**Reforma: Jurnal Pendidikan dan Pembelajaran**
**Faculty of Teacher Training and Education**
**Universitas Islam Lamongan**
https://jurnalpendidikan.unisla.ac.id/index.php/reforma

*November 2026 Vol. 16, No. 2*
*e-ISSN*: 2621-4172
*p-ISSN*: 2503-1228
*pp. 151-167*

# Bridging Policy and Practice in Competency-Based Islamic History Education: Multi-Phase Alignment and Pedagogical Reconstruction

**[1*]Marwal Blantara Susetyo, [2]Kharisma Pratama Nurhidayat, [3]Muhammad Fahri Ali Husaeni, [4]Mifta Ghina Dzakiya, [5]Ridwan Maulana**

[1]Master of Islamic Religious Education, Faculty of Social Sciences Education, Universitas Pendidikan Indonesia, Jl. Dr. Setiabudhi No. 229, Bandung, West Java, Indonesia. Postal Code: 40154
[2,3,4] Master of Islamic Religious Education, Universitas Islam Negeri Sunan Gunung Djati Bandung, Jl. Cimencrang, Panyileukan, Gedebage, Bandung, West Java, Indonesia. Postal Code 40292
[5] Islamic Religious Education Study Program, Majalengka University, Jl. Raya K.H. Abdul Halim No. 103, Majalengka, West Java, Indonesia. Postal Code: 45418

*Corresponding Author e-mail: marwalblantara@upi.edu

Received: 01 August 2026; Revised: 21 August 2026; Accepted: 09 October 2026; Published: 14 November 2026

**Abstract**

*This study investigates the vertical alignment and pedagogical translation of competency-based curricula across developmental phases, addressing the critical research gap between normative policy mandates and actual classroom enactment. Employing a qualitative document analysis and policy-practice synthesis, we evaluated a targeted corpus of national policy directives (Capaian Pembelajaran) alongside peer-reviewed empirical evaluations (2020–2026) using Bloom and Marzano's cognitive taxonomies. The analysis reveals a profound, systemic policy-practice gap across three dimensions: a cognitive mismatch between mandated higher-order thinking and entrenched rote-memorization, pedagogical barriers due to teacher-centered transmission, and assessment inertia. While the policy outlines a progressive, multi-phase cognitive continuum—transitioning from lower-secondary dynastic narratives to upper-secondary global civilizational synthesis—actual instruction remains decoupled from these standards. To resolve this, we propose a reconstructed pedagogical framework that structurally aligns instructional models with curriculum phases, prioritizing Project-Based Learning (PjBL) for digital artifact creation in lower-secondary levels and Problem-Based Learning (PBL) for historical inquiry in upper-secondary levels. This study advances curriculum scholarship by offering a replicable, phase-aligned model for religious humanities education, demonstrating how structured active learning can bridge universal policy-practice divides in global competency-based reforms.*

**Keywords**: *Competency-based education; Curriculum alignment; Islamic history education; Pedagogical reconstruction; Merdeka Curriculum.*

**How to Cite:** Susetyo, M. B., Nurhidayat, K. P., Husaeni, M. F. A., Dzakiya, M. G., & Maulana, R. (2026). Bridging policy and practice in competency-based Islamic history education: Multi-phase alignment and pedagogical reconstruction. *Reforma: Jurnal Pendidikan dan Pembelajaran*, 16(2), 151–167. https://doi.org/10.30736/reforma.v16i2.1444

https://doi.org/10.30736/reforma.v16i2.1444

Copyright© 2026, Susetyo, Marwal Blantara
This is an open-access article under the CC-BY-SA License.

## INTRODUCTION

Globally, translating competency-based policy mandates into classroom realities remains a persistent, systemic challenge in secondary education (Majani, 2026). In the field of religious humanities, this policy-practice friction is acutely visible

in Islamic History Education (*Sejarah Peradaban Islam* or SPI) under Indonesia's recently implemented *Kurikulum Merdeka* (Suari et al., 2023). While national curriculum reforms demand a swift transition from passive knowledge transmission toward active, student-centered historical reasoning, the operational enactment of these mandates faces immediate pedagogical hurdles (Cheema et al., 2026). The urgency of addressing this gap in SPI is paramount; SPI is not merely an academic subject but a primary vehicle for moral, civic, and moderate religious character formation (*rahmatan lil-ʿālamīn*) in Indonesian secondary schools (Suari et al., 2023; Wardani, 2023). When top-down national standards demand a progressive continuum of cognitive synthesis, yet actual classroom instruction remains stuck in static memorization, the core developmental and character-building promise of the curriculum reform is profoundly compromised (Shi et al., 2026).

SPI represents a conceptually rich and meaningful site for investigating global curriculum reform because it sits at the complex intersection of critical historical inquiry, ethical reflection, and religious value internalization (Wardani, 2023). Unlike technical or scientific subjects where competencies are primarily procedural or technical, competency-based SPI requires a dual-layered translation. Students are expected to transition from memorizing historical timelines to evaluating complex socio-political cause-and-effect relationships, while simultaneously developing moderate religious values and global citizenship (Suari et al., 2023; Wardani, 2023). Analyzing how educators navigate these multidimensional expectations offers critical, generalizable insights into the limits and possibilities of competency-based humanities education globally (Cheema et al., 2026).

In SPI classrooms, this central implementation gap manifests as a concrete structural misalignment across cognitive, pedagogical, and assessment dimensions. Empirically, while the national Learning Outcomes (*Capaian Pembelajaran* or CP) mandate that students "analyze" complex historical developments such as the socio-political dynamics of Islamic empires or the cultural adaptation strategies of the Wali Songo (Maylawati et al., 2025) actual classroom observations reveal that SPI instruction is heavily reduced to passive factual recall. Students are typically assessed on their ability to memorize the names of rulers, dates of battles, and chronological lists of treaties (Mpate, 2025). Pedagogically, instead of designing phase-appropriate inquiry or active projects, teachers frequently rely on traditional, teacher-centered lectures due to deficits in Technological Pedagogical Content Knowledge (TPACK) and a lack of engaging historical learning media (Kabir, 2023; Mwangi et al., 2026). Assessment practices further entrench this gap; traditional multiple-choice tests focused on low-level recall consistently override the performance-based and portfolio evaluations required to measure authentic historical empathy and critical values (Mujiburrahman et al., 2025; Shi et al., 2026).

