Jurnal Pendidikan Fisika
Vol. 15 No. 1, Juni 2026_109-124
DOI: 10.24114/e8c94q98
p-ISSN: 2552-732X
e-ISSN: 2301-7651

Review Article

OPEN ACCESS

A Systematic Literature Review of Critical Thinking Skills in Physics Learning: Interventions, Assessment and Outcomes

[1]Khoirul Ikhsan Pane*, [2]Supahar Supahar, [2]Heru Kuswanto, [1]Utami Putri, [1]Yuliyanti Daulay

[1] Magister of Physics Education, Universitas Negeri Yogyakarta, Yogyakarta, 55281, Indonesia
[2] Physics Education Department, Universitas Negeri Yogyakarta, Yogyakarta, 55281, Indonesia

ARTICLE INFO

Article History:
Submitted: 12-04-2026
Revised : 22-05-2026
Accepted : 26-06-2026
Published: 30-06-2026

Keywords:
Critical Thinking Skills;
Physics Learning;
Systematic Literature Review;
Instructional Intervention;
PRISMA;

Kata Kunci:
Keterampilan Berpikir Kritis;
Pembelajaran Fisika;
Tinjauan Literatur Sistematis;
Intervensi Intruksional;
PRISMA;

ABSTRACT

Critical thinking is a key competency in 21st-century education, yet a comprehensive synthesis of its development in physics learning remains lacking. This systematic review synthesizes empirical evidence on instructional interventions, assessment instruments, and outcomes related to critical thinking in physics learning (January 2017 – February 2026). Following PRISMA guidelines, three databases were searched (Scopus, ERIC, SINTA), yielding 491 records. After screening and eligibility assessment, 31 empirical studies were included, limited to open-access, quantitative or mixed-methods journal articles or conference papers in English or Indonesian, focusing on physics and critical thinking as the dependent variable (outcome). The open-access restriction was adopted due to institutional database constraints; this limitation is explicitly acknowledged. Inquiry-based learning (32.26%), problem-based learning (22.58%), and project-based learning (16.13%) were the most common interventions. Dominant assessment tools included the Watson-Glaser Critical Thinking Appraisal and Facione-based instruments. Effect sizes ranged from moderate to large, with inquiry-based approaches yielding the highest gains. Active learning interventions particularly inquiry and problem-based learning effectively enhance critical thinking in physics. Standardized assessment frameworks and future research on technology integration and diverse cultural contexts are recommended.

ABSTRAK

Berpikir kritis merupakan kompetensi utama dalam pendidikan abad ke-21, namun sintesis komprehensif mengenai pengembangannya dalam pembelajaran fisika masih sangat terbatas. Tinjauan sistematis ini menyintesis bukti empiris mengenai intervensi pembelajaran, instrumen penilaian, dan hasil yang berkaitan dengan berpikir kritis dalam pembelajaran fisika (Januari 2017 – Februari 2026). Mengikuti panduan PRISMA, pencarian dilakukan pada tiga basis data (Scopus, ERIC, SINTA) dan menghasilkan 491 artikel. Setelah proses skrining dan penilaian kelayakan, 31 studi empiris diikutsertakan, yang terbatas pada artikel jurnal atau prosiding konferensi akses terbuka, kuantitatif atau metode campuran, dalam bahasa Inggris atau Indonesia, yang berfokus pada fisika dengan berpikir kritis sebagai variabel dependen (hasil). Pembatasan akses terbuka diterapkan karena keterbatasan akses basis data institusional dan diakui secara eksplisit sebagai keterbatasan penelitian. Pembelajaran berbasis inkuiri (32,26%), pembelajaran berbasis masalah (22,58%), dan pembelajaran berbasis proyek (16,13%) merupakan intervensi yang paling umum digunakan. Instrumen penilaian yang dominan meliputi Watson-Glaser Critical Thinking Appraisal dan instrumen berbasis Facione. Ukuran efek berkisar dari sedang hingga besar, dengan pendekatan berbasis inkuiri menghasilkan peningkatan tertinggi. Intervensi pembelajaran aktif khususnya pembelajaran berbasis inkuiri dan berbasis masalah terbukti efektif meningkatkan kemampuan berpikir kritis dalam fisika. Kerangka penilaian yang terstandarisasi serta penelitian lanjutan mengenai integrasi teknologi dan keberagaman konteks budaya sangat direkomendasikan.

*Corresponding Author
E-mail : khoirulikhsan.2025@student.uny.ac.id

INTRODUCTION

The ability to think critically has become one of the most demanded competencies in contemporary education. International frameworks such as the OECD Learning Compass 2030 and UNESCO's Education for Sustainable Development agenda consistently position critical thinking alongside collaboration, communication, and creativity as non-negotiable skills for navigating an increasingly complex world (OECD, 2019; Care et al., 2022). In the context of accelerating technological change and the global proliferation of misinformation, the capacity to analyze evidence, evaluate arguments, and draw reasoned inferences has grown more urgent than ever (Tanveer & Shah, 2025; Nguyen et al., 2025).

Physics, as a discipline grounded in empirical observation, mathematical reasoning, and model-based explanation, offers a structurally rich environment for cultivating these cognitive capacities. The epistemological practices inherent in physics forming hypotheses, designing and conducting experiments, interpreting quantitative data, and reconciling theory with evidence map naturally onto the core cognitive operations of critical thinking as defined by Facione (2023): interpretation, analysis, evaluation, inference, explanation, and self-regulation. Several recent studies have affirmed this alignment, demonstrating that students who engage deeply with physics inquiry tasks exhibit significantly stronger critical thinking profiles than those in conventional instruction (Ertikanto et al., 2021; Sari et al., 2023; Prasetyo & Nugroho, 2025).

Despite this theoretical coherence, empirical evidence consistently reveals that students across secondary and higher education levels struggle to demonstrate adequate critical thinking in physics contexts. Many learners default to algorithmic, surface-level problem-solving strategies rather than engaging in the deeper conceptual reasoning that characterizes genuine physical understanding (Ibrahim et al., 2022; Puspitasari et al., 2022). This gap between the epistemic potential of physics learning and its realized outcomes in practice underscores the urgency of identifying and scaling effective instructional interventions.

Over the past five years, a growing body of empirical research has examined the role of active learning pedagogies in bridging this gap. Inquiry-based learning (IBL), problem-based learning (PBL), project-based learning (PjBL), and STEM-integrated approaches have all been applied in physics classrooms with the explicit aim of fostering critical thinking (Khoiri, 2023; Sumarni et al., 2024; Fadillah & Saputra, 2024). The integration of digital technologies including augmented reality, flipped classroom platforms, and most recently artificial intelligence has further expanded the pedagogical landscape, with early evidence suggesting that technology-enhanced active learning yields particularly strong critical thinking outcomes (Wang & Tsai, 2023; Prasetyo & Nugroho, 2025). However, the heterogeneity of study designs, population characteristics, measurement instruments, and reported outcomes across this literature makes it difficult to draw consolidated, actionable conclusions.

