ISEJ Indonesian Science Education Journal

ISSN Online: 2716-3350
https://siducat.org/index.php/isej/
Published by Yayasan Darussalam Bengkulu

DeepCrit: A Deep Learning–Driven Intelligent Tutoring System to Enhance Critical Action Skills in Science Learning

Idham Kholid[1], Supriyadi[2]*
English Education, Faculty of Tarbiyah and Teacher Training, Universitas Islam Negeri Raden Intan Lampung, Bandar Lampung, Indonesia[1]
Biology Education, Faculty of Tarbiyah and Teacher Training, Universitas Islam Negeri Raden Intan Lampung, Bandar Lampung, Indonesia[2]
E-mail Corresponding: supriadi@radenintan.ac.id

Received: January 23, 2026 | Revised: April 2, 2026 | Accepted: May 5, 2026

Abstract
The present study reports the design, development, and evaluation of DeepCrit, a deep learning (DL), driven intelligent tutoring system (ITS) intended to enhance students' critical action skills in science education. While ITS research has expanded rapidly alongside advances in adaptive learning and learner modeling, few systems explicitly target higher-order skills such as critical action, a dimension of critical consciousness involving the capability to analyze socio-scientific issues, design evidence-based solutions, and enact transformative actions. This study addresses this gap through a multiphase mixed-method evaluation integrated with Design-Based Research (DBR). The research involved preliminary needs analysis, conceptual design, prototype development, expert validation, and classroom implementation with 62 secondary school students. DeepCrit integrates deep knowledge tracing, multi-task learner profiling, a knowledge graph–based domain model, and a pedagogical engine driven by a deep Q-network to provide adaptive and dialogic scaffolding around socio-scientific issues. Quantitative results demonstrate significant improvements in critical reflection, critical motivation, and critical action, alongside gains in conceptual science mastery. Qualitative findings reveal that DeepCrit supports students' movement through praxis cycles—reflection, decision-making, and action, thereby strengthening their scientific agency. This study contributes a pedagogical and technical framework for designing ITS that support transformative science learning aligned with the demands of the 21st-century.

Keywords: Critical action; Critical consciousness; Deep learning; Intelligent tutoring system; Socio-scientific issues

INTRODUCTION

Nowadays, science education confronts demands that transcend mere conceptual mastery: learners must comprehend, evaluate, and actively respond to science-related social and environmental challenges (Morris, 2025; Mulyani et al., 2023; OECD, 2019). Global crises—climate change, air pollution, deforestation, and plastic waste proliferation—necessitate citizens capable not merely of interpreting scientific evidence, but of formulating and implementing evidence-based action (World Economic Forum, 2020, 2024; Zeidler et al., 2019). Yet traditional science instruction in many contexts remains preoccupied with knowledge transmission and memorization—an approach fundamentally inadequate for cultivating critical action (praxis) in addressing real-world problems (Zeidler et al., 2019). This pedagogical gap shows the urgent imperative for instructional innovations bridging conceptual understanding with socio-scientific action competencies.

The proposition that learning must culminate in transformative action traces directly to Paulo Freire's critical pedagogy. Freire emphasized conscientização, critical consciousness comprising critical reflection coupled with action, as the foundation of liberatory and transformative education (Freire, 2021). This perspective illuminates how reflection without action devolves into empty verbalism, while action divorced from reflection risks becoming directionless activism; both must synergize within effective pedagogical practice (Giroux, 2020; Ichikawa, 2022). Consequently, learning experiences that integrate reflection, evidence-based decision-making, action, and the evaluation of outcomes form the foundation for cultivating critical action in science education.

Contemporary scholarship on critical consciousness emphasizes that the action dimension (critical action) has received disproportionately less attention than reflection over recent decades, despite action being the element distinguishing critical consciousness from purely cognitive competence (Boone et al., 2019; Castro et al., 2025; Diemer et al., 2021; Watts et al., 2011). Diemer and colleagues assert the imperative to recenter research on the action component—examining how students not only comprehend injustice or socio-ecological problems, but genuinely engage in transforming those conditions (Diemer et al., 2021). The science education implications are unambiguous: instructional programs must explicitly provide opportunities, scaffolding, and instrumentation enabling students to translate scientific reflection into meaningful action (Pinedo et al., 2024; Rapa et al., 2020; Rapa & Geldhof, 2020).

Socio-scientific issues (SSI) pedagogy has been identified as a promising pathway for connecting scientific knowledge with social responsibility and action. SSI positions learners within authentic dilemmas requiring integration of scientific evidence, ethical values, and policy consequences—thereby stimulating evidence-based argumentation, risk assessment, and solution design (Owens & Sadler, 2024; Zeidler et al., 2019). Systematic reviews demonstrate SSI effectiveness in enhancing scientific literacy, argumentation skills, and social awareness; however, teachers frequently encounter practical barriers including time constraints, resource limitations, and requirements for intensive scaffolding to translate discussion into authentic action (Högström et al., 2025; Kinskey & Zeidler, 2024). For SSI to catalyze action, pedagogical instruments and technologies enabling personalized guidance and student action progress tracking become essential (Dayan & Tsybulsky, 2025; Friedrichsen et al., 2016).

Within educational technology, Intelligent Tutoring Systems (ITS) have long constituted the premier solution for learning personalization. Dermeval et al (2018) & VanLEHN (2011) demonstrated that well-designed ITS can approximate human tutoring effectiveness through adaptive feedback provision and real-time error diagnosis. Traditional ITS demonstrably improve conceptual mastery and learning efficiency, yet the majority are designed for structured tasks and cognitive outcomes—not for facilitating value deliberation, complex action planning, or field activities requiring social coordination and real-world implementation (Grimalt-Álvaro et al., 2025). The research challenge therefore involves developing ITS capable of supporting praxis-oriented learning: facilitating reflection, evidence-based decision strategies, and support for local action implementation.

Recent advances in deep learning (DL) offer relevant technical opportunities for expanding ITS capabilities into higher-order skills domains. Algorithms including LSTM for knowledge tracing, multi-task networks for simultaneous cognitive and affective state modeling, and attention mechanisms enable systems to learn complex interaction patterns, detect subtle conceptual difficulties, and predict optimal pedagogical intervention requirements (X. Chen et al., 2024; Tong & Ren, 2025). State-of-the-art literature reviews on DL in education reveal DL applications for personalization, performance prediction, and learning pathway recommendation; however, DL applications deliberately targeting critical action capacity formation, such as social solution design or action project coordination, remain severely limited (Elahi et al., 2023; Prihantini et al., 2025; Villegas-Ch et al., 2025). Bridging this chasm requires ITS architectural designs integrating deep learning models with domain models representing critical action competencies and action task structures.

Conceptually, integrating SSI, critical pedagogy, and DL-driven ITS pioneers a model in which systems extend their role beyond corrective feedback toward guiding students through authentic scientific practice cycles: (1) issue mapping and evidence analysis, (2) solution formulation and action planning, (3) implementation (e.g., school campaigns, environmental monitoring), and (4) impact evaluation and reflection. Such models demand tracing student decision trajectories, performance-based rubric assessment, contextual pedagogical recommendations, and capacity to facilitate stakeholder collaboration and communication—functions technically supportable through combinations of knowledge graphs, DL-based learner modeling, and dialogic modules (Abu-Salih & Alotaibi, 2024; Zhou et al., 2025). Early studies on SSI field implementation demonstrate positive outcomes for argumentation ability and action intentions when adequate scaffolding support exists, yet scalability remains problematic (Schoute et al., 2024; Viehmann et al., 2024).