The educational consequences of this prolonged misalignment are highly detrimental to student development. When daily SPI instruction remains decoupled from policy standards, students suffer from weak historical reasoning and limited cognitive agency, viewing history as a static repository of facts rather than an active process of inquiry (Mpate, 2025). Furthermore, a reliance on rote learning leads to superficial value internalization, where the ethical lessons of history fail to shape student character or civic identity in meaningful ways (Wardani, 2023). This

misalignment also creates an acute disconnect between assessment and actual learning outcomes, where students are evaluated on factual recall rather than their capacity to synthesize complex global dynamics (Mujiburrahman et al., 2025). Addressing this gap is therefore not merely an administrative necessity, but an essential step to ensure that competency-based reforms deliver their intended developmental outcomes (Maylawati et al., 2025; Mujiburrahman et al., 2025).

To contextualize these challenges, previous scholarship on the *Kurikulum Merdeka* has extensively documented broad school readiness, administrative hurdles, and general stakeholder perceptions (Maylawati et al., 2025; Wardani, 2023). Despite these valuable insights, a critical gap remains in the literature. Existing studies have failed to examine how specific cognitive progressions and vertical curriculum alignments behave across multi-phase developmental continuums in religious humanities (Kasuga & Kalolo, 2025). Consequently, how high-level operational verbs in national policies translate into actual classroom enactment remains unresolved. This study addresses this specific gap by conducting a rigorous, document-based analysis of the vertical coherence between policy mandates and practical instructional pathways in SPI (Christopher & Jones, 2025; Kasuga & Kalolo, 2025).

The novelty of this study lies in its focus on multi-phase curriculum alignment and pedagogical reconstruction. Investigating multi-phase alignment is crucial because competency-based learning is inherently cumulative. Without structured cognitive progression, the transition from lower-secondary descriptive learning (Phase D) to upper-secondary analytical synthesis (Phase F) becomes fragmented, leading to severe learning loss and instructional disjointedness (Derouich, 2025; Kasuga & Kalolo, 2025). By illustrating how learning objectives can scale progressively, this study demonstrates why systemic curriculum coherence is vital for preventing the decoupling of policy from classroom enactment.

Rather than presenting a broad or procedural analysis, this study systematically analyzes the vertical cognitive progression of Learning Outcomes (*Capaian Pembelajaran*) across Phase D, Phase E, and Phase F of the national SPI curriculum, and reconstructs these policy directives into practical, phase-aligned instructional models. Specifically, we reconstruct the SPI pedagogical framework by mapping Project-Based Learning (PjBL) for digital artifact creation to Phases D and E, and Problem-Based Learning (PBL) for critical historical inquiry to Phase F, integrated with a sustainable, school-embedded teacher professional development (TPD) pathway (Nugraha et al., 2023; Pratami et al., 2024). This precise reconstruction provides educators with actionable, phase-specific pedagogical strategies to bridge the policy-practice gap directly in the classroom.

To address these critical issues, this study is structured around three interrelated lines of inquiry. First, it investigates how the cognitive progression and vertical alignment of learning outcomes are structured across Phase D, Phase E, and Phase F of the national Islamic History Education curriculum. Building upon this structural foundation, the study uncovers the specific, empirical policy-practice gaps that persist between these high-level national mandates and actual secondary classroom enactment. Ultimately, it explores how instructional models specifically Project-Based Learning (PjBL) and Problem-Based Learning (PBL) alongside systemic teacher professional development frameworks, can be reconstructed to effectively

bridge these alignment gaps and establish constructive curriculum coherence in daily practice.

## METHOD

### Research Design and Methodological Justification

This study employs a qualitative document analysis (QDA) integrated with a policy-practice synthesis design. Methodologically, a document-based synthesis is far more suitable than localized, field-based inquiry for evaluating vertical curriculum alignment across national developmental phases. While a field-based study (such as classroom observations in a few select schools) offers deep idiographic insights, it is geographically bounded and cannot capture the systemic, nationwide vertical coherence of curriculum mandates. By pairing QDA of national policy decrees with a systematic synthesis of 28 peer-reviewed empirical studies documenting classroom realities across diverse Indonesian provinces (2020–2026), this design achieves a macro-to-micro evaluation of curriculum coherence that is generalizable to national policy-practice dynamics.

The boundary conditions of this desk-based approach are acknowledged: because this study relies on official policy texts and published empirical evaluations, it is bounded by the transparency of state-mandated documentation and the methodological scopes of existing research. It does not capture real-time, unrecorded pedagogical improvisations or individual teacher motivations that primary qualitative fieldwork might reveal.

### Research Setting and Pedagogical Context: Developmental Phases D, E, and F

Rather than treating curriculum phases as mere administrative labels, this study approaches them as a developmental curriculum continuum designed to scaffold students' cognitive and moral progression in Islamic History Education (SPI) across secondary education (Table 1).

**Table 1.** Pedagogical and Cognitive Framework of SPI Across Curriculum Phases

| Curriculum Phase | Educational Level & Age | Core Pedagogical Focus | Cognitive & Moral Progression Target |
|---|---|---|---|
| Phase D | Lower Secondary (Grades 7–9; Ages 12–15) | Narrative and Descriptive Foundations | Transitioning from concrete historical storytelling (e.g., local Islamic expansion, dynastic narratives) to identifying basic cause-and-effect relationships. Focuses on developing historical empathy and initial moral reflection. |
| Phase E | Transitional Upper Secondary (Grade 10; Ages 15–16) | Conceptual and Structural Mapping | Moving toward structural historical reasoning. Students analyze the institutional, cultural, and socio-political dimensions of Islamic civilizations, mapping local events to global historical shifts. |
| Phase F | Advanced Upper Secondary (Grades 11–12; Ages 16–18) | Analytical Synthesis and Critical Inquiry | High-level historiographical inquiry. Students critically evaluate multiple historical perspectives, reconstruct complex socio-political dynamics, and internalize universal moderate values (*rahmatan lil-ʿālamīn*) to form active civic identities. |

Analyzing this complete secondary continuum is methodologically critical. Competency-based learning is inherently cumulative; therefore, tracking progression across Phases D, E, and F allows us to identify the precise points of "cognitive fracture" where vertical coherence fails, particularly during transitional years (Phase E) when students are expected to pivot from descriptive recall to complex analytical synthesis (Christopher & Jones, 2025; Derouich, 2025).

