

Design and Evaluation of a Culturally Responsive Flipbook E-Module on Quantities and Units for Improving Students' Critical Thinking Skills

Ellianawati Ellianawati^{1*}, Figi Anggun Yossindy¹, Bambang Subali¹, Suharto Linuwih¹, & Lolly Jean C. Simbulas²

¹Physics Education Study Program, Universitas Negeri Semarang, Indonesia

²University of the Immaculate Conception, Philippines

*Corresponding email: ellianawati@unnes.ac.id

Received: 08 December 2025

Accepted: 19 January 2026

Published: 05 February 2026

Abstract: This study reports the development and empirical evaluation of a culturally responsive flipbook e-module for teaching quantities and units in Indonesian senior secondary physics. Developed using the ADDIE framework and grounded in Culturally Responsive Teaching (CRT), the module integrates local measurement systems, *mud, tala, lot, and deben*, as epistemic resources for scientific reasoning. A quasi-experimental field trial involved 117 Grade 10 students across three intact classes: two experimental groups (digital: $n = 41$; print: $n = 40$) and one control group (conventional worksheets: $n = 36$). Expert validation confirmed strong content validity (Aiken's $V = 0.87$), particularly in cultural authenticity and alignment with critical thinking constructs. Student surveys indicated high practicality (89.55%) and positive perceptions (85.75%), with learners highlighting how culturally familiar contexts, such as *zakat* unit analysis or land measurement using *lots*, enhanced relevance and cognitive engagement. To evaluate learning impact, a pretest–posttest design was implemented, with critical thinking measured via a validated open-response instrument. Analysis of covariance (ANCOVA), controlling for pretest scores, revealed a statistically significant effect of instructional modality on posttest performance, $F(2, 113) = 58.34, p < 0.001$, with a large effect size (*partial* $\eta^2 = 0.513$). Adjusted means showed that the experimental groups significantly outperformed the control group ($M_{\text{adj}} = 79.14$ (digital); $M_{\text{adj}} = 80.72$ (print); $M_{\text{adj}} = 42.13$ (control); $p < 0.001$), while there was no meaningful difference between the delivery modes ($p = 0.682$). These findings suggest that the CRT-based design, rather than the medium, influences learning outcomes, supporting scalable implementation across diverse infrastructural contexts. When cultural knowledge is structurally integrated into disciplinary reasoning, students engage more deeply, affirming the module's potential as an equitable, contextually relevant tool for improving rigorous science education.

Keywords: critical thinking skills, e-module, culturally responsive teaching.

Article's DOI: <https://doi.org/10.23960/jpmipa.v27i1.pp175-193>

■ INTRODUCTION

Contemporary education systems face growing demands to prepare learners not only with disciplinary knowledge but also with higher-order competencies, particularly the ability to navigate complexity, ambiguity, and conflicting evidence. Among the core 21st-century skills, critical thinking serves as a cognitive cornerstone. It empowers individuals to deconstruct information, interrogate assumptions, assess empirical validity, and formulate reasoned

conclusions (Soleh & Arifin, 2021; Maryam et al., 2020). Ennis (1987) defines this capacity as “reasonable, reflective thinking focused on deciding what to believe or do,” emphasizing justification as its hallmark. Building on this, Arifah et al. (2021) describe critical thinking as an emergent result of dynamic interactions among problems, arguments, and statements, highlighting integral role in disciplined, evidence-based inquiry.

Physics education, by its very nature, provides a fertile ground for cultivating such

reasoning. The discipline routinely requires students to align conceptual models with empirical observations, evaluate the plausibility of proposed solutions, and refine interpretations in light of new evidence (Sugiarti & Dwikoranto, 2021). These cognitive practices are especially crucial in Grade X, where foundational ideas, such as physical quantities and units, lay the groundwork for the entire subject. As Sandari (2021) notes, this system is far more than a set of notational conventions; it functions as the formal grammar through which physical laws are articulated, tested, and communicated.

Despite its importance, students often struggle with its core components. Empirical studies show ongoing misconceptions, including confusion between physical quantities and mere numerical values, misreading of unit symbols, inconsistent use of dimensional analysis, and a procedural rather than conceptual approach to units in problem-solving (Widodo et al., 2022; Nguyen & Riconsolido, 2024). These difficulties stem not from computational errors but from fragmented conceptual understanding, which in turn hampers students' ability to engage in the sophisticated reasoning required for scientific literacy.

Contemporary science education increasingly recognizes the need to ground abstract scientific concepts in learners' lived realities and sociocultural contexts. In response to persistent challenges such as conceptual fragmentation and disengagement, pedagogical approaches have shifted toward frameworks that honor students' cultural identities as intellectual resources rather than barriers to learning. Culturally Responsive Teaching (CRT) offers a robust foundation for this shift by affirming students' linguistic repertoires, community knowledge, and everyday experiences as integral to meaningful knowledge construction (Cruz et al., 2020; Rahmawati & Taylor, 2020). When local measurement systems, such as *mud*, *tala*,

or *lot*, are authentically embedded into physics instruction, they function not as superficial cultural tokens but as epistemic bridges that make disciplinary practices more accessible, coherent, and relevant (Widodo et al., 2022; Buchori & Harun, 2020). Empirical studies show that culturally grounded strategies can significantly improve conceptual understanding, foster classroom engagement, and support the development of critical analytical skills across varied science learning settings (Özüdođru, 2022; Miskiyyah, 2023).

Meanwhile, advances in digital pedagogy create new opportunities to apply these principles. Flipbook-style electronic modules (e-modules), when designed according to cognitive load theory and Mayer's (2009) multimedia learning principles (Sriyanti et al., 2021), provide a structured yet adaptable platform for presenting complex ideas. Their multimodal architecture, integrating visual sequences, explanatory text, and interactive cues, facilitates self-regulated learning while reducing extraneous cognitive load. This format is especially well-suited for guiding students through the iterative reasoning processes central to physics problem-solving. Emerging evidence suggests that thoughtfully designed e-modules can indeed nurture critical thinking (Prasasti & Anas, 2023; Ainy et al., 2024). However, a notable limitation persists: most research treats technological innovation and cultural responsiveness as separate domains, rarely exploring their synergistic potential within a unified instructional design.

This separation reveals a critical gap in the literature. To date, there is limited empirical research on flipbook-based e-modules that deliberately integrate CRT principles into foundational physics content at the secondary level, particularly when student outcomes are assessed using validated, multidimensional frameworks of critical thinking. CRT, as conceptualized by Vavrus (2008) and Gay (2002), emphasizes that academic rigor and

cultural relevance are not mutually exclusive; instead, learning becomes most powerful when knowledge is connected to learners' frames of reference. Such an approach not only validates students' cultural heritage but also enhances motivation, academic achievement, and classroom belonging (Nasution et al., 2023; Wahira et al., 2024). Furthermore, it promotes inclusive learning environments where diverse cultural expressions are recognized as assets for meaning-making (Mumpuniarti et al., 2020; Yasmin, 2024).

Existing studies offer partial insights but fall short of a holistic integration. For example, Sya'bana et al. (2024) demonstrated that CRT-enhanced instruction improves science learning outcomes by linking content to students' daily lives, yet their intervention did not involve digital tools. Buchori & Harun (2020) developed a CRT-based flipbook e-module, but their focus was on geometric transformations in vocational mathematics rather than physics. Latifah et al. (2020) created a physics e-module to promote critical thinking, though it relied on a 7-step meaningful learning model rather than a culturally responsive lens. Similarly, Simatupang (2024) applied CRT in history education to build cultural awareness but did not use interactive digital formats such as flipbooks.

