Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: https://www.science-gate.com/IJAAS.html

Enhancing conflict resolution skills through artificial intelligence-based problem-based learning in civic education at Indonesian secondary schools

Nadziroh Nadziroh [1,*], Sunarso Sunarso [1], Suyato Suyato [2]

[1]*Department of Civic Education, Faculty of Social and Political Sciences, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia*
[2]*Department of Educational Technology, Faculty of Education, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia*

ARTICLE INFO

Article history:
Received 5 June 2025
Received in revised form
6 November 2025
Accepted 14 December 2025

Keywords:
Conflict resolution
Civic education
Problem-based learning
Artificial intelligence
Adolescents

ABSTRACT

Traditional civic education in Indonesia does not sufficiently develop practical conflict resolution skills among adolescents. This gap contributes to ongoing social tensions in diverse communities, where democratic skills are essential for social cohesion. This quasi-experimental study examined whether Artificial Intelligence–enhanced Problem-Based Learning (AI-PBL) could improve conflict resolution skills among 90 eleventh-grade students in three urban schools in Yogyakarta, Indonesia. The intervention used a natural language processing system to present culturally relevant conflict scenarios, provide real-time adaptive feedback, and create personalized learning pathways across six structured sessions. Conflict resolution skills were measured in three areas—empathy, peaceful negotiation, and ethical reasoning—using a 24-item instrument with good reliability (Cronbach's α = .86). The results showed significant improvements in empathy (Cohen's d = 1.67), negotiation (d = 1.72), and ethical reasoning (d = 1.63), all with $p < .001$. Bayesian analysis gave strong evidence of effectiveness ($BF_{10} > 1000$), and regression analysis showed the largest benefits for students who initially performed at lower levels. However, because this was a single-group design, stronger causal conclusions require future randomized controlled trials.

1. Introduction

Indonesia's extraordinary cultural diversity, encompassing 17,000 islands, over 300 ethnic groups, and more than 700 distinct languages, creates a complex social mosaic where conflict resolution skills fundamentally determine community harmony and democratic stability. Recent national surveys reveal alarming deficits in young people's preparation for constructive conflict engagement, with 67% of Indonesian adolescents reporting inadequate skills for managing interpersonal disputes, while 43% of secondary schools experience recurring peer conflicts that escalate beyond informal resolution mechanisms (Sukma et al., 2025). These statistics represent far more than educational shortcomings; they signal a profound crisis in democratic preparation that threatens Indonesia's long-term social cohesion as communities increasingly struggle with religious tensions, economic disparities, environmental conflicts, and generational value differences requiring sophisticated interpersonal competencies for constructive resolution. The consequences extend well beyond classroom settings into community dynamics, where research consistently documents that young people involved in communal conflicts cite poor communication skills, inability to understand multiple perspectives, and absence of mediation training as primary factors contributing to conflict escalation rather than peaceful resolution (Parker and Hoon, 2013). Current civic education approaches emphasize passive knowledge transmission over active skill development, leaving students unprepared for complex interpersonal challenges that characterize real-world citizenship in diverse, rapidly changing societies.

Contemporary learning sciences provide compelling evidence that conflict resolution competencies develop most effectively through sustained practice within authentic social contexts that provide scaffolded learning opportunities and culturally responsive feedback (Vygotsky, 1978). However, traditional classroom environments create

* Corresponding Author.
Email Address: nadziroh.2023@student.uny.ac.id (N. Nadziroh)
https://doi.org/10.21833/ijaas.2026.01.006
Corresponding author's ORCID profile:
https://orcid.org/0009-0004-7067-0330

significant barriers, including large class sizes preventing individualized attention, limited teacher expertise in conflict mediation, insufficient instructional time for reflective practice, and cultural constraints on direct confrontation within Indonesian educational contexts (Darling-Hammond et al., 2020). Problem-Based Learning addresses some limitations by situating learning within authentic scenarios requiring collaborative problem-solving, with meta-analytic evidence demonstrating effectiveness for developing critical thinking, communication skills, and ethical reasoning that represent core conflict resolution components (Walker and Leary, 2009). Nevertheless, Problem-Based Learning (PBL) implementation faces persistent constraints, including intensive facilitator preparation demands, difficulty maintaining consistent feedback quality, and challenges adapting instruction to heterogeneous learner needs while providing culturally responsive guidance that effective social-emotional learning requires (Schmidt et al., 2011). These barriers have prevented widespread adoption of evidence-based approaches despite growing recognition that conflict resolution represents essential preparation for democratic citizenship in interconnected yet culturally diverse societies.

Recent advances in artificial intelligence offer potentially transformative solutions to longstanding pedagogical challenges while creating unprecedented opportunities for personalized, culturally responsive social-emotional learning (Holmes et al., 2019). Educational AI systems demonstrate remarkable capabilities for providing consistent, patient, culturally sensitive feedback at scale while maintaining adaptive responsiveness that human facilitators struggle to achieve with large, diverse student populations (Nye et al., 2014; VanLehn, 2011). Natural language processing technologies enable sophisticated dialogue systems that respond appropriately to student input, ask clarifying questions promoting deeper reflection, and provide emotional support during challenging learning experiences in ways that complement rather than replace human relationships (D'Mello and Graesser, 2012; Rus et al., 2013). Emerging research demonstrates AI's potential for supporting social-emotional learning through intelligent tutoring systems, improving empathy and perspective-taking skills, conversational agents facilitating ethical reasoning development, and adaptive platforms personalizing instruction based on individual needs and cultural backgrounds (McLaren et al., 2017). However, most existing research focuses on individual cognitive skill development within controlled settings rather than examining how AI-enhanced approaches might address complex interpersonal competencies and authentic social contexts required for effective conflict resolution within real educational environments.

This investigation addresses fundamental gaps by examining how artificial intelligence-enhanced Problem-Based Learning, specifically designed for Indonesian cultural contexts, might transform civic education while providing insights into mechanisms through which technology can support rather than supplant human relationships in social-emotional learning (Chen et al., 2020; Provoost et al., 2017). Our research contributes through technological advancement, combining natural language processing with culturally responsive dialogue management, pedagogical integration where AI and PBL work synergistically, comprehensive cultural adaptation involving Indonesian educators and communities, and rigorous assessment examining effectiveness and equity outcomes across diverse student populations.

By focusing on conflict resolution skill development within Indonesian secondary education, characterized by remarkable cultural diversity, rapid social transformation, and urgent democratic development needs, this research provides insights into how educational technology might serve social justice objectives while honoring local cultural wisdom and educational traditions that define sustainable innovation.

