. Accordingly, no sensitivity analysis excluding cases was warranted. Full diagnostic results are provided in Supplementary Material Data Analysis.

RESULTS AND DISCUSSION

Among the 218 respondents in this survey, the predominant group (87%) consisted of educators in mathematics and science. In contrast, the remaining participants were from the fields of technology and engineering (**Table 1**). This ratio illustrates the prevalence of scientific and mathematics educators in primary and secondary education systems, especially in developing nations, where enhancing STEM competencies is a fundamental objective of educational policy (Marginson et al., 2013; UNESCO, 2017).

Table 1

Demographic and AI usage summary

Variable	Frequency (%)
STEM Subject Taught	

Science and Mathematics	190 (87%)
Technology and Engineering	28 (13%)
Prior Use of AI-based Tools	
Yes	140 (64%)
No	78 (36%)
Self-reported confidence in understanding AI principles	
Yes	59 (27%)
No	159 (73%)

AI use and confidence in basic AI principles

Of the 218 respondents, 140 (64%) reported having used AI-based tools, whereas 78 (36%) reported no prior use. Confidence in understanding basic AI principles was less common: 59 respondents (27%) answered affirmatively and 159 (73%) did not. At the aggregate level, exposure to AI was therefore more widespread than confidence in its foundational principles.

This pattern portrays a teaching population encountering AI in practice while conceptual confidence remains uneven. Because the two indicators were summarized separately rather than cross-tabulated, the result describes the sample as a whole; it does not assign the difference to particular individuals. Even with that boundary, the coexistence of relatively broad use and limited confidence is educationally consequential because it distinguishes access to a tool from confidence in explaining the system behind it. Long & Magerko (2020) conceptualized AI literacy as a broader constellation of competencies, and Sperling et al. (2024) showed that teacher-education scholarship frames AI literacy in multiple ways. UNESCO similarly differentiates a human-centered mindset, ethics, AI foundations and applications, pedagogy, and continuing professional learning (Miao & Cukurova, 2024). The two binary indicators used here illuminate early engagement with AI, while the composite measures examine how teachers perceive its operational value and ethical demands.

The 64% use rate is consistent with the rapid normalization of AI-assisted information processing and lesson preparation in teachers' work. TALIS 2024 found that teachers used AI more often for information-related and planning activities than for grading or analyzing student performance (OECD, 2025). Although that international evidence is not a direct benchmark for this purposively selected Indonesian sample, it provides a plausible professional context: educators can encounter AI first as a convenient planning and production tool. In the present study, reported use denotes exposure rather than intensity. The survey did not differentiate particular platforms, frequency, duration, or instructional purpose. The result therefore captures the breadth of contact with AI-based tools and establishes that a majority of respondents reported some prior exposure to AI-based tools.

The 27% confidence figure adds a distinct conceptual layer. An affirmative response may arise from training, repeated use, independent study, or familiarity with basic terminology; a negative response may indicate limited knowledge, uncertainty about what qualifies as an AI principle, or a deliberately cautious self-assessment. The result is therefore best understood as the distribution of perceived conceptual confidence. It neither equates nonconfidence with inability nor treats confidence as verified mastery. This precision is valuable because self-confidence can influence willingness to experiment, ask critical questions, and seek support, even when it does not perfectly correspond to objective knowledge. The finding thus identifies conceptual confidence as a salient dimension of teachers' engagement with AI without conflating it with a complete measure of AI literacy.

Two indicators reveal an asymmetry in the sample's AI engagement: practical exposure was common, whereas confidence in foundational principles was reported by a minority. This asymmetry contributes to the literature by showing why teacher engagement with AI cannot be represented by a single adoption indicator. Studies centered only on whether teachers use AI may overlook differences in how securely they understand its operation and limitations. Chen et al. (2020) and Zawacki-Richter et al. (2019) identified a broader imbalance between the expansion of AI applications and attention to educators and educational theory. The present data locate that issue in an Indonesian STEM-teaching context and establish the rationale for examining operational perceptions alongside critical ethical awareness.