Crucially, no published study has empirically developed and validated a CRT-integrated flipbook e-module specifically for teaching *quantities and units*, a conceptually dense and epistemologically foundational topic in early physics education (Sandari, 2021; Widodo et al., 2022). Misconceptions in this domain often stem from treating units as mere labels rather than as constitutive elements of physical meaning, thereby hindering higher-order reasoning (Nguyen & Riconsolido, 2024). Compounding this issue, many studies assess critical thinking using narrow or teacher-designed instruments, overlooking established multidimensional models, such as

those by Facione (1990) and Halpern (2014), that delineate distinct cognitive dimensions, including interpretation, analysis, evaluation, inference, and metacognitive regulation.

In response, this study proposes a novel synthesis of three underexplored yet interdependent dimensions: (1) Disciplinary specificity, by targeting quantities and units, a fragile yet pivotal concept in physics literacy; (2) Pedagogical-cultural integration, by positioning local measurement practices as authentic epistemic resources that scaffold scientific reasoning, not as illustrative add-ons; and (3) Techno-pedagogical innovation, by delivering this integration through a flipbook e-module grounded in evidence-based design principles (Mayer, 2009), enabling sequential, multimodal, and learner-centered engagement. This triadic convergence, culturally responsive content, discipline-specific foundational knowledge, and theoretically informed digital delivery have not been realized or evaluated in prior research. By bridging this gap, the study contributes both theoretically, by extending CRT into digital science education, and practically, by offering a scalable, context-sensitive model for physics instruction in culturally diverse classrooms.

Critical thinking is operationalized through a validated instrument comprising twelve indicators, analytically organized into five dimensions: interpretation, analysis, evaluation, inference, and explanation, ensuring both theoretical fidelity and empirical tractability. The following questions guide the research:

1. To what extent does the CRT-based flipbook e-module meet content and construct validity criteria?
2. How practical is the e-module from students' perspectives?
3. Does its implementation yield significantly greater gains in critical thinking compared to conventional instruction?

4. How do students describe its influence on their reasoning processes in physics learning?

■ **METHOD**

Participants

The study was conducted in a public senior high school in Indonesia, with the target population comprising all Grade 10 students enrolled during the academic year of implementation. Given the constraints inherent in school-based intervention research, particularly the infeasibility of disrupting existing class structures or reassigning students, a non-random, criterion-guided sampling strategy was employed. Three intact classes were selected in consultation with physics instructors, who served as informants (not selectors) to identify cohorts meeting three pre-established criteria: (1) alignment in curriculum pacing for the quantities and unit's topic; (2) comparable weekly instructional time allocation; and (3) relative homogeneity in prior academic performance, as indicated by end-of-semester assessment records from the preceding term. This process yielded a final sample of 117 students, distributed as follows: Experimental Group 1 (digital e-module, Class X.6, $n = 41$), Experimental Group 2 (print e-module, Class X.7, $n = 40$), and Control Group (conventional worksheet instruction, Class X.8, $n = 36$).

Such an approach reflects standard practice in quasi-experimental educational research, where methodological rigor is maintained not through randomization, which is often pedagogically and ethically impractical in real classrooms, but through systematic control of confounding variables at the design and analysis stages (e.g., pretest covariation, matched instructional duration, and teacher consistency). While relying on intact classes from a single institution inevitably limits external validity, it improves ecological fidelity: the findings directly relate to how the intervention works under normal teaching conditions, rather than in artificially controlled

environments. As Shadish et al. (2002) caution, generalizability in such contexts should be grounded in the depth of contextual description and mechanism articulation rather than in statistical methods. Accordingly, detailed profiles of the school, student demographics, and instructional context are provided in the supplementary materials to support readers' judgments of transferability.

Research Design and Procedures

The CRT-based flipbook e-module emerged from four iterative cycles within the ADDIE framework, grounded in design-based co-creation with teachers and cultural informants. The stages described above are summarized in Figure 1.

The analysis phase, the initial step in ADDIE-based development research, involves identifying problems and gathering data for a needs assessment. In this study, preliminary investigations were conducted at SHS Tugumulyo through classroom observations, interviews with physics teachers, and questionnaires administered to Grade X students. Findings revealed two key aspects: (1) the school holds potential for developing culture-integrated teaching materials enhanced by technology, yet currently relies primarily on printed textbooks, basic PowerPoint slides, and Quizizz for practice, with no existing flipbook-based digital resources; and (2) students perceive the topic of magnitude and measurement as challenging. Although teachers occasionally link physics concepts to local cultural contexts, such integration remains superficial. To address these gaps, the study aims to design a CRT-informed flipbook e-module to foster students' critical thinking skills.

Following the analysis of identified problems and potential solutions, the research proceeds to the design phase, the stage dedicated to planning the teaching materials. The first step in this phase involves creating a storyboard, which serves as