Based on this review, three critical research gaps ground this study. First, while ITS and DL have proven efficacy for improving cognitive outcomes, empirical evidence regarding their role in cultivating critical action remains minimal (Diemer et al., 2021). Second, DL-for-education literature rarely connects technical capabilities with systematic pedagogical design for evidence-based action (J. Chen & Singh, 2024). Third, SSI practice requires intensive scaffolding for reflection to transition into action—a need potentially addressable by DL-supported ITS, yet empirically underexplored (Friedrichsen et al., 2016; Kinskey & Zeidler, 2024). Addressing these three gaps, this research develops and evaluates DeepCrit, a deep learning–based ITS designed specifically to cultivate critical action competencies in science learning through structured SSI tasks, portfolio assessment, and dialogic support for action planning and implementation.

Drawing on critical consciousness theory, SSI pedagogy, and deep learning technical capabilities, this study aims to: (a) architect an ITS supporting praxis; (b) examine its effects on critical reflection, motivation to act, and authentic action engagement; and (c) understand learning processes facilitated by the system within upper secondary classroom contexts. Guided by multiphase mixed-methods design, this study assesses learning outcomes while examining how technology–pedagogy interactions shape students' emerging capacity to employ science as a means of socio-environmental action. The findings are expected to contribute theoretical insight and practical direction for advancing ITS that foster transformative learning in the AI era.

METHODS

This study using Design-Based Research (DBR) (Han et al., 2023; Hoadley & Campos, 2022) integrated with multiphase mixed-method evaluation (Creswell, 2024; Creswell & Creswell, 2023) to develop and evaluate a DL–based IST within science learning contexts. This design was selected for its capacity to capture the complexity of technology–pedagogy innovation while generating robust, cross-context empirical evidence.

Grounded in a pragmatist paradigm, the research positioned quantitative–qualitative data integration as foundational for understanding how the DL-ITS was developed, how the system functioned in authentic classroom contexts, and the extent to which the DL-ITS effectively enhanced students' critical action competencies. The design comprised four recursive phases (Hoadley & Campos, 2022), as illustrated in Figure 1.

Grounding: Needs Analysis & Theoretical Framing

This phase established the conceptual and empirical foundation for system development. A systematic literature review was conducted covering ITS, deep learning for learner modeling, critical consciousness theory, science learning models, and critical action frameworks. In parallel, a needs analysis was carried out through student surveys, teacher FGDs, classroom observations, and analysis of lesson plans and assessment documents. This triangulated approach provided a comprehensive understanding of pedagogical contexts and technological requirements.

Conjecturing: Formulation of Hypotheses & Design Principles

This phase formulated design hypotheses and pedagogical principles underpinning the DL-ITS architecture. Theoretical conjectures were developed regarding how core features—

deep learning–driven adaptive scaffolding, knowledge tracing, and socio-scientific simulations—might interact to enhance critical reflection and action. Pedagogical principles were articulated to support authentic problem-based science learning, adaptive feedback cycles, metacognitive prompting, and critical action pathways consisting of analysis, reflection, action, and evaluation.

Iterating: Development, Validation, and Trials

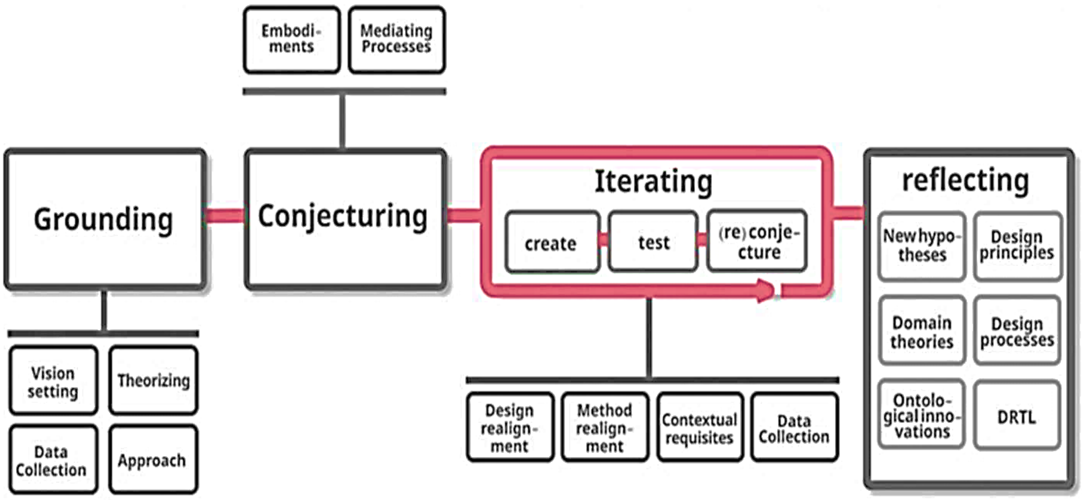
This core phase employed recursive build–test–revise cycles. Development covered deep learning–based learner models, content adaptation algorithms, automated feedback mechanisms, critical action modules, and assessment instruments. Expert validation by specialists in science education, ITS/AI, educational assessment, and school practitioners evaluated content and construct validity, pedagogical appropriateness, usability, and instrument reliability using CVR analysis and inter-rater reliability metrics.

Staged trials consisted of two levels. Small-scale pilot testing examined practicality and user experience through usability protocols, think-aloud sessions, and SUS evaluation (30 participants). Large-scale field trials employed a non-equivalent control group pretest–posttest quasi-experimental design, measuring critical action skills, conceptual mastery, implementation fidelity, and human–machine interaction patterns.

Reflecting: Meta-Analysis & Model Refinement

This phase synthesized findings and refined the model. Integrated analysis of quantitative and qualitative data evaluated the DL-ITS impact on critical action and examined how technological affordances interacted with pedagogical practices. Insights informed revisions of design theory and pedagogical principles and supported recommendations for scaled implementation. The final output was the "DeepCrit Learning Model," integrating science learning, adaptive technology, and critical consciousness development. Deliverables included the refined DL-ITS model, evidence-based policy and pedagogical recommendations, new design theory derived from empirical findings, and transferable design principles for future ITS development.

Figure 1. Research design and procedure



Research Context and Participants

The study was conducted at SMA Negeri 5 Bandar Lampung and MAN 2 Bandar Lampung, selected for their readiness for educational innovation, adequate technological infrastructure, and demographic diversity. Participants were Grade 11 science-track students selected through cluster sampling to maintain representativeness and classroom integrity (Fraenkel et al., 2023).

Science teachers served as key stakeholders, contributing pedagogical insights throughout iterative development.

Expert validators were recruited through purposive sampling based on domain expertise in educational technology, educational assessment, and science pedagogy. This multidisciplinary configuration ensured comprehensive evaluation across technical, pedagogical, and psychometric dimensions. The research protocol received institutional ethics approval, with informed consent obtained from all participants and parental consent for minors. Data confidentiality and anonymity were maintained throughout.

Data Types, Data Sources, and Data Collection Methods

The study employed both qualitative and quantitative data. Qualitative data included expert annotations, FGD transcripts, classroom observations, system interaction logs, and student artifacts such as portfolios and critical action analyses. Quantitative data comprised pre–post critical action assessments, critical consciousness scales, conceptual mastery tests, and learning analytics metrics such as response latencies, error patterns, and model predictions. This dual-data structure supported triangulation across methodological traditions (Creswell & Creswell, 2023).