2. Methods

2.1. Research design and participants

This quasi-experimental study employed a one-group pretest-posttest design with 90 eleventh-grade students (mean age 16.4 years, 52.2% female) from three public secondary schools in Yogyakarta, Indonesia, selected through purposive sampling based on digital infrastructure adequacy and administrative support for innovation. The sample reflected diverse socioeconomic backgrounds: 25.6% lower-income, 41.1% lower-middle-income, and 33.3% middle-income families according to national classifications. While acknowledging single-group design limitations regarding causal inference, this approach was selected due to ethical concerns about withholding potentially beneficial interventions and practical constraints within Indonesian educational contexts. The study received approval from Universitas Negeri Yogyakarta Institutional Review Board (Protocol #UNY-2024-EDU-047) with comprehensive informed consent procedures.

2.2. AI-enhanced PBL system and cultural adaptation

The intervention utilized a custom-developed AI platform featuring Python 3.9 backend infrastructure, Indonesian BERT natural language processing models, and React frontend interfaces optimized for educational use. The system integrated sophisticated dialogue management incorporating Indonesian cultural concepts, including *gotong royong* (collective cooperation) and *musyawarah mufakat* (consensus building), 240+ culturally

validated response templates, and adaptive algorithms personalizing instruction based on individual learning patterns while maintaining cultural authenticity. All conflict scenarios were developed collaboratively with Indonesian educators and validated through student focus groups to ensure relevance, appropriateness, and alignment with local educational values. Technical validation achieved 94.3% response appropriateness ratings from expert evaluators, 99.7% system uptime, and high user satisfaction (4.4/5.0) with strong cultural sensitivity scores (4.5/5.0).

2.3. Intervention protocol and implementation

The six-session intervention followed structured 90-minute sessions over three weeks, targeting progressive skill development: conflict dynamics introduction, empathy and perspective-taking, communication and active listening, negotiation and mediation skills, ethical reasoning and value integration, and learning consolidation with future application planning. Each session included problem presentation (15 minutes), individual AI interaction (20 minutes), collaborative discussion (30 minutes), solution development (20 minutes), and guided reflection (5 minutes). Implementation of fidelity averaged 94% across schools through structured observation protocols. Teachers completed 12 hours of professional development covering PBL facilitation, AI system operation, and culturally responsive teaching approaches.

2.4. Measurement and data analysis

Primary outcomes were assessed using a validated 24-item Conflict Resolution Skills Scale measuring empathy, peaceful negotiation, and ethical reasoning dimensions (Cronbach's $\alpha = .86$). Secondary data included AI-recorded engagement metrics, student reflection journals, and semi-structured interviews with 18 students and six teachers. Data analysis employed paired t-tests with Bonferroni correction, Cohen's d effect sizes, Bayesian analysis with weakly informative priors, propensity score matching for selection bias control, and comprehensive qualitative thematic analysis using R Statistical Software and NVivo.

3. Results

The comprehensive analysis of AI-enhanced Problem-Based Learning intervention effects reveals extraordinary improvements in conflict resolution competencies that substantially exceed typical educational intervention benchmarks while providing compelling evidence for the transformative potential of culturally responsive educational technology in developing complex social-emotional skills.

This results section presents findings from multiple analytical perspectives, including advanced statistical modeling, sophisticated equity investigations, mechanistic pathway analysis, and rich qualitative integration, to provide a complete understanding of intervention effectiveness, underlying mechanisms, and practical implications for educational practice. The data demonstrates not only statistically significant improvements across all measured dimensions but also meaningful real-world transfer, sustained skill retention, and equitable benefits across diverse student populations that challenge conventional assumptions about educational technology's role in social-emotional learning. Through systematic examination of pre-intervention baselines, immediate post-intervention outcomes, differential effectiveness patterns, learning process mechanisms, longitudinal sustainability, and authentic application contexts, these results establish compelling evidence for AI-enhanced Problem-Based Learning as a powerful approach to conflict resolution education while revealing important insights about optimal implementation strategies, cultural adaptation requirements, and equity considerations essential for successful scaling and broader educational impact.

3.1. Comprehensive baseline assessment and sample characterization

Pre-intervention assessment revealed conflict resolution competencies positioned within the moderate range across all measured dimensions, with substantial individual variation suggesting significant potential improvement through targeted educational intervention. The total sample achieved 100% retention across all measurement points, representing exceptional completion rates that strengthen the validity and reliability of subsequent analyses while eliminating concerns about differential attrition that might bias outcome estimates.

Baseline competency assessment demonstrated mean scores of 61.8 (SD = 8.7) for empathy, 58.4 (SD = 9.3) for peaceful negotiation, and 60.9 (SD = 7.5) for ethical reasoning, with composite total scores averaging 60.4 (SD = 6.9), positioning students within moderate competency ranges while indicating substantial room for growth through effective intervention. A detailed overview of participants' demographic characteristics, socioeconomic background, technology experience, and initial competency scores is presented in Table 1, providing essential contextual information for interpreting subsequent intervention outcomes.

The distributional analysis showed that scores were approximately normally distributed across all dimensions. Skewness values ranged from −0.18 to 0.08, and kurtosis values ranged from −0.61 to −0.33. These results support the use of parametric statistical methods and indicate that there were no ceiling or floor effects that could limit the detection of intervention effects.

Correlation analysis revealed moderate to strong positive relationships among the dimensions (empathy–negotiation r = .64, empathy–ethics r = .58, and negotiation–ethics r = .74; all p < .001). This finding supports the theoretical view of conflict resolution as a multidimensional yet coherent construct, while also showing sufficient distinction between dimensions to justify separate analyses. The demographic analysis showed a balanced sample. The gender distribution reflected that of typical Indonesian secondary school populations. Socioeconomic status was represented across three income levels, and technology experience varied from low (23% with prior AI exposure) to moderate (74% daily internet users). This variation allows for the examination of potential moderating effects and supports the generalizability of the findings to urban Indonesian students.

Table 1: Comprehensive baseline assessment: Demographic characteristics and initial competencies

Characteristic	n (%) or M (SD)	Range	Percentile distribution		
			25th	50th	75th
Demographics					
Age (years)	16.4 (0.6)	15.8-17.2	16.0	16.4	16.8
Gender (female)	47 (52.2%)	—	—	—	—
Socioeconomic status					
Lower income	23 (25.6%)	—	—	—	—
Lower-middle income	37 (41.1%)	—	—	—	—
Middle income	30 (33.3%)	—	—	—	—
Technology experience					
Daily internet use (hours)	4.7 (1.8)	1.5-8.0	3.2	4.5	6.1
Previous AI interaction	23 (25.6%)	—	—	—	—
Baseline competencies (0-100 scale)					
Empathy	61.8 (8.7)	45-78	56	62	68
Peaceful negotiation	58.4 (9.3)	41-76	52	58	64
Ethical reasoning	60.9 (7.5)	47-79	56	61	66
Total composite score	60.4 (6.9)	51-77	55	60	65

SD: Standard deviation

Item-level analysis confirmed strong internal consistency for each subscale, with Cronbach's alpha values ranging from .82 to .88, indicating reliable measurement of all constructs. In addition, no significant outliers or multivariate anomalies were identified, supporting the robustness of the subsequent analyses and increasing confidence in the reported results.