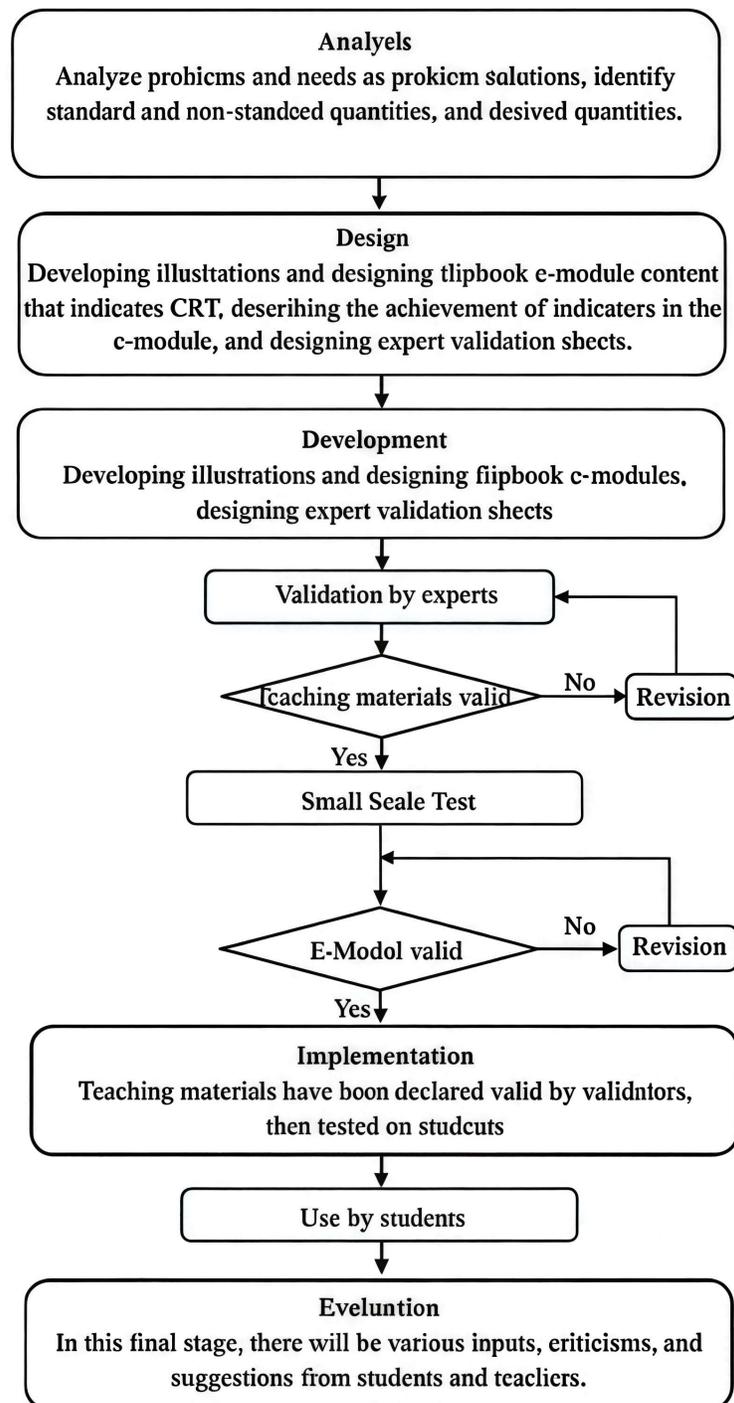


Figure 1. The ADDIE iteration cycle with key revisions

a visual and structural blueprint to guide the development of the e-module and clarify its intended learning flow. This storyboard is aligned with the study's predefined learning objectives.

Concurrently, the core content of the e-module is drafted, ensuring conceptual coherence and relevance. Additionally, supporting components, such as formative and summative evaluation

items, illustrative images, and other multimedia elements, are prepared to enrich the instructional design and enhance student engagement.

At the development stage, the initial version of the e-module is constructed according to the instructional sequence outlined in the storyboard. The process begins with preparing the core content, specifically the topic of quantity and units, refined to align with the study's learning objectives. Visual and structural elements are then designed: the front and back covers are created using Canva, while the background, header, and footer are formatted directly in Microsoft Excel. The e-module is further enriched with multimedia components, including animations, videos, audio clips, hyperlinks, QR codes, illustrative images, worked examples, and practice questions, all tailored to support culturally responsive teaching and the development of critical thinking.

Once all components are integrated, the complete e-module is converted into an interactive, web-based flipbook using an HTML5 flipbook platform. Before implementation, the e-module undergoes expert validation to assess its feasibility. This validation serves as a quality control checkpoint; if the material does not meet predetermined criteria, revisions are made based on expert feedback. Validity is determined through two parallel assessments: (1) evaluation of the e-module itself via an expert judgment sheet covering three domains, content accuracy, grammatical correctness, and language clarity; and (2) validation of the accompanying critical thinking assessment instrument. The e-module is deemed valid only when it satisfies the established validation standards across all specified criteria.

The implementation stage involves the direct application of the developed e-module flipbook in the target classroom. Prior to implementation, the teaching material had already undergone expert validation and was revised based on validator feedback. Following the classroom trial,

students completed two instruments: a practicality sheet and a student response questionnaire. These tools were used to assess the ease of use, accessibility, and overall usability of the e-module flipbook. The responses collected also provided valuable input and suggestions for further refinement, leading to subsequent revisions of the e-module flipbook teaching material to enhance its effectiveness and user-friendliness. The study adopted a non-equivalent control group pretest–posttest design, a well-established approach for evaluating instructional interventions in authentic school settings where randomization is neither feasible nor educationally appropriate (Shadish et al., 2002). Given the institutional reality of fixed class rosters, participants remained in their naturally constituted Grade 10 cohorts throughout the study period.

The implementation spanned four consecutive weeks, focusing on the foundational topic of quantities and units. Before any instruction began, all participants completed a validated critical-thinking pretest to establish a common baseline for comparison. The intervention phase then proceeded with three distinct instructional approaches assigned to different groups. The first experimental group interacted with a culturally responsive teaching-based flipbook e-module that integrated local measurement systems, such as *mud*, *tala*, and *lot*, as meaningful contexts to foster scientific reasoning. The second experimental group used printed materials derived from the same content structure and learning objectives as the digital module; these materials were adapted following an evidence-informed protocol designed for low-connectivity environments, in line with recommendations from Lee et al. (2023) and UNESCO (2023). Meanwhile, the control group continued with the school's standard physics curriculum, receiving conventional teacher-led instruction supplemented by textbook-based worksheets, thereby reflecting regular classroom practice without any experimental alterations.

Instructional time was carefully equated across groups (six 45-minute sessions), and the same teacher facilitated all three classes to minimize instructor-related variance. Upon completion of the treatment phase, a parallel-form posttest, matched in difficulty, format, and construct coverage to the pretest, was administered to assess shifts in critical thinking performance. This design enabled both *within-group* (pre–post) and *between-group* (ANCOVA-controlled) comparisons, strengthening causal inference despite the absence of random assignment.

Evaluation in this study is not confined to the end of the development cycle but is embedded throughout each phase of the ADDIE model. During the analysis phase, field data are continuously evaluated to identify key characteristics and contextual needs, ensuring the resulting instructional design aligns with the research objectives. In the design phase, evaluation focuses on verifying that the teaching materials adequately address students' learning needs, fulfill curriculum requirements, and meaningfully integrate local wisdom as a contextual anchor for physics concepts. At the development stage, evaluation is guided by expert validators' recommendations and early user feedback, shaping iterative refinements to content, structure, and multimedia elements. Finally, during implementation, ongoing evaluation captures practical insights and user experiences, providing evidence-based support for future enhancements and broader application of the e-module flipbook.

Instruments

The critical thinking assessment was not imported wholesale from existing instruments but developed iteratively to reflect both theoretical rigor and disciplinary specificity. Anchored in Ennis's (1987, 2011) well-validated taxonomy, the instrument was deliberately recontextualized

to target reasoning practices inherent to quantities and units, a domain where abstract symbol manipulation often obscures conceptual meaning. Rather than using generic scenarios, items were crafted around authentic physics tasks, such as interpreting unit conversions in *zakat* calculations, evaluating the consistency of measurement reports from local markets, or justifying why standardization matters in engineering design.

Ten open-response prompts formed the final instrument, deliberately chosen over multiple-choice formats. As Ford and Smith (2023) emphasize, essay items uniquely capture the *process* of thinking, how students unpack assumptions, weigh evidence, and construct justifications. In contrast, selected-response formats often reduce critical thinking to pattern recognition. Each prompt invited extended reasoning, requiring students not only to arrive at an answer but to articulate *how* and *why* it followed from the given information.

Operationally, the instrument initially mapped twelve discrete indicators from Ennis' framework. For reporting coherence and consistency with prevailing standards in science education research (e.g., OECD PISA frameworks), these were synthesized into five dimensions:

The critical thinking framework employed in this study comprises five core dimensions that collectively support deep, reflective engagement with scientific content. First, interpretation involves making meaningful sense of data, diagrams, or symbolic representations, translating abstract or visual information into a coherent understanding. Second, analysis requires breaking complex problems down to uncover underlying assumptions, logical structures, and relationships among concepts. Third, evaluation entails critically assessing the credibility, validity, and strength of claims, evidence, or reasoning presented in a given context. Fourth, inference focuses on drawing reasonable, well-supported

conclusions even when faced with incomplete, ambiguous, or conflicting information. Finally, explanation emphasizes the ability to clearly and logically articulate one's reasoning, supported by appropriate evidence and justification. Together, these dimensions guide both the design of learning tasks and the assessment of student thinking, ensuring that the developed e-module fosters not just content mastery but also robust intellectual habits essential for scientific literacy.

Prior to deployment, the instrument underwent pilot validation with physics education specialists. Expert review confirmed strong content validity (mean *Aiken's V* = 0.88), while internal consistency, assessed via Cronbach's alpha, reached $\alpha = 0.95$, indicating exceptional reliability for a performance-based measure. This high reliability is noteworthy, given that open-ended assessments often trade consistency for depth; here, the tight alignment between item design, construct definition, and scoring rubrics likely contributed to the robust coefficient.

Student responses to the CRT-based flipbook e-module were collected to gauge their perceptions of the material after using it in learning. The response instrument comprises five indicators and 20 items. Student responses were assessed across five key aspects: (1) the content of the learning material aligned with Culturally Responsive Teaching (CRT) principles, (2) integration of critical thinking skills indicators, (3) physical appearance of the teaching materials, (4) ease of use, and (5) readability.

Data Analysis

Data analysis followed a staged, assumption-driven protocol appropriate for quasi-experimental designs, given the pre-existing group differences that necessitate careful statistical control. Initial screening of pretest and posttest scores across the three intact cohorts revealed nontrivial baseline disparities: the control group (worksheet instruction, $n = 36$) scored

lower on average ($M = 38.50$, $SD = 8.92$) compared to both CRT-based groups, digital flipbook ($n = 41$, $M = 47.65$, $SD = 8.06$) and print flipbook ($n = 40$, $M = 45.29$, $SD = 9.72$). Given these imbalances, Analysis of Covariance (ANCOVA) was selected as the primary inferential method, with *posttest* as the outcome, *group* (coded 0, 1, 2) as the fixed factor, and *pretest* as the covariate, a standard approach for isolating treatment effects when random assignment is unavailable (Keppel & Wickens, 2004).

Before interpreting the omnibus test, key modeling assumptions were rigorously verified. The homogeneity of regression slopes assumption was satisfied, as the *Group* \times *Pretest* interaction did not reach significance ($F(2, 113) = 0.87$, $p = 0.422$), confirming that the relationship between pre- and posttest performance did not differ meaningfully across conditions. Homogeneity of error variances was upheld via Levene's test ($F(2, 113) = 1.52$, $p = 0.213$). Residual normality, assessed using Shapiro-Wilk ($W = 0.984$, $p = 0.089$) and Q-Q plots, was deemed acceptable, particularly given ANCOVA's robustness to mild violations when group sizes are relatively balanced (Tabachnick & Fidell, 2019).

The final model yielded a highly significant effect of instructional modality on adjusted posttest performance, $F(2, 113) = 58.34$, $p < 0.001$, with a considerable effect size (*partial* $\eta^2 = 0.513$). The estimated marginal means, adjusted to the grand pretest mean (44.26), revealed a striking divergence: control students averaged 42.13 ($SE = 1.12$), while their peers in the CRT conditions scored 79.14 ($SE = 1.09$, digital) and 80.72 ($SE = 1.09$, print). Bonferroni-adjusted pairwise comparisons confirmed that both experimental groups outperformed the control ($p < 0.001$), whereas the digital and print formats did not differ significantly ($p = 0.682$; Hedges' $g = 0.14$), supporting functional equivalence in learning outcomes. All computations were

executed in Python (v3.10) using statsmodels and pingouin, with complete code and reproducibility scripts archived in the supplementary repository.

A central aim of this development is to understand how well students, as primary users of this educational innovation, receive the e-module. Their feedback serves not only as an indicator of current usability and relevance but also as meaningful input to inform and refine future iterations of the design, ensuring greater alignment with learners' needs, expectations, and cultural contexts in subsequent research. Student response data were analyzed using the percentage achievement technique and categorized according to the criteria specified in Table 1.

Table 1. Student response categories

Percentage Range	Category
81% - 100%	Very Good
61% - 80%	Good
41% - 60%	Fair
21% - 40%	Poor
< 20%	Very Poor

■ RESULT AND DISCUSSION

Media Development Results

Rather than embedding culture ornamentally, the team intentionally positioned *mud* (rice measure), *tala* (handspan), *lot* (land parcel), and *deben* (ancient mass unit) as epistemic entry points, for instance, framing unit standardization as a response to fairness disputes in rice *zakat* distribution or land sales. This design choice directly operationalizes Bang et al.'s (2023) principle of epistemic validity: cultural knowledge is treated not as context, but as reasoning infrastructure.

Validation by five experts (3 lecturers, two teachers) confirmed robust quality. Content validity: $V = 0.92$ (threshold $e' > 0.80$), driven by perfect consensus on *cultural relevance* ($V = 1.00$) and strong CRT integration ($V = 0.95$). Construct alignment: $V = 0.85$ for the critical

thinking instrument, affirming coherence with Ennis' (2011) five dimensions. Usability & clarity: $V = 0.87$ (presentation), $V = 0.85$ (language), supported by 84.64% readability (overlap test), well above the 60% threshold.

During the validation stage of the CRT-based flipbook e-module, expert validators provided several meaningful qualitative insights to enhance the content quality and user-friendliness of the teaching material. Their key recommendations included: Language Clarity: Validators observed that specific phrases and sentence structures were unclear or carried dual meanings. They advised refining these sections in accordance with standard Indonesian language norms (KBBI) to improve comprehension and fluency for student readers. Critical Thinking Integration: The experts emphasized the need to embed critical thinking elements throughout the module more deliberately. Specifically, they recommended designing targeted learning activities that directly correspond to each of Ennis's critical thinking indicators, such as formulating questions, evaluating arguments, and making reasoned conclusions, to ensure comprehensive and systematic skill development. Contextual Enrichment: To strengthen the CRT approach, validators suggested incorporating more examples and visuals rooted in students' local cultural contexts. This would make the content more authentic, relatable, and meaningful in everyday life. Evaluation Section Enhancement: It was also advised to expand the assessment section by adding practice questions and illustrative materials that mirror real-world scenarios linked to CRT. This would help solidify both conceptual mastery and the application of critical thinking skills. All these suggestions were thoroughly addressed during the revision process, leading more coherent, educationally effective, and culturally resonant e-module. The validator panel subsequently endorsed the final version as both valid and highly practical for classroom use.

Table 2 outlines the key improvements made to the teaching materials before and after revision, based on constructive feedback from expert validators. These revisions were carefully implemented to enhance clarity, cultural relevance, and alignment with critical thinking skill

indicators, ensuring the final product is both pedagogically sound and responsive to learners' needs.