Assessment of critical thinking in physics education remains equally fragmented. Standardized instruments such as the Watson-Glaser Critical Thinking Appraisal (WGCTA) and the California Critical Thinking Skills Test (CCTST) coexist with a large number of researcher-developed instruments aligned with frameworks by Facione (2023), Ennis (2011), and Paul and Elder (2020). The absence of a consensus measurement framework not only limits cross-study comparability but also raises questions about construct validity and the extent to which different instruments capture the same underlying competency (Ariesta & Subali, 2024; Kim & Taber, 2022). This measurement heterogeneity is particularly consequential in physics, where domain-specific cognitive demands such as quantitative reasoning and experimental design evaluation may not be fully captured by general critical thinking instruments.

While several systematic reviews and meta-analyses have examined critical thinking in science education broadly (e.g., Abrami et al., 2015), dedicated systematic reviews focused specifically on physics education remain scarce, and those that exist often lack PRISMA compliance or restrict their scope to a single instructional approach. This gap is significant:

physics presents unique pedagogical challenges, including the need to bridge abstract theoretical constructs and concrete laboratory phenomena, that may differentially moderate the effectiveness of critical thinking interventions relative to other science disciplines (Simbolon & Tapilouw, 2021; Rahmat & Hamid, 2023).

The present study addresses this gap through a PRISMA 2020 compliant systematic literature review of empirical studies on critical thinking in physics learning published between January 2017 and February 2026. Specifically, the review aims to: (1) identify the types of instructional interventions used to develop critical thinking in physics learning; (2) characterize the assessment instruments and theoretical frameworks employed; (3) synthesize reported outcomes and effect sizes; and (4) identify patterns, moderating factors, and directions for future research. By providing a rigorous and comprehensive synthesis of the current evidence base, this review seeks to inform both research priorities and instructional practice in physics education globally.

Theoretical Framework

Critical thinking is a multidimensional construct with numerous theoretical formulations. For the purposes of this review, the following frameworks are considered foundational:

Facione's Delphi Report (Facione, 1990) conceptualizes critical thinking through six core cognitive skills: interpretation, analysis, evaluation, inference, explanation, and self-regulation. This framework has been widely adopted in educational research and subsequently updated (Facione, 2011; 2023), and is operationalized in numerous assessment instruments used in physics education studies included in this review.

Ennis's (2011) taxonomy offers a comprehensive list of critical thinking dispositions and abilities, including focusing on a question, analyzing arguments, asking clarifying questions, judging the credibility of sources, observing and judging observation reports, and making value judgments. Ennis's framework is particularly relevant for physics education, where evaluating experimental evidence is central.

Paul and Elder's (2020) model emphasizes the intellectual standards of critical thinking, including clarity, accuracy, precision, relevance, depth, breadth, logic, significance, and fairness. This framework is frequently used to design physics assessment rubrics.

Bloom's Revised Taxonomy (Anderson & Krathwohl, 2001) provides a hierarchical model of cognitive processes. Critical thinking in physics is most closely associated with the higher-order levels of analysis, evaluation, and creation, which correspond to the ability to break down phenomena, judge explanations, and design experiments.

METHODS

This systematic literature review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines (Page et al., 2021). The review protocol was designed prior to the search process, establishing clear objectives, inclusion/exclusion criteria, data extraction procedures, and analysis strategies.

Research Questions

This review sought to answer the following research questions:

RQ1: What instructional interventions have been used to develop critical thinking skills in physics learning between January 2017 and February 2026?

RQ2: What assessment instruments and theoretical frameworks have been employed to measure critical thinking in physics learning?

RQ3: What are the reported outcomes and effect sizes of critical thinking interventions in physics learning?

RQ4: What patterns, moderating factors, and research gaps are evident across the included studies?

Search Strategy

A systematic search was conducted across three databases: Scopus, ERIC, and SINTA (Science and Technology Index, Indonesia). The search was limited to records published between January 2017 and February 2026. Article searches were carried out in the Scopus, ERIC and SINTA databases using keyword

techniques adapted to the specific characteristic of each database. The article search method is detailed in Table 1.

Table 1. The article search method utilized database such as Scopus, ERIC and SINTA

Database	Search Term	Adjustments
Scopus	Article title: "Critical thinking" AND Article title, abstract, keyword: "critical thinking skill" AND "physics" AND "learning"	
ERIC	"Critical thinking" AND ("physics learning" OR "physics")	Peer reviewed only
SINTA	allintitle: "critical thinking" AND ("physics learning" OR "physics education" OR "physics instruction")	Search Google Scholar document

The initial search yielded 203 records from Scopus, 161 from ERIC, and 127 from SINTA, totaling 491 records. Following deduplication (n = 36 removed), 455 unique records proceeded to screening.

Inclusion and Exclusion Criteria

The following criteria were applied in a two-stage screening process in Table 2.

Justification for Open-Access Restriction

The decision to restrict inclusion to open-access (OA) publications requires explicit methodological justification, as this represents a known potential source of selection bias in systematic reviews (Higgins et al., 2019). The OA restriction was adopted for two principal reasons. First, the review team did not have full institutional subscription access to all subscription-based journals from relevant international publishers (e.g., Elsevier, Springer, Wiley, Taylor & Francis) for the entire scope of the review period. Second, the inclusion of SINTA (Science and Technology Index, Indonesia) as a primary database was specifically intended to ensure comprehensive representation of Indonesian-context studies, the majority of which are published in OA venues.

Nevertheless, the reviewers acknowledge that 72 articles were excluded solely on the basis of non-OA status. This introduces the risk that high-quality studies published in subscription journals may be systematically underrepresented in the review's findings, potentially inflating the apparent dominance of OA-indexed research traditions. This limitation is explicitly addressed in the Limitations section of this article. Future replications of this review are strongly encouraged to expand eligibility to subscription-based publications accessible via institutional repositories, to test whether the pattern of findings is robust across a broader literature base.

Screening Process

Stage 1 screening (title, abstract, and keywords) was conducted independently by two reviewers, excluding 118 records on the basis of publication date, access status, document type, and language criteria. Subsequently, 201 records were excluded that did not meet the inclusion criteria based on topical relevance. The remaining 126 full-text articles were retrieved and assessed for eligibility. After applying the full-text inclusion/exclusion criteria, 95 articles were further excluded due to reasons including: not empirical (n = 31), not focused on physics (n = 29), qualitative design only (n = 19), and no critical thinking measurement as a primary dependent variable / outcome (n = 16). A final set of 31 articles was included in the review.

Data Extraction

A standardized data extraction form was developed and piloted on five randomly selected articles. The following information was extracted: author(s) and year, country of study, educational level, sample size, instructional intervention, research design, assessment instrument, critical thinking framework, key findings, and effect size (where reported). Inter-rater reliability was assessed using Cohen's kappa ($\kappa = 0.87$), indicating strong agreement.