**Data Corpus and Selection Criteria**

The data corpus is bifurcated into two distinct, highly specified components: the Policy Document Corpus and the Empirical Literature Corpus (Table 2).

**Table 2.** Overview of the Analyzed Data Corpus

| Corpus Component | Document Type & Source | Year / Version | Specific Selection Criteria / Focus | Quantity |
|---|---|---|---|---|
| Policy Document Corpus | National Learning Outcomes (*Capaian Pembelajaran* / CP) for SPI; issued by the Ministry of Education, Culture, Research, and Technology (MoECRT). | Decree No. 033/H/KR/2022 and updated Decree No. 032/H/KR/2024. | Official national decrees mandating the learning targets, 3 operational verbs, and content areas for SPI across Phases D, E, and F. | 2 Core Decrees (encompassing 3 developmental phases) |
| Policy Document Corpus | Institutional Learning Objectives (*Tujuan Pembelajaran* / TP) and Exemplar Lesson Plans (*Modul Ajar*). | MoECRT Curriculum Center Exemplars (2023–2025). | Official exemplar operationalizations showing how CP is translated into lesson-level objectives. | 15 Exemplar Modules |
| Empirical Literature Corpus | Peer-reviewed journal articles, conference proceedings, and empirical evaluations. | Published between 2020 and 2026. | Empirical studies specifically investigating SPI or secondary religious humanities classroom implementation, teacher readiness, pedagogical challenges, and assessments under the *Kurikulum Merdeka*. | 28 Empirical Studies |

**Systematic Literature Search and Screening Procedure**

To construct the Empirical Literature Corpus with maximum transparency and reproducibility, we executed a systematic search and screening protocol mirroring the PRISMA guidelines (Page et al., 2021). The search was conducted across four major academic databases: Scopus, Google Scholar, SINTA (Science and Technology Index Indonesia), and ERIC.

1. Search Strings: The primary search string was: ("Kurikulum Merdeka" OR "Merdeka Curriculum") AND ("Sejarah Kebudayaan Islam" OR "Islamic

History Education" OR "SPI") AND ("implementasi" OR "tantangan" OR "classroom" OR "practice" OR "learning gap")

2. Inclusion Criteria: Articles were included if they: (a) were peer-reviewed and published between 2020 and 2026; (b) focused explicitly on SPI or secondary-level Islamic religious education (*Pendidikan Agama Islam* or PAI) with a dedicated historical curriculum component; and (c) provided empirical evidence of classroom enactment, pedagogical practices, teacher challenges, or assessment methods.
3. Exclusion Criteria: Articles were excluded if they: (a) were speculative opinion pieces, news reports, or non-peer-reviewed blog posts; (b) focused on primary education (elementary school) or higher education; or (c) only discussed curriculum policies conceptually without any classroom-level or teacher-readiness empirical data.
4. Screening Flow: The initial database search yielded 142 records. After removing duplicates (n = 38), 104 records were screened based on title and abstract. At this stage, 52 records were excluded for failing to meet the inclusion criteria (predominantly focusing on other subjects like Science or general administrative management). The remaining 52 full-text articles were assessed for eligibility. Of these, 24 were excluded because they lacked empirical classroom data or focused solely on generic curriculum socialization. This screening process resulted in a final corpus of 28 highly relevant empirical studies.

**Coding Scheme, Taxonomy, and Analytical Procedures**

Following the compilation of the corpus, data analysis proceeded through three progressive, auditable stages: Document Mapping, Gap Analysis, and Pedagogical Reconstruction (Figure 1).

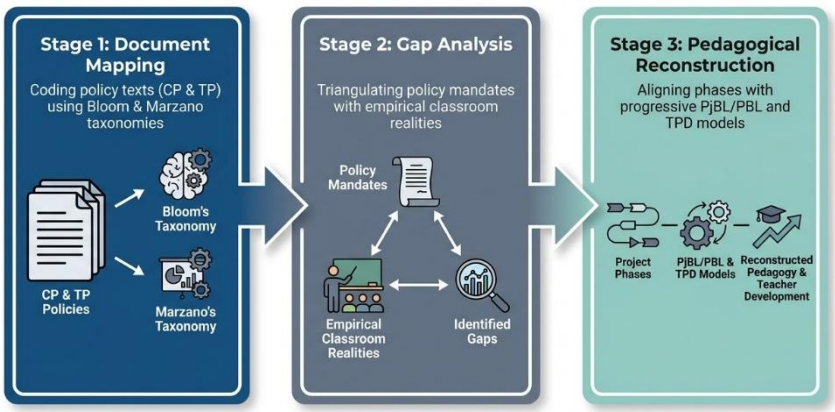



**Figure 1.** The three-stage qualitative analytical workflow.

To ensure transparency and reliability, all documents were coded using a structured deductively-derived coding scheme based on Bloom's Revised Taxonomy and Marzano's Cognitive Taxonomy. Coding was executed using NVivo 14 qualitative analysis software. Table 3 presents the coding categories, explicit decision rules, and examples of coded excerpts from the policy corpus.