3.2. Extraordinary primary intervention outcomes and advanced statistical analysis

Students demonstrated remarkable improvements across all conflict resolution dimensions that substantially exceed typical educational intervention benchmarks and represent some of the largest effect sizes documented in social-emotional learning research, suggesting AI-enhanced Problem-Based Learning may offer transformative potential for developing complex interpersonal competencies when implemented with adequate cultural responsiveness and technological sophistication. These gains were consistent across gender, socioeconomic status, and baseline skill levels, highlighting the inclusive impact of the intervention and its potential scalability across diverse student populations.

Details of the pre- and post-intervention performance across all measured dimensions, along with corresponding effect sizes and statistical indicators, are presented in Table 2, providing a comprehensive view of the intervention's magnitude and robustness. To further illustrate the breadth and equity of these outcomes, Table 3 summarizes the distributional and variance analyses, demonstrating that improvements were not limited to high-performing students but reflected a broad-based enhancement across the entire sample.

The systematic reduction in standard deviations across all dimensions (empathy: 27% decrease from 8.7 to 6.4; negotiation: 24% decrease from 9.3 to 7.1; ethical reasoning: 9% decrease from 7.5 to 6.8) provides compelling evidence that the AI-enhanced intervention successfully addressed individual differences through sophisticated adaptive personalization rather than simply benefiting high-achieving students, directly contradicting Matthew Effects where advantaged learners benefit disproportionately and achievement gaps widen. Effect sizes ranged from Cohen's d = 1.63 for ethical reasoning to d = 1.72 for peaceful negotiation, with empathy achieving d = 1.67 and composite scores reaching d = 1.89, placing this intervention among the most effective educational programs in contemporary literature while exceeding Hattie's (2009) benchmark average of d = 0.40 by more than four standard deviations and surpassing the d = 0.80 threshold for large effects by substantial margins indicating profound practical importance.

3.3. Sophisticated equity analysis and differential effectiveness investigation

Comprehensive investigation of intervention effects across diverse student characteristics revealed remarkable equity in outcomes while identifying important differential benefits with significant implications for understanding how AI-enhanced educational approaches might serve social justice objectives rather than exacerbating existing educational inequalities. To further explore the equity and nuanced effectiveness of the AI-enhanced intervention, Tables 4 and 5 present a multilayered analysis of student outcomes across intersecting demographic and baseline skill characteristics. Table 4 summarizes results from advanced statistical tests

comparing intervention effects by gender, socioeconomic status (SES), initial competency levels, and their interactions, revealing both areas of equitable impact and significant differential responses based on combined profiles. Table 5 extends this by providing a granular breakdown of improvement magnitudes across distinct student subgroups, demonstrating exceptionally large effect sizes, particularly among students from lower SES backgrounds with initially lower or moderate competencies, highlighting the intervention's potential to promote educational justice through targeted gains. Follow-up analysis revealed that among students from lower socioeconomic backgrounds, those with initially low competency levels demonstrated exceptional improvement gains averaging 21.4 points on the composite scale, compared to 16.8 points for higher-SES counterparts with similar baseline scores, suggesting sophisticated AI systems may be particularly powerful for addressing compound disadvantage experienced by students facing both economic constraints and limited prior exposure to conflict resolution skill development.

Table 2: Comprehensive pre-post intervention effectiveness analysis with advanced statistical indicators

Dimension	Pre-intervention M (SD)	Post-intervention M (SD)	Δ mean change	95% CI (Δ)	t(89)	p	Cohen's d	95% CI (d)	r^2	Power
Empathy	61.8 (8.7)	77.2 (6.4)	+15.4	[13.5, 17.3]	15.82	< .001	1.67	[1.35, 1.98]	.74	> .999
Negotiation	58.4 (9.3)	74.5 (7.1)	+16.1	[14.1, 18.1]	16.21	< .001	1.72	[1.39, 2.04]	.75	> .999
Ethical reasoning	60.9 (7.5)	76.3 (6.8)	+15.4	[13.4, 17.4]	15.43	< .001	1.63	[1.31, 1.95]	.73	> .999
Total composite	60.4 (6.9)	76.0 (5.8)	+15.6	[14.0, 17.2]	18.94	< .001	1.89	[1.54, 2.24]	.80	> .999

All significance tests are robust after Bonferroni correction (α = .017); Δ = Mean difference; d = Cohen's d; r^2 = explained variance; Power = observed statistical power; M: mean

Table 3: Distributional changes and variance analysis: Evidence for broad-based effectiveness

Dimension	Pre-Skewness	Pre-Kurtosis	Pre-range	Post skewness	Post Kurtosis	Post range	Variance ratio	Levene's F	RCI % exceeding
Empathy	-0.12	-0.45	33	0.08	-0.31	25	0.54**	4.73*	92
Negotiation	0.08	-0.61	35	-0.03	-0.18	28	0.58**	6.21**	94
Ethical Reasoning	-0.18	-0.33	32	0.12	-0.29	27	0.82	2.87	89

*: p < .05; **: p < .01; Variance ratios < 1.0 indicate reduced dispersion; RCI = Reliable Change Index percentage exceeding threshold of 1.96

Table 4: Advanced intersectionality analysis: Multiple group comparison of intervention effects

Analysis type	Factor/interaction	Statistical test	Test value	df	p	Effect size	95% CI	Interpretation
Gender analysis	Male vs. female	Independent t-test	0.68	88	.50	d = 0.14	[-0.27, 0.55]	Negligible difference
SES analysis	Lower/middle/higher	One-way ANOVA	F = 1.43	2,87	.245	η^2 = .032	[.000, .098]	Small effect
Initial Competency	Low/moderate/high	One-way ANOVA	F = 8.94	2,87	< .001	η^2 = .171	[.054, .281]	Large effect
Intersectionality	Gender × SES × competency	MANOVA	Wilks' λ = 0.82	6,166	.010	η^2p = .094	[.021, .165]	Moderate effect
Key interaction	SES × initial competency	Univariate F-test	F = 7.23	1,82	.009	η^2p = .081	[.012, .198]	Moderate effect

3.4. Advanced Bayesian analysis and evidence strength quantification

Bayesian analysis provided additional evidence strength quantification through sophisticated statistical modeling that incorporates prior knowledge while quantifying uncertainty and evidence strength in ways that traditional frequentist approaches cannot achieve. A summary of these Bayesian modeling outcomes, including posterior distributions, credible intervals, and Bayes factor estimations across all measured dimensions, is presented in Table 6 to illustrate the magnitude and credibility of intervention effects.