Perceived operational effectiveness and critical ethical awareness

Table 2 presents the composite-score results. Perceived operational effectiveness averaged 4.18 (SD = 0.63), while critical ethical awareness averaged 3.42 (SD = 0.57), with both constructs measured on five-point response scales. The operational-effectiveness distribution was concentrated toward favorable judgments of AI's capacity to support instructional work. Critical ethical awareness also occupied the upper half of the response scale, while displaying a different central tendency. Because the composites represent distinct constructs with different item content, the two means are interpreted separately rather than as a direct numerical gap. This construct-specific reading preserves the meaning of each scale: one captures perceived usefulness and efficiency, and the other captures attention to accuracy, privacy, bias, transparency, accountability, and human oversight.

Table 2

Descriptive statistics of perceptions toward AI in STEM education

Variable	Mean (M)	SD	Scale
Perceived Operational Effectiveness of AI	4.18	0.63	1-5
Critical Ethical Awareness of AI	3.42	0.57	1-5

The operational-effectiveness mean demonstrates that respondents generally perceived AI as capable of assisting the practical work of teaching. These perceptions may include faster preparation, generating examples, adapting explanations, producing feedback, or supporting routine instructional tasks. At the same time, operational effectiveness in this study is a belief-based construct: it records teachers' judgments of usefulness rather than independently observed gains in lesson quality or student outcomes. Research on AI-supported teaching similarly presents educational value as contingent on pedagogical design and professional judgment (Celik et al., 2022; Holmes et al., 2019; Roll & Wylie, 2016). In educational practice, AI-based systems may support the preparation of instructional materials and activities, feedback provision, learning analytics, and selected forms of automated assessment. However, empirical evidence indicates that these benefits depend on pedagogical appropriateness, output accuracy, and teachers' capacity to review system-generated recommendations (Celik et al., 2022; Holmes et al., 2019; Morris et al., 2025; Uğraş et al., 2025). The favorable mean therefore represents a strong orientation toward practical value within the measured domain.

The critical-ethical-awareness mean shows that respondents also attended to the responsibilities accompanying AI use. Its items address the need to examine accuracy, protect privacy, recognize bias, maintain transparency and accountability, and preserve human oversight. These concerns are not signs of general resistance to AI; they represent the evaluative work required when automated output can affect teaching materials, assessment, feedback, and student information. Holmes et al. (2021) positioned ethics in educational AI as a community-wide responsibility, while Floridi et al. (2018) articulated the principles of beneficence, non-maleficence, autonomy, justice, and explicability. The present finding connects these high-level principles with teachers' reported attention to concrete professional risks. It indicates that ethical reflection was present alongside favorable operational perceptions rather than emerging only among educators who rejected AI.

The standard deviations indicate some variability in respondents' perceptions. Teachers did not hold identical views of AI's practical value or ethical demands. Variation can plausibly arise from differences in experience, disciplinary tasks, school level, training, institutional policy, digital infrastructure, and the purposes for which AI is used (Azzam & Charles, 2024; Morris et al., 2025; Popenici & Kerr, 2017). These factors were not modeled in the present analysis, but the observed dispersion establishes that the sample cannot be reduced to a uniformly enthusiastic or uniformly cautious group. Instead, the descriptive results portray a differentiated professional field in which teachers are negotiating usefulness and responsibility from distinct positions.

Association between perceived operational effectiveness and critical ethical awareness

Pearson's product-moment correlation identified a positive association of moderate magnitude between perceived operational effectiveness and critical ethical awareness, $r(216) = 0.49$, 95% CI [0.38, 0.59], $p < 0.001$ (Table 3). The interval estimate supports a consistently positive relationship under the analytical assumptions. The squared correlation, $r^2 = 0.24$, indicates that the two composite scores shared approximately 24% of their variance in this sample. This degree of overlap is substantively meaningful while leaving substantial construct-specific variance. Operational judgments and ethical awareness are therefore connected but remain distinguishable dimensions of teachers' engagement with AI.