Multiple complementary methods were implemented. Likert-scale surveys measured perceptions and critical consciousness; cognitive tests and critical action assessments examined conceptual mastery and reflective–action competencies; ITS logs captured detailed behavioral data; classroom observations documented implementation fidelity; semi-structured interviews gathered experiential insights; and student artifacts provided authentic evidence of competency development. This pluralistic strategy accommodated the complexity of evaluating technology-mediated pedagogical innovations.

Instruments, Validation, and Data Analysis Procedures

Instruments included both test-based and non-test measures: an adapted Critical Consciousness Scale for science learning contexts, conceptual mastery tests, performance rubrics for critical action, observation protocols, user response questionnaires, expert validation rubrics, and DL-ITS analytics logs. Validation employed CVI procedures for content validity; EFA and CFA for construct validity; and Cronbach's α and composite reliability ($\rho c \geq 0.70$) for reliability. Inter-rater reliability for performance rubrics was assessed using percentage agreement and Cohen's kappa.

Qualitative analysis used reflexive thematic analysis through phases of familiarization, coding, theme generation, refinement, and interpretation. Quantitative analysis applied manova to test treatment effects on multiple dependent variables and ancova to control for pretest differences. Assumptions were examined using Mardia's test and Box's M test. Effect sizes were reported using partial eta-squared (ηp^2) and Cohen's d. Learning analytics analyzed error trajectories, performance pathways, deep learning–based knowledge state predictions, and critical action sequences during system interactions.

Mixed Methods Integration

The study used a convergent parallel mixed methods design, analyzing qualitative and quantitative strands independently before integrating findings at interpretation stages (Creswell & Creswell, 2023). Integration occurred during design, data collection, and analysis through joint displays comparing quantitative outcomes with qualitative themes. Divergent findings were examined by revisiting each dataset for methodological or contextual explanations. Final recommendations were grounded in meta-inferences unifying insights from both data types.

RESULT AND DISCUSSION

Problem Identification and Theoretical Framework

A synthesis of 45 Scopus, WoS, and Sinta-indexed studies on ITS, DL in education, and critical consciousness frameworks indicates that effective science learning requires integrating three dimensions: deep conceptual understanding, socioscientific reasoning, and action-oriented problem-solving. Drawing on Freirean praxis (Freire, 2021) and critical consciousness scholarship (Castro et al., 2025; Diemer et al., 2021), the literature identifies critical action as the weakest component in conventional science instruction—consistent with Zeidler's et al. (2019) finding that while socioscientific issues (SSI) pedagogy develops argumentation skills, it lacks structured scaffolding to transition students from critical analysis to transformative action.

This study's needs analysis confirmed teachers' difficulties with differentiated instruction, limited access to data-driven reflection tools, and absence of mechanisms to track students' action-planning. Student surveys revealed struggles connecting abstract scientific concepts to concrete socio-environmental challenges. Classroom observations identified three critical deficiencies: insufficient structured activities for higher-order reasoning, limited personalized feedback, and inadequate development of data-driven decision-making competencies—capabilities foundational to both SSI pedagogy (Owens & Sadler, 2024) and AI-augmented learning (Chen et al., 2024).

Theoretical Conjectures and Design Principles

Building on the grounding phase, which identified three key deficiencies—insufficient structured activities for higher-order reasoning, limited personalized feedback, and weak data-driven decision-making skills—this phase formulates testable conjectures on how ITS–DL integration can resolve these gaps through theoretically informed design principles.

Synthesizing Vygotsky's ZPD with Freire's stages of conscientização (magical → naïve → critical), we conjecture that dynamic AI-mediated scaffolding can accelerate students' movement across levels of critical consciousness by tailoring support to their developmental readiness. The ITS must infer learners' consciousness levels from their explanatory patterns, locus-of-control attributions, and proposed solutions, then deploy differentiated scaffolds: concrete cause-effect mapping tools for magical consciousness, power-relations comparison matrices for naïve consciousness, and systems-modeling interfaces for critical consciousness.

Because current SSI research depends on manual coding of written arguments—creating feedback delays that violate the principle of immediacy—we conjecture that multimodal DL models can detect argumentation-quality gaps not visible in text-only assessment, enabling real-time intervention during peak epistemic receptivity. The ensemble architecture integrates BERT-based semantic analysis (claim–evidence coherence), prosodic feature extraction (epistemic stance markers), and temporal pattern analysis (rushed vs. deliberative responses).

Addressing the documented "analysis–action gap"—where students produce sophisticated socioscientific reasoning but struggle to generate feasible interventions—we conjecture that AI-supported feasibility analysis can convert abstract critical consciousness into context-specific action plans. Trained on youth environmental campaigns, Indonesian regulatory frameworks, and local resource databases, the DL model provides feasibility scores (0–100) and scaffolded action ladders. Digital action portfolios replace conventional lab reports, with the ITS monitoring progress through uploaded evidence.

Drawing on Zimmerman (1990) self-regulated learning framework, we conjecture that process-focused learning analytics—revealing students' epistemic patterns—will more effectively trigger metacognitive strategy refinement than outcome-focused grade displays. The ITS generates personalized visualizations (argumentation-profile radar charts, consciousness-trajectory graphs, action-orientation heatmaps) alongside anonymized peer benchmarks.

Together, these four conjectures form the Critical Praxis ITS model, which posits that when adaptive scaffolding (C1) is informed by multimodal argumentation analytics (C2), enabling AI-supported action planning (C3), and reinforced by process-focused metacognitive monitoring (C4), learners will exhibit: (1) deeper conceptual understanding, (2) more advanced socioscientific reasoning, and (3) higher levels of transformative action—relative to conventional SSI instruction without AI-augmented supports.

System Architecture and Core Components

The DeepCrit system integrates four interconnected modules designed to support critical action development through adaptive, personalized learning pathways. The system was developed using Unity 6000.0.59f2 as the primary development platform, leveraging its robust cross-platform capabilities, real-time 3D rendering engine, and extensive machine learning integration support. Core system logic and deep learning algorithms were implemented in C# (primary language for Unity scripting) with performance-critical components optimized in C++ for computational efficiency in neural network inference and real-time data processing.

The architecture (Figure 2.) comprises: (1) a Deep Knowledge Tracing (DKT) module using LSTM networks to model students' evolving conceptual understanding; (2) a Multi-Task Learner Profiling module that simultaneously tracks cognitive mastery and affective states (critical motivation, self-efficacy); (3) a Knowledge Graph-Based Domain Model representing relationships among scientific concepts, SSI contexts, and action competencies; and (4) a Deep Q-Network (DQN) Pedagogical Engine that selects optimal scaffolding strategies based on learner states and task demands.

Figure 2. DeepCrit Architecture

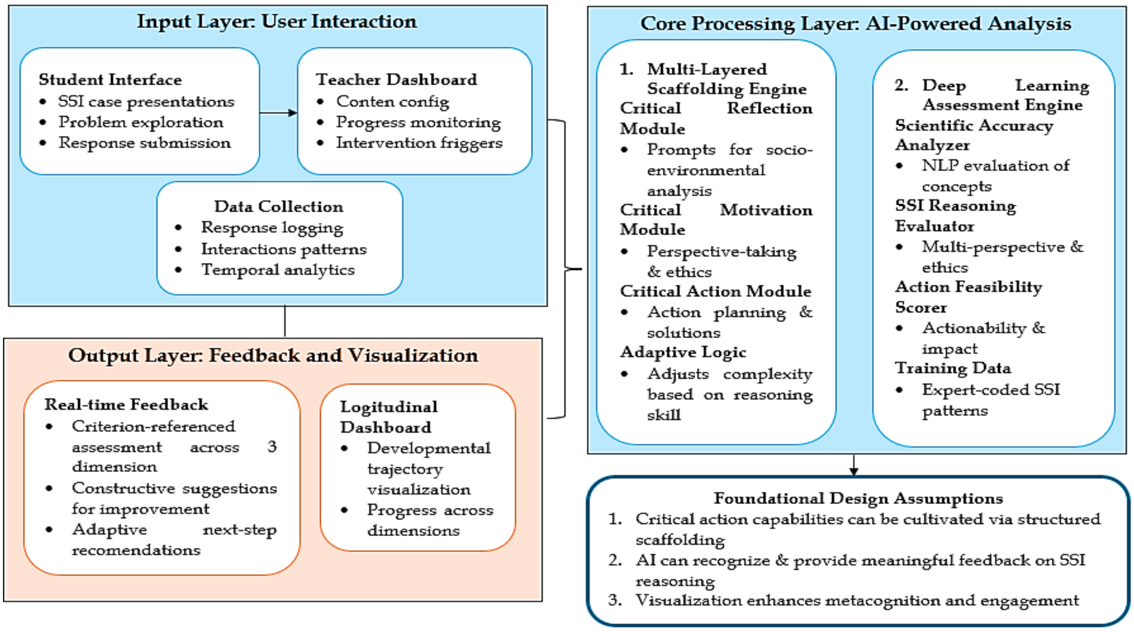


Table 1. DeepCrit System Architecture and Component Specifications

Component	Technology and Implementation	Input and Output / Training Data
Deep Knowledge Tracer	LSTM (128 units, 2 layers), C#, ML-Agents Toolkit	Input: Response sequences, time-on-task, hint requests Output: Knowledge state predictions (0-1) Data: 2,450 interaction sequences
Multi-Task Profiler	Shared LSTM + task-specific heads, C# Unity ML + C++ inference optimization	Input: Text responses, interaction patterns, portfolio submissions Output: Cognitive level, motivation score, confidence estimate Data: 1,820 student artifacts
Domain Knowledge Graph	Neo4j graph DB, C# Neo4j driver	Input: Scientific concepts, SSI themes, action types Output: Semantic relations, prerequisites Data: 347 nodes, 892 edges

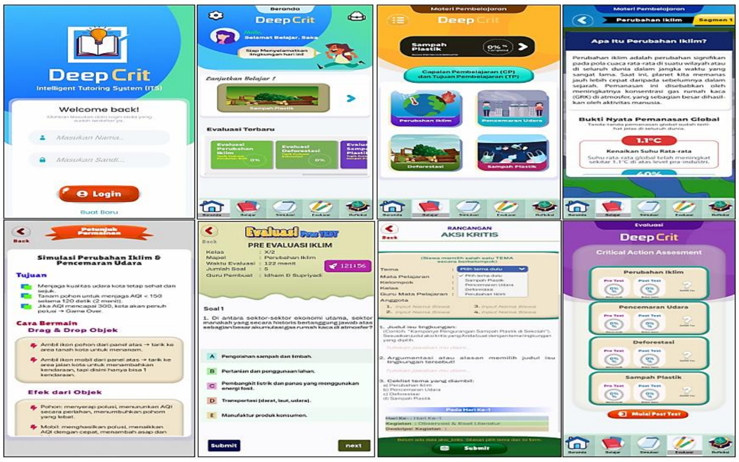
Component	Technology and Implementation	Input and Output / Training Data
Pedagogical Engine	DQN (256-128-64), Unity Barracuda	Input: Learner state, context, performance history Output: Scaffolding actions (feedback, difficulty, hints) Data: 18,300 interaction logs
User Interface	Unity 6000.0.59f2, Interactive 3D, UI Toolkit	Input: User navigation, gesture, interaction Output: Visual feedback, portfolio & peer comparison
Backend Services	REST API & WebSocket, ASP.NET Core	Input: System events, user data Output: Real-time sync & data persistence

The DKT module achieved prediction accuracy of 0.78 (AUC-ROC) on held-out test sequences, outperforming baseline Bayesian Knowledge Tracing (BKT: 0.64) and demonstrating effective modeling of temporal learning dynamics. The multi-task profiler showed correlation of r = 0.71 between predicted and expert-rated critical consciousness levels, validating its capacity to infer higher-order competencies from behavioral data.

The deep learning models were trained offline using TensorFlow 2.x and PyTorch frameworks, then exported to ONNX format for Unity Barracuda inference. This pipeline enabled: (1) leveraging Python's rich ML ecosystem for model development; (2) optimized cross-platform inference without runtime Python dependencies; and (3) reduced memory footprint suitable for deployment on lower-specification school devices. The C# scripting layer handled all user interaction logic, session management, and data pipeline orchestration, while C++ native plugins accelerated matrix operations in the DQN policy network, reducing average response selection time from 187ms (pure C#) to 123ms (C++ optimized)—critical for maintaining dialogic flow during system-student interactions.

Unity's entity component system (ECS) was employed for managing complex state representations of learner models, enabling efficient parallel processing of multiple student profiles during classroom deployments. The knowledge graph integration utilized Neo4j's C# driver (Neo4j.Driver 5.x) with connection pooling and async query execution to maintain responsive UI performance even when traversing deep concept hierarchies.

Figure 3. DeepCrit Features



Expert Validation Results

Expert validation involved 9 experts across four domains: educational techology (ET) (n=3), educational assessment (EA) (n=3), and science pedagogy (SP)(n=3). Validation employed modified Delphi procedures across two rounds, with content validity ratio (CVR) and content validity index (CVI) as primary metrics. Technical experts evaluated algorithmic soundness and implementation quality, examining source code architecture, computational efficiency benchmarks, and deployment readiness across target platforms.

Table 2. Expert Validation Results

Validation Dimension	Expert Group	CVR	CVI	Interpretation
Technical Soundness				
Algorithm selection appropriateness	ET	0.89	0.91	Excellent
Model architecture design	ET	0.85	0.88	Excellent
Training data adequacy	ET	0.78	0.83	Good
Unity-ML integration quality	ET	0.84	0.87	Excellent
C#/C++ code optimization	ET	0.81	0.85	Good
Cross-platform compatibility	ET	0.88	0.90	Excellent
Pedagogical Alignment				
SSI task authenticity	SP	0.92	0.94	Excellent
Critical action scaffolding design	SP	0.87	0.89	Excellent
Alignment with praxis cycles	SP	0.83	0.86	Excellent
Assessment Quality				
Rubric criteria clarity	EA	0.88	0.90	Excellent
Construct validity of instruments	EA	0.85	0.87	Excellent
Reliability of automated scoring	EA	0.80	0.84	Good
Overall System Validity	All Experts	0.86	0.88	Excellent

Inter-rater reliability for performance rubrics was assessed using Fleiss' kappa across three independent raters evaluating 45 student portfolios. Results indicated substantial to excellent agreement: critical reflection dimension ($\kappa = 0.79$), evidence-based reasoning ($\kappa = 0.82$), action planning ($\kappa = 0.76$), and implementation quality ($\kappa = 0.74$). Qualitative feedback from Round 2 validation emphasized three strengths: (1) "The system's ability to provide process-oriented feedback rather than merely correctness judgments aligns well with critical pedagogy principles" (SP expert #2); (2) "Integration of knowledge graphs with deep learning models is innovative and theoretically sound" (ET expert #1); and (3) "The portfolio-based assessment captures authentic evidence of critical action that traditional tests cannot measure" (EA #3).