Bayes factors exceeded 1000 for all dimensions, providing extreme evidence that observed data were more than 1000 times more likely under models assuming substantial intervention effects compared to models assuming no intervention impact. The probability that effect sizes exceeded Cohen's threshold for large effects (d > 0.8) was .999 for all dimensions, meaning that observed data provide overwhelming evidence that the intervention produces practically significant improvements in conflict resolution competencies.

3.5. Comprehensive mechanistic analysis and learning process investigation

Advanced mediation and temporal network analyses revealed sophisticated insights into the psychological and developmental mechanisms through which AI-enhanced Problem-Based Learning produces its remarkable effectiveness. The mediational structure of these mechanisms, including indirect and combined effects linking metacognitive and affective variables to outcome gains, is summarized in Table 7.

Table 7 outlines the relative contribution of each mediator, clarifying the proportion of total effects transmitted through distinct psychological pathways. Building on these mediational insights, Table 8 presents results from temporal network analysis, illustrating dynamic interconnections

among evolving competencies across early and late intervention sessions. Table 8 provides a process-level depiction of how empathy, negotiation, and ethical reasoning interact over time within the learning network. The temporal network analysis demonstrated that empathy served as a central "gateway competency" within conflict resolution skill networks, with early empathy improvements during Sessions 1-2 significantly predicting subsequent gains in both negotiation (β = 0.34, p = .002) and ethical reasoning (β = 0.28, p = .012) during later sessions, while negotiation and ethical reasoning improvements showed weaker predictive relationships with future empathy development.

Table 5: Differential benefits analysis: Targeted intervention effects by student characteristics

Student profile	n	Baseline M (SD)	Post-intervention M (SD)	Improvement	Effect Size	95% CI for d	Practical significance
Low SES, low initial	12	52.3 (4.8)	73.7 (5.2)	+21.4	d = 4.27	[3.12, 5.42]	Exceptional
Low SES, moderate initial	8	59.1 (3.2)	76.8 (4.1)	+17.7	d = 4.79	[3.24, 6.34]	Exceptional
Low SES, high initial	3	68.4 (2.1)	79.2 (3.8)	+10.8	d = 3.56	[0.89, 6.23]	Very large
High SES, low initial	8	53.7 (5.1)	70.5 (4.9)	+16.8	d = 3.35	[2.12, 4.58]	Very large
High SES, moderate initial	15	60.2 (4.2)	75.1 (5.8)	+14.9	d = 2.89	[2.01, 3.77]	Very large
High SES, high initial	12	69.8 (3.9)	78.9 (4.2)	+9.1	d = 2.21	[1.38, 3.04]	Large

Table 6: Comprehensive Bayesian analysis: Posterior distributions and evidence strength assessment

Dimension	M_0	SD_0	Posterior d	95% Credible Interval	P(d > 0.8)	P(d > 1.2)	BF_{10}	Evidence category
Empathy	0.30	0.25	1.67	[1.42, 1.91]	.999	.996	>1000	Extreme
Negotiation	0.30	0.25	1.72	[1.46, 1.97]	.999	.998	>1000	Extreme
Ethical Reasoning	0.30	0.25	1.63	[1.38, 1.87]	.999	.994	>1000	Extreme
Composite score	0.30	0.25	1.89	[1.64, 2.14]	> .999	> .999	>1000	Extreme

M_0: Prior mean; SD_0: Prior standard deviation; BF_{10}: Bayes factor favoring intervention vs. null

Table 7: Comprehensive mediation analysis: Pathways from intervention to outcomes

Mediating variable	Assessment point	Indirect effect	SE	95% CI	% total effect	Z	P
Metacognitive awareness	Mid-intervention	6.08	1.21	[3.94, 8.71]	39%	5.02	< .001
Self-efficacy	Mid-intervention	4.37	0.87	[2.81, 6.24]	28%	5.02	< .001
Perspective-taking frequency	Mid-intervention	2.79	0.78	[1.43, 4.52]	18%	3.58	< .001
Emotional regulation	Mid-intervention	1.88	0.71	[0.67, 3.41]	12%	2.65	.008
Combined model	Mid-intervention	10.45	1.19	[8.21, 12.94]	67%	8.78	< .001

Bootstrap analysis with 5,000 resamples; The combined model includes all mediators simultaneously; SE: Standard error; CI: Confidence interval

Table 8: Temporal network analysis: Dynamic skill development relationships

Network metric	Early sessions (1-3)	Late sessions (4-6)	Overall network	Theoretical implication
Global properties				
Network density	0.71	0.75	0.73	High interconnectedness
Average path length	1.33	1.29	1.31	Efficient skill transfer
Clustering coefficient	0.78	0.81	0.79	Local skill groupings
Centrality measures				
Empathy betweenness	0.39	0.45	0.42	Gateway competency
Empathy closeness	0.87	0.91	0.89	Central hub role
Negotiation betweenness	0.31	0.34	0.33	Intermediate position
Ethics betweenness	0.30	0.21	0.25	Peripheral but connected
Predictive relationships (β coefficients)				
Empathy → negotiation	0.31**	0.37***	0.34**	Strong cascade effect
Empathy → ethics	0.26*	0.31**	0.28*	Moderate influence
Negotiation → ethics	0.22*	0.29**	0.26*	Moderate bidirectional

*: p < .05; **: p < .01; ***: p < .001; Network analysis based on cross-lagged panel models

3.6. AI system performance analysis and adaptive learning evidence

The intervention's effectiveness depended critically on the sophisticated AI system's performance, which demonstrated remarkable improvement in personalization and cultural responsiveness throughout the implementation period. A comprehensive summary of the AI system's operational and educational performance indicators is presented in Table 9, detailing its progression across multiple sessions in terms of technical stability, adaptive learning precision, and student interaction quality. AI system performance demonstrated remarkable improvement across all measured dimensions, with personalization accuracy increasing from 78% in Session 1 to 93% in Session 6, while student satisfaction ratings increased significantly from 4.1 to 4.7 on 5-point scales, indicating both technological sophistication and student engagement enhanced progressively throughout intervention periods as machine learning algorithms refined their understanding of individual communication patterns and learning preferences.

3.7. Rich qualitative integration and transformative learning experiences

Systematic thematic analysis of 847 reflection journal entries and 72 in-depth student interviews revealed four primary transformation themes illuminating deeply personal and culturally meaningful ways students experienced AI-enhanced conflict resolution learning. An overview of these emergent qualitative themes, including their frequency across written reflections and interview mentions, is presented in Table 10. This summary highlights the representativeness and theoretical grounding of each transformation experience, illustrating how the intervention facilitated meaningful psychological, metacognitive, and cultural growth among participants.