Table
Pearson correlation matrix

Variable	Perceived Operational Effectiveness of AI	Critical Ethical Awareness of AI
Perceived Operational Effectiveness of AI	1.00	0.49
Critical Ethical Awareness of AI	0.49	1.00
Note: N = 218. 95% CI for r = [0.38, 0.59]; p < 0.001		

Substantively, respondents who viewed AI as more operationally useful also tended to report stronger attention to ethical risks and professional responsibilities. This pattern challenges the assumption that enthusiasm for AI necessarily weakens critical scrutiny. In the observed sample, favorable judgments of usefulness coexisted with attention to accuracy, privacy, bias, transparency, accountability, and the preservation of teacher judgment. Human-centered approaches conceptualize such oversight as a constituent of responsible innovation, not as an obstacle to it (Holmes et al., 2021; Miao & Cukurova, 2024). The correlation thus supports an integrated interpretation of teacher engagement: operational optimism and ethical vigilance can coexist.

The relationship can be understood through several complementary mechanisms. Direct encounters with AI may expose educators to both its conveniences and its limitations, producing more differentiated judgments of usefulness and risk. Teachers who begin with a strong sense of professional responsibility may also evaluate AI outputs more carefully and, through that scrutiny, identify appropriate operational uses. In addition, professional learning, digital expertise, and institutional support may strengthen both dimensions. Du et al. (2024) identified a positive relationship between AI literacy and ethical awareness in teacher learning, providing a related account of how knowledge and ethical reflection may develop together. The current correlation concerns two perceptual composites and does not establish temporal order, but its moderate magnitude provides empirical grounds for treating operational and ethical dimensions as related components of responsible engagement.

The construct specificity of this result is central to its contribution. The association links perceived operational effectiveness with critical ethical awareness; it is not an analysis of objective AI knowledge, readiness, acceptance, classroom implementation, or student outcomes. Nor was the binary confidence indicator correlated with ethical awareness. By retaining the measured construct labels, the study offers a precise account of what varies together: teachers' judgments that AI can assist their work and their reported vigilance toward the ethical conditions of that assistance. This precision makes the finding theoretically interpretable and prevents broader concepts from obscuring the empirical relationship.

Integrated interpretation and contribution to AI-enabled STEM teaching

Taken as a whole, the findings advance an integrated account of educators' perceptions related to responsible AI engagement in STEM education. Perceived operational effectiveness and critical ethical awareness are not competing endpoints. Teachers need to determine whether an AI response saves time or expands instructional options and, simultaneously, whether it is accurate, equitable, transparent, privacy-preserving, and subject to professional review. This integration is especially visible in discipline-specific tasks: evaluating an AI-generated scientific explanation, checking a mathematical derivation, reviewing automated feedback, or deciding whether student information can be entered into a system. The positive association observed here provides empirical support for positioning these judgments within the same professional practice.

For educators, verification is both disciplinary and ethical. Scientific explanations require conceptual accuracy and evidentiary adequacy; mathematical solutions require valid assumptions, procedures, and results; data interpretations require attention to provenance and uncertainty; and engineering recommendations require consideration of constraints, safety, and trade-offs. Fluent or visually convincing AI output can fail on any of these dimensions. Celik et al. (2022) emphasized teacher knowledge for critical AI use, while Raza et al. (2025) connected accuracy, validity, and ethics in AI-enhanced education. The present results extend this literature by showing that teachers' perceived practical value is positively, rather than negatively, aligned with reported critical awareness.

The results also locate responsible AI engagement within an institutional setting. Teachers exercise judgment inside systems that define approved tools, data-protection expectations, intellectual-property practices, disclosure requirements, and procedures for human review. Secure access and coherent policy can determine whether ethical awareness becomes routine practice. OECD (2025) and UNESCO guidance (Miao et al., 2021; Miao & Cukurova, 2024) similarly connect capacity building with governance and professional roles. Although institutional policy was not measured directly, the coexistence of favorable operational perceptions, ethical concern, and substantial unexplained variance is compatible with a multilevel understanding in which individual judgment and organizational conditions jointly shape practice.