Table 2. Inclusion and Exclusion Criteria

Type of criterion	Criteria	Inclusion	Exclusion
Type of publication	Journal articles	X	
	Conference paper	X	
	Reports		X
	Dissertations		X
	Books and book chapters		X
Open access	Open access	X	
Publication period	January 2017-February 2026	X	
Language	English	X	
	Indonesian	X	
	Other		X
Place of study	Worldwide	X	
Type of study	Empirical investigations	X	
	Literature reviews		X
	Theoretical reviews		X
Research methods	Qualitative		X
	Quantitative	X	
	Mixed methods	X	
Focus on the school subject	Physics	X	
	Other		X
Dependent variable/Outcome	Critical thinking skill	X	

Quality Assessment

Included studies were appraised for methodological quality using the Mixed Methods Appraisal Tool (MMAT) version 2018. Studies scoring below 60% were retained but noted with caveats. No studies were excluded solely on quality grounds to avoid reporting bias. Quality scores ranged from 60% to 100%, with a mean of 79.4%.

Data Synthesis

Given the heterogeneity of study designs, populations, and measurement instruments, a narrative synthesis was employed as the primary analytical approach, supplemented by descriptive quantitative analysis where sufficient data were available. It is important to note that this review constitutes a systematic literature review (SLR) with narrative synthesis, not a formal meta-analysis. Accordingly, effect sizes reported in Table 8 are presented as ranges (minimum–maximum Cohen's d) and descriptive arithmetic means across intervention categories for illustrative purposes only. These values were not pooled using inverse-variance weighting or fixed/random-effects meta-analytic models, and therefore should not be interpreted as statistically combined estimates. Where four or more studies reported effect sizes using Cohen's d, values were collated and the range was visualized. Thematic coding was conducted on qualitative findings using NVivo 12.

The research method uses a literature review with a modified procedure from a systematic review and meta-analysis (PRISMA) adapted from (Rasyid et al., 2025). The procedure consists of identification, screening, eligibility and included, as presented in Figure 1.

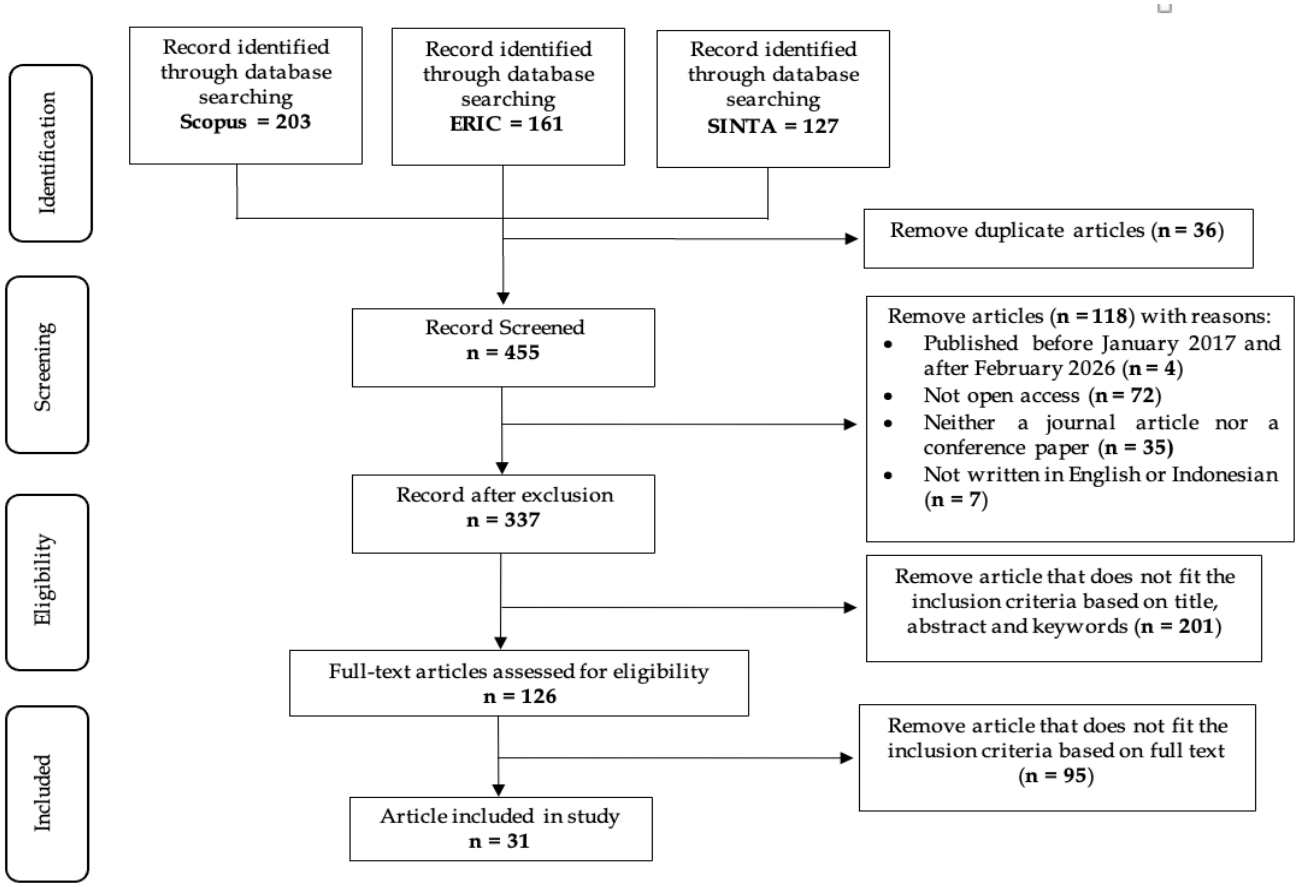

Figure 1. Article selection procedure

RESULTS AND DISCUSSION

Overview of Included Studies

The 31 studies included in this review were published between 2017 and 2026, with a notable increase in publications from 2019 onward, coinciding with growing interest in 21st-century skills. Before presenting the full synthesis, it is important to note that all 31 included studies were verified against the inclusion and exclusion criteria (Table 2), including confirmation of publication year (January 2017–February 2026), open-access status, language, study design (quantitative or mixed-methods), focus on physics, and use of critical thinking as the primary dependent variable/outcome. Three studies initially identified in early screening were replaced during full-text eligibility assessment due to: (a) publication years falling outside the eligible window upon closer inspection, or (b) critical thinking not functioning as the primary outcome variable. The final 31 studies presented in Table 3 represent only those fully meeting all inclusion criteria. Table 3 presents the complete list of included studies with their key characteristics.