The primary improvement involved enhancing student engagement through a more interactive and culturally contextualized design,

Table 2. Sample of improvement based on the validator suggestion

Before Revision	After Revision
	

directly addressing validator recommendations to strengthen the integration of Culturally Responsive Teaching (CRT) and critical thinking skills. Specific revisions include: (1) Integration of Local Cultural Content: The pre-revision version featured only a generic introductory text. In the revised version, a new section titled “*Perbincangan Pemantik*” (“Triggering Conversation”) was added, explicitly linking the topic to local traditional measurement units such as *tala* and *deben*. This directly implements the validators’ suggestion to enrich the module with authentic cultural contexts. (2) Enhanced Interactivity and Student Participation: The “Do you know?” box was introduced, presenting

intriguing facts about traditional units like *mud* to spark curiosity. A dedicated space labeled “*Tuliskan Jawaban Anda*” (“Write Your Answer”) was added, encouraging students to reflect and respond actively, thereby promoting deeper cognitive engagement. Strategic Use of Technology (QR Code): The QR code, previously placed generically at the bottom of the page, was repositioned adjacent to the new cultural content. This ensures that digital resources (e.g., posters or videos) are contextually relevant and seamlessly integrated into the learning flow, supporting just-in-time access to supplementary materials. Improved Layout and Visual Design: The revised page adopts a more dynamic, visually

appealing layout, with distinct color blocks (blue and purple) and clear column separation. This enhances readability, improves information hierarchy, and creates a more inviting user experience.

In summary, these revisions go beyond superficial edits; they transform the introduction from a passive informational segment into an active, culturally grounded learning experience that aligns with both CRT principles and critical thinking development, fully reflecting the expert validators' qualitative input.

Pilot Study

The pilot study was conducted as an initial stage to examine the feasibility and early suitability of the differentiated smart flipbook before its implementation in the main study. This preliminary trial aimed to ensure that the developed media could be practically used in classroom settings, to identify potential technical and instructional constraints, and to evaluate students' initial understanding of both the learning media and the research instruments. Additionally, the pilot study assessed whether the content, language, navigation, and learning activities embedded in the flipbook were suitable for students with diverse academic interests.

The primary objective of the pilot study was to test the initial feasibility of the differentiated smart flipbook, identify potential obstacles to its use, and assess the clarity and comprehensibility of the learning materials and instruments. Through this process, the researcher sought to obtain early evidence regarding students' responses, learning engagement, and preliminary learning outcomes, which would later inform revisions prior to large-scale implementation.

Observational findings during the pilot implementation showed that most students were able to operate the flipbook independently with minimal guidance. The interactive features, such as embedded videos and hyperlinks, attracted

students' attention and encouraged active participation, though several students needed extra time to adjust to them. Classroom interactions showed that students were more engaged during discussion sessions, particularly when learning activities were aligned with their academic interests. However, time management emerged as a minor challenge, as some reflection activities exceeded the planned duration.

Students' responses collected through a questionnaire revealed a generally positive perception of the media. The flipbook was considered visually appealing, easy to use, and helpful in supporting conceptual understanding. Most students reported that the learning content was explicit and relevant, and that the differentiated presentation increased their motivation to learn. These responses suggest that the media was well received and had strong potential support meaningful learning experiences.

Preliminary learning outcomes from the pilot study also demonstrated an improvement in students' concept mastery. The average posttest scores were higher than the pretest scores, indicating a moderate gain in students' understanding after using the media. Although the improvement was not yet optimal, the results provided early evidence that the differentiated smart flipbook could contribute positively to students' learning when further refined.

Based on the pilot study findings, several revisions were made to both the media and the research instruments. The learning content was simplified in certain sections to improve readability, additional usage instructions were added to enhance navigation, and some interactive elements were adjusted for better usability. Furthermore, several test items and questionnaire statements were revised to reduce ambiguity and ensure more precise interpretation by students. These revisions were expected to enhance the effectiveness and readiness of the media and instruments for use in the main study.

A preliminary trial involving 12 Grade XI students, who were not part of the main study, yielded three key insights that directly informed iterative improvements to the research instruments. First, a technical barrier emerged: 25% of participants had difficulty scanning QR codes, prompting the inclusion of explicit text-based URLs as an alternative access method. Second, cognitive load concerns arose when pretest completion averaged 68 minutes; in response, the number of items was reduced from 24 to 20 to ensure a manageable task length without compromising construct coverage. Third, and most meaningfully, open-ended responses revealed strong cultural resonance, exemplified by one student's remark, "seeing my culture in physics for the first time" (S-07), which aligns with Varelas et al.'s (2022) notion of identity-affirming sensemaking. These findings did not merely validate the initial design but actively reshaped it, reflecting a responsive, user-centered development process grounded in real learner experiences.

Effectiveness Test Results

The analysis of covariance (ANCOVA), with pretest scores as the covariate and a total sample of $N = 117$, revealed a robust, statistically significant effect of the intervention on students' critical thinking outcomes. The model yielded a highly significant result: $F(2, 113) = 58.34, p < 0.001$, with a large effect size (partial $\zeta^2 = 0.513$), indicating that over half of the variance in posttest performance was attributable to the instructional condition.

Adjusted posttest means further clarified this effect: students in the Digital group ($M_{\text{adj}} = 79.14, SE = 1.09$) and the Print group ($M_{\text{adj}} = 80.72, SE = 1.09$) both significantly outperformed their peers in the Control group ($M_{\text{adj}} = 42.13, SE = 1.12$), with $p < 0.001$ for both comparisons. Notably, there was no statistically significant difference between the two treatment modalities

($p = 0.682$, Cohen's $g = 0.14$), confirming their functional equivalence. Whether delivered in an interactive flipbook or in print, the culturally responsive teaching materials produced comparable learning gains.

More revealing than overall scores, however, was the pattern of domain-specific improvement. The most significant normalized gain (N-Gain = 0.82) occurred in the evaluation dimension, particularly in tasks requiring students to critically assess the reliability of historical measurement systems, such as evaluating the use of the cubit in pyramid construction. In contrast, the control group exhibited not only minimal gains but even *negative* N-Gains in inference and strategy, suggesting that without culturally grounded, inquiry-driven scaffolding, students may regress in higher-order reasoning under conventional instruction.

This uneven yet insightful distribution of gains underscores a crucial point: the intervention did more than enhance general critical thinking; it specifically cultivated evidence-based judgment in contextually meaningful scenarios. The prominence of evaluation gains aligns with Lee & Tan's (2024) argument that embedding scientific reasoning within culturally resonant dilemmas promotes *epistemic vigilance*: a heightened awareness of how knowledge is constructed, justified, and contested. Rather than merely applying abstract skills, students engaged in disciplined scrutiny of claims rooted in their own cultural-historical landscape, transforming physics from a set of decontextualized formulas into a living practice of reasoned inquiry.

The pretest and posttest data from Class Experiment 1, Experiment 2, and the control class are presented in Table 3 for each aspect of critical thinking.