The descriptive profile further indicates that teachers enter AI-related professional learning from different starting points. Some have no prior use, some have practical experience without conceptual confidence, and others may combine experience with confidence. These positions imply different learning trajectories, but all converge on a common professional task: linking tool operation to verification and accountable decision making. A coherent program therefore joins foundational concepts, authentic STEM applications, ethical scenarios, and reflection on professional responsibility. In that structure, conceptual knowledge explains why outputs require scrutiny, practical tasks develop the capacity to scrutinize them, and ethical reasoning guides the final instructional decision.

This integrated interpretation also sharpens the outcomes by which professional learning can be evaluated. Satisfaction and confidence capture participants' perceptions, whereas objective knowledge tests, performance tasks, lesson artifacts, verification behavior, scenario-based ethical decisions, and evidence of data safeguards capture demonstrated capability. The distinction mirrors the present results: self-report is informative about how educators position themselves toward AI, but it is not identical to observed competence. Bringing perceptual and performance evidence together would provide a fuller account of responsible AI capacity in STEM teaching.

CONCLUSION

This study examined Indonesian educators' reported use of artificial intelligence (AI), confidence in understanding basic AI principles, perceived operational effectiveness of AI, and critical ethical awareness. Among the 218 participants, 64% reported having used AI in educational contexts, whereas 27% reported confidence in their understanding of basic AI principles. These findings indicate an aggregate discrepancy between reported AI use and confidence in understanding AI, although they do not establish that this discrepancy occurred within the same individuals.

Perceived operational effectiveness was positively associated with critical ethical awareness, $r(216) = 0.49$, 95% CI [0.38, 0.59], $p < 0.001$. The squared correlation indicated that approximately 24% of the variance was statistically shared between the two composite scores. Thus, educators who perceived AI as more operationally effective also tended to report greater awareness of its ethical implications. This relationship should not be interpreted as causal because the study employed a cross-sectional design and relied on self-reported data.

The findings suggest that professional development for educators should address both the practical use of AI and critical reflection on issues such as bias, transparency, privacy, accountability, and responsible decision-making. Nevertheless, the purposive sampling procedure, reliance on self-reported measures, unequal representation across STEM disciplines, and absence of objective measures of AI knowledge limit the generalizability of the findings. Future research should employ more representative samples, objective assessments of AI knowledge and practice, confirmatory validation of the measurement model, and longitudinal or mixed-method designs to clarify how educators' perceptions and ethical awareness develop over time.

Acknowledgment

The authors thank PPAPT, BPI, LPDP, all supervisors, and Universitas Negeri Yogyakarta and Universitas Ngudi Waluyo for supporting this research.

Funding Sources

The authors wish to convey profound appreciation to the Indonesian Education Scholarship, Center for Higher Education Funding and Assessment, and Indonesian Endowment Fund for Education Agency, which fully funded this article.

Ethics Statement

Ethical approval for this study was granted by the Research Ethics Committee of Universitas Ngudi Waluyo on May 20, 2025 (Approval No. 245/KEP/EC/UNW/2025). Before data collection, all participants were informed of the study's purpose and procedures, the voluntary nature of participation, the confidentiality of their data, and their right to decline participation or withdraw without penalty. Informed consent was obtained from each participant before questionnaire completion. All data were anonymized and used exclusively for research and scholarly publication purposes. We stored electronic data in password-protected files accessible only to the research team. Identifying information was not collected or was removed before analysis.

Conflict of Interest

We declare that there are no conflicts of interest.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

AI disclosure

The authors used ChatGPT by OpenAI solely to assist with English-language editing and clarity. All AI-assisted suggestions were critically reviewed and revised by the authors, who take full responsibility for the accuracy, originality, interpretations, and final content of the manuscript. AI was not used to generate, fabricate, modify, or analyze the research data.