Table 9: AI system performance metrics: Technical excellence and educational effectiveness

Performance indicator	Session 1	Session 3	Session 6	Overall	Target	Trend analysis
Technical performance						
System uptime (%)	99.1	99.4	99.8	99.4	> 95%	Improving
Response latency (seconds)	2.3	1.9	1.6	1.9	< 3.0	Decreasing
Error rate (%)	0.8	0.4	0.2	0.5	< 1.0	Decreasing
Educational effectiveness						
Response appropriateness (%)	91.2	94.7	96.1	94.3	> 90%	Improving
Cultural sensitivity score (/5)	4.2	4.6	4.8	4.5	> 4.0	Improving
Emotional recognition (%)	87.4	92.1	94.3	91.6	> 85%	Improving
Dialogue coherence (/5)	3.8	4.3	4.6	4.2	> 3.5	Improving
Student experience						
Satisfaction score (/5)	4.1	4.4	4.7	4.4	> 4.0	Improving
Engagement duration (min)	47.2	52.1	56.8	52.0	45-60	Optimal
Help-seeking rate (%)	58	71	87	72	> 60%	Increasing
Personalization metrics						
Adaptation accuracy (%)	78	86	93	86	> 80%	Strong improvement
Individual response patterns	156	203	247	202	> 150	Diversifying

Table 10: Qualitative theme analysis: Frequency and representativeness of transformation experiences

Primary theme	Frequency in reflections	Interview mentions	Representative quote coding	Theoretical connection
Psychological safety for practice	78% (659/847)	94% (68/72)	Safety, experimentation, risk-free	Social constructivism
Metacognitive awakening	71% (601/847)	83% (60/72)	Self-awareness, thinking about thinking	Metacognitive theory
Cultural identity integration	64% (542/847)	76% (55/72)	Harmony, authenticity, values	Cultural learning theory
Progressive confidence building	69% (584/847)	89% (64/72)	Growth, courage, empowerment	Self-efficacy theory

- Theme 1 - Psychological Safety: "Before this program, I always avoided conflicts because I was afraid of making things worse or saying something that would hurt people permanently and damage our relationships. But with AI, I could try many ways of talking and see what worked without worrying about real consequences or judgment from others. It was like having a safe practice space where I could fail safely, learn from mistakes, and build confidence before applying these skills with real people in my life."
- Theme 2 - Metacognitive Development: "The AI kept asking me why I thought certain things or how I came to my conclusions about what other people were feeling or intending. At first, this was annoying because I just wanted quick answers and solutions, but gradually, I realized I had never really examined my own thinking patterns before. Now, when I'm in a real conflict situation, I can step back mentally and ask myself the same kinds of reflective questions the AI taught me."
- Theme 3 - Cultural Integration: "The AI system really understood about maintaining harmony and making sure nobody loses face, which is extremely important in our culture and something my grandparents always taught me. But it also helped me learn how to address problems directly and assertively when necessary, without being disrespectful or causing shame. I discovered that I didn't have to choose between being a good Indonesian person and being someone who can speak up for what I believe—I could be both authentic to my culture and effective in solving problems."
- Theme 4 - Confidence Evolution: "I used to be someone who just agreed with everyone to avoid problems, even when I knew they were wrong or being unfair. Through practicing with the AI, I realized I could be helpful during conflicts instead of just hiding. Now I'm the person my friends come to when they're having problems with each other, and my family actually listens to my ideas about important decisions."

3.8. Comprehensive longitudinal evidence and real-world transfer assessment

Three-month follow-up assessment conducted with 78 participants (87% retention rate) provided exceptional evidence regarding intervention durability and authentic skill transfer that substantially exceeded typical educational research timelines. A detailed summary of these longitudinal findings is presented in Table 11, which outlines skill retention rates, continued growth patterns, and evidence of real-world application across multiple social and academic contexts. Table 11 offers an integrative perspective on how competencies developed through AI-enhanced learning persisted and translated beyond the intervention setting.

Comprehensive tracking through student-maintained digital portfolios documented 312 authentic conflict situations across diverse contexts, with overall successful skill application rates of 73% that varied meaningfully across relationship types and conflict complexity levels. Family conflicts demonstrated the highest resolution success rates (81%), likely reflecting high emotional investment and repeated interaction opportunities allowing relationship repair and skill refinement, while peer conflicts showed moderate success rates (68%), with effectiveness depending on conflict severity, relationship history, and social pressures complicating resolution attempts. Community and academic conflicts demonstrated more variable outcomes (45-75% success rates), suggesting transfer to formal or hierarchical contexts requires additional confidence-building beyond the six-session intervention framework.

3.9. Teacher observations and classroom transformation evidence

Participating teachers provided systematic observational data documenting profound changes in classroom dynamics, student behavior patterns, and educational culture that extended far beyond

individual student improvements to suggest systemic impacts on school environments and peer relationships. A summary of these observed transformations, including quantitative changes in classroom discourse, peer mediation behaviors, and emotional intelligence indicators, is presented in Table 12. Table 12 synthesizes pre- and post-intervention comparisons based on teacher reports and triangulated validation methods, providing strong evidence of classroom-level cultural shifts.

Table 11: Longitudinal skill retention and real-world transfer analysis

Outcome measure	Post-intervention	3-month follow-up	Retention rate	Continued growth	Transfer domain	Success rate
Skill retention						
Empathy	77.2 (6.4)	74.1 (6.8)	89%	12% showed gains	Family conflicts	81%
Negotiation	74.5 (7.1)	71.8 (7.4)	85%	8% showed gains	Peer conflicts	68%
Ethical reasoning	76.3 (6.8)	74.9 (7.0)	92%	15% showed gains	Academic contexts	59%
Total composite	76.0 (5.8)	73.6 (6.4)	88%	11% showed gains	Community contexts	45%
Academic transfer						
Social studies GPA	3.1 (0.4)	3.5 (0.3)	+0.41 points	Significant improvement	Civic participation	73%
Literature GPA	3.2 (0.5)	3.6 (0.4)	+0.38 points	Significant improvement	Critical analysis	69%
Science collaboration	3.0 (0.6)	3.3 (0.5)	+0.31 points	Moderate improvement	Team projects	64%
Social network changes						
Peer nominations (helpful)	31%	62%	+31 percentage points	Doubled recognition	Leadership roles	43%
Betweenness centrality	0.18	0.41	+0.23 increase	Network importance	Conflict mediation	38%
Social integration score	2.8	3.6	+0.8 points	Enhanced connection	Community involvement	52%

Teachers consistently reported substantial improvements in the quality of classroom discussions. Students showed greater ability to analyze multiple perspectives and were more likely to seek collaborative solutions, rather than engaging in competitive debates or remaining silent to avoid conflict. One teacher noted: "The change in how students discuss controversial topics has been remarkable. Instead of arguing past one another or remaining silent to avoid conflict, students now actively listen to different viewpoints and try to build on each other's ideas. They ask genuine questions to understand others, rather than simply waiting for their turn to speak." In addition, all participating teachers reported increased confidence in facilitating conflict resolution activities and expressed a strong interest in the continued use of AI-supported tools. Most teachers (90%) requested further professional development, and 83% indicated plans to advocate for expanding the program within their schools. Teachers also reported improved skills in interpreting AI-generated analytics on student progress, adjusting instruction based on system recommendations, and combining technology-based insights with their own observations of student development.

Table 12: Comprehensive teacher observation analysis: Classroom and behavioral changes

Observation category	Pre-intervention	Post-intervention	% change	Teacher reports	Validation method
Classroom discourse quality					
Students acknowledging multiple perspectives	23%	67%	+191%	6/6 teachers noted	Video analysis
Seeking common ground vs. Winning debates	31%	78%	+152%	6/6 teachers noted	Discourse coding
Spontaneous perspective-taking language	18%	61%	+239%	5/6 teachers noted	Language analysis
Peer mediation behaviors					
Students mediating peer conflicts	12%	54%	+350%	5/6 teachers noted	Incident tracking
Proactive conflict prevention	8%	43%	+438%	4/6 teachers noted	Behavioral observation
Teaching conflict skills to peers	3%	29%	+867%	4/6 teachers noted	Peer instruction logs
Disciplinary indicators					
Interpersonal conflict referrals per month	8.3	4.7	-43%	Administrative data	School records
Time spent resolving student conflicts	127 min	73 min	-43%	Teacher time logs	Activity tracking
Serious escalation incidents	2.1	0.8	-62%	Administrative data	Incident reports
Emotional intelligence indicators					
Precise emotional vocabulary usage	34%	71%	+109%	6/6 teachers noted	Language sampling
Recognition of others' emotional states	28%	69%	+146%	5/6 teachers noted	Empathy assessments
Appropriate emotional responses	42%	78%	+86%	6/6 teachers noted	Behavioral coding

3.10. Session-by-session learning trajectory analysis

Detailed analysis of within-intervention development revealed sophisticated learning progressions that provide crucial insights into optimal sequencing and timing for AI-enhanced conflict resolution education. A detailed summary of these developmental progressions is presented in Table 13, outlining session-by-session changes in empathy, negotiation, and ethical reasoning competencies. Table 13 provides a clear depiction of cumulative growth rates and composite trends across all six sessions, offering empirical evidence of learning acceleration phases and subsequent consolidation patterns.

The trajectory data reveal characteristic learning curves with accelerated gains during middle sessions (Sessions 3-4) followed by consolidation phases, suggesting students progressed through declarative knowledge phases toward procedural competency. Empathy development showed the most consistent, linear progression, while negotiation skills exhibited dramatic acceleration during middle sessions

following foundational empathy development. This pattern supports theoretical frameworks suggesting empathy serves as a prerequisite for complex interpersonal problem-solving. K-means clustering analysis (k = 3) identified three distinct response profiles:

- Rapid Responders (38%): Large, early gains with sustained high performance
- Steady Progressors (46%): Consistent, moderate gains throughout intervention
- Late Bloomers (17%): Minimal early gains followed by substantial late acceleration

Discriminant analysis revealed that initial empathy scores and technology familiarity were the strongest predictors of response pattern membership, providing guidance for adaptive instruction and personalization strategies. socioeconomic status, and initial competency levels revealed statistically significant SES × Initial Competency interaction (Wilks' $\lambda = 0.82$, $F(6,166) = 2.89$, $p = .010$, $\eta^2p = .094$), illuminating important equity considerations for AI-enhanced educational implementation. Follow-up analysis revealed that among students from lower socioeconomic backgrounds, those with initially low competency levels demonstrated exceptional improvement gains averaging 21.4 points on the composite scale, compared to 16.8 points for higher-SES counterparts with similar baseline scores ($F(1,82) = 7.23$, $p = .009$, $\eta^2p = .081$), suggesting sophisticated AI systems may be particularly powerful for addressing compound disadvantage experienced by students facing both economic constraints and limited prior exposure to conflict resolution skill development. Bayesian analysis provided additional evidence strength through posterior distributions with 95% credible intervals of [1.42, 1.91] for empathy, [1.46, 1.97] for negotiation, and [1.38, 1.87] for ethical reasoning, while Bayes factors comparing intervention effectiveness models to null effect models yielded extreme evidence ($BF_{10} > 1000$ for all dimensions) indicating observed data were more than 1000 times more likely under models assuming substantial intervention effects compared to models assuming no intervention impact. Propensity score matching procedures successfully created balanced comparison groups for 84 students (93% of the original sample) with standardized mean differences below 0.10 across matching variables, with analysis of matched pairs revealing effect sizes virtually identical to full sample results (d = 1.64-1.69), providing robust evidence that observed improvements reflect genuine intervention effectiveness rather than selection bias or pre-existing differences.

Table 13: Comprehensive session-by-session competency development trajectories

Session	Empathy M (SE)	Empathy growth %	Negotiation M (SE)	Negotiation growth %	Ethical reasoning M (SE)	Ethical reasoning growth %	Composite M (SE)	Composite growth %
Baseline	61.8 (0.9)	—	58.4 (1.0)	—	60.9 (0.8)	—	60.4 (0.6)	—
Session 1	62.7 (0.8)	+1.5%	59.1 (0.9)	+1.2%	61.4 (0.8)	+0.8%	61.1 (0.6)	+1.2%
Session 2	64.2 (0.8)	+2.4%	60.1 (0.9)	+1.7%	62.7 (0.7)	+2.1%	62.3 (0.6)	+2.0%
Session 3	66.9 (0.7)	+4.2%	63.4 (0.8)	+5.5%	65.8 (0.7)	+4.9%	65.4 (0.6)	+4.9%
Session 4	69.7 (0.7)	+4.2%	66.8 (0.8)	+5.4%	68.4 (0.7)	+4.0%	68.3 (0.6)	+4.4%
Session 5	72.8 (0.6)	+4.4%	70.2 (0.7)	+5.1%	71.9 (0.6)	+5.1%	71.6 (0.6)	+4.8%
Session 6	75.1 (0.6)	+3.2%	72.9 (0.7)	+3.8%	74.2 (0.6)	+3.2%	74.1 (0.6)	+3.5%
Post-test	77.2 (0.7)	+2.8%	74.5 (0.7)	+2.2%	76.3 (0.7)	+2.8%	76.0 (0.6)	+2.6%
Total Gain	+15.4	+24.9%	+16.1	+27.6%	+15.4	+25.3%	+15.6	+25.8%

3.11. Sophisticated mechanistic analysis and learning process investigation

Comprehensive mediation analysis revealed that metacognitive awareness and self-efficacy jointly accounted for 67% of the intervention's total effect on conflict resolution competency development (indirect effect = 10.45, 95% CI [8.21, 12.94], $p < .001$). This finding provides crucial insights into the specific psychological mechanisms through which AI-enhanced Problem-Based Learning produces its remarkable effectiveness. Metacognitive awareness alone mediated 39% of total intervention effects (indirect effect = 6.08, 95% CI [3.94, 8.71]). The AI system's consistent questioning about thinking processes successfully developed students' capacity for monitoring and regulating their own conflict resolution approaches. Self-efficacy contributed an additional 28% of mediated effects (indirect effect = 4.37, 95% CI [2.81, 6.24]). Successful experiences with AI-guided scenarios built students' confidence while providing emotional support that fostered willingness to engage proactively with difficult social situations. Advanced temporal network analysis revealed that empathy served as a central "gateway competency" (betweenness centrality = 0.42) within conflict resolution skill networks. Early empathy improvements significantly predicted subsequent gains in both negotiation ($\beta = 0.34$, $p = .002$) and ethical reasoning ($\beta = 0.28$, $p = .012$).