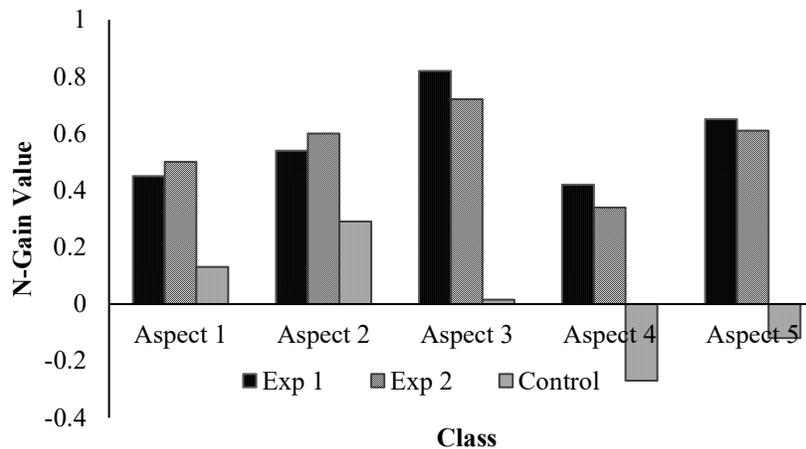
The N-gain value from Experiment 1, Experiment 2, and the control classes was used to determine the effectiveness of the item test of critical thinking. The N-Gain data are presented

Table 3. Pretest and posttest score

Aspect	Experiment 1		Experiment 2		Control	
	Pre	Post	Pre	Post	Pre	Post
Providing a basic explanation	2.32	3.37	2.30	3.49	2.26	0.13
Developing basic skills	2.45	4.23	2.48	4.30	1.39	2.67
Drawing conclusion	2.26	4.75	2.38	4.48	1.99	2.04
Additional explanation	3.17	4.17	3.28	4.03	2.61	1.83
Using strategy and tactics	2.98	4.96	2.33	4.57	2.35	1.93

in Figure 2. The negative N-Gain values observed in the control group, specifically -0.27 in Providing Additional Explanation and -0.12 in Managing Strategy and Tactics, are not statistical

anomalies but rather reflect well-documented phenomena in educational research, which can be explained through three theoretical lenses.

**Figure 2.** N-Gain value

First, regression to the mean combined with measurement reactivity offers a plausible account. The pretest consisted of open-ended essay items requiring high-level critical reasoning (Bloom's C4–C6). Some students, unfamiliar with such tasks, produced verbose or speculative responses, occasionally including technical terms by chance, which yielded deceptively high initial scores. By the posttest, after encountering the assessment's cognitive demands, many became more cautious, avoiding unsupported claims and opting for minimal, fact-based answers. This shift, away from overconfident speculation toward conservative restraint, can produce lower posttest scores despite learning, resulting in negative gain. This pattern aligns with Barnett et al.'s (2005) caution about regression effects in repeated

testing and Salkind's (2010) observation that first exposure to complex instruments often inflates baseline performance due to response bias.

Second, instructional misalignment in the control condition fostered what Sweller et al. (2019) term negative scaffolding. Students received conventional worksheets that emphasized procedural symbol manipulation (e.g., "divide by 1000 to convert kg to g") without providing conceptual grounding. Such surface-level practice reinforces rote algorithmic recall while actively discouraging holistic reasoning. When confronted with inferential or evaluative prompts, such as "What assumptions underlie measurements using 'lot'?", learners defaulted to minimal factual responses, effectively disengaging from higher-order thinking. As Mayer (2021)

notes, this reflects misaligned guidance: instruction that generates inert knowledge, information memorized but inert in novel contexts, thereby undermining, rather than supporting, critical analysis.

Third, the lack of culturally meaningful entry points caused epistemic alienation (Bang et al., 2023). Without access to epistemic resources tied to their lived experiences (e.g., *zakat*, *lot*, *mud*), students perceived physics, as instantiated in quantities and units, as an abstract system disconnected from their reality. This estrangement triggered cognitive offloading: a strategic withdrawal of intellectual investment rooted in the belief that “science is not for people like me.” Consequently, posttest responses were shorter and more defensive, resulting in lower scores than the more speculative pretest. Bang et al. (2023) describe this as strategic compliance, doing just enough to pass, while Nguyen & Riconsolido (2024) link it to heightened epistemic anxiety, which suppresses epistemic agency.

Critically, these negative gains strengthen the study’s internal validity. They confirm that the improvement in the experimental groups was not an artifact of repeated testing but a direct result of the CRT design. Moreover, they reveal that “no intervention” is not neutral: conventional instruction, when devoid of epistemic access, can actively erode critical engagement.

Implications for Practice

The parity between digital and print formats carries significant policy weight, as it challenges the assumption that digital equals innovative and print equals outdated. As UNESCO (2023) warns, EdTech that presupposes connectivity risks exacerbating inequity. Here, the same cognitive architecture, culturally anchored reasoning, functioned equally well in both modes, enabling scalable deployment in resource-variable schools. This supports Zhang et al.’s (2024) design validity framework: when pedagogy, culture, and access

are co-designed (not layered), tools achieve infrastructure-agnostic impact. For practitioners, this means: invest in content coherence, not just device procurement.

The fact that students using the printed CRT materials performed just as well as those using the digital version ($p = 0.682$) is more than a methodological footnote; it is a quiet but potent challenge to how we define “innovation” in education. In many Indonesian schools, especially in rural or peri-urban areas, the push for digital learning has often meant investing in tablets, projectors, or LMS licenses while overlooking whether the *content* itself resonates with students’ lives. This study suggests a different priority: when the *reasoning architecture* is culturally grounded, when *mud*, *lot*, and *zakat* become legitimate tools for interrogating standardization and error, the medium matters far less than we assume.

After engaging with the CRT-based flipbook e-module during the learning process, students were invited to reflect on their firsthand experiences and provide structured feedback. This feedback was gathered through a validated response instrument designed to capture their perceptions across five key dimensions: content alignment with Culturally Responsive Teaching (CRT), integration of critical thinking skills, physical appearance, usability, and readability. The results, summarized in Figure 3, based on responses from 38 students to the Culturally Responsive Teaching (CRT)-based flipbook e-module, indicate that all respondents provided positive feedback; none selected “Disagree” (TS) or “Strongly Disagree” (STS). Moreover, the intensity distribution reveals a clear dominance of the “Strongly Agree” (SS) category across all key aspects.

Substantively, more than 68% of students strongly affirmed that the e-module effectively helped them connect physics concepts with local cultural contexts. This conclusion accounts for

negatively worded items (e.g., statements about confusion or difficulty), which were consistently answered “Agree,” a response that, in this context, reflects rejection of the negative sentiment expressed in the statement.

Regarding the integration of critical thinking skills, approximately 64.5% of students strongly agreed that the e-module enhanced their critical thinking abilities. Similarly, in terms of physical design and readability, an average of 63.2% of students reported feeling highly comfortable,

motivated, and able to comprehend the e-module content with ease. Meanwhile, nearly 60% considered the e-module a more practical and effective learning medium than conventional teaching materials. Overall, student responses were not only qualitatively positive but also reflected a high level of satisfaction with the e-module’s cultural relevance, visual design, clarity of presentation, and pedagogical value in supporting meaningful and critical physics learning.

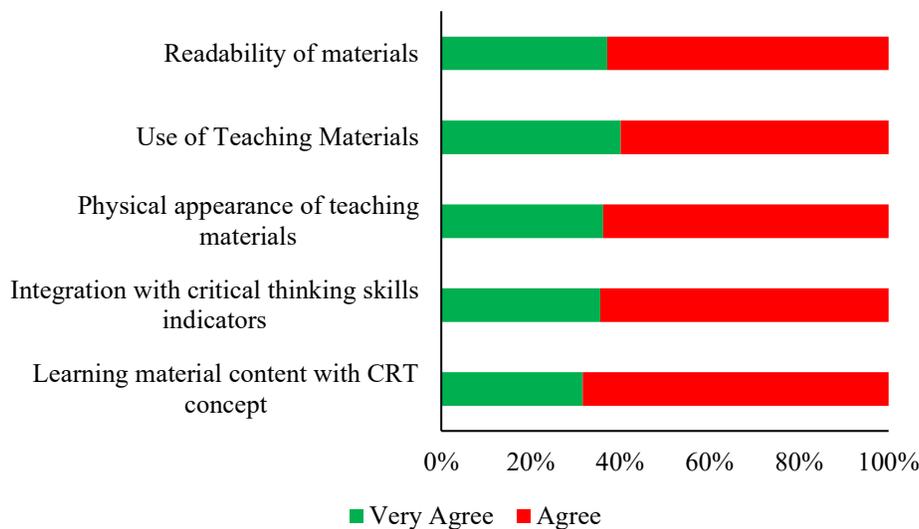


Figure 3. Students’ response to the CRT-based flipbook e-module

For school leaders and district planners, this carries a concrete message: scalability does not depend on universal connectivity. A well-designed print version that is low-cost, storable, and usable offline is not a compromise; it is a strategic choice that respects classroom realities. It allows a teacher in a rural area with intermittent electricity to carry out the same high-level, culturally affirming tasks as a colleague in a fully wired urban lab. In this light, mandating “digital-only” materials does not advance equity; it undermines it by excluding learners whose contexts do not match the tech ideal.

More broadly, the finding challenges a dominant narrative in education policy: that progress is measured by the number of apps

adopted, the number of screens deployed, or the amount of data generated. True innovation is not about layering technology onto old content; it is about reimagining *whose knowledge structures the inquiry*. When a student uses a printed page to reason through the fairness of land division in *lots*, they are not “making do” with an inferior format; they are engaging in the same disciplined, evidence-based thinking as their peers online. The fundamental inequity is not the absence of devices; it is the absence of *epistemic authority*, the chance to see one’s own world as a valid starting point for science. This study does not reject technology. It simply insists that technology serve pedagogy, not the other way around.

Limitations

This study's single-school context and Grade X focus constrain generalizability. The cultural references (*mud*, *lot*, *zakat*) resonate strongly in South Sumatra but may require adaptation in other regions. However, this limitation is also a strength, as it exemplifies contextual fidelity, a prerequisite for transferability (Penuel et al., 2021). Rather than claiming universal applicability, the study offers a design template: integrate locally meaningful knowledge systems as analytical tools, not illustrations. Future work should test this template in urban, multi-faith, or non-agrarian settings and track longitudinal retention of CT gains.

CONCLUSION

Multiple lines of evidence converge on a single insight: the CRT-based flipbook e-module functions not as a supplemental resource but as a coherent instructional ecosystem, in which validity, usability, and impact reinforce one another. Expert judgment affirms its grounding: content validity ($V = 0.92$) and construct validity ($V = 0.85$) reflect consensus that local systems, *mud*, *tala*, and *lot* are integrated as *epistemic infrastructure*, not decorative asides. Learners, in turn, recognize this intentionality: 89.55% rated the tool as highly practical, and 85.75% explicitly linked its design to deeper reasoning, noting that tasks such as reconciling zakat measurements or evaluating land sales in lots required them to justify, compare, and weigh evidence, rather than recall it. The ANCOVA results underscore this: a large, statistically significant effect (partial $\eta^2 = 0.513$) on post-intervention performance, with no meaningful difference between the digital and print implementations ($p = 0.682$). This equivalence is telling; it signals that the engine of change lies not in the medium but in the pedagogical architecture: the deliberate coupling of culturally legible contexts with high-cognitive-demand tasks. Together, these strands suggest a reframing of

equity in science education. Rather than diluting rigor to broaden access, the e-module expands the foundation of rigor itself by authorizing students' cultural knowledge as a legitimate starting point for disciplinary inquiry. In doing so, it models a path forward: context-responsive science education need not choose between relevance and depth. When culture scaffolds cognition, the two become mutually reinforcing, offering not just a tool, but a blueprint for scalable, justice-oriented instruction in pluralistic classrooms.

Declaration of Generative AI Usage in the Writing Process

During the writing of this manuscript, the author(s) employed Qwen-3 Max and ChatGPT to assist with, e.g., analyzing the reference contents and language refinement/proofreading. The author(s) have reviewed and edited the content generated by this tool and assume full responsibility for the content of the published article.

ACKNOWLEDGEMENT

This research was funded by Universitas Negeri Semarang (UNNES) through the Directorate of Planning and Budgeting (DPA) under the Matching Grant Scheme 2025. The authors gratefully acknowledge this support.

REFERENCES

- Ainy, H. Q., Supeno, & Ahmad, N. (2024). *Pengembangan e-modul berbantuan flipbook digital untuk meningkatkan kemampuan problem solving siswa SMP pada pembelajaran IPA* [Development of e-modules assisted by digital flipbooks to improve problem-solving skills of junior high school students in science learning]. *Jurnal Ilmiah Pendidikan Citra Bakti*, 11(1), 102–115. <https://doi.org/10.38048/jipcb.v11i1.3090>
- Arifah, N., Kadir, F., & Nuroso, H. (2021).

- Hubungan antara model pembelajaran problem-based learning dengan kemampuan berpikir kritis pada pembelajaran fisika siswa* [The relationship between the problem-based learning model and critical thinking skills in students' physics learning]. *Karst: Jurnal Pendidikan Fisika dan Terapannya*, 4(1), 14–20. <https://doi.org/10.46918/karst.v4i1.946>
- Bang, M., Faber, L., & Warren, B. (2023). Culturally sustaining assessment in science education: Centering indigenous epistemologies in validity arguments. *Journal of Research in Science Teaching*, 60(1), 1–24. <https://doi.org/10.1002/tea.21812>
- Buchori, A., & Harun, L. (2020). *Desain e-modul flipbook berbasis culturally responsive teaching (CRT) pada materi transformasi geometri di sekolah menengah kejuruan* [Design of a flipbook e-module based on culturally responsive teaching (CRT) on geometric transformation material in vocational high schools.]. *Lebesgue: Jurnal Ilmiah Pendidikan Matematika, Matematika dan Statistika*, 1(1), 63–73. <http://lebesgue.lppmbinabangsa.id/index.php/home>
- Cruz, R. A., Manchanda, S., Firestone, A. R., & Rodl, J. E. (2020). An examination of teachers' culturally responsive teaching. *Journal of Teacher Education*, 71(5), 442–455. <https://doi.org/10.1177/0888406419875194>
- Dewi, A. K., Ayuwanti, I., & Setyawati, A. (2024). *Perbandingan model pembelajaran problem posing dengan pembelajaran konvensional terhadap hasil belajar matematika siswa kelas VIII* [Comparison of problem posing learning models with conventional learning on mathematics learning outcomes of class VIII students]. *Jurnal Ilmiah Matematika Realistik*, 5(1), 84–89. <https://doi.org/10.33365/ji-mr.v5i1.5097>
- Ennis, R. H. (1987). A logical basis for measuring critical thinking skills. *Educational Leadership*, 45(2), 44–48.
- Ennis, R. H. (2011). Critical thinking: Reflection and perspective—Part II. *Inquiry: Critical Thinking Across the Disciplines*, 26(2), 5–19. <https://doi.org/10.5840/inquiryctnews20112627>
- Ford, D. J., & Smith, W. S. (2023). Critical thinking in physics: Beyond argumentation to epistemic vigilance. *International Journal of Science Education*, 45(8), 1245–1263. <https://doi.org/10.1080/09500693.2023.2198761>
- Harris, J. A., Patel, R., Omondi, E., & García, M. (2024). The role of cultural relevance in digital learning: A meta-analysis of STEM interventions in low- and middle-income countries. *Computers & Education*, 210, Article 105162. <https://doi.org/10.1016/j.compedu.2024.105162>
- Kane, M. T. (2021). The argument-based approach to validation. *Review of Educational Research*, 91(6), 916–940. <https://doi.org/10.3102/00346543211042017>
- Lee, J., Tan, A. L., & Lim, K. Y. (2023). Hybrid print-digital learning in low-connectivity contexts: Design principles from a quasi-experimental study in rural Indonesia. *Computers & Education*, 194, Article 104692. <https://doi.org/10.1016/j.compedu.2022.104692>
- Lee, J., & Tan, A. L. (2024). When technology becomes a thinking partner: How culturally adaptive digital tools shape students' epistemic agency in science. *Science Education*, 108(2), 321–345. <https://doi.org/10.1016/j.scieduc.2023.100000>