Biographical sketch

Ela Suryani is a lecturer in the Department of Elementary Teacher Education, Faculty of Computer and Education (FKP), Universitas Ngudi Waluyo, Indonesia. She holds a doctoral degree in education and specializes in elementary science education and STEAM education. Her research interests include elementary science teaching and learning, STEAM-integrated learning in primary schools, innovative pedagogical approaches for meaningful elementary education, as well as students' scientific literacy, creativity, and environmental literacy.

Kartika Yuni Purwanti is a PhD student in Elementary Education at the Faculty of Education, Universitas Negeri Yogyakarta (UNY), Indonesia, and a lecturer in the Department of Elementary Teacher Education, Faculty of Computer and Education, Universitas Ngudi Waluyo, Indonesia. Her primary research field is elementary social studies education. Her research interests include social studies learning in primary schools, the development and use of instructional media, digital learning media, and social and cultural literacy among elementary school students.

Mochamad Rizqi Adhi Pratama is an English lecturer at Universitas Ngudi Waluyo, Indonesia, and a PhD candidate at Universitas Negeri Semarang (UNNES), Indonesia. He obtained his master's degree at UNNES and has written books such as Speaking for Formal Interactions and Basic English Grammar. He has been interested in Teaching English as a Foreign Language (TEFL) and Computer-Assisted Language Learning (CALL). He has written several journal articles, including articles on innovation in language teaching and the utilization of technology in language teaching.

Wulan Aulia Azizah is a Lecturer and Researcher in Primary Education at the Department of Elementary School Teacher Education Program, Faculty of Education and Psychology (FIPP), Universitas Negeri Semarang (UNNES), Indonesia. She earned a Doctoral Degree in Primary Education from Yogyakarta State University. Her primary field of expertise is Primary Education, particularly Science and STEAM Education and Teacher Professional Development. Her research interests include Science and STEAM Education, Teacher Professional Development and Pedagogical Knowledge, Technology-Enhanced Learning, and Culturally Responsive/ Ethnoscience Education.

Disclaimer/Publisher's Note

REFERENCES

Azzam, A., & Charles, T. (2024). A review of artificial intelligence in K-12 education. *Open Journal of Applied Sciences, 14*(8), 2088–2100. https://doi.org/10.4236/ojapps.2024.148137

Baker, R., & Siemens, G. (2014). Educational data mining and learning analytics. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 253–272). Cambridge University Press. https://doi.org/10.1017/CBO9781139519526.016

Celik, I., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research. *TechTrends, 66*(4), 616–630. https://doi.org/10.1007/s11528-022-00715-y

Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 1*, Article 100002. https://doi.org/10.1016/j.caeai.2020.100002

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.

Du, H., Sun, Y., Jiang, H., Islam, A. Y. M. A., & Gu, X. (2024). Exploring the effects of AI literacy in teacher learning: An empirical study. *Humanities and Social Sciences Communications, 11*, Article 559. https://doi.org/10.1057/s41599-024-03101-6

Ertmer, P. A., & Ottenbreit-Leftwich, A. T. (2010). Teacher technology change: How knowledge, confidence, beliefs, and culture intersect. *Journal of Research on Technology in Education, 42*(3), 255–284. https://doi.org/10.1080/15391523.2010.10782551

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education, 32*(3), 504–526. https://doi.org/10.1007/s40593-021-00239-1

Li, Y., Wang, K., Xiao, Y., & Froyd, J. E. (2020). Research and trends in STEM education: A systematic review of journal publications. *International Journal of STEM Education, 7*, Article 11. https://doi.org/10.1186/s40594-020-00207-6

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

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.

Marginson, S., Tytler, R., Freeman, B., & Roberts, K. (2013). *STEM: Country comparisons*. Australian Council of Learned Academies.

Miao, F., & Cukurova, M. (2024). *AI competency framework for teachers*. UNESCO. https://doi.org/10.54675/ZJTE2084

Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). *AI and education: Guidance for policy-makers*. UNESCO. https://doi.org/10.54675/PCSP7350

Morris, W., Holmes, L., Choi, J. S., & Crossley, S. (2025). Automated scoring of constructed response items in math assessment using large language models. *International Journal of Artificial Intelligence in Education, 35*(2), 559–586. https://doi.org/10.1007/s40593-024-00418-w

OECD. (2021). *AI and the future of skills, Volume 1: Capabilities and assessments*. OECD Publishing. https://doi.org/10.1787/5ee71f34-en

OECD. (2025). *Results from TALIS 2024: The state of teaching*. OECD Publishing. https://doi.org/10.1787/90df6235-en

Pedersen, C., Aagaard, T., Daus, S., Nagel, I., Amdam, S. H., Vika, K. S., Røkenes, F. M., & Andreasen, J. K. (2024). Profiling teacher educators' strategies for professional digital competence development. *Teachers and Teaching: Theory and Practice, 30*(4), 417–436. https://doi.org/10.1080/13540602.2024.2336612

Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning, 12*, Article 22. https://doi.org/10.1186/s41039-017-0062-8

Raza, F. A., Singh, A. D., Kovilpillai, J. J. S., Hamdan, A., & Rajaratnam, V. (2025). Safeguarding integrity in AI-enhanced education: Stakeholder perspectives on accuracy, validity, and ethics in ASEAN. *European Journal of STEM Education, 10*(1), Article 22. https://doi.org/10.20897/ejsteme/17307

Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education, 26*(2), 582–599. https://doi.org/10.1007/s40593-016-0110-3

Saavedra, A. R., & Opfer, V. D. (2012). Learning 21st-century skills requires 21st-century teaching. *Phi Delta Kappan, 94*(2), 8–13. https://doi.org/10.1177/003172171209400203

Saifudin, N. A. M. E., Anwar, N., Hashim, H., & Kurniawan, Y. (2024). Artificial intelligence in education: A preliminary study. *Open Journal of Social Sciences, 12*(11), 561–578. https://doi.org/10.4236/jss.2024.1211039

Sperling, K., Stenberg, C. J., McGrath, C., Åkerfeldt, A., Heintz, F., & Stenliden, L. (2024). In search of artificial intelligence (AI) literacy in teacher education: A scoping review. *Computers and Education Open, 6*, Article 100169. https://doi.org/10.1016/j.caeo.2024.100169

Suryani, E., Kun, Z., & Haryanto, H. (2023). The implementation of STEM approach (science, technology, engineering, and mathematics) on science learning at elementary school. *Proceedings Series on Social Sciences & Humanities, 12*, 315–322. https://doi.org/10.30595/pssh.v12i.814

Suryani, E., Prasetyo, Z. K., Hermanto, H., & Purwanti, K. Y. (2025). A comparative study of inquiry, STEAM, and STEAM-based guided inquiry (GI-STEAM). *European Journal of STEM Education, 10*(1), Article 16. https://doi.org/10.20897/ejsteme/17191

Uğraş, H., Uğraş, M., Papadakis, S., & Kalogiannakis, M. (2025). Innovative early childhood STEM education with ChatGPT: Teacher perspectives. *Technology, Knowledge and Learning, 30*, 809–831. https://doi.org/10.1007/s10758-024-09804-8

UNESCO. (2017). *Cracking the code: Girls' and women's education in science, technology, engineering and mathematics (STEM).*

Williamson, B., & Piattoeva, N. (2019). Objectivity as standardization in data-scientific education policy, technology and governance. *Learning, Media and Technology, 44*(1), 64–76. https://doi.org/10.1080/17439884.2018.1556215

Wu, H., Wang, Y., & Wang, Y. (2024). "To use or not to use?" A mixed-methods study on the determinants of EFL college learners' behavioral intention to use AI in the distributed learning context. *International Review of Research in Open and Distributed Learning, 25*(3), 158–178. https://doi.org/10.19173/irrodl.v25i3.7708

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? *International Journal of Educational Technology in Higher Education, 16*, Article 39. https://doi.org/10.1186/s41239-019-0171-0