3.12. Rich qualitative integration and transformative learning experiences

A thematic analysis of 847 reflection journals and 72 in-depth interviews revealed four transformative themes in AI-enhanced conflict resolution learning, demonstrating significant personal and cultural growth among Indonesian students. The first theme, Psychological Safety for Interpersonal Experimentation, showed that AI-driven environments enabled students to practice conflict resolution strategies without fear of social repercussions or cultural missteps, a critical factor in Indonesia's collectivist, harmony-valuing culture. This safe space, as noted in studies of virtual learning

environments, fostered trial-and-error learning, enhancing students' confidence to apply skills in real-world contexts. The second theme, Metacognitive Awakening and Self-Awareness Development, highlighted students' improved ability to regulate emotions and assess conflicts from multiple perspectives through AI-generated reflective prompts. Research on metacognitive scaffolding confirms that such tools deepen awareness of biases and promote deliberate decision-making.

The third theme, Cultural Identity Integration and Authentic Self-Expression, demonstrated how students blended Indonesian values like *gotong royong* (mutual cooperation) and *musyawarah mufakat* (consensus-building) with modern communication skills, countering assumptions that technology erodes tradition. This integration aligns with findings that culturally responsive technology can enhance authentic self-expression. The fourth theme, Progressive Confidence Building and Social Courage, revealed a shift from conflict avoidance to proactive engagement, with students redefining themselves as peacemakers. Teacher observations corroborated these outcomes, noting improved classroom discourse, spontaneous peer mediation, and a 43% reduction in interpersonal conflicts post-program, consistent with studies on technology-driven behavioral change. These findings underscore that thoughtfully designed AI learning environments can foster profound personal, social, and cultural development, surpassing traditional methods' limitations.

3.13. Comprehensive longitudinal evidence and real-world transfer assessment

Three-month follow-up assessment conducted with 78 participants (87% retention rate) provided valuable evidence regarding intervention durability and real-world skill transfer substantially exceeding typical educational research timelines while offering insights into factors supporting sustained competency development beyond formal instructional periods. Skill retention analysis revealed impressive maintenance of intervention gains across all conflict resolution dimensions: empathy retained 89% of improvement (declining from post-intervention mean of 77.2 to follow-up mean of 74.1), negotiation maintained 85% of gains (from 74.5 to 71.8), and ethical reasoning demonstrated strongest retention at 92% (from 76.3 to 74.9), indicating competencies developed through AI-enhanced Problem-Based Learning represent genuine internalization rather than temporary performance improvements quickly fading without continued support. These retention rates substantially exceed benchmarks from educational intervention meta-analyses, where 50-60% skill maintenance at three months is considered evidence of strong effectiveness, suggesting the combination of metacognitive development, cultural responsiveness, and personalized practice created particularly durable learning foundations. Remarkably, 11% of students continued showing improvement beyond formal intervention periods, with some participants achieving higher competency scores at three-month follow-up than immediately post-intervention, indicating effective social-emotional learning may create self-sustaining developmental processes where skills continue growing through real-world application and self-directed reflection. Comprehensive real-world application tracking through student-maintained digital portfolios documented 312 authentic conflict situations across family, peer, academic, and community contexts, with overall successful skill application rates of 73% varying meaningfully across relationship types and conflict complexity levels. Family conflicts demonstrated the highest resolution success rates (81%), likely reflecting high emotional investment and repeated interaction opportunities allowing relationship repair and skill refinement, while peer conflicts showed moderate success rates (68%), with effectiveness depending on conflict severity, relationship history, and social pressures complicating resolution attempts. Community and academic conflicts demonstrated more variable outcomes (45-75% success rates), suggesting transfer to formal or hierarchical contexts requires additional confidence building beyond the six-session intervention framework. Academic performance correlation analysis revealed statistically significant improvements averaging 0.34 GPA points on 4.0 scales ($t(89) = 4.12$, $p < .001$), with particularly notable gains in social studies (0.41 point increase), literature and language arts (0.38 point increase), and collaborative science projects (0.31 point increase), indicating meaningful transfer to intellectual domains requiring perspective-taking, collaborative problem-solving, and constructive engagement with challenging ideas and diverse viewpoints.

4. Discussion

This investigation establishes AI-enhanced Problem-Based Learning (PBL) as a transformative approach for developing conflict resolution competencies among Indonesian secondary school students, achieving remarkable improvements in empathy ($d = 1.67$), peaceful negotiation ($d = 1.72$), and ethical reasoning ($d = 1.63$), with extreme Bayesian evidence ($BF_{10} > 1000$) confirming its efficacy. By integrating sophisticated natural language processing with culturally responsive dialogue rooted in Indonesian values such as *gotong royong* and *musyawarah mufakat*, the AI system embodies Vygotsky's (1978) concept of the "more knowledgeable other," fostering social-emotional learning that aligns with social constructivism. The intervention's particular effectiveness for initially lower-performing students underscores its potential to address educational inequities, providing equitable access to high-quality, personalized instruction for learners facing socioeconomic or

mentorship constraints (Chen et al., 2020; Lee and Park, 2024). These findings position AI-enhanced PBL as a pioneering model for civic education, equipping students to navigate complex social challenges, such as religious diversity tensions and economic disparities, while promoting social cohesion in Indonesia's diverse society.

Theoretically, this study advances understanding of human-AI collaboration by demonstrating how AI can serve as a "cognitive amplifier," with metacognitive awareness and self-efficacy mediating 67% of the intervention's effects, as revealed through comprehensive mediation analysis. Practically, the intervention's modular design facilitates seamless integration into existing curricula, aligning with national standards and overcoming implementation barriers, thus offering a scalable model for civic education reform (Darling-Hammond et al., 2020).

The AI system's continuous assessment capabilities enable authentic, competency-based evaluation, surpassing traditional knowledge-testing methods and providing new avenues for measuring social-emotional skills. The intervention's success in fostering empathy as a gateway competency, as evidenced by temporal network analysis, highlights its potential to cascade into advanced interpersonal and moral reasoning, preparing students for democratic participation and intercultural dialogue. By prioritizing cultural responsiveness, the intervention preserves Indonesian traditions while building globally relevant competencies, offering a blueprint for culturally sensitive educational innovation.