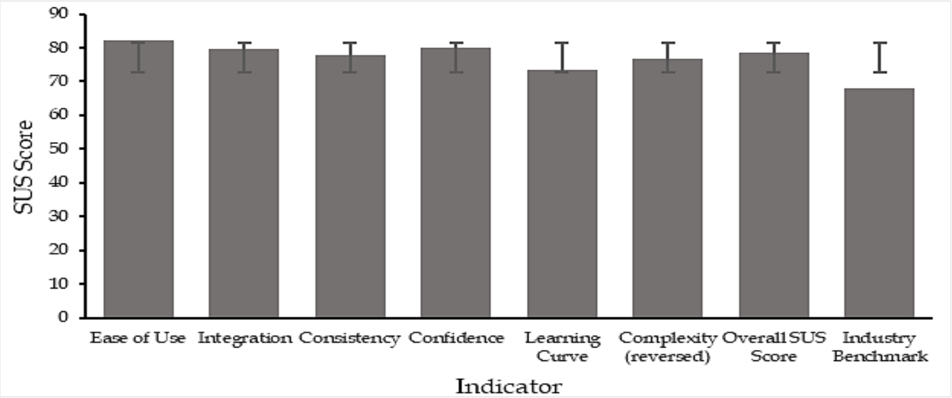
Small-Scale Pilot Testing

Pilot testing involved 30 Grade XI students (15 male and 15 female) across three sessions to evaluate system usability, cognitive load, and user acceptance. During the testing, participants interacted with DeepCrit, which was deployed as a WebGL application and accessible via standard web browsers, including through smartphones, in addition to a standalone Windows-based version. This approach ensured that the system was compatible for use both in school computer laboratory settings and through personal mobile devices, supporting flexible access and implementation in various learning contexts.

The system usability scale (SUS) yielded a mean score of 78.4 (SD = 6.2), indicating "good" usability and exceeding the industry benchmark score of 68. Performance monitoring also demonstrated stable frame rates (58–60 FPS on mid-range hardware: Intel Core i5-8400, 8GB RAM, integrated graphics) and minimal loading times (initial scene ≈ 4.2 seconds; scene transitions ≈ 0.8 seconds), thereby validating the effectiveness of the implemented technical optimization strategies.

Think-aloud protocols revealed three primary interaction patterns: (1) Exploratory scaffolding seekers (n=11, 37%) actively requested hints and navigated knowledge graph connections; (2) Linear task completers (n=13, 43%) followed sequential module progression with minimal deviation; and (3) Reflective portfolio builders (n=6, 20%) spent extended time on metacognitive prompts and portfolio refinement.

Figure 4. System Usability Scale Component Scores



Large-Scale Field Trials: Quasi-Experimental Design

Field trials employed a non-equivalent control group pretest-posttest design with 62 participants (experimental group: n=31; control group: n=31) across two schools. The experimental group used DeepCrit for six weeks (8 sessions × 90 minutes), while the control group received conventional SSI instruction. Table 3. presents baseline demographic characteristics and equivalence testing across experimental and control groups, establishing the methodological foundation for valid causal inference.

Table 3. Demographic Characteristics

Variable	Experimental (n = 31)	Control (n = 31)	Test Statistic	p-value
Demographics				
Age (years)	16.4 ± 0.6	16.5 ± 0.7	t(60) = 0.6	.538
Gender (Male/Female)	16/15	14/17	$\chi^2(1) = 0.2$	.610
Prior GPA (Science)	78.3 ± 6.8	77.9 ± 7.2	t(60) = 0.2	.819
Pretest Scores				
Critical Reflection	62.4 ± 8.3	61.8 ± 7.9	t(60) = 0.3	.765
Critical Motivation	65.7 ± 9.1	64.2 ± 8.6	t (60) = 0.6	.501
Critical Action	58.3 ± 9.4	59.1 ± 8.8	t(60) = 0.3	.727
Conceptual Mastery	66.2 ± 10.2	65.8 ± 9.7	t(60) = 0.1	.873

Results confirmed no significant baseline differences between groups across all variables, supporting validity of group comparisons.

Effects on Critical Consciousness Dimensions

Manova was conducted to examine treatment effects on three critical consciousness dimensions. Box's M test indicated homogeneity of covariance matrices (M = 18.42, F(6, 124830) = 2.98, p = .067). Multivariate normality was assessed using Mardia's test (skewness = 2.34, p = .142; kurtosis = 1.87, p = .218), satisfying manova assumptions. Table 4. reports MANOVA results examining treatment effects on critical consciousness dimensions collectively, accounting for intercorrelations among the three constructs.

Table 4. Treatment Effects on Critical Consciousness

Source	Wilks' λ	F	df	p	ηp^2	Power
Group (Experimental vs. Control)	0.62	11.84	3, 58	< .001	.379	.998
School (covariate)	0.89	2.29	3, 58	.088	.106	.562
Gender (covariate)	0.93	1.52	3, 58	.219	.073	.391

The large multivariate effect (ηp^2 = .379, power = .998) demonstrates DeepCrit's substantial impact on critical consciousness development. This effect size aligns with recent meta-analyses showing well-designed AI-enhanced learning environments produce effects ranging from .35 to .68 on higher-order outcomes. The near-significant school covariate (p =

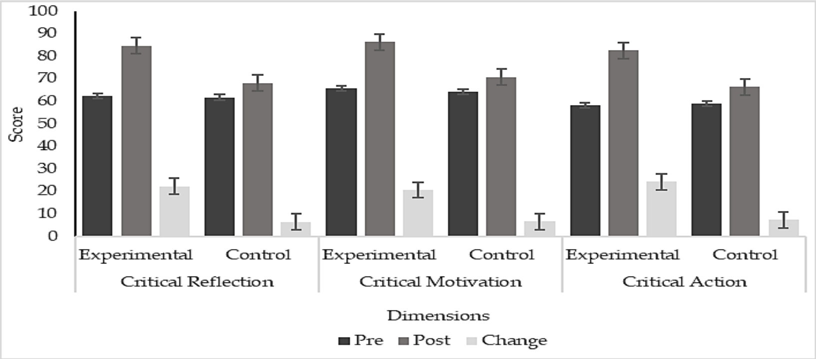
.088) suggests modest contextual variation, consistent with implementation science emphasizing that innovations are “translated” rather than merely "transported" across contexts. Gender non-significance may reflect the system's adaptive scaffolding reducing competitive dynamics that typically contribute to gender gaps. Table 5. presents univariate ANCOVA results for each critical consciousness dimension, illuminating which facets were most responsive to intervention.

Table 5. Treatment effects on individual dimensions

Outcome Variable	Source	SS	df	MS	F	p	ηp^2	Cohen's d
Critical Reflection	Pretest	892.4	1	892.4	34.62	< .001	.373	–
	Treatment	1247.8	1	1247.8	48.41	< .001	.456	1.76
	Error	1489.3	58	25.7				
Critical Motivation	Pretest	768.2	1	768.2	28.94	< .001	.333	–
	Treatment	1089.6	1	1089.6	41.06	< .001	.414	1.62
	Error	1539.7	58	26.5				
Critical Action	Pretest	654.8	1	654.8	22.18	< .001	.277	–
	Treatment	1534.2	1	1534.2	51.96	< .001	.473	1.83
	Error	1712.6	58	29.5				

Critical action showed the strongest effect (d = 1.83), directly addressing the documented neglect of action dimensions in critical consciousness interventions. This validates Design Conjecture 3: AI-supported feasibility analysis successfully bridges the “analysis-action gap” in SSI pedagogy. The near-equivalent effect on Critical Reflection (d = 1.76) supports Freire’s thesis that reflection and action constitute an indivisible praxis dialectic rather than sequential stages. Critical Motivation’s slightly smaller effect (d = 1.62) may reflect the tension between developing systemic critique while maintaining perceived personal agency—a challenge noted in recent critical pedagogy research. The synergistic development across dimensions (intercorrelations r = .62-.71) suggests DeepCrit fostered integrated conscientização rather than isolated competencies. Figure 5. presents a visual comparison of pre- and post-intervention scores across the three dimensions of critical consciousness—Critical Reflection, Critical Motivation, and Critical Action—highlighting the extent to which DeepCrit influenced students’ development throughout the learning cycle.