**Table 3.** Deductive Coding Scheme and Operationalization

| Coding Category | Operational Definition / Decision Rule | Exemplar Coded Excerpt (SPI Policy Corpus) | Taxonomy Classification |
|---|---|---|---|
| Lower-Order Thinking Skills (LOTS) | Verbs and objectives requiring factual recall, simple comprehension, or chronological listing without analytical processing. | "*Peserta didik mampu menyebutkan nama-nama khalifah Daulah Umayyah...*" (Phase D) | Remembering / Understanding (Bloom) |
| Higher-Order Thinking Skills (HOTS) | Verbs demanding critical analysis, evaluation, comparison of historical sources, or synthesizing socio-political cause-and-effect. | "*Peserta didik menganalisis faktor penyebab runtuhnya Daulah Abbasiyah dan dampaknya terhadap peradaban dunia...*" (Phase F) | Analyzing / Evaluating / Creating (Bloom) |
| Performance-Based Learning | Objectives requiring students to actively produce, construct, or design digital/non-digital artifacts demonstrating historical knowledge. | "*Peserta didik membuat media digital (vlog/infografis) tentang kontribusi ilmuwan Muslim...*" (Phase E) | Knowledge Utilization / Execution (Marzano) |
| Value Internalization | Objectives focusing on historical empathy, moral reflection, and adopting moderate religious values (*rahmatan lil-ʿālamīn*). | "*Peserta didik merefleksikan nilai-nilai keteladanan perjuangan Wali Songo dalam konteks kebinekaan global...*" (Phase E/F) | Metacognitive System / Internalization (Marzano) |

To validate the coding scheme and ensure analytical reliability, two independent coders evaluated a random sample of 20% of the document corpus using a pre-established Formal Coding Guidebook (Anderson & Krathwohl, 2001; Marzano & Kendall, 2007). This guidebook detailed precise linguistic and grammatical parameters for classifying verbs into cognitive levels. An initial inter-coder agreement check yielded a Cohen's Kappa coefficient of 0.84, indicating a high level of coding reliability (Miles et al., 2014). Discrepancies in coding were resolved through iterative peer-debriefing sessions until 100% consensus on the operational definitions and coding classifications was reached.

**Integration of Analytical Frameworks and Gap Classification**

The analytical strength of this study lies in the direct, structural integration of the Bloom/Marzano cognitive taxonomies with the subsequent policy-practice gap analysis.

First, Stage 1 (Document Mapping) uses Bloom and Marzano taxonomy codes to establish a baseline of policy intent—empirically identifying the cognitive level, pedagogical mode, and value expectations mandated by national CP and TP texts. Second, in Stage 2 (Gap Analysis), we apply these identical codes to the Literature Corpus to extract the empirical reality of classroom enactment.

Finally, we directly overlay the two coded datasets. A "policy-practice gap" is systematically identified and classified when a code representing a high-level policy mandate directly collides with an empirical code indicating lower-order or traditional execution in the classroom (Table 4).

**Table 4.** Classification Criteria for the Policy-Practice Gap Taxonomy

| Gap Category | Operational Definition | Empirical Evidence Required for Classification |
|---|---|---|
| Cognitive Gap | Disconnect between mandated high-level cognitive verbs and actual classroom cognitive demands. | Mismatch where policy baseline is coded as "HOTS" (e.g., analyze, evaluate), but classroom reality is coded as "LOTS" (e.g., memorizing names, dates, and facts). |
| Methodological Gap | Mismatch between mandated active learning instructional models and actual classroom delivery. | Mismatch where policy baseline is coded as "Performance-Based Learning" or "Inquiry," but classroom reality is coded as "Passive Lecturing." |
| Assessment Gap | Disconnection between mandated authentic performance-based rubrics and actual grading mechanisms. | Mismatch where policy baseline is coded as "Value Internalization" or "Product-Creation," but classroom assessments are coded as "Traditional recall-based testing." |

**Trustworthiness, Dependability, and Audit Trail**

To ensure the trustworthiness and dependability of our qualitative findings, we implemented a rigorous peer examination and triangulation protocol.

We achieved document-to-empirical triangulation by validating national policy directives against empirical findings extracted from 28 distinct studies conducted by different researchers across multiple geographic locations in Indonesia (including Java, Sumatra, and Sulawesi) and employing various methods (surveys, case studies, and classroom action research). This extensive triangulation ensures that the identified gap patterns are structural and systemic, rather than isolated to specific regional contexts.

The preliminary coding, gap analysis, and reconstructed pedagogical models were subjected to a formal peer examination by two external educational experts: one senior scholar specializing in curriculum theory and design, and one expert in Islamic History Education pedagogy. These experts reviewed the entire raw coding sheet, the mapped policy excerpts, and the gap taxonomy classifications. Their feedback led to critical adjustments: they recommended merging several overlapping sub-categories of "Value Internalization" and advised clarifying the transition of cognitive demands between Phase E and Phase F.

To demonstrate analytical dependability, a comprehensive audit trail was maintained. The kept records include: (a) raw PDF/Word documents of the official policy decrees; (b) the systematic literature screening log showing the justification for each article's inclusion or exclusion; (c) NVivo project files containing the complete node hierarchy, coding stripes, and codebooks; and (d) detailed minutes of the peer-examination sessions. This transparent trail ensures that any external researcher can audit the journey from raw document corpus to our finalized pedagogical models (Lincoln & Guba, 1985).

**Ethical Considerations**

Because this study did not involve human participants and relied strictly on publicly available government curriculum documents and published, peer-reviewed academic literature, traditional institutional review board (IRB) approval was exempt. To maintain rigorous academic ethics, all analyzed materials are properly cited, and

secondary studies are represented objectively without intellectual distortion or plagiarism.

## RESULTS AND DISCUSSION

### Results

### Document-Based and Analytical Findings

This section presents the analytical findings from the qualitative document analysis (QDA) of national curriculum texts and the systematic synthesis of 28 peer-reviewed empirical studies (2020–2026). It evaluates the vertical coherence of policy expectations and the synthesized realities of secondary classroom enactment.

### Multi-Phase Cognitive and Material Mapping of SPI Learning Outcomes

To establish an analytical baseline, official *Capaian Pembelajaran* (CP) and *Tujuan Pembelajaran* (TP) documents were coded using a hybrid taxonomy derived from Bloom's Revised Taxonomy (Anderson & Krathwohl, 2001) and Marzano's Cognitive Taxonomy (Marzano & Kendall, 2007). This framework categorizes mandated policy demands into four hierarchical domains (Table 2).

**Table 2.** Hierarchical Cognitive and Material Mapping of SPI Across Educational Phases

| Phase & Grade | Subject Topics | LOTS Verbs | HOTS Verbs | Performance Verbs (Doing) | Value Verbs (Being) | Policy Source |
|---|---|---|---|---|---|---|
| Phase D *(Grades 7–9)* | Umayyad, Abbasid, Ottoman, Safavid Empires; Local Sultanates. | *Describe*, *Explain* (chronologies). | *Analyze* (factors of dynastic growth/decline). | *Construct* (visual maps/diagrams). | *Internalize* (scientific zeal/tolerance). | Decree No. 032/H/KR/2024; Maylawati et al. (2025) |
| Phase E *(Grade 10)* | Peaceful Islamization; Cultural adaptation of the *Wali Songo*. | *Identify* (routes); *List* (kingdoms). | *Synthesize* (local adaptation & Islamic values). | *Create* (digital timelines/infographics). | *Adopt* (moderate attitudes / *tawassuṭ*). | Decree No. 032/H/KR/2024; Suari et al. (2023) |
| Phase F *(Grades 11–12)* | Global Islamic civilizations; Modern reformists; National organizations. | *Outline* (biographies/structures). | *Evaluate* (reformist thoughts); *Critically reason*. | *Formulate* (policy briefs); *Debate* (complex causes). | *Demonstrate* (global citizenship/ethical leadership). | Decree No. 032/H/KR/2024; Wardani (2023) |

Coding decisions followed strict linguistic rules: policy verbs demanding factual identification (e.g., "*menyebutkan*") were classified as LOTS, while those requiring structural analysis or multi-perspective source evaluation (e.g., "*menganalisis*") were coded as HOTS. Verbs mandating the translation of concepts into concrete outputs (e.g., "*membuat karya*") were coded as Performance-Based, and those targeting moderate values (*rahmatan lil-ʿālamīn*) were categorized as Value-Oriented.

This progression reveals a pedagogical tension. While vertically coherent, the transition from Phase D to Phase E constitutes a "cognitive cliff" (Suari et al., 2023). In Phase D, students primarily describe and explain, but in Phase E, they are immediately expected to execute complex socio-cultural synthesis and independent digital production. Without structural scaffolding, this sudden escalation causes a

"cognitive fracture," separating factual recall from critical historical inquiry (Christopher & Jones, 2025).

**Synthesized Policy-Practice Gaps in SPI Implementation**

Because this study did not involve direct fieldwork, SPI classroom realities were synthesized from reported implementation barriers across the 28 empirical studies. The synthesis reveals a persistent decoupling between policy mandates (*Idealita CP*) and classroom realities (*Realita Field*) across three core dimensions (Table 3).

**Table 3.** Synthesized Policy-Practice Gap Taxonomy in SPI Education

| Gap Dimension & Literature Frequency | Normative Policy Mandates (*Idealita CP*) | Synthesized Classroom Realities (*Realita Field*) | Underlying Structural Drivers | Supporting Literature |
|---|---|---|---|---|
| **Cognitive Domain** *(82% frequency; n=23)* | HOTS analysis, critical historical reasoning, source evaluation. | Dominance of rote memorization (LOTS) of timelines, names, and battle dates. | Rigid exam-oriented cultures; heavy textbook reliance; lack of cognitive scaffolding. | Mpate (2025); Wardani (2023) |
| **Instructional Method** *(75% frequency; n=21)* | Active student-centered PjBL, PBL, and digital artifact production. | Traditional teacher-centered lectures; passive textbook reading. | Deficits in teacher TPACK; classroom time limits; administrative burdens. | Kabir (2023); Mwangi et al. (2026) |
| **Assessment System** *(89% frequency; n=25)* | Authentic performance rubrics, portfolios, product-based defenses. | Reliance on paper-and-pencil or digital multiple-choice recall tests. | Systemic high-stakes testing pressures; administrative grading overload. | Mujiburrahman et al. (2025); Shi et al. (2026) |

These structural gaps directly compromise historical understanding in SPI. In lessons on the Treaty of Hudaibiyah (Phase E), where policy mandates *analysis* of diplomacy to internalize conflict resolution, actual instruction is reduced to memorizing the treaty's signing date and emissaries (Mpate, 2025). Similarly, in studying the Abbasid Golden Age (Phase D), policy mandates *analyzing* the scientific synergy of *Bayt al-Ḥikmah* (Maylawati et al., 2025), yet classroom practices remain confined to lecturing and memorizing static lists of scholars and books (Wardani, 2023).

**Reconstructed PjBL and PBL Pedagogical Framework for SPI Education**

To bridge these gaps, a reconstructed, phase-aligned pedagogical framework was developed, matching instructional models directly to the coded CP verbs and identified gap dimensions.

- PjBL for Phases D and E: Mapped to the CP performance verb "*membuat karya*" (create products). Lower secondary students require concrete, highly structured tasks. PjBL structurally embeds the creation of tangible historical artifacts, converting passive content into active production (Nugraha et al., 2023).
- PBL for Phase F: Mapped to Phase F's analytical and evaluative verbs ("*evaluate*", "*critically reason*"). Upper secondary students possess the cognitive

maturity for dialectical historical inquiry. PBL forces students to analyze complex historical controversies using conflicting primary sources (Nabilah & Suyanto, 2025).

Table 4 outlines the instructional syntax and historical reasoning targets of this reconstructed model.

**Table 4.** Reconstructed, Phase-Aligned PjBL and PBL Pedagogical Model

| Phase & Topic Focus | Target CP Competency & Code | Reconstructed Model & Syntax | Target Deliverable | Historiographical Reasoning Target (De-coding the Artifact) |
|---|---|---|---|---|
| Phase D *(Class VIII: Abbasid Era)* | Analyze scientific contributions & create visual artifacts. *(HOTS Performance)* | PjBL: Essential Question → Source Investigation → Draft Design → Peer Feedback → Public Defense. | Digital Canva Infographic depicting interfaith collaboration at *Bayt al-Ḥikmah*. | Source Synthesis: Moving beyond software literacy; students must evaluate translation movement phases and visually defend historical claims. |
| Phase E *(Class X: Wali Songo)* | Analyze peaceful Islamization & construct timeline. *(HOTS Performance)* | PjBL: Historical Inquiry → Local Cultural Mapping → Timeline Design → Class Presentation. | Interactive Digital Timeline of localized Nusantara Islamization. | Causality Reasoning: Students trace the specific historical points of accommodation between Islamic values and local cultures. |
| Phase F *(Class XI-XII: Abbasid Collapse)* | Evaluate multi-causal factors, critically reason, debate. *(HOTS Dialectical)* | PBL: Problem Presentation → Multi-Perspective Source Gathering → Structured Debate → Synthesis Report. | Multi-perspective Case Report and Structured Oral Debate Performance. | Multi-Perspective Historiography: Evaluating biases in primary sources and comparing competing internal and external collapse theories. |

To ensure that digital artifact creation facilitates deep historical inquiry rather than superficial design, student deliverables are evaluated using a Multi-Dimensional Assessment Rubric (Table 5).

**Table 5.** Multi-Dimensional Assessment Rubric for SPI Deliverables

| Rubric Indicator | Unsatisfactory (1 Point) | Developing (2 Points) | Proficient (3 Points) | Advanced (4 Points) |
|---|---|---|---|---|
| Historical Accuracy | Contains major factual errors or chronological anachronisms; relies entirely on unverified blog posts. | Chronological sequence is present but fragmented; minor factual discrepancies in names or key dates. | Fully accurate historical facts, timelines, and actors; references standard secondary textbooks. | Flawless historical accuracy; successfully integrates and cross-references multiple primary and secondary sources. |
| Analytical Depth | Purely narrative or descriptive listing of events; fails to explain "why" or "how" historical shifts occurred. | Identifies basic causes of historical events but fails to link them to broader social or civilizational consequences. | Explains cause-and-effect relationships clearly; analyzes institutional, political, and cultural factors. | Sophisticated analysis of multi-causal factors; critically evaluates conflicting historical perspectives and socio-political dynamics. |
| Technical Creativity | Unstructured text dump onto a digital template; poor visual | Visual elements are present but serve as mere decoration rather | Strong visual hierarchy; effective and clean design in | Exemplary layout; design choices (color contrast, icons, visual |

| Rubric Indicator | Unsatisfactory (1 Point) | Developing (2 Points) | Proficient (3 Points) | Advanced (4 Points) |
|---|---|---|---|---|
| | hierarchy and no effective use of digital tools. | than enhancing the historical narrative. | Canva or Sutori; chronological flow is intuitive. | timelines) actively strengthen the historical argument. |
| Ethical Reflection | No connection made between historical events and contemporary ethical values or moral lessons. | Mentions values superficially; moral lessons read like rhetorical cliches disconnected from the historical evidence. | Clearly connects historical adaptation strategies to values of religious tolerance and moderation (*tawassuṭ*). | Deep, critical reflection linking historical narratives directly to modern civic responsibility and global citizenship. |

**Strategic Teacher Professional Development (TPD) Pathway as a Structural Enabler**

Rather than a speculative recommendation, the TPD framework emerged as a direct structural result of our gap analysis, which identified teacher TPACK deficits and assessment inertia as the primary drivers of policy-practice decoupling (Kabir, 2023; Mujiburrahman et al., 2025). The model is structured across three continuous pillars:

1. Subject-Specific TPACK Mastery: Equipping SPI teachers to access digital archives, utilize primary sources, and facilitate student-led digital design in Canva or Sutori (Rahayu et al., 2026).
2. Authentic Rubric-Based Assessment Design: Training educators to implement the newly formulated Multi-Dimensional Rubric (Table 5) to evaluate infographics and debate performances based on analytical depth and value internalization (Mujiburrahman et al., 2025).
3. Collaborative School-Based Learning Communities: Utilizing local Subject Teacher Consultations (*MGMP*) as collaborative hubs where teachers draft modules, conduct peer-teaching simulations, and audit student learning artifacts (Pratami et al., 2024).

## Discussion

### Theoretical Evaluation of SPI as a Spiral Curriculum

The progression of learning outcomes mapped in Section 4.1 structurally attempts to mirror the foundational principles of a spiral curriculum model, a theoretical concept pioneered by Jerome Bruner (1960) and further conceptualized by Kelly (2023). This theory posits that any subject can be taught effectively to students at any stage of development, provided the ideas are translated into an appropriate cognitive format and revisited iteratively over time with increasing levels of depth and complexity (Bruner, 1960). In the national SPI curriculum, this spiral architecture is evident in the repeated exploration of civilizational development across Phase D (lower secondary), Phase E (transitional upper secondary), and Phase F (advanced upper secondary).

However, analyzing the text of the official *Capaian Pembelajaran* (CP) decrees reveals a structural anomaly that contradicts spiral curriculum theory. A healthy spiral design requires a continuous, carefully scaffolded cognitive transition (Kelly, 2023). Yet, Decree No. 032/H/KR/2024 mandates an abrupt, non-scaffolded

escalation in cognitive demands. In Phase D, the CP text focuses heavily on basic narrative comprehension, requiring students to merely "*menceritakan sejarah nabi dan perkembangan daulah*" (narrate history and dynastic growth). Upon entering Phase E, however, the CP suddenly demands that students "*menganalisis dan mensintesis masuknya Islam secara damai*" (analyze and synthesize peaceful Islamization) and translate this into independent digital production.

This sudden cognitive leap constitutes a "cognitive cliff" (Suari et al., 2023). Because the curriculum fails to provide intermediate transitional scaffolding between lower-secondary descriptive recall and upper-secondary analytical demands, the spiral progression fractures (Kelly, 2023). This structural misalignment explains why, as synthesized from the empirical literature, actual classroom instruction consistently struggles to progress past basic factual memorization; teachers are presented with ambitious cognitive goals without the developmental scaffolding necessary to guide students up the spiral (Christopher & Jones, 2025; Derouich, 2025).

Systemic Deconstruction of Misalignment through Constructive Alignment

To understand why this cognitive fracture persists in secondary classrooms, these implementation gaps must be evaluated through John Biggs' (1996, 2011) framework of Constructive Alignment. Biggs posits that effective, deep learning occurs only when there is structural coherence between three critical pillars: the intended learning outcomes (policy CP), the teaching methods (classroom enactment), and the assessment tasks (Rodas, 2020). If these components are not aligned, students will default to a surface approach to learning, focusing strictly on memorizing isolated facts to pass unaligned assessments (Rodas, 2020).

Our analytical findings reveal systemic non-alignment across these three pillars, which is driven by a unique professional friction experienced by SPI educators. Unlike teachers of secular humanities, SPI teachers operate at a challenging intersection of confessional-dogmatic value transmission and academic-critical historical inquiry (Suari et al., 2023). SPI teachers often feel a pastoral responsibility to protect students' faith and transmit moral lessons as sacred, absolute truths (Wardani, 2023). Consequently, when national policies mandate active student-centered inquiry, teachers often retreat to passive monologue lecturing of safe, dogmatic narratives, fearing that open-ended student debates on controversial historical shifts or civilizational declines might undermine students' religious conviction or moral internalization (Kabir, 2023; Wardani, 2023).