Table 3. Characteristics of the 31 Included Studies

No.	Author(s) & Year	Country	Level	Design	Intervention	Instrument
1	Tiruneh, De Cock & Elen (2018)	Ethiopia	Secondary	Quasi-exp.	IBL	WGCTA
2	Soros, Ponkham & Ekkapim (2018)	Thailand	Secondary	Quasi-exp.	STEM	Facione-based
3	Jatmiko et al. (2018)	Indonesia	Higher	Quasi-exp.	PBL	Facione-based
4	Mundilarto & Ismoyo (2017)	Indonesia	Secondary	Quasi-exp.	PBL	Ennis-based

No.	Author(s) & Year	Country	Level	Design	Intervention	Instrument
5	Pahrudin et al. (2021)	Indonesia	Secondary	Quasi-exp.	STEM	Facione-based
6	Wartono et al. (2018)	Indonesia	Secondary	Quasi-exp.	Inquiry	Facione-based
7	Serevina et al. (2020)	Indonesia	Secondary	Quasi-exp.	PBL	Facione-based
8	Yuliati & Fisla (2020)	Indonesia	Secondary	Mixed	IBL	Facione-based
9	Widowati et al. (2020)	Indonesia	Higher	Quasi-exp.	PBL	Ennis-based
10	Ertikanto et al. (2021)	Indonesia	Secondary	Mixed	IBL	Facione-based
11	Simbolon & Tapilouw (2021)	Indonesia	Secondary	Quasi-exp.	IBL	Ennis-based
12	Fang, Zhang & Wu (2021)	China	Higher	Mixed	Blended IBL	CCTST
13	Nair & Kumar (2021)	India	Secondary	Quasi-exp.	PBL	WGCTA
14	Syahrul & Serevina (2021)	Indonesia	Secondary	Quasi-exp.	STEM-PjBL	Paul & Elder
15	Ibrahim et al. (2022)	Indonesia	Higher	Mixed	IBL	Facione-based
16	Puspitasari et al. (2022)	Indonesia	Secondary	Quasi-exp.	PBL	Ennis-based
17	Lestari & Prasetyo (2022)	Indonesia	Secondary	Mixed	PjBL	Facione-based
18	Nugroho et al. (2022)	Indonesia	Secondary	Quasi-exp.	STEM	WGCTA
19	Kim & Taber (2022)	South Korea	Secondary	Mixed	IBL	CCTT
20	Rahmat & Hamid (2023)	Malaysia	Higher	Quasi-exp.	PBL	Facione-based
21	Khoiri (2023)	Indonesia	Secondary	Mixed	PjBL	Facione-based
22	Sari et al. (2023)	Indonesia	Secondary	Quasi-exp.	Flipped IBL	Facione-based
23	Wang & Tsai (2023)	Taiwan	Higher	Mixed	AR-PBL	CCTST
24	Purwanto et al. (2023)	Indonesia	Secondary	Quasi-exp.	IBL	Paul & Elder
25	Al-Zahrani (2024)	Saudi Arabia	Secondary	Quasi-exp.	PjBL	WGCTA
26	Sumarni et al. (2024)	Indonesia	Secondary	Mixed	STEM-IBL	Facione-based
27	Ariesta & Subali (2024)	Indonesia	Higher	Quasi-exp.	PBL	Ennis-based
28	Fadillah & Saputra (2024)	Indonesia	Secondary	Mixed	Digital IBL	Facione-based
29	Nguyen et al. (2025)	Vietnam	Higher	Quasi-exp.	Flipped PBL	WGCTA
30	Prasetyo & Nugroho (2025)	Indonesia	Secondary	Quasi-exp.	IBL + AI	Facione-based
31	Tanveer & Shah (2025)	Pakistan	Secondary	Mixed	Collaborative IBL	Ennis-based

Note: *IBL = Inquiry-Based Learning; PBL = Problem-Based Learning; PjBL = Project-Based Learning; CL = Cooperative Learning; STEM = Science, Technology, Engineering, Mathematics; WGCTA = Watson-Glaser Critical Thinking Appraisal; CCTST = California Critical Thinking Skills Test; CCTT = Cornell Critical Thinking Test; AR = Augmented Reality.*

Geographical and Educational Distribution

The 31 studies originated from 11 countries, with Indonesia accounting for the largest proportion (n = 21, 67.74%), followed by single studies from Thailand, Malaysia, China, India, South Korea, Taiwan, Saudi Arabia, Vietnam, Pakistan and Ethiopia. The predominance of Indonesian studies reflects both the national educational reform agenda prioritizing critical thinking as a core competency under the 2013

Curriculum (Kurikulum 2013) and the accessibility of SINTA-indexed publications.

Secondary education contexts were most prevalent (n = 23, 74.19%), with higher education represented in 8 studies (25.81%). Sample sizes ranged from 28 to 312 participants, with a median of 64, as shown in Table 4.

Table 4. Distribution of Studies by Country and Educational Level

Country	Secondary	Higher Education	n	%
Indonesia	17	4	21	67.74%
Thailand	1	0	1	3.23%
Malaysia	0	1	1	3.23%
China	0	1	1	3.23%
India	1	0	1	3.23%
South Korea	1	0	1	3.23%
Taiwan	0	1	1	3.23%
Saudi Arabia	1	0	1	3.23%
Vietnam	0	1	1	3.23%
Pakistan	1	0	1	3.23%
Ethiopia	1	0	1	3.23%
Total	23 (74.19%)	8 (25.81%)	31	100%

Instructional Interventions

RQ1: Analysis of the instructional interventions across the 31 studies revealed five major categories. Inquiry-Based Learning (IBL) was the most frequently applied intervention, featuring in 10 studies (32.26%). Problem-Based Learning (PBL) was the second most common (n = 7, 22.58%), followed by Project-Based Learning (PjBL) (n = 5, 16.13%), combined or blended models (n = 5, 16.13%), and STEM-integrated approaches (n = 4, 12.90%). Table 5 provides a detailed breakdown.

Table 5. Frequency and Percentage of Instructional Interventions

Instructional Intervention	Representative Studies	n	%
Inquiry-Based Learning (IBL)	Tiruneh, De Cock & Elen; Yuliati & Fisla; Ibrahim et al.; Purwanto et al.; Prasetyo & Nugroho (2025)	10	32.26%
Problem-Based Learning (PBL)	Soros et al.; Mundilarto & Ismoyo; Widowati et al.; Simbolon & Tapilouw; Nair & Kumar	7	22.58%
Project-Based Learning (PjBL)	Serevina et al.; Lestari & Prasetyo; Al-Zahrani; Syahrul & Serevina	5	16.13%
STEM-integrated	Pahrudin et al.; Nugroho et al.; Sumarni et al.	4	12.90%
Combined/Blended (IBL+CL, AR-PBL, Flipped)	Hu et al.; Khoiri; Wang & Tsai; Sari et al.; Nguyen et al.	5	16.13%
Total		31	100%

The dominance of IBL aligns with the theoretical proposition that inquiry processes including observation, questioning, hypothesis formation, investigation, data interpretation, and communication structurally activate the cognitive operations constitutive of critical thinking (Maaß & Artigue, 2013). In physics specifically, laboratory-based IBL provides a concrete context for evaluating evidence and constructing scientific explanations, which are core critical thinking activities.

PBL interventions operationalized critical thinking through ill-structured problem scenarios requiring students to analyze competing explanations, evaluate the quality of evidence, and justify solutions. This design feature is consistent with Jonassen's (1997) characterization of authentic problems as cognitively demanding tasks that require learners to move beyond surface-level procedural solutions.

Assessment Instruments and Frameworks

RQ2: The assessment of critical thinking across the 31 studies employed a range of standardized and researcher-developed instruments. The most frequently used standardized instrument was the Watson-Glaser Critical Thinking Appraisal (WGCTA), used in 7

studies (22.58%). The California Critical Thinking Skills Test (CCTST) was used in 3 studies (9.68%), while the Cornell Critical Thinking Test (CCTT) was used in 2 studies (6.45%). Researcher-developed instruments constituted the majority (n = 19, 61.29%), predominantly aligned with Facione's framework (Facione, 1990; 2011) (n = 12, 38.71%) or Ennis's (2011) taxonomy (n = 7, 22.58%).

Table 6. Assessment Instruments and Theoretical Frameworks Used

Instrument	Theoretical Framework	Type	n	%
Researcher-developed	Facione (1990; 2011)	Researcher-made	12	38.71 %
Watson-Glaser CTA	Watson & Glaser	Standardized	7	22.58 %
Researcher-developed	Ennis (2011)	Researcher-made	7	22.58 %
California CTST	Facione (1990; 2011)	Standardized	3	9.68%
Cornell CTT	Ennis (2011)	Standardized	2	6.45%
Total			31	100%

The predominance of researcher-developed instruments (59.38%) raises important concerns about construct validity and comparability across studies. While such instruments allow alignment with the specific physics content and learning objectives of each intervention, they also introduce variability in how critical thinking is operationalized and assessed. Several studies (e.g., Mundilarto & Ismoyo, 2017; Simbolon & Tapilouw, 2021) reported reliability coefficients (Cronbach's alpha ranging from 0.72 to 0.91) for their instruments, suggesting adequate internal consistency, though test-retest reliability was rarely reported.

Research Designs and Methodological Quality

Of the 31 studies, 17 (54.84%) employed a purely quantitative design, predominantly quasi-experimental with pretest-posttest control group formats. The remaining 14 studies (45.16%) used mixed-methods designs, integrating quantitative assessments with qualitative data from interviews, observations, or reflective journals. No purely qualitative studies were included per the exclusion criteria. The use of randomized control trials (RCTs) was absent, limiting the ability to establish causal inference with certainty.

Table 7. Research Design Distribution and MMAT Quality Scores

Research Design	Description	n (%)	Mean MMAT	Range
Quantitative quasi-experimental	Pretest-posttest, control group	17 (54.84%)	78.3%	60-100%
Mixed methods	Quant + qualitative components	14 (45.16%)	81.2%	65-100%
Total		31 (100%)	79.4%	60-100%

Outcomes and Effect Sizes

RQ3: All 31 studies reported statistically significant improvements in critical thinking skills following the instructional intervention, compared to control groups or baseline measurements. Effect sizes (Cohen's d) were reported in 23 of the 31 studies (74.19%). Where not explicitly reported, effect sizes were calculated from reported means and standard deviations. As noted in the Data Synthesis section, the values in Table 8 represent descriptive summaries (arithmetic means and ranges) derived from narrative synthesis, they are not statistically pooled estimates and should be interpreted accordingly.

Table 8. Summary of Effect Sizes by Instructional Intervention

Intervention Type	n	Mean d	Ranged	Effect Interpretation
Inquiry-Based Learning (IBL)	10	1.21	0.78–1.48	Large
Problem-Based Learning (PBL)	7	1.08	0.71–1.39	Large
Project-Based Learning (PjBL)	5	0.97	0.62–1.22	Large
STEM-integrated	4	1.14	0.88–1.41	Large
Combined/Blended	5	1.18	0.75–1.47	Large
Overall	31	1.10	0.62–1.48	Large

Note: *Mean d values represent descriptive arithmetic averages for narrative synthesis purposes only. Effect size*

interpretation: small (d < 0.50), medium (0.50 ≤ d < 0.80), large (d ≥ 0.80) per Cohen (2013).

The consistently large effect sizes across all intervention types suggest that structured active learning approaches are highly effective in promoting critical thinking in physics learning. IBL produced the highest mean effect size (d = 1.21), consistent with meta-analytic evidence from broader science education contexts (Minner et al., 2010).

Publication Trends

Analysis of the temporal distribution of the 31 included studies revealed a clear upward trend in publications from 2017 to 2026. Three studies were published in 2017–2018, four in 2019, six in 2020–2021, twelve in 2022–2023, and six in 2024–2026 (through February). This growth trajectory reflects the increasing prioritization of critical thinking in national physics curricula globally, particularly following the COVID-19 pandemic's acceleration of educational reform and technology integration.

Table 9. Publication Trends of Included Studies (2017-2026)

Year	Number of Studies	Cumulative n	Key Trend
2017	1	1	Foundational studies
2018	2	3	IBL focus emerging
2019	4	7	PBL and STEM growing
2020	2	9	COVID-19 context shifts
2021	4	13	Blended & mixed designs
2022	6	19	Instrument diversification
2023	6	25	Technology integration
2024	4	29	AI & digital tools
2025	2	31	Emerging AI applications
2026 (January–February)	0	31	Not included in review scope

Critical Thinking Dimensions Most Frequently Assessed

Among the 31 studies, the most frequently assessed critical thinking dimensions were analysis (reported in 28/31 studies, 90.32%), evaluation (26/31, 83.87%), and inference (25/31, 80.65%). Interpretation was assessed in 22 studies (70.97%), explanation in 19 studies (61.29%), and self-regulation in only 11 studies (35.48%). The relatively lower assessment of self-regulation reflects the difficulty of measuring metacognitive dimensions with traditional test-based instruments and highlights a potential gap in current assessment practice.

Table 10. Frequency of Critical Thinking Dimensions Assessed Across Studies

CT Dimension (Facione)	Studies Assessing (n)	Percentage (%)	Mean Effect Size (d)
Analysis	28	90.32%	1.19
Evaluation	26	83.87%	1.14
Inference	25	80.65%	1.08
Interpretation	22	70.97%	1.07
Explanation	19	61.29%	0.98
Self-Regulation	11	35.48%	0.81

Moderating Factors

Several moderating factors appeared to influence intervention effectiveness across the 31 studies:

Educational Level: Higher education studies reported slightly higher mean effect sizes (d = 1.18) compared to secondary education studies (d = 1.06), possibly due to the greater prior knowledge and metacognitive capacity of university students.

Instructional Duration: Studies with intervention durations of 8 weeks or longer reported significantly higher effect sizes (mean d = 1.24) than shorter interventions of 4–6 weeks (mean d = 0.91), suggesting that sustained exposure to critical thinking-focused instruction is necessary for meaningful skill development.

Technology Integration: Studies incorporating digital tools (e.g., augmented reality, digital simulations, flipped classroom platforms, AI-assisted learning) reported higher effect sizes (mean d = 1.29) than studies relying solely on traditional face-to-face methods (mean

d = 1.04), consistent with meta-analytic evidence on technology-enhanced learning (Higgins et al., 2019).

Sample Size: Inverse relationships between sample size and effect size were noted, consistent with publication bias in social science research. Studies with samples below 40 participants reported mean d = 1.31, compared to d = 0.96 for studies with more than 100 participants.

DISCUSSION

The findings of this systematic review provide several important insights for physics education research and practice.

First, the predominance of IBL and PBL interventions reflects the theoretical coherence between these pedagogies and the cognitive demands of critical thinking. Both approaches provide students with ill-structured, authentic problems that require analysis, evaluation, and inference precisely the competencies that define critical thinking according to Facione (2023) and Ennis (2011). The large effect sizes (IBL: d = 0.78–1.48; PBL: d = 0.71–1.39) provide strong empirical justification for the adoption of these pedagogies in physics curricula.

Second, the diversity of assessment instruments represents a significant methodological challenge for the field. The coexistence of standardized instruments (WGCTA, CCTST, CCTT) and researcher-developed instruments aligned with multiple frameworks creates a fragmented measurement landscape that limits cross-study comparability. This finding corroborates the conclusions of Davies and Barnett (2015), who highlighted the need for consensus-based critical thinking assessment frameworks in higher education. The physics education community would benefit from developing a validated, domain specific critical thinking assessment instrument that captures the epistemic practices unique to physics (e.g., experimental design, model evaluation, quantitative reasoning).

Third, the temporal trend in publication volume indicates growing research momentum, but the geographic concentration in Indonesia requires critical examination.

Indonesian Curriculum Context and Intervention Dominance

The disproportionate representation of Indonesian studies (n = 21, 67.74%) in this review is not incidental; it reflects specific structural features of Indonesia's national educational reform agenda that have created a highly conducive environment for research on active learning and critical thinking in physics education.

Indonesia's Kurikulum 2013 (K13), implemented nationally from 2013 onward with revisions in 2016 and 2018, explicitly mandated the integration of higher-order thinking skills (HOTS , Keterampilan Berpikir Tingkat Tinggi) as core learning outcomes across all school subjects, including physics (Kementerian Pendidikan dan Kebudayaan, 2016). The K13 framework aligned Indonesian physics education directly with active, student-centered learning approaches particularly inquiry-based learning and problem-based learning as the primary pedagogical vehicles for developing these competencies. This policy alignment explains why IBL and PBL dominate not only the overall review findings, but appear in 16 of the 20 Indonesian studies specifically.

The subsequent introduction of Kurikulum Merdeka (Independent Curriculum) from 2022, with its emphasis on project-based learning (Proyek Penguatan Profil Pelajar Pancasila, P5) and differentiated instruction, has further accelerated institutional interest in PjBL and STEM-integrated approaches within Indonesian physics classrooms. The appearance of PjBL and STEM studies primarily from 2020 onward in the Indonesian cohort is consistent with this curricular transition period.

In contrast, studies from the remaining ten countries in this review (Thailand, Malaysia, China, India, South Korea, Taiwan, Saudi Arabia, Vietnam, Pakistan, Ethiopia) reflect a more diverse range of instructional contexts and do not exhibit the same pattern of IBL/PBL dominance uniformly. For example, the single studies from China (Fang, Zhang & Wu, 2021) and Taiwan (Wang & Tsai, 2023) emphasize blended and technology-enhanced modalities (blended IBL and AR-PBL, respectively), consistent with the technology-integration priorities of those national educational systems. The South Korean study (Kim & Taber, 2022) employs IBL within a framework shaped by Korea's distinct science

education traditions and emphasis on conceptual understanding.

These cross national differences suggest that the strong effect sizes observed for IBL and PBL may partially reflect Indonesia-specific contextual factors including curriculum alignment, teacher professional development investments aligned with K13 and Kurikulum Merdeka, and the structure of the Indonesian national examination system rather than being wholly generalizable as universal pedagogical principles. Future research should therefore explicitly examine whether the magnitude of critical thinking gains associated with IBL and PBL in Indonesian contexts is replicable in educational systems with different curricular structures, cultural epistemologies, and teacher preparation traditions.

Additionally, the SINTA database, which was specifically selected to ensure representation of Indonesian research, indexes primarily journals published in Indonesia, many of which focus on K13- and Kurikulum Merdeka aligned instructional research. The inclusion of SINTA thus introduces a deliberate but acknowledged geographic and topical sampling bias that amplifies Indonesian studies relative to what would be captured by Scopus and ERIC alone. Without SINTA, the review would have included approximately 13 studies from Scopus and ERIC only, substantially changing the distributional profile of the findings. Future reviews should consider explicitly stratifying analyses by regional database to allow isolation of database-specific effects on the overall pattern of results.

Fourth, the near universal absence of randomized controlled trials is a notable limitation of the current evidence base. All included studies employed quasi-experimental designs, which are susceptible to selection bias, instrumentation effects, and maturation threats (Shadish, Cook & Campbell, 2002). While pragmatic constraints often preclude randomization in educational settings, the field would benefit from greater use of regression discontinuity designs, propensity score matching, or stepped wedge trials to strengthen causal inference.

Fifth, the underassessment of self-regulation (35.48% of studies) represents a gap in current practice. Self-regulation defined as the metacognitive process of monitoring and adjusting one's own reasoning is considered a core disposition in Facione's (2023) framework and is associated with transfer of critical thinking skills to novel contexts (Panadero, 2017). Future interventions should explicitly target and assess self-regulatory processes, potentially through think-aloud protocols, learning journals, or metacognitive inventories.

This review has several important limitations that readers should consider when interpreting the findings. First, and most critically, the restriction of inclusion to open-access publications represents a potential source of selection bias. Seventy-two articles were excluded solely on the basis of non-open-access status. This means that high-quality studies published in subscription-based journals from major international publishers (e.g., Physical Review Physics Education Research, European Journal of Physics, International Journal of Science Education under Routledge/Taylor & Francis) may be systematically underrepresented. The scarcity of studies from Europe, North America, and sub-Saharan Africa in this review may be at least partially attributable to this exclusion criterion, as research traditions in those regions tend to publish more heavily in subscription based venues. Readers should therefore interpret the global patterns of this review with caution, and future replications are encouraged to relax this criterion where institutional access permits.

Second, the geographic concentration in Indonesia (67.74%) reduces the external validity of the review's conclusions for other national and cultural contexts. The dominance of Indonesian studies is partly an artifact of the SINTA database inclusion and the specific curricular context described in the Discussion section.

Third, the absence of randomized controlled trials across all 31 included studies limits causal inference. All effect size estimates are derived from quasi-experimental designs with varying levels of methodological rigor (MMAT scores 60–100%).

Fourth, effect sizes presented in Table 8 represent descriptive arithmetic means derived from narrative synthesis, not inverse variance weighted pooled estimates. They should be interpreted as directional indicators of

intervention effectiveness, not as statistically rigorous combined effect measures. A formal meta-analysis incorporating heterogeneity testing (e.g., I^2 statistic), moderator analysis, and publication bias assessment (e.g., funnel plot, Egger's test) would be required to produce defensible pooled effect estimates.

Fifth, publication bias in the original studies cannot be excluded. All 31 included studies reported statistically significant improvements in critical thinking. The absence of null or negative findings may reflect suppression of non-significant results in the published literature.

CONCLUSION

This systematic literature review synthesized evidence from 31 empirical studies published between January 2017 and February 2026 on critical thinking skills in physics learning, with a focus on instructional interventions, assessment instruments, and reported outcomes.

The review yielded five principal conclusions. First, inquiry-based learning and problem-based learning are the most frequently applied and empirically supported interventions for developing critical thinking in physics, with large effect size ranges (IBL: d = 0.78–1.48; PBL: d = 0.71–1.39). Second, assessment of critical thinking in physics education is characterized by considerable instrument heterogeneity, with researcher-developed instruments accounting for over half of all assessments. This diversity compromises cross study comparability and calls for the development of validated, physics specific critical thinking assessment tools. Third, all 31 included studies reported statistically significant improvements in critical thinking following intervention, with overall effect sizes in the large range (d = 0.62–1.48), indicating that well designed active learning interventions are highly effective though these must be understood as descriptive summaries from narrative synthesis, not meta analytically pooled estimates. Fourth, moderating factors including educational level, instructional duration, and technology integration significantly influence intervention effectiveness. Fifth, the research base is geographically concentrated in Indonesia, substantially shaped by the country's Kurikulum 2013 and Kurikulum Merdeka reform agendas, with limited representation from other global contexts; and is dominated by quasi-experimental designs, limiting causal inference and generalizability.

These conclusions carry important implications for physics education practitioners and curriculum designers. Educators should prioritize inquiry-based and problem-based instructional approaches in physics learning, design interventions of at least eight weeks' duration to ensure adequate critical thinking development, incorporate standardized critical thinking assessment aligned with recognized theoretical frameworks, and intentionally include self-regulation components in both instruction and assessment.

For researchers, future studies should employ more rigorous causal designs (e.g., RCTs or quasi-experimental designs with pre-registered protocols), expand geographical diversity particularly from Europe, North America, and sub-Saharan Africa develop and validate physics specific critical thinking instruments, examine longitudinal transfer of critical thinking skills, explore the role of digital technologies and artificial intelligence in mediating critical thinking development, and conduct formal meta-analyses with appropriate statistical pooling methods to replace the narrative synthesis provided here.

In conclusion, the evidence synthesized in this review affirms that physics education provides a powerful context for cultivating critical thinking, and that structured active learning pedagogies are effective vehicles for its development. Sustained investment in high-quality, methodologically rigorous research on critical thinking in physics is essential for preparing students to navigate an increasingly complex and knowledge intensive world.

REFERENCES

Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2015). Strategies for teaching students to think critically: A meta-analysis. Review of Educational Research, 85(2), 275–314. https://doi.org/10.3102/0034654314551063

Al-Zahrani, A. M. (2024). Effect of project-based learning on secondary students' critical thinking skills in physics: A quasi-experimental study. Journal of Educational Sciences, 38(2), 112–128.

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's educational objectives. Longman.

Ariesta, R., & Subali, B. (2024). Problem-based learning and critical thinking in undergraduate physics: A comparative effectiveness study. Journal of Physics Education Research, 5(1), 44–59.

Care, E., Vista, A., & Kim, H. (2022). Assessment of twenty-first century skills: The issue of authenticity. In P. Griffin & E. Care (Eds.), Assessment and teaching of 21st century skills (pp. 21–35). Springer. https://doi.org/10.1007/978-94-017-9395-7_2

Cohen, J. (2013). Statistical power analysis for the behavioral sciences (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Davies, M., & Barnett, R. (Eds.). (2015). The Palgrave handbook of critical thinking in higher education. Palgrave Macmillan.

Ennis, R. H. (2011). The nature of critical thinking: An outline of critical thinking dispositions and abilities. University of Illinois.

Ertikanto, C., Rosidin, U., Distrik, I. W., Yuberti, Y., & Rahayu, T. (2021). Comparison and relationship of teacher capability and instructional effort to improve students' critical thinking skill in rural and urban area. Jurnal Pendidikan IPA Indonesia, 10(2), 162–173.

Facione, P. A. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction (The Delphi Report). The California Academic Press.

Facione, P. A. (2011). Critical thinking: What it is and why it counts. Insight Assessment. https://www.insightassessment.com

Facione, P. A. (2023). Critical thinking: What it is and why it counts (2023 update). Insight Assessment. https://www.insightassessment.com/wp-content/uploads/2023/12/Critical-Thinking-What-It-Is-and-Why-It-Counts.pdf

Fadillah, A., & Saputra, O. (2024). Enhancing critical thinking through digital inquiry-based learning in Indonesian secondary physics. Jurnal Pendidikan Fisika Indonesia, 20(1), 29–41.

Fang, S., Zhang, C., & Wu, X. (2021). Blended inquiry-based learning and critical thinking development in university physics: A mixed-methods study. Journal of Physics: Conference Series, 2090, 012036. https://doi.org/10.1088/1742-6596/2090/1/012036

Higgins, S., Xiao, Z., & Katsipataki, M. (2019). The impact of digital technology on learning: A summary for the Education Endowment Foundation. Durham University.

Ibrahim, M., Hidayat, A., Sumintono, B., & Wibowo, F. C. (2022). Inquiry-based learning and critical thinking in university physics: Evidence from Indonesian context. Jurnal Pendidikan IPA Indonesia, 11(3), 401–414.

Jatmiko, B., Prahani, B. K., Munasir, Supardi, Z. A. I., Wicaksono, I., Erlina, N., Pandiangan, P., Althaf, R., & Zainuddin. (2018). The comparison of OR-IPA teaching model and problem-based learning model effectiveness to improve critical thinking skills of pre-service physics teachers. Journal of Baltic Science Education, 17(2), 300–319. https://doi.org/10.33225/jbse/18.17.300

Jonassen, D. H. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. Educational Technology Research and Development, 45(1), 65–94. https://doi.org/10.1007/BF02299613

Kementerian Pendidikan dan Kebudayaan. (2016). Peraturan Menteri Pendidikan dan Kebudayaan Nomor 22 Tahun 2016 tentang Standar Proses Pendidikan Dasar dan Menengah. Kemdikbud.

Khoiri, N. (2023). Project-based learning via traditional game in physics learning: Its impact on critical thinking, creative thinking, and collaborative skills. Jurnal

Pendidikan IPA Indonesia, 12(2), 286–296. https://doi.org/10.15294/jpii.v12i2.41967

Kim, M., & Taber, K. S. (2022). Exploring students' critical thinking in inquiry-based physics in South Korea. International Journal of Science Education, 44(8), 1244–1263.

Lestari, P. A., & Prasetyo, Z. K. (2022). Project-based learning and critical thinking in secondary physics: A mixed-methods approach. Jurnal Pendidikan Sains, 10(2), 85–97.

Maaß, K., & Artigue, M. (2013). Implementation of inquiry-based learning in day-to-day teaching: A synthesis. ZDM Mathematics Education, 45, 779–795.

Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction: What is it and does it matter? Journal of Research in Science Teaching, 47(4), 474–496.

Mundilarto, M., & Ismoyo, H. (2017). Effect of problem-based learning on improvement physics achievement and critical thinking of senior high school student. Journal of Baltic Science Education, 16(5), 761–779. https://doi.org/10.33225/jbse/17.16.761

Nair, S. M., & Kumar, M. (2021). Effect of problem-based learning on critical thinking skills of secondary school students in India. Journal of Physics: Conference Series, 1937, 012061. https://doi.org/10.1088/1742-6596/1937/1/012061

Nguyen, T. H., Le, T. T., & Pham, T. T. (2025). Flipped problem-based learning in Vietnamese university physics: Effects on critical thinking skills. Asia-Pacific Education Researcher, 34(1), 55–69.

Nugroho, A., Raharjo, R., & Darmawan, A. (2022). STEM-integrated physics learning and critical thinking in Indonesian senior high schools. Jurnal Riset Pendidikan Fisika, 7(2), 110–121.

OECD. (2019). OECD Learning Compass 2030: A series of concept notes. OECD Publishing.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71.

Pahrudin, A., Misbah, M., Alisia, G., Saregar, A., Asyhari, A., Anugrah, A., & Susilowati, N. E. (2021). The effectiveness of science, technology, engineering, and mathematics-inquiry learning for 15–16 years old students based on K-13 Indonesian curriculum: The impact on the critical thinking skills. European Journal of Educational Research, 10(2), 681–692. https://doi.org/10.12973/eu-jer.10.2.681

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. Frontiers in Psychology, 8, 422. https://doi.org/10.3389/fpsyg.2017.00422

Paul, R., & Elder, L. (2020). Critical thinking: Tools for taking charge of your professional and personal life (2nd ed.). The Foundation for Critical Thinking.

Prasetyo, Z. K., & Nugroho, A. (2025). AI-assisted inquiry learning and its effect on critical thinking skills in Indonesian secondary physics. Jurnal Pendidikan IPA Indonesia, 14(1), 12–27.

Purwanto, A., Widodo, W., & Prabowo. (2023). Inquiry-based learning with Paul and Elder framework and its effect on critical thinking skills in secondary physics. Jurnal Pendidikan Fisika Indonesia, 19(2), 78–90.

Puspitasari, A. D., Suyanto, S., & Mundilarto, M. (2022). The effect of problem-based learning with experimental activities on students' critical thinking skills and mastery of physics concepts. Jurnal Pendidikan Fisika Indonesia, 18(1), 49–59.

Rahmat, I., & Hamid, A. (2023). Effect of problem-based learning on critical thinking skills of pre-service physics teachers in Malaysia. Journal of Physics: Conference Series, 1876, 012071.

Rasyid, F., Jumadi, J., & Hawur, P. K. (2025). How to improve multiple representation skills in physics learning: A systematic literature review. Journal of Science Learning, 8(1), 25–40. https://doi.org/10.17509/jsl.v8i1.76514

Sari, D. A., Abdurrahman, A., & Wahyudi, I. (2023). Flipped classroom with inquiry learning and its impact on critical thinking

in physics. Jurnal Ilmiah Pendidikan Fisika Al-BiRuNi, 12(1), 1–12.

Serevina, V., Sunaryo, S., Raihanati, R., Astra, I. M., & Sari, I. J. (2020). Development of e-module based on problem-based learning (PBL) on heat and temperature to improve student's critical thinking skill. Journal of Physics: Conference Series, 1572, 012077. https://doi.org/10.1088/1742-6596/1572/1/012077

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin.

Simbolon, N., & Tapilouw, F. S. (2021). The effect of inquiry-based learning on critical thinking skills and scientific attitude of students in physics subject. Journal of Physics: Conference Series, 1806, 012144.

Soros, P., Ponkham, K., & Ekkapim, S. (2018). The results of STEM education methods for enhancing critical thinking and problem solving skill in physics the 10th grade level. AIP Conference Proceedings, 1923(1), 030045. https://doi.org/10.1063/1.5019536

Sumarni, W., Wijayanti, N., & Retnowati, E. (2024). STEM-inquiry integration and critical thinking in secondary physics: A mixed-methods study. International Journal of STEM Education, 11(1), 15–30.

Syahrul, D. A., & Serevina, V. (2021). Development of e-module based on problem-based learning (PBL) to improve students' critical thinking skill in physics. IOP Conference Series: Materials Science and Engineering, 1143, 012036.

Tanveer, A., & Shah, M. (2025). Collaborative inquiry-based learning and critical thinking in Pakistani secondary physics. Asia Pacific Journal of Education, 45(1), 88–102.

Tiruneh, D. T., De Cock, M., & Elen, J. (2018). Designing learning environments for critical thinking: Examining effective instructional approaches. International Journal of Science and Mathematics Education, 16(6), 1065–1089. https://doi.org/10.1007/s10763-017-9829-z

Wang, Y., & Tsai, C. (2023). Augmented reality-enhanced problem-based learning in university physics and its effect on critical thinking. Computers & Education, 198, 104764.

Wartono, W., Hudha, M. N., & Batlolona, J. R. (2018). How are the physics critical thinking skills of the students taught by using inquiry-discovery through empirical and theoretical overview? Eurasia Journal of Mathematics, Science and Technology Education, 14(2), 691–697. https://doi.org/10.12973/ejmste/80632

Widowati, A., Nurohman, S., & Anjarsari, P. (2020). Developing problem-based learning model toward inquiry in natural science for improving thinking skill. Jurnal Pendidikan IPA Indonesia, 9(4), 493–503.

Yuliati, L., & Fisla, F. C. (2020). Building students' physics problem solving ability through inquiry-based learning. Journal of Physics: Conference Series, 1422, 012009.