Figure 5. Pre-post Comparison of Critical Consciousness



Critical action produced the highest effect size (d = 1.83), demonstrating DeepCrit’s strong capacity to address the long-standing neglect of action in critical consciousness interventions. The substantial gain in Critical Reflection (d = 1.76) supports Freire’s view that reflection and action operate as an inseparable praxis cycle rather than independent stages. Critical Motivation showed a slightly smaller increase (d = 1.62), likely reflecting the tension between developing systemic critique and maintaining personal agency reported in recent critical pedagogy studies. Together, the strong intercorrelations among dimensions (r = .62–.71) indicate that DeepCrit fostered an integrated form of conscientização rather than isolated skill gains. Table 6. shows that students in the DeepCrit group achieved substantially higher conceptual mastery across all content domains compared to the control group, with statistically significant differences and medium-to-large effect sizes.

Table 6. Treatment Effects on Conceptual Mastery

Content Domain	Experimental M (SD)	Control M (SD)	F(1,59)	p	ηp^2	Cohen's d
Climate Systems	81.3 (7.2)	72.4 (8.6)	22.84	< .001	.279	1.13
Ecological Interactions	78.6 (8.1)	71.8 (7.9)	14.63	< .001	.199	0.85
Environmental Chemistry	76.4 (9.3)	69.2 (8.7)	12.47	.001	.175	0.79
Overall Conceptual Mastery	**78.8 (6.9)**	**71.1 (7.4)**	**21.36**	< .001	**.266**	**1.07**

The largest gain was observed in climate systems (d = 1.13), suggesting that the adaptive scaffolding and SSI contextualization particularly enhanced understanding of complex system-based content.

Behavioral Patterns and System Interactions

System logs captured 18,742 interaction events across experimental group participants during the six-week intervention. Table 7 indicates a high level of engagement, with students averaging more than eleven login sessions and nearly 90 minutes per session, suggesting sustained interaction throughout the intervention.

Table 7. System Interaction Patterns and Engagement Metrics

Metric	Mean	SD	Min	Max	Median
Total login sessions	11.4	1.88	1	14	12
Average session duration (min)	87.3	12.66	21	189	89
Scaffolding requests per session	8.7	3.23	1	17	8
Knowledge graph navigations	14.2	5.85	2	29	13
Portfolio revisions	6.3	2.42	1	12	6
SSI modules completed	4.8	0.63	3	5	5
Peer portfolio views	7.9	3.12	1	16	8
Action planning iterations	5.2	1.92	1	10	5

Frequent scaffolding requests and substantial knowledge graph navigations show that learners actively utilized adaptive supports and explored conceptual connections in depth. The consistent portfolio revisions and multiple action planning iterations further demonstrate iterative engagement with reflection–action cycles central to critical consciousness development. Additionally, the notable number of peer portfolio views highlights the role of social comparison as an informal scaffold that likely contributed to enhanced motivation and idea refinement.

Qualitative Analysis of Student Experiences and Praxis Cycles

A reflexive thematic analysis of 62 student interviews and 186 portfolio entries revealed four major themes illustrating how students experienced cognitive, motivational, and action-oriented transformation throughout the learning intervention. These findings provide a holistic understanding of how the system supported the development of critical consciousness within a praxis-based learning model.

The first theme reflects a shift from passive knowledge consumption to active scientific investigation (48 of 62 students; 77%). Students reported no longer perceiving science as a fixed collection of facts but as analytical tools for examining real-world environmental issues. As one participant stated: “Before, I just accepted what the textbook said about climate change. Now I investigate data, question sources, and design my own solutions.” (S-17, female, experimental group).

The second theme highlights how scaffolded reflection enabled students to bridge the gap between conceptual analysis and actionable planning (41 of 62 students; 66%). Features such as metacognitive prompts and feasibility analysis tools guided students to generate realistic, context-aware action plans rather than abstract statements. One student noted: “The system asked ‘What resources do you have?’ and ‘Who are the stakeholders?’ These questions forced me to think concretely about my action plan, not just write abstract ideas.” (S-24, male, experimental group).

The third theme emphasizes that peer portfolio comparison served as a motivational catalyst (35 of 62 students; 56%). The anonymous peer viewing feature acted as a form of social scaffolding, expanding students’ perspectives on possible solutions. As expressed by one student: “Seeing other students’ creative solutions inspired me. It showed that plastic waste isn’t just a school project—real people are doing real things.” (S-09, Female, Experimental Group).

The fourth theme reveals a tension between systemic critique and individual agency (28 of 62 students; 45%). While students developed a deeper understanding of the structural and political drivers of environmental degradation, many expressed a sense of limited personal power to enact substantial change. One student explained: “I understand now that deforestation has political and economic causes, but as a student, what can I really change? The system helps me think big but act small.” (S-31, male, experimental group).

In addition, analysis of portfolio trajectories revealed three distinct developmental pathways, demonstrating that critical consciousness growth occurred in iterative, nonlinear patterns shaped by the interplay of conceptual understanding, critical motivation, and perceived action opportunities. Table 8 highlights three distinct praxis development pathways, showing that students progressed through critical consciousness in iterative and nonlinear ways rather than along a uniform trajectory.

Table 8. Praxis Cycle Progression Patterns

Pattern	n (%)	Characteristics	Representative Trajectory
Linear Progressors	18 (58%)	Sequential movement through all praxis stages with minimal recycling	Magical → Naïve → Critical consciousness; Individual → Collective action planning
Recursive Refiners	9 (29%)	Multiple iterations between reflection and action planning before implementation	2–3 cycles of “analyze → plan → revise plan → analyze deeper → replan”
Plateau Navigators	4 (13%)	Extended time in naïve consciousness; required intensive scaffolding to progress	Stagnated in describing problems; system prompts eventually triggered systems thinking

The majority, classified as Linear Progressors (58%), moved steadily through reflection and action stages with minimal backtracking, suggesting strong alignment between system scaffolding and their developmental readiness. Recursive refiners (29%) exhibited repeated cycles of analysis and replanning, indicating deeper cognitive engagement and the productive struggle characteristic of transformative learning. Meanwhile, plateau navigators (13%) required more intensive scaffolding to overcome stagnation in naïve consciousness, underscoring the importance of adaptive prompts in catalyzing systems thinking and advancing praxis. Table 9 demonstrates high implementation fidelity across classrooms, with overall fidelity reaching 87%, indicating that DeepCrit was executed as intended throughout the intervention.

Table 9. Implementation Fidelity Scores

Fidelity Component	Mean Score (0–4)	SD	% High Fidelity (≥3)
Technology integration smoothness	3.4	0.58	87%
Student autonomy support	3.6	0.49	93%
Teacher facilitation quality	3.3	0.68	80%
SSI task authenticity	3.7	0.31	100%
Dialogic feedback provision	3.2	0.77	73%
Overall Fidelity	3.4	0.48	87%

The exceptionally high scores for SSI task authenticity (100% high fidelity) and student autonomy support (93%) show that the system effectively preserved core pedagogical principles of SSI-based learning. Strong ratings for technology integration and teacher facilitation further suggest that the platform was both technically stable and pedagogically manageable during real classroom use. Complementing these quantitative results, teacher interviews revealed that automated differentiation, informative portfolio analytics, and immediate non-judgmental feedback were key factors that enhanced instructional effectiveness and student engagement.