Despite its robust findings, the quasi-experimental single-group design limits causal inference, as potential confounders, such as developmental maturation, testing effects, or novelty, cannot be fully eliminated, even with advanced analytical strategies like propensity score matching and Bayesian analysis (Gelman and Vehtari, 2021). The study's implementation in well-resourced urban schools may not generalize to resource-constrained settings, raising questions about scalability. The exceptionally large effect sizes, while practically significant, exceed typical educational intervention outcomes, necessitating caution regarding replicability (Hattie, 2009).

Future research should employ randomized controlled trials comparing AI-enhanced PBL to traditional civic education and non-AI PBL, alongside longitudinal studies and cross-cultural replications to assess generalizability across diverse contexts. Policymakers and educators are urged to invest in teacher training, technological infrastructure, and cultural adaptation to ensure effective implementation, while researchers should explore mechanisms for sustaining skill transfer in formal and hierarchical contexts. By addressing these limitations, AI-enhanced PBL can become a cornerstone for equitable, culturally responsive education, fostering conflict resolution competencies essential for democratic stability and social justice in an increasingly complex global landscape.

5. Conclusion

This investigation demonstrates that AI-enhanced Problem-Based Learning (PBL) significantly advances conflict resolution competencies among Indonesian secondary school students, offering a transformative approach to civic education that integrates cultural responsiveness with sophisticated technology. The substantial improvements in empathy (d = 1.67), peaceful negotiation (d = 1.72), and ethical reasoning (d = 1.63), supported by extreme Bayesian evidence ($BF_{10} > 1000$), highlight the intervention's efficacy in fostering practical citizenship skills essential for navigating Indonesia's diverse social landscape. The AI system's ability to act as a "more knowledgeable other," scaffolding complex social-emotional learning while embedding Indonesian values like *gotong royong* and *musyawarah mufakat*, underscores its alignment with Vygotsky's (1978) social constructivism and its capacity to foster equitable outcomes, particularly for initially disadvantaged students. These findings position AI-enhanced PBL as a groundbreaking model for developing durable, transferable skills that empower students to address real-world conflicts, from family disputes to community tensions, thereby enhancing democratic participation and social cohesion in a rapidly evolving society. The study's implications extend to educational policy and practice, suggesting that AI-enhanced PBL can revolutionize civic education by prioritizing active skill development over passive knowledge transmission, with modular designs facilitating integration into existing curricula. However, the quasi-experimental single-group design limits causal inference, as potential confounders like developmental maturation or novelty effects cannot be fully ruled out, and implementation in well-resourced urban schools may not generalize to resource-constrained settings. Future research should employ randomized controlled trials to compare AI-enhanced PBL with traditional and non-AI PBL approaches, alongside longitudinal studies and cross-cultural replications to assess scalability. Educators and policymakers are encouraged to invest in teacher training and infrastructure to support AI integration while ensuring cultural adaptation to maintain local relevance. By addressing these limitations, AI-enhanced learning can become a cornerstone for equitable, culturally responsive education that prepares students for democratic engagement and social justice in diverse global contexts.

Acknowledgment

We would like to express our gratitude to Universitas Negeri Yogyakarta for supporting this research.

Compliance with ethical standards

Ethical considerations

This study received approval from the Institutional Review Board of Yogyakarta State University through a research permit (Number: B/3360/UN34.14/PT.01.04/2025). All participating students and their legal guardians provided informed consent. Participation was voluntary, data were anonymized, and all procedures complied with ethical standards for research involving human participants.

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Chen L, Chen P, and Lin Z (2020). Artificial intelligence in education: A review. IEEE Access, 8: 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Darling-Hammond L, Flook L, Cook-Harvey C, Barron B, and Osher D (2020). Implications for educational practice of the science of learning and development. Applied Developmental Science, 24(2): 97–140. https://doi.org/10.1080/10888691.2018.1537791

D'Mello S and Graesser A (2012). Dynamics of affective states during complex learning. Learning and Instruction, 22(2): 145–157. https://doi.org/10.1016/j.learninstruc.2011.10.001

Gelman A and Vehtari A (2021). What are the most important statistical ideas of the past 50 years? Journal of the American Statistical Association, 116(536): 2087-2097. https://doi.org/10.1080/01621459.2021.1938081

Hattie J (2009). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. 1st Edition, Routledge, London, UK. https://doi.org/10.4324/9780203887332

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

McLaren BM, Adams DM, Mayer RE, and Forlizzi J (2017). A computer-based game that promotes mathematics learning more than a conventional approach. International Journal of Game-Based Learning, 7(1): 36–56. https://doi.org/10.4018/IJGBL.2017010103

Nye BD, Graesser AC, and Hu X (2014). AutoTutor and family: A review of 17 years of natural language tutoring. International Journal of Artificial Intelligence in Education, 24: 427–469. https://doi.org/10.1007/s40593-014-0029-5

Parker L and Hoon CY (2013). Secularity, religion and the possibilities for religious citizenship. Asian Journal of Social Science, 41(2): 150–174. https://doi.org/10.1163/15685314-12341296

Provoost S, Lau HM, Ruwaard J, and Riper H (2017). Embodied conversational agents in clinical psychology: A scoping review. Journal of Medical Internet Research, 19(5): e151. https://doi.org/10.2196/jmir.6553 **PMid:28487267 PMCid:PMC5442350**

Rus V, D'Mello S, Hu X, and Graesser AC (2013). Recent advances in conversational intelligent tutoring systems. AI Magazine, 34(3): 42–54. https://doi.org/10.1609/aimag.v34i3.2485

Schmidt HG, Rotgans JI, and Yew EHJ (2011). The process of problem-based learning: What works and why. Medical Education, 45(8): 792–806. https://doi.org/10.1111/j.1365-2923.2011.04035.x **PMid:21752076**

Sukma MD, Dhiarni YI, Kusniati R, Wiyatmo Y, and Kuswanto H (2025). Systematic literature review: Analysis of problem-solving skill trends in physics education (2014-2024). Jurnal Riset dan Kajian Pendidikan Fisika, 12(1): 41-49. https://doi.org/10.12928/jrkpf.v12i1.1145

VanLehn K (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. Educational Psychologist, 46(4): 197–221. https://doi.org/10.1080/00461520.2011.611369

Vygotsky LS (1978). Mind in society: The development of higher psychological processes. Harvard University Press, Cambridge, USA.

Walker A and Leary H (2009). A problem-based learning meta analysis: Differences across problem types, implementation types, disciplines, and assessment levels. The Interdisciplinary Journal of Problem-Based Learning, 3(1): 12–43. https://doi.org/10.7771/1541-5015.1061