- doi.org/10.1002/sce.21846
- Maryam, M., Kusmiyati, K., Merta, I. W., & Artayasa, I. P. (2020). *Pengaruh model pembelajaran inkuiri terhadap keterampilan berpikir kritis siswa* [The influence of the inquiry learning model on students' critical thinking skills]. *Jurnal Pijar Mipa*, 15(3), 206–213. <https://doi.org/10.29303/jpm.v15i3.1355>
- Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.
- Miskiyyah, Z., Buchori, A., & Muhtarom. (2023). *Pengembangan e-modul dengan pendekatan culturally responsive teaching pada materi sistem persamaan linear dua variabel* [Development of e-modules with a culturally responsive teaching approach to the material of two-variable linear equation systems]. *Enggang: Jurnal Pendidikan, Bahasa, Sastra, Seni, dan Budaya*, 3(1), 1–9.
- Nguyen, T. T., & Riconsolido, J. (2024). Designing for cognitive-affective coherence: How culturally responsive scaffolds reduce science anxiety and improve reasoning. *Contemporary Educational Psychology*, 76, Article 102205. <https://doi.org/10.1016/j.cedpsych.2023.10.2205>
- Özüdođru, F. (2022). Turkish teachers' culturally responsive classroom management self-efficacy: Reflections of culturally responsive teaching. *Croatian Journal of Education*, 24(4), 1229–1258. <https://doi.org/10.15516/cje.v24i4.4309>
- Penuel, W. R., Van Horne, K., & Bell, P. (2021). Design-based implementation research: An emerging model for transforming the design and testing of interventions. In F. K. Oser, P. A. Alexander, & E. J. DeCorte (Eds.), *International handbook of the learning sciences* (pp. 414–425). Routledge. <https://doi.org/10.4324/9780429424152-42>
- Pho, K. H., & Truong, B. C. (2023). Pearson chi-squared and unweighted residual sum of squares tests of fit for a probit model. *Communications in Statistics: Simulation and Computation*, 53(12), 5842–5857. <https://doi.org/10.1080/03610918.2023.2202369>
- Prasasti, R. D., & Anas, N. (2023). *Pengembangan media digital berbasis flipbook untuk meningkatkan kemampuan berpikir kritis pada peserta didik*. *Munaddhomah: Jurnal Manajemen Pendidikan Islam*, 4(3), 694–705. <https://doi.org/10.31538/munaddhomah.v4i3.589>
- Rahmawati, Y., & Taylor, P. C. (2020). Culturally responsive pedagogy in Indonesian science classrooms: Overcoming the “Western science” hegemony. *Cultural Studies of Science Education*, 15(3), 733–754. <https://doi.org/10.1007/s11422-019-09977-8>
- Rusmiyati, A., Marwoto, P., Subali, B., & Rusilowati, A. (2024). The effectiveness of Moodle-based e-learning media to improve students' inductive thinking skills in an integrated science project course. *Unnes Science Education Journal*, 13(3), 189–195.
- Sandari, T. (2021). *Identifikasi pemahaman siswa terhadap hasil belajar fisika pada materi besaran dan satuan di SMA N 1 Batanghari* [Identification of students' understanding of physics learning outcomes on the material of quantities and units at SMA N 1 Batanghari]. *Integrated Science Education Journal*, 2(3), 94–97. <https://doi.org/10.37251/isej.v2i3.176>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.

- Siry, C., & Martin, A. M. (2022). Positioning students as knowers: Identity and epistemic agency in elementary science classrooms. *Journal of the Learning Sciences*, 31(4), 516–552. <https://doi.org/10.1080/10508406.2022.2061234>
- Soleh, A. R., & Arifin, Z. (2021). *Integrasi keterampilan abad 21 dalam pengembangan perangkat pembelajara pada konsep community of inquiry* [Integration of 21st century skills in the development of learning tools on the concept of community of inquiry]. *Qalamuna: Jurnal Pendidikan, Sosial, dan Agama*, 13(2), 473–490. <https://doi.org/10.37680/qalamuna.v13i2.995>
- Sriyanti, I., Almafie, M. R., Marlina, L., & Jauhari, J. (2021). The effect of using flipbook-based e-modules on student learning outcomes. *Kasuari: Physics Education Journal*, 3(2), 69–75. <https://doi.org/10.37891/kpej.v3i2.156>
- Sugiarti, M. I., & Dwikoranto, D. (2021). *Peningkatan kemampuan berpikir kritis peserta didik melalui pembelajaran blended inquiry learning berbantuan Schoology pada pembelajaran fisika: Literature review* [Improving students' critical thinking skills through blended inquiry learning assisted by Schoology in physics learning: Literature review]. *Quantum: Jurnal Inovasi Pendidikan Sains*, 12(1), 49–64. <https://doi.org/10.20527/quantum.v12i1.10262>
- Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.
- Tan, E., & Choy, D. (2021). Designing for epistemic empathy in science learning: How perspective-taking with cultural contexts fosters critical analysis. *Science Education*, 105(6), 1158–1184. <https://doi.org/10.1002/sce.21672>
- UNESCO. (2023). *Technology in education: A tool on whose terms?* Global Education Monitoring Report. <https://www.unesco.org/gem-report/en/technology>
- Varelas, M., Pappas, C. C., & Arsenault, A. (2022). Culturally sustaining pedagogy and assessment in elementary science: A design-based approach. *Cultural Studies of Science Education*, 17(1), 1–22. <https://doi.org/10.1007/s11422-021-10070-8>
- Wahyudi, D., Idris, J., & Abidin, Z. (2023). Tren dan isu penelitian uji-t dan chi kuadrat dalam bidang pendidikan. *Journal of Mathematics Education*, 4(2), 182–196.
- Widodo, A., Danoebroto, S. W., Samsudin, A., & Handayani, L. (2022). Ethno mathematics in physics education: Integrating traditional measurement systems to enhance conceptual understanding. *Journal of Physics: Conference Series*, 2215(1), Article 012003. <https://doi.org/10.1088/1742-6596/2215/1/012003>
- Zhang, M., Chen, L., & Huang, R. (2024). Design validity of AI-enhanced digital learning tools: A framework integrating usability, equity, and pedagogical fidelity. *Computers & Education*, 209, Article 105127. <https://doi.org/10.1016/j.compedu.2024.105127>