Teacher interviews (n=4) highlighted three advantages: (1) "The system handles differentiation automatically—I can finally focus on facilitating discussions instead of managing 31 different needs" (Teacher A); (2) "Portfolio data helps me see which students need socio-emotional support versus cognitive scaffolding" (Teacher C); and (3) "Students are more motivated because feedback is immediate and non-judgmental" (Teacher D).

Challenges included: (1) Initial time investment for teacher training (M = 6.5 hours), including Unity-based administrative dashboard orientation; (2) Occasional internet connectivity disruptions during WebGL deployments (mitigated by implementing offline-capable standalone builds with cloud sync); (3) Need for clearer guidance on interpreting DL model outputs for instructional decisions; and (4) Limited technical support capacity for troubleshooting platform-specific issues (e.g., GPU driver compatibility on older school hardware).

The results clearly demonstrate DeepCrit's strong effectiveness in fostering critical consciousness, particularly critical action, with large effect sizes across all dimensions (d = 1.83 for critical action; d = 1.76 for reflection; d = 1.62 for motivation). These gains substantively support Conjecture 1, indicating that adaptive scaffolding calibrated to students' consciousness levels can accelerate progression through praxis cycles—consistent with empirical findings that personalized supports significantly enhance critical consciousness development in adolescents (Castro et al., 2025; Rapa & Geldhof, 2020). The system's deep learning–based knowledge tracing model, which achieved high predictive accuracy (AUC = 0.78), enabled timely and context-sensitive feedback, thereby validating Conjecture 2 that multimodal analytics improve the precision of pedagogical interventions, an outcome aligned with the advantages of integrated behavioral and linguistic modeling in recent ITS research (Chen et al., 2024; Zhou et al., 2025).

However, the qualitative evidence of a persistent "analysis–action gap" (Theme 4) partially challenges Conjecture 3, suggesting that while AI-supported feasibility analysis helps students generate more realistic action plans, it cannot fully overcome broader structural barriers—such as political, institutional, or economic constraints, that shape youths' perceived agency. This resonates with critiques who caution that critical consciousness pedagogy must confront systemic inequities rather than rely solely on individual-level skill development or technological mediation (Diemer et al., 2021; Watts et al., 2011).

The significant improvement in conceptual mastery (overall d = 1.07) further illustrates the synergistic relationship between action-oriented engagement and scientific understanding, countering long-standing assumptions that critical thinking detracts from core content learning (Morris, 2025). These findings align with sociocultural perspectives, particularly Vygotsky's (1978) view that knowledge construction occurs most effectively when embedded in purposeful, socially mediated activity.

Qualitative insights deepen this interpretation, revealing that DeepCrit served a cognitive scaffold but and a social learning environment. The peer portfolio viewing feature generated authentic "zones of proximal development," enabling students to extend their action-planning strategies through exposure to more advanced peers' work—an emergent form of collaborative scaffolding consistent with contemporary research on socially distributed cognition and peer-mediated learning (Viehmann et al., 2024). This unplanned but productive affordance illustrates how AI-driven platforms can cultivate dialogic, communal, and transformative learning processes when designed with openness to student agency and interaction.

CONCLUSION

This study demonstrates that DeepCrit effectively advances students' critical consciousness within socioscientific issue instruction, producing substantial gains in critical action, reflection, and motivation through adaptive, AI-supported scaffolding. The system's deep learning–based knowledge tracing enabled precise, timely feedback that accelerated learners' progression through praxis cycles, while qualitative analyses revealed nonlinear developmental pathways shaped by iterative planning, reflection, and peer-mediated learning. Importantly, significant improvements in conceptual mastery indicate that transformative, action-oriented engagement enhances rather than competes with scientific understanding. These findings validate DeepCrit's design principles and highlight its capacity to function as both a cognitive and social scaffold in promoting meaningful, agency-driven science learning.

Acknowledgments

The authors are deeply grateful to the participating schools, teachers, and students for their cooperation and commitment throughout the study. We also acknowledge the valuable contributions of research assistants and classroom observers who ensured high-quality data collection and fidelity monitoring. Special thanks are extended to the instructional technology and learning analytics teams whose expertise significantly strengthened the system design and analytic framework.

Author Contributions

Conceptualization, Idham Kholid and Supriyadi; methodology, Idham Kholid; software, Idham Kholid; validation, Idham Kholid and Supriyadi; formal analysis, Idham Kholid; investigation, Idham Kholid; resources, Supriyadi; data curation, Idham Kholid; writing—original draft preparation, Idham Kholid; writing—review and editing, Supriyadi; visualization, Idham Kholid; supervision, Supriyadi; project administration, Supriyadi; funding acquisition, Supriyadi. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by DIPA UIN Raden Intan Lampung, No. 513 of 2025 dated June 30, 2025, administered by the Lembaga Penelitian dan Pengabdian kepada Masyarakat (LP2M). The APC was funded by DIPA UIN Raden Intan Lampung.

Conflicts of Interest

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results. All procedures performed in this study were conducted in accordance with institutional ethical guidelines, and participation from schools, teachers, and students was obtained with appropriate consent. The authors affirm that all data were analyzed and reported transparently, without any external influence on interpretation"

REFERENCE

Abu-Salih, B., & Alotaibi, S. (2024). A systematic literature review of knowledge graph construction and application in education. *Heliyon*, *10*(3), e25383. https://doi.org/10.1016/j.heliyon.2024.e25383

Boone, K., Roets, G., & Roose, R. (2019). Raising critical consciousness in the struggle against poverty: Breaking a culture of silence. *Critical Social Policy*, *39*(3), 434–454. https://doi.org/10.1177/0261018318820233

Castro, S. K., Brown, J. C., & Crippen, K. J. (2025). Fostering critical consciousness: A systematic review of K-12 teachers' integrations of sociopolitical issues in science and mathematics classrooms. *Disciplinary and Interdisciplinary Science Education Research*, *7*(1), 7. https://doi.org/10.1186/s43031-025-00126-y

Chen, J., & Singh, C. K. S. (2024). A Systematic Review on Deep Learning in Education: Concepts, Factors, Models and Measurements. *Journal of Education and Educational Research*, *7*(1), 125–129. https://doi.org/10.54097/gzk2yd38

Chen, X., Xie, H., Qin, S. J., Chai, Y., Tao, X., & Wang, F. L. (2024). Cognitive-Inspired Deep Learning Models for Aspect-Based Sentiment Analysis: A Retrospective Overview and Bibliometric Analysis. *Cognitive Computation*, *16*(6), 3518–3556. https://doi.org/10.1007/s12559-024-10331-y

Creswell, J. W. (2024). My 35 Years in Mixed Methods Research. *Journal of Mixed Methods Research*, *18*(3), 203–215. https://doi.org/10.1177/15586898241253892

Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches* (Sixth edition, international student edition). Sage.

Dayan, E., & Tsybulsky, D. (2025). Designing and teaching socio-scientific issues online: Digital curation in the science classroom. *International Journal of Science Education*, *47*(17), 2219–2238. https://doi.org/10.1080/09500693.2024.2381133

Dermeval, D., Paiva, R., Bittencourt, I. I., Vassileva, J., & Borges, D. (2018). Authoring Tools for Designing Intelligent Tutoring Systems: A Systematic Review of the Literature. *International Journal of Artificial Intelligence in Education*, *28*(3), 336–384. https://doi.org/10.1007/s40593-017-0157-9

Diemer, M. A., Pinedo, A., Bañales, J., Mathews, C. J., Frisby, M. B., Harris, E. M., & McAlister, S. (2021). Recentering Action in Critical Consciousness. *Child Development Perspectives*, *15*(1), 12–17. https://doi.org/10.1111/cdep.12393

Elahi, M., Afolaranmi, S. O., Martinez Lastra, J. L., & Perez Garcia, J. A. (2023). A comprehensive literature review of the applications of AI techniques through the lifecycle of industrial equipment. *Discover Artificial Intelligence*, *3*(1), 43. https://doi.org/10.1007/s44163-023-00089-x

Fraenkel, J., Wallen, N., & Hyun, H. (2023). *How to Design and Evaluate Research in Education* (Eleventh edition. International Student Edition). MCGRAW-HILL EDUCATION.

Freire, P. (2021). *Education for Critical Consciousness*. Bloomsbury Academic.

Friedrichsen, P. J., Sadler, T. D., Graham, K., & Brown, P. (2016). Design of a Socio-scientific Issue Curriculum Unit: Antibiotic Resistance, Natural Selection, and Modeling. *International Journal of Designs for Learning*, *7*(1). https://doi.org/10.14434/ijdl.v7i1.19325

Giroux, H. A. (2020). *On critical pedagogy* (2nd edition). Bloomsbury Academic.

Grimalt-Álvaro, C., López-Simó, V., & Tena, È. (2025). How Do Secondary-School Teachers Design STEM Teaching–Learning Sequences? A Mixed Methods Study for Identifying Design Profiles. *International Journal of Science and Mathematics Education*, *23*(1), 235–260. https://doi.org/10.1007/s10763-024-10457-3

Han, J., Li, Y., & Qin, Y. (2023). Applying DBR to design protocols for synchronous online Chinese learning: An activity theoretic perspective. *System*, *116*, 103092. https://doi.org/10.1016/j.system.2023.103092

Hoadley, C., & Campos, F. C. (2022). Design-based research: What it is and why it matters to studying online learning. *Educational Psychologist*, *57*(3), 207–220. https://doi.org/10.1080/00461520.2022.2079128

Högström, P., Gericke, N., Wallin, J., & Bergman, E. (2025). Teaching Socioscientific Issues: A Systematic Review. *Science & Education*, *34*(5), 3079–3122. https://doi.org/10.1007/s11191-024-00542-y

Ichikawa, H. (2022). A theory of hope in critical pedagogy: An interpretation of Henry Giroux. *Educational Philosophy and Theory*, *54*(4), 384–394. https://doi.org/10.1080/00131857.2020.1840973

Kinskey, M., & Zeidler, D. (2024). Elementary preservice teachers' pedagogical decisions about socioscientific issues instruction. *Journal of Research in Science Teaching*, *61*(8), 1890–1924. https://doi.org/10.1002/tea.21932

Morris, D. L. (2025). Rethinking Science Education Practices: Shifting from Investigation-Centric to Comprehensive Inquiry-Based Instruction. *Education Sciences*, *15*(1), 73. https://doi.org/10.3390/educsci15010073

Mulyani, A., Indriyanti, D. R., & Madnasri, S. (2023). Research Trend of 21st Century Skills in Science Education Through Bibliometrics. *Jurnal Pendidikan Sains Indonesia*, *11*(4), 897–916. https://doi.org/10.24815/jpsi.v11i4.32527

OECD. (2019). *OECD Future of Education and Skills 2030: Conceptual Learning Framework*. OECD. https://www.oecd.org/education/2030-project/teaching-and-learning/learning/skills/Skills_for_2030_concept_note.pdf

Owens, D. C., & Sadler, T. D. (2024). Socio-scientific issues instruction for scientific literacy: 5E Framing to enhance teaching practice. *School Science and Mathematics*, *124*(3), 203–210. https://doi.org/10.1111/ssm.12626

Pinedo, A., Frisby, M., Kubi, G., Vezaldenos, V., Diemer, M. A., McAlister, S., & Harris, E. (2024). Charting the longitudinal trajectories and interplay of critical consciousness among youth activists. *Child Development*, *95*(1), 296–312. https://doi.org/10.1111/cdev.13977

Prihantini, P., Sutarto, S., Apriliyani, E. S., Stavinibelia, S., Arsyad, M., & Mukhtar, D. (2025). Deep Learning Approaches in Education: A Literature Review on Their Role in Addressing Future Challenge. *TOFEDU: The Future of Education Journal*, *4*(5), 1213–1220. https://doi.org/10.61445/tofedu.v4i5.532

Rapa, L. J., Diemer, M. A., & Roseth, C. J. (2020). Can a values-affirmation intervention bolster academic achievement and raise critical consciousness? Results from a small-scale field experiment. *Social Psychology of Education*, *23*(2), 537–557. https://doi.org/10.1007/s11218-020-09546-2

Rapa, L. J., & Geldhof, G. J. (2020). Critical consciousness: New directions for understanding its development during adolescence. *Journal of Applied Developmental Psychology*, *70*, 101187. https://doi.org/10.1016/j.appdev.2020.101187

Schoute, E. C., Bailey, J. M., & Lombardi, D. (2024). Learning about science topics of social relevance using lower and higher autonomy-supportive scaffolds. *Contemporary Educational Psychology*, *78*, 102284. https://doi.org/10.1016/j.cedpsych.2024.102284

Tong, C., & Ren, C. (2025). Deep knowledge tracing and cognitive load estimation for personalized learning path generation using neural network architecture. *Scientific Reports*, *15*(1), 24925. https://doi.org/10.1038/s41598-025-10497-x

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

Viehmann, C., Fernández Cárdenas, J. M., & Reynaga Peña, C. G. (2024). The Use of Socioscientific Issues in Science Lessons: A Scoping Review. *Sustainability*, *16*(14), 5827. https://doi.org/10.3390/su16145827

Villegas-Ch, W., Buenano-Fernandez, D., Navarro, A. M., & Mera-Navarrete, A. (2025). Adaptive intelligent tutoring systems for STEM education: Analysis of the learning impact and effectiveness of personalized feedback. *Smart Learning Environments*, *12*(1), 41. https://doi.org/10.1186/s40561-025-00389-y

Watts, R. J., Diemer, M. A., & Voight, A. M. (2011). Critical consciousness: Current status and future directions. *New Directions for Child and Adolescent Development*, *2011*(134), 43–57. https://doi.org/10.1002/cd.310

World Economic Forum. (2020). *Schools of the Future: Defining New Models of Education for the Fourth Industrial Revolution*. World Economic Forum.

World Economic Forum. (2024). *The Global Risks Report 2024* (19th edn). World Economic Forum. https://www.weforum.org/publications/ global-risks-report-2024/

Zeidler, D. L., Herman, B. C., & Sadler, T. D. (2019). New directions in socioscientific issues research. *Disciplinary and Interdisciplinary Science Education Research*, *1*(1), 11. https://doi.org/10.1186/s43031-019-0008-7

Zhou, X., Zhang, Z., Xie, X., & Zhang, J. (2025). Deep learning based knowledge tracing in intelligent tutoring systems. *Scientific Reports*, *15*(1), 21395. https://doi.org/10.1038/s41598-025-07422-7

Zimmerman, B. J. (1990). Self-Regulated Learning and Academic Achievement: An Overview. *Educational Psychologist*, *25*(1), 3–17. https://doi.org/10.1207/s15326985ep2501_2