This ideological tension interacts with structural barriers to create a self-reinforcing negative loop. Time constraints and rigid, high-stakes school assessment cultures force teachers to prioritize covering extensive historical timelines before exams (Kabir, 2023; Mwangi et al., 2026). Because school assessments remain rigidly confined to traditional multiple-choice recall tests, teachers use passive lecturing as the most time-efficient delivery mechanism (Mujiburrahman et al., 2025). This assessment inertia completes the loop: students adopt surface memorization strategies, which in turn discourages teachers from using active pedagogical models like PjBL or PBL (Mpate, 2025; Shi et al., 2026). Thus, breaking this loop requires aligning instructional activities and assessment rubrics directly with the cognitive policy verbs (Rodas, 2020).

**Universal Implications for Global Competency-Based Humanities Education**

This study contributes to global curriculum scholarship by offering a novel conceptual angle on competency-based education (CBE) in the religious humanities. While current international CBE literature focuses heavily on procedural or technical competencies (Majani, 2026), the SPI case illustrates the dual-layered complexity of implementing competency reforms in value-laden, faith-aligned subjects.

In secular humanities, curriculum coherence typically requires a single-layered alignment: matching critical cognitive inquiry with analytical assessments (Christopher & Jones, 2025). However, religious history education demands a double-layered constructive alignment. It must simultaneously align:

1. Critical Historiographical Reasoning: Transitioning students from descriptive chronological recall to evaluating multi-perspective socio-political cause-and-effect (the cognitive dimension).
2. Normative Value Internalization: Guiding students to internalize universal ethical values, historical empathy, and moderate religious character (*rahmatan lil-ʿālamīn*) to form active, tolerant civic identities (the moral/being dimension).

Our proposed, phase-aligned PjBL/PBL model does not claim to be a panacea, but rather serves as a structured framework that helps minimize this cognitive and moral misalignment. By requiring students to design digital infographics representing historical interfaith collaboration (Phase D) or engage in structured debates on civilizational collapse (Phase F), this model reduces cognitive mismatch by transforming values from passive, dogmatic dictations into active, critically reasoned ethical choices (Nabilah & Suyanto, 2025; Pratami et al., 2024). This double-layered approach demonstrates to curriculum designers globally how to maintain critical scholarly rigor and robust moral education within unified, competency-based reforms.

**Methodological Scope and Limitations**

The findings and discussion in this study must be interpreted within its methodological boundary conditions. Because this study is desk-based and did not collect primary field data, all assertions regarding classroom practices, teacher TPACK deficits, and assessment behaviors are strictly attributed to the synthesized findings of the 28 peer-reviewed empirical studies in our reviewed literature corpus (2020–2026).

Crucially, because our proposed, phase-aligned PjBL/PBL pedagogical framework has not yet been empirically tested or validated in live SPI classrooms, it remains a conceptual-pedagogical proposition. While the framework is theoretically grounded in our cognitive mapping and gap analysis, its actual ecological validity, classroom feasibility, and direct impact on student historical reasoning await empirical validation through future classroom action research.

## CONCLUSION

This study synthesizes a persistent, systemic misalignment between the ambitious competency-based policy mandates of the *Kurikulum Merdeka* and the synthesized classroom realities of Islamic History Education (SPI) in secondary schools. The qualitative document analysis and systematic synthesis of existing

literature reveal that while national directives mandate a progressive, multi-phase cognitive continuum of historical analysis and active performance, daily classroom practice as documented across empirical studies remains largely decoupled from these standards due to deep-seated teacher TPACK deficits, structural time constraints, and assessment inertia. To bridge this divide, this study moves beyond generic educational recommendations by conceptualizing a phase-aligned pedagogical framework. This framework matches Project-Based Learning (PjBL) for digital artifact creation in lower secondary levels (Phases D and E) and Problem-Based Learning (PBL) for dialectical inquiry in senior secondary levels (Phase F), directly supported by a school-embedded Teacher Professional Development (TPD) pathway that addresses the root structural drivers of classroom decoupling.

The core scholarly contribution of this study lies in its dual conceptual and pedagogical value to the fields of curriculum theory, policy-practice alignment, and religious humanities education. Conceptually, it introduces the model of "double-layered constructive alignment," illustrating how competency-based reforms in value-laden subjects must simultaneously align critical historical inquiry with moral-ethical internalization without sacrificing scholarly rigor. Pedagogically, it provides curriculum designers and educators with a concrete, vertically coherent blueprint to systematically scaffold cognitive and value-oriented progressions across developmental phases. Ultimately, this research demonstrates that bridging the policy-practice gap in religious history education is not merely an administrative adjustment, but a structural requirement to prevent learning loss and ensure that competency-based curriculum reforms can realistically achieve their transformative potential in secondary classrooms.

## RECOMMENDATION

Based on the study's insights, future research should focus on longitudinal, empirical field evaluations and quasi-experimental trials across diverse secondary schools to measure the real-world impact of reconstructed Project-Based Learning and Problem-Based Learning frameworks on student higher-order thinking skills, historical literacy, and character formation. Teacher education institutions and policymakers are recommended to embed subject-specific Technological Pedagogical Content Knowledge (TPACK) and authentic assessment design directly into pre-service training and school-based professional development networks.

However, several critical barriers may influence research and implementation outcomes, including severe resource and technological infrastructure disparities between urban and rural schools, entrenched high-stakes examination cultures that prioritize lower-order factual recall, limited instructional time, and heavy administrative workloads that restrict teachers' capacity to adopt student-centered methodologies. Addressing these structural constraints through targeted funding, workload realignment, and systemic assessment reform remains essential for the successful scaling and sustainability of future instructional innovations.

**Author Contributions**

We collaboratively conceived and designed the study. The first author led the overall study conceptualization, research design, qualitative document collection and coding, analytical

synthesis, and initial manuscript drafting. The second and fifth authors contributed to the development of the theoretical framework, validation of qualitative data analysis, and review of supporting literature. The third and fourth authors contributed to the formulation of the research methodology, data and reference management, and critical analysis of study findings. All authors participated in critical revisions, manuscript editing, and gave final approval for the final version to be submitted.

**Acknowledgment**
We express sincere gratitude to the course lecturers and academic advisors at the School of Postgraduate Studies, Universitas Pendidikan Indonesia (UPI), for their valuable guidance, insightful feedback, and continuous support throughout the development of this research. Deep appreciation is also extended to the Department of Islamic Religious Education at Universitas Pendidikan Indonesia, Universitas Islam Negeri Sunan Gunung Djati Bandung, and Universitas Majalengka for providing the necessary academic resources and institutional support that enabled the completion of this study.

**Conflict of interests**
The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## REFERENCES

Anderson, & Krathwohl. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*.

Cheema, G. K., Rather, M. A., Aggarwal, V., Mandloo, S. A., & Faustino, A. (2026). Competency-Based Curricula and Technology: Game Changers for Literature Subjects in Secondary Schools in Arusha City, Tanzania. *Journal of Interdisciplinary Studies in Education*, *15*(5), 195–214. https://doi.org/10.32674/3yqn3z42

Christopher, M., & Jones, M. (2025). Evaluating contexts for implementing a competency-based curriculum in south and eastern Africa: A meta-narrative review. *Curriculum Journal*, *36*(3), 368–383. https://doi.org/10.1002/curj.301

Derouich, M. (2025). Ensuring outcome-based curriculum coherence through systematic CLO–PLO alignment and feedback loops. *Discover Education*, *4*(1). https://doi.org/10.1007/s44217-025-00915-7

Kabir, S. M. A. (2023). English Language at Secondary Level in Bangladesh: (Dis)connections Between Policy and Practice of Oral Skills. *Qualitative Report*, *28*(1), 301–322. https://doi.org/10.46743/2160-3715/2023.5761

Kasuga, W., & Kalolo, J. F. (2025). Competency-based curriculum in Tanzania: Charting dilemmas for successful implementation. *Cogent Education*, *12*(1). https://doi.org/10.1080/2331186X.2025.2563711

Kelly, K. T. (2023). The Logic of Success. In *Philosophy of Science Today* (pp. 11–38). Oxford University Press. https://doi.org/10.1093/oso/9780199250547.003.0002

Majani, W. P. (2026). Demystifying the policy-practice gap in Tanzania educational change landscape. *African Identities*. https://doi.org/10.1080/14725843.2026.2653772

Marzano, & Kendall. (2007). *New Taxonomy of Educational Objectives Adapted from Marzano & Kendall*.

Maylawati, D. S., Mulyana, R., Rizqullah, N., Nurjaman, I., & Ramdhani, M. A. (2025). Assessing Indonesian Islamic schools' readiness for the independent curriculum using text analytics. *Multidisciplinary Reviews*, *8*(10).

https://doi.org/10.31893/multirev.2025336

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative Data Analysis: A Methods Sourcebook* (3rd ed.). Sage publications.

Mpate, H. (2025). Biology teachers' implementation of the competence based curriculum in Tanzania: challenges and opportunities. *Journal of Biological Education*, *59*(1), 16–36. https://doi.org/10.1080/00219266.2023.2282428

Mujiburrahman, Ratnaya, I. G., Parwata, I. G. L. A., & Budhyani, I. D. A. M. (2025). Evaluation of the Merdeka Curriculum in Educational Institutions: A Systematic Literature Review of Strategies, Challenges, and Student Learning Outcomes. *Jurnal Paedagogy: Jurnal Penelitian Dan Pengembangan Pendidikan*, *12*(4), 1255–1268.

Mwangi, N. M., Karuku, S., & Obura, E. (2026). Teacher preparedness for competency-based curriculum in Kenyan schools: training and perceptions prior to implementation. *International Journal of Evaluation and Research in Education*, *15*(1), 714–724. https://doi.org/10.11591/ijere.v15i1.35742

Nabilah, N., & Suyanto, S. (2025). A CIPP-based evaluation of critical thinking in science : Implementation of the Merdeka curriculum in Mataram public junior high schools. *STEM Education*, *5*(May), 587–616. https://doi.org/10.3934/steme.2025028

Nugraha, I. R. R., Supriadi, U., & Firmansyah, M. I. (2023). Efektivitas Strategi Pembelajaran Project Based Learning dalam Meningkatkan Kreativitas Siswa. *Jurnal Penelitian Dan Pendidikan IPS (JPPI)*, *17*(1), 39–47.

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Cynthia, D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mcdonald, S., Mcguinness, L. A., … Mckenzie, J. E. (2021). *PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews*.

Pratami, D., Akhmal, N. H., Maulana, M. I. I. M., & Hassan, S. A. H. S. (2024). Introducing Project-Based Learning Steps To The Preschool Teachers In Bandung, Indonesia. *Journal of Technology and Science Education*, *14*(3), 883–902.

Rahayu, B. M., Muhyidin, A., & Jamaludin, U. (2026). Integrating local wisdom into digital teaching materials for next-generation competencies in future education: A systematic literature review. *Multidisciplinary Reviews*, *9*(8). https://doi.org/10.31893/multirev.2026366

Rodas, J. J. M. (2020). Competences evaluation in knowledge fields, a look at the curriculum. *Revista de Humanidades (SPAIN)*, *39*, 175–204. https://doi.org/10.5944/RDH.39.2020.22617

Shi, F., Yang, Q., Zhang, M.-X., & Wang, X.-T. (2026). Translating policy into practice: teacher agency amid cognitive and ecological constraints in enacting competency-based assessment. *Frontiers in Psychology*, *17*. https://doi.org/10.3389/fpsyg.2026.1836710

Suari, S., Indrawadi, J., & Syahriani, F. (2023). Exploration of Learning Strategies in the Social Studies Based on the Merdeka Curriculum. *Al-Ta'lim Journal*, *30*(1), 195–210. https://doi.org/10.15548/jt.v30i2.920

Wardani, H. K. (2023). Analysis of the Impact of the Merdeka Curriculum Policy on Stakeholders at Primary School. *The Indonesian Journal of the Social Sciences*, *11*(2).