MUMTAZ
JURNAL PENDIDIKAN AGAMA ISLAM
Publisher:
Program Studi Pendidikan Agama Islam
UNIVERSITAS ISLAM IBRAHIMY BANYUWANGI
E-ISSN: 2809-204X
P-ISSN: 2809-2139

MUMTAZ : Jurnal Pendidikan Agama Islam
Volume 6 Number 2 (2026) : 621-643
Available online at https://ejournal.uniib.ac.id/index.php/mumtaz

FROM USERS TO DESIGNERS: RESEARCH-BASED PjBL FOR MASTER'S STUDENTS DESIGNING ISLAMIC CULTURAL HISTORY INSTRUCTION

Masykurillah*[1], A. Qomarudin[2]
Master's Programme in Islamic Religious Education, UIN Jurai Siwo Lampung, Indonesia
e-mail: masykurillah@metrouniv.ac.id, masqomar@uinjusila.ac.id

DOI: http://doi.org/10.69552/mumtaz.v6i2.3944

Received:	Revised:	Accepted:	Published:
July 27, 2026	August 24, 2026	August 29, 2026	September, 2026

*Corresponding author

Abstract

This study developed and conceptually validated a research-based Project-Based Learning (PjBL) model for preparing master's students in Islamic Religious Education to investigate instructional problems and use evidence to design Islamic Cultural History (SKI) instruction. Educational research and development used a qualitatively dominant multimethod design because the central task was interpretive: explaining how documented gaps, artefacts, stakeholder accounts, and disciplinary requirements should be translated into design decisions. Descriptive quantitative data were used to indicate salience and expert consensus rather than population effects. Needs evidence showed weak alignment among outcomes, project milestones, research activity, historical reasoning, assessment, feedback, and revision. The integrated analysis generated eleven design requirements, five competence dimensions, and a nine-stage course syntax linking alignment, authentic diagnosis, SKI-context analysis, theory-evidence synthesis, project planning, concept selection, prototype construction, academic review-revision, and reflective communication. Four experts rated 400 items; 355 received a score of 5 and 45 a score of 4, producing a mean of 4.89 and 97.75% conceptual feasibility. The study contributes a traceable postgraduate design architecture, while practicality, fidelity, assessment reliability, equity, and effectiveness remain to be tested empirically.

Keywords: *Research-Based PjBL; Instructional-Design Competence; Islamic Cultural History; Postgraduate Islamic Religious Education; Design Traceability.*

Abstrak

Penelitian ini mengembangkan dan memvalidasi secara konseptual model Project-Based Learning (PjBL) berbasis riset untuk membekali mahasiswa Magister

Copyright © 2026 The Author(s).
This article is distributed under the terms of the Creative Commons Attribution-ShareAlike 4.0 International License (CC BY-SA 4.0).

Pendidikan Agama Islam meneliti masalah pembelajaran dan menggunakan bukti tersebut dalam mendesain pembelajaran Sejarah Kebudayaan Islam (SKI). Penelitian dan pengembangan pendidikan menggunakan desain multimetode dominan kualitatif karena persoalan utama penelitian bersifat interpretif, yaitu menjelaskan bagaimana kesenjangan terdokumentasi, artefak, pandangan pemangku kepentingan, dan tuntutan disipliner dikonversi menjadi keputusan desain. Data kuantitatif deskriptif digunakan untuk menunjukkan tingkat kebutuhan dan konsensus ahli, bukan untuk mengestimasi efek populasi. Analisis terpadu menemukan lemahnya keselarasan antara capaian, tahapan proyek, aktivitas riset, penalaran historis, asesmen, umpan balik, dan revisi. Temuan menghasilkan sebelas persyaratan desain, lima dimensi kompetensi, dan sintaks sembilan tahap yang menghubungkan penyelarasan capaian, diagnosis autentik, analisis konteks SKI, sintesis teori-bukti, perencanaan proyek, pemilihan konsep, konstruksi prototipe, telaah-revisi akademik, serta komunikasi reflektif. Empat ahli menilai 400 butir; 355 memperoleh skor 5 dan 45 memperoleh skor 4, dengan rerata 4,89 dan kelayakan konseptual 97,75%. Kontribusi penelitian berupa arsitektur desain perkuliahan pascasarjana yang terlacak, sedangkan kepraktisan, fidelitas, reliabilitas asesmen, keadilan akses, dan efektivitas memerlukan pengujian empiris lanjutan.

Kata Kunci: *PjBL Berbasis Riset; Kompetensi Desain Pembelajaran; Sejarah Kebudayaan Islam; Pendidikan Agama Islam Pascasarjana; Keterlacakan Desain.*

INTRODUCTION

The pedagogical competence expected of master's students in Islamic Religious Education should extend beyond selecting and applying established teaching models. Graduate students are expected to diagnose instructional problems, interpret curriculum outcomes, analyse learners and subject matter, justify design mechanisms, and produce coherent learning systems that can be evaluated and revised. Recent competency research treats professional design as integrated performance rather than a collection of isolated technical procedures (Yalçın et al., 2021), while teacher-competence research similarly shows that knowledge becomes educationally consequential through situation-specific interpretation and decision making (Blömeke et al., 2022). This expectation is particularly relevant for practising teachers: professional experience provides access to authentic classroom problems, but experience does not by itself produce an evidence-based design argument.

The challenge is sharper in Islamic Cultural History, or Sejarah Kebudayaan Islam (SKI). Across school, madrasah, and higher-education settings, SKI should not be reduced to names, dates, dynasties, and chronological recall. Recent history-education studies emphasise source appraisal, contextualisation, causation, change

and continuity, multiperspectivity, and evidence-based argumentation as the disciplinary work through which historical understanding develops (Claravall & Irey, 2022; Wilke et al., 2023). Inquiry with primary sources, heritage, local figures, and project work can strengthen historical literacy and thinking when learners actively investigate and communicate interpretations (Lionar et al., 2025; Roro Wilis & Dyah Kumalasari, 2025; Sariyatun et al., 2025; Van Doorsselaere, 2025). For SKI, those operations must also support responsible reflection on values and ibrah without allowing predetermined moral conclusions to replace historical evidence.

Project-Based Learning (PjBL) appears relevant because the target competence culminates in a complex artefact developed through sustained inquiry. Recent synthesis reports positive associations between PjBL and learning, engagement, collaboration, and higher-order competencies (Rehman et al., 2024; Yusri et al., 2024; Zhang & Ma, 2023). Yet the evidence should not be read uncritically. A 2026 umbrella review of 15 PjBL meta-analyses found a consistently positive direction but substantial heterogeneity in effect magnitude and rated all included meta-analyses as critically low in methodological quality (Farshad & Fortin, 2026). This weakens the assumption that merely assigning a project is sufficient. The more consequential question for postgraduate design education is which project structures, evidence requirements, quality controls, and feedback mechanisms make the work educationally and epistemically defensible.

Research-based learning addresses a different but complementary weakness. Longitudinal and pre-post studies indicate that participation in research can strengthen students' research competences and their capacity to connect disciplinary knowledge with inquiry (Thiem et al., 2023; Wessels et al., 2021). A systematic review of research-methodology teaching also shows that research learning is most meaningful when methodology is integrated with authentic disciplinary activity rather than taught as disconnected procedures (Matos et al., 2023). However, these studies primarily examine research competence, methodology learning, or inquiry participation. They do not specify how research evidence should govern each decision in the construction of an instructional model. Thus, research-based learning can still remain parallel to design work unless the course explicitly connects problem evidence, literature synthesis, design rationale, prototype decisions, review, and revision.

Recent history-education research narrows the disciplinary gap but does not eliminate the design gap. Digital PjBL, research-based learning, source inquiry, and historical-method integration have been associated with historical concepts, historical consciousness, and historical thinking (Matitaputty et al., 2024; Rahmawati & Hadi, 2026; Universitas Riau Indonesia et al., 2025). These studies

demonstrate that disciplinary inquiry matters, yet they largely evaluate learners' historical outcomes or a particular instructional intervention. They do not prepare postgraduate teachers to construct, defend, and revise an entire SKI learning model. Likewise, contextual Islamic-education research shows the importance of connecting knowledge, lived problems, and knowledge-based virtues (Parhan et al., 2024), but it does not operationalise an auditable chain from empirical need to instructional-design decision. The literature is therefore rich in component solutions but fragmented at the level of postgraduate course architecture.

The preliminary investigation in the Master's Programme in Islamic Religious Education at UIN Jurai Siwo Lampung made this fragmentation visible. Curriculum outcomes called for analysis, research, innovation, and instructional development, but these expectations were not consistently translated into weekly artefacts, milestones, quality gates, aligned assessment, and documented revision. Students could identify familiar models and produce creative sequences, yet their products varied in distinguishing a model from a method or module, tracing decisions to needs evidence, synthesising theory, completing all model components, and operationalising historical reasoning, HOTS, values, and ibrah. In one course context, the written expectation to produce instructional designs still yielded mainly papers, slides, and preliminary activity sketches; in another, a project orientation produced more explicit artefacts, but their mean product score remained 60 and conceptual coherence varied. These data were diagnostic rather than evidence of the developed model's effectiveness.

The research gap is therefore not simply the absence of another PjBL variation. Three unresolved problems intersect. First, recent PjBL research establishes possible benefits but provides limited guidance on how to engineer a postgraduate instructional-design project when evidence quality and implementation conditions vary. Second, research-based learning develops inquiry competence but rarely makes research evidence the mandatory decision infrastructure of a design project. Third, history and Islamic-education studies develop disciplinary inquiry and contextual values but do not integrate them with the full competence of diagnosing needs, synthesising theory, constructing model components, defending assessment decisions, and documenting revision. No recent study located for this project operationalised these requirements as one traceable postgraduate SKI course model with explicit transition criteria, feedback-to-revision records, and the rule that historical validity must precede ibrah.

Accordingly, this study aimed to develop and conceptually validate a Research-Based PjBL-SKI Model that supports master's students in moving from users of established models toward evidence-informed designers of SKI instruction. Its

novelty lies in the architecture of integration rather than in juxtaposing two labels: authentic problems are converted into needs evidence; evidence is synthesised with recent research; findings are transformed into design requirements; disciplinary SKI analysis constrains the model; project stages generate auditable artefacts; quality gates regulate progression; and academic defence, feedback uptake, and revision become assessed components of competence. The intended contribution is a postgraduate design hypothesis that can subsequently be tested for readability, practicality, fidelity, reliability, equity, and effectiveness.

RESEARCH METHODS

This study employed educational research and development to produce and establish the conceptual feasibility of a Research-Based PjBL-SKI course model. The unit of development was the postgraduate course architecture that guides students in designing SKI instruction, not the school-level models produced by students. The development logic connected five operations: needs analysis, recent-literature synthesis, prototype construction, multidisciplinary expert appraisal, and design refinement. Analysis, design, and development were treated as connected evidence-to-decision processes so that every major model component could be traced back to a documented need or theoretical mechanism.

The study was conducted in the Master's Programme in Islamic Religious Education at UIN Jurai Siwo Lampung from March to July 2026. Purposively selected participants comprised ten students enrolled in the SKI learning-development course, one lecturer, one programme manager, two PAI/SKI teachers as contextual informants, and four expert validators, totalling eighteen participants. The validators consisted of two specialists in educational research/instructional-model design and two specialists in Islamic education, PAI, and SKI. The multimethod design was deliberately qualitatively dominant. The central research questions were developmental and interpretive: what the identified gaps meant, how the gaps interacted, why particular design responses were warranted, and how disciplinary and pedagogical evidence should be converted into model specifications. Those questions required close analysis of documents, artefacts, interview accounts, and expert reasoning rather than statistical estimation. Qualitatively driven multimethod research is appropriate when one interpretive logic provides the core inference and quantitative components are used to elaborate, compare, or strengthen that inference rather than operate as an equal strand (Love & Corr, 2022; Morgan & Hoffman, 2021).

The quantitative component therefore had a bounded supporting role. Questionnaire means and needs gaps indicated the salience and relative priority of

problems, while expert ratings summarised consensus about conceptual feasibility. They were not used to estimate population parameters, test causal hypotheses, or compare the effectiveness of teaching conditions. The sample was purposive and small, the artefacts were heterogeneous, and the product had not yet been prospectively implemented. Treating the quantitative data as co-dominant would therefore have implied a level of inferential precision that the design did not support. Integration occurred at two explicit points: in needs analysis, numerical gaps were interpreted alongside documentary and interview evidence; in validation, item ratings were interpreted together with expert comments and the location of the design evidence. Qualitative interpretation governed the final design decision whenever a numerical summary could not explain the nature or consequence of a gap.

Data were collected through document analysis, semi-structured interviews, needs questionnaires, structured artefact review, recent-literature synthesis, and expert validation. Documents included curriculum and course plans, outcomes, assignment specifications, learning-management-system records, presentations, drafts, feedback notes, rubrics, and student products. Semi-structured interviews were used because they maintain topical consistency while allowing participants to explain context, reasoning, and unanticipated issues in depth (Adeoye-Olatunde & Olenik, 2021). The literature synthesis prioritised peer-reviewed empirical studies and systematic or meta-analytic reviews from 2021-2026 that had a clear design function for PjBL, research-based learning, instructional design, feedback, self-regulation, historical reasoning, and Islamic education.

Qualitative data were analysed through familiarisation, coding, category development, theme review, and interpretation of design implications. Reflexive thematic analysis was used as a transparent interpretive practice rather than a mechanical coding formula (Braun & Clarke, 2021). Quantitative questionnaire and expert-validation data were summarised using frequencies, percentages, means, and needs gaps. Integration was performed through joint displays that placed quantitative signals beside documentary or interview patterns and then recorded the resulting design implication. Trustworthiness was strengthened by source and method triangulation, comparison of convergent and discrepant evidence, an audit trail from evidence to design requirement, and independent expert appraisal. Participation was voluntary, identities were coded, and student products were analysed only after course assessment and permission. Because the product was conceptually validated before prospective classroom implementation, claims were limited to documented needs, design coherence, traceability, and expert-assessed feasibility.

RESULTS AND DISCUSSION

Results

The findings are organised around the multimethod needs analysis, the competence construct and design requirements, the architecture of the Research-Based PjBL-SKI Model, and multidisciplinary expert appraisal. Because the model was developed and validated conceptually before prospective classroom implementation, the results concern documented needs, design coherence, and expert-assessed feasibility rather than causal improvement in students' competence.

Empirical needs and the design problem

The needs analysis combined two course contexts in 2025/2026 with fourteen substantive information sources. The quantitative patterns were not interpreted in isolation; each numerical signal was examined against course artefacts, documentary evidence, and stakeholder accounts. Table 1 presents this integration and makes explicit how the qualitative strand explains what the numerical gaps actually represented.

Table 1. Integrated multimethod evidence profile of the needs analysis

Evidence source	Quantitative signal	Qualitative/documentary pattern	Integrated inference
Class A students	Current condition = 1.78; urgency = 4.90; gap = 3.12 (very high priority)	Written expectation to produce an instructional-design product, yet outputs were dominated by papers, slides, and preliminary activity sketches.	A curriculum-level expectation was not sufficiently operationalised into a complete design process.
Class B students	Current condition = 3.06; urgency = 4.90; gap = 1.84 (high priority); five product scores = 55-65, M = 60	A project orientation produced products, but milestones, quality gates, research requirements, domain-specific criteria, and revision records were inconsistent.	A project assignment alone did not ensure conceptual depth or internal coherence.
Lecturer	Needs gap = 2.96 (very high priority)	Guidance and feedback existed, but project phases, scaffolding, evidence standards, and moderation required stronger standardisation.	The course needed an explicit facilitation and quality-control architecture.
Programme manager	Needs gap = 2.37 (very high priority)	Programme expectations emphasised analysis, research, innovation, and development, but cross-document alignment was not fully visible.	Shared standards were needed across outcomes, course activities, artefacts, and assessment.

Evidence source	Quantitative signal	Qualitative/documentary pattern	Integrated inference
Documents and artefacts	No population statistic was inferred	Traceability from problem evidence to theory, complete model components, assessment, review, and revision was fragmented across products.	The central problem concerned the architecture of the course, not one isolated teaching technique.

The integrated evidence revealed a twofold problem. Without a project architecture, an outcome statement did not reliably generate a complete instructional model. With a project assignment but without standardised milestones and epistemic controls, students could still produce artefacts whose depth depended heavily on prior experience and informal guidance. Most participants were practising teachers with access to authentic school problems, but experience was often expressed as anecdotal authority rather than converted into systematically collected evidence. The developmental problem was therefore how to transform professional experience into inquiry, and inquiry into defensible design decisions.

Table 2. Five need clusters and their corresponding design responses

Need cluster	Empirical problem	Required design response	Primary design requirement(s)
Curricular alignment	Outcomes were not fully linked to products, stage artefacts, and assessment.	Outcome-stage-artefact-assessment matrix and shared standards.	DR-01, DR-11
Project and assessment	Milestones, gates, rubrics, feedback uptake, moderation, and revision varied.	Phased project, compulsory artefacts, analytic rubrics, quality gates, and revision records.	DR-04, DR-08, DR-09
SKI disciplinary specificity	Source work, historical reasoning, HOTS, and ibrah were inconsistently operationalised.	Source appraisal, historical-reasoning criteria, HOTS tasks, and evidence-based ibrah.	DR-03, DR-06
Student characteristics	Professional experience was strong, but research/design literacy and time regulation varied.	Diagnostic support, contingent scaffolding, peer mentoring, differentiated pathways, and common standards.	DR-07, DR-10
Course architecture	Problem, research, theory, construction, review, and revision were fragmented.	Integrated Research-Based PjBL-SKI model with traceable transitions.	DR-02, DR-05, DR-11

Competence construct and design requirements

The competence to design SKI learning models was formulated as an integrated capacity comprising both process quality and product quality. D1 concerns analysing authentic instructional problems and needs. D2 concerns aligning curriculum, learners, context, and the distinctive characteristics of SKI. D3 concerns appraising and synthesising theory and research evidence into a design rationale. D4 concerns constructing complete and internally coherent model elements. D5 concerns designing assessment, defending decisions, using feedback, and revising the design transparently. This formulation prevents competence from being reduced to the visual completeness of a final document; the evidentiary pathway through which the product is built becomes part of what is assessed.

The empirical findings were converted into eleven design requirements. DR-01 requires auditable alignment among programme outcomes, course outcomes, project stages, artefacts, and assessment. DR-02 requires ethically collected needs evidence. DR-03 requires source appraisal, theoretical synthesis, and a traceable data-theory-decision relationship. DR-04 requires milestones, schedules, quality gates, and minimum evidence before transition. DR-05 requires complete model components and coherence checks. DR-06 embeds historical reasoning, HOTS, source responsibility, values, and evidence-based ibrah. DR-07 specifies roles, academic adab, probing, contingent scaffolding, and peer mentoring. DR-08 requires analytic and moderated assessment. DR-09 requires review forms, feedback uptake, version comparison, and a revision audit trail. DR-10 addresses workload, ethics, contextual adaptation, low-technology alternatives, and claim boundaries. DR-11 requires cross-class standardisation, shared course documents, repositories, metadata, and quality assurance.

Recent research was then used to test and refine the mechanisms behind those requirements. Instructional-design competency research supported integrated design performance (Yalçın et al., 2021); constructive-alignment studies supported explicit links among outcomes, activity, and assessment while cautioning against bureaucratic rigidity (Hailikari et al., 2022; Loughlin et al., 2021); self-regulated learning evidence justified project planning and monitoring support (Theobald, 2021); and recent feedback research justified making uptake, judgement, and action visible rather than treating feedback as lecturer transmission (Carless, 2023; De Kleijn, 2023; Winstone et al., 2022). Historical inquiry studies supplied the disciplinary mechanisms for source work and reasoning, while Islamic-education research informed contextual and value-oriented reflection (Parhan et al., 2024).

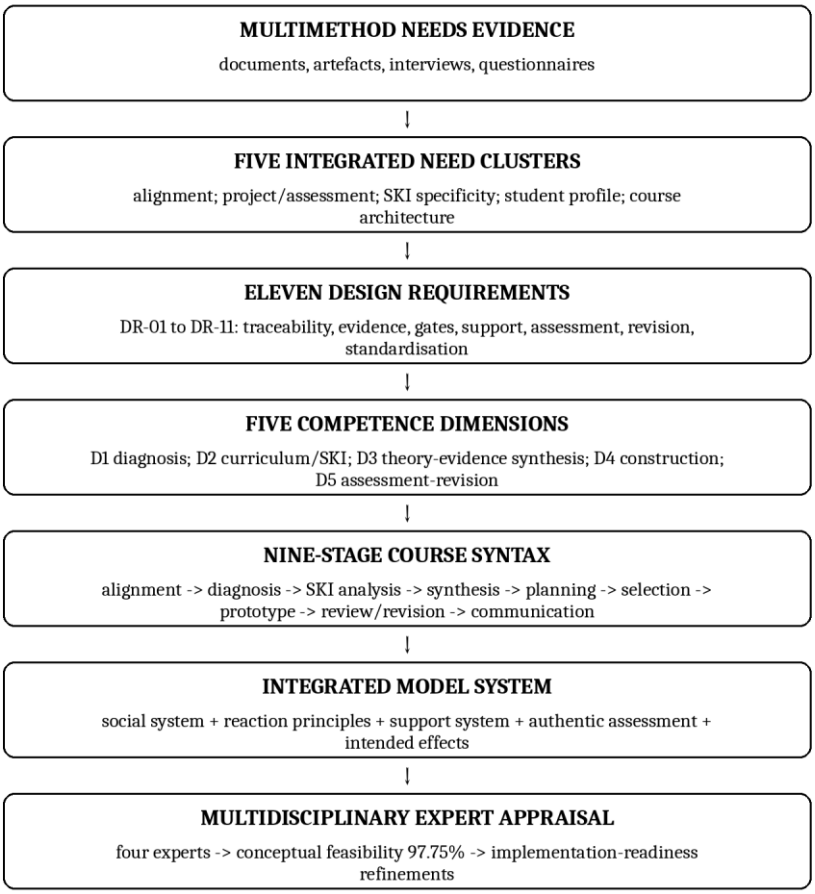

Figure 1. Evidence-to-design derivation pathway of the Research-Based PjBL-SKI Model

Architecture of the Research-Based PjBL-SKI Model

The resulting product was named the Research-Based Project-Based Learning Model for SKI Instructional Design, abbreviated as the Research-Based PjBL-SKI Model. PjBL provides the organisational structure of an extended authentic project; research provides the epistemic basis for decisions; and SKI specifies the disciplinary domain and the nature of the designed instruction. The research product is the course model that organises this activity, whereas students' SKI designs are learning artefacts and evidence of competence.

The model adopts a constructivist, contextual, inquiry-oriented, research-based, design-oriented, and reflective-value approach. Its methods include diagnosis, document and source analysis, literature synthesis, design studios, decision matrices, prototyping, peer review, defence, reflection, and dissemination. Methods are assigned according to function rather than added as classroom

variation. Each stage produces a compulsory artefact and has a transition rule, making progression dependent on evidence rather than attendance or completion alone.

Table 3. Nine-stage syntax of the Research-Based PjBL-SKI Model

Stage	Primary purpose	Key artefact	Main dimension
1. Alignment	Clarify outcomes, product, evidence, and standards.	Alignment matrix	D2, D4, D5
2. Authentic diagnosis	Establish an ethical, evidence-based instructional problem.	Needs dossier	D1
3. SKI-context analysis	Analyse curriculum, learners, resources, sources, and disciplinary characteristics.	Context and SKI map	D2
4. Theory-evidence synthesis	Build mechanisms, design rationale, and requirements.	Synthesis matrix	D3, D4
5. Project planning	Set roles, schedule, risks, milestones, and quality gates.	Project charter	D4, D5
6. Ideation and selection	Compare alternatives and justify one concept transparently.	Decision matrix	D3, D4, D5
7. Prototype construction	Construct the complete model and minimum support system.	Prototype + traceability matrix	D4
8. Review, defence, revision	Test evidence and reasoning; document improvement.	Review + revision log	D1-D5
9. Reflection and communication	Integrate portfolio, limits, integrity, and dissemination.	Final portfolio	D5

The social system is a collaborative design studio governed by evidence, academic dialogue, adab, transparency, and individual accountability. The lecturer safeguards historical, ethical, and design quality while students act as researchers, designers, reviewers, defenders, and reflective practitioners. Reaction principles include feed-up, feedback, and feed-forward; evidence-focused probing; contingent scaffolding; temporary suspension of evaluation during ideation; direct correction for critical historical or ethical errors; and remediation when transition criteria are not met. Eight support-system groups cover alignment resources, needs-research instruments, SKI/source-analysis tools, theory-synthesis templates, project-management tools, model-construction templates, assessment/feedback

instruments, and documentation/ethics/repository resources. Low-technology alternatives are included so that access constraints do not remove evidence, feedback, or individual accountability.

Expert appraisal and design refinement

Four experts independently appraised the prototype: two in educational research/instructional-model design and two in Islamic education/PAI/SKI. Across 400 item-level judgements, 355 ratings were 5 and 45 were 4; no item received a score from 1 to 3. The research-and-design experts awarded 978 of 1,000 possible points, while the Islamic-education/SKI experts awarded 977 of 1,000. The combined mean was 4.8875, rounded to 4.89, and conceptual feasibility was 97.75%. All validators concluded that the model could proceed without substantive conceptual revision.

Table 4. Integrated expert-appraisal results

Expert group	Combined score	Mean	Feasibility	Decision
R&D/instructional-model design (n = 2)	978/1,000	4.89	97.8%	Proceed without substantive revision
Islamic education/PAI/SKI (n = 2)	977/1,000	4.89	97.7%	Proceed without substantive revision
Overall (n = 4)	1,955/2,000	4.89	97.75%	Highly feasible conceptually

Qualitative comments converged on four strengths: traceability from empirical need to assessment; coherence among syntax, social system, reaction principles, support system, and intended outcomes; disciplinary specificity through historical-source responsibility, reasoning, HOTS, adab, and ibrah; and standardised flexibility for different schedules, technologies, experience levels, and institutional partners. Suggestions addressed accessibility and implementation readiness rather than conceptual reconstruction. The nine-stage syntax, D1-D5 outcomes, assessment construct, model identity, and the principle of historical validity before ibrah were therefore retained. Refinements concerned low-technology examples, visual navigation, differentiated support, dissent and revision logs, moderation anchors, a condensed sixteen-meeting guide, flexible clinic scheduling, repository navigation, and terminology consistency.

Discussion

From model use to evidence-informed design

The findings reposition the transition from model user to model designer as a change in epistemic responsibility rather than a change in vocabulary. A user may adapt an established procedure; a designer must justify why a problem requires intervention, why a mechanism is credible, how disciplinary constraints shape the solution, which evidence demonstrates quality, and how criticism changes the product. This interpretation converges with Yalçın et al. (2021), who conceptualise instructional design competence as integrated performance, and with Blömeke et al. (2022), who show that professional knowledge is mediated by situation-specific interpretation and decision making. The present study extends those accounts by adding an auditable design chain and by making SKI disciplinary reasoning part of the competence itself. In other words, competence is not only what the student knows or produces, but whether the route from problem evidence to design decision can be inspected.

Why PjBL requires research-based decision infrastructure

The needs data support PjBL as an organisational structure but also expose a limitation that recent PjBL literature can obscure when it focuses mainly on average outcomes. Zhang and Ma (2023) reported positive meta-analytic effects, and Yusri et al. (2024) identified inquiry, collaboration, communication, creativity, and problem solving as recurrent elements. Farshad and Fortin (2026), however, showed that the meta-analytic evidence is methodologically fragile: all 15 reviewed meta-analyses were rated critically low, and effect magnitudes varied markedly. The present study therefore does not treat PjBL as an inherently effective label. Its Class B evidence is instructive: a project condition produced artefacts, but product quality remained modest when milestones, quality gates, research requirements, and domain-specific reasoning were not fully engineered. This finding qualifies optimistic PjBL claims by suggesting that project presence is a weak proxy for project quality.

The model responds by integrating research-based learning as the decision infrastructure of the project. This converges with Wessels et al. (2021) and Thiem et al. (2023), who found that research-based learning can strengthen research competence, and with Matos et al. (2023), who showed the importance of connecting methodology to meaningful educational work. The present study differs in the target of the research process. Research is not the final product and is not treated as a parallel methods exercise; needs evidence justifies the problem, recent scholarship justifies the mechanism, historical sources constrain claims, and review

evidence justifies revision. This changes the function of research from 'learning to conduct research' to 'using research to govern professional design decisions.'

Disciplinary specificity: historical validity before ibrah

The model also diverges from generic PjBL by requiring a disciplinary grammar for SKI. Recent studies show that PjBL and inquiry can strengthen historical concepts, literacy, and thinking when students work with sources, local figures, heritage, or historical methods (Lionar et al., 2025; Matitaputty et al., 2024; Rahmawati & Hadi, 2026; Sariyatun et al., 2025). Research-based history learning likewise demonstrates the value of designing inquiry around disciplinary evidence (Universitas Riau Indonesia et al., 2025). The present study does not replicate those interventions; it shifts the unit of competence from learning history to designing history learning. Master's students must therefore demonstrate not only that historical inquiry is desirable, but exactly how provenance, chronology, causality, context, change-continuity, multiperspectivity, and argumentation are encoded in objectives, tasks, scaffolds, sources, and assessment.

The principle of historical validity before ibrah is a further extension. Contextual Islamic education can connect knowledge with virtue and relevant action (Parhan et al., 2024), but value integration can become superficial if moral conclusions are predetermined before evidence is examined. The developed model reverses that order. Students establish what can responsibly be claimed, recognise uncertainty and competing interpretations, and only then formulate ibrah that is historically defensible and pedagogically meaningful. The result is not a separation of Islamic values from history, but a sequencing rule intended to protect both: religious reflection is strengthened because it emerges from disciplined historical interpretation rather than replacing it.

Scaffolding, formative assessment, and observable feedback uptake

The heterogeneity of the master's students required common standards without assuming identical starting points. Self-regulated-learning evidence indicates that university students benefit from explicit support for goal setting, monitoring, strategy use, and reflection (Theobald, 2021). Accordingly, the model includes diagnostic entry points, project charters, milestone dashboards, peer mentoring, clinics, and contingent scaffolding that is reduced as competence develops. This differs from a template-heavy approach in which support becomes a form-filling substitute for judgement. The lecturer's task is to intervene just enough to make evidence and criteria visible, then return responsibility to the student.

Feedback is treated similarly as observable action rather than information delivered by the lecturer. Discipline-specific feedback literacy requires learners to understand standards, make judgements, manage responses, and use feedback in

context (Winstone et al., 2022). Teacher feedback literacy likewise depends on designing feedback regimes that create iterative change rather than one-off comments (Carless, 2023), and instructional models of feedback emphasise that students need structured opportunities to act on information (De Kleijn, 2023). The current model makes that action auditable through accept-modify-reject decisions, reasoned responses, revision logs, and version comparisons. In this sense, D5 does not sit at the end of D1-D4; it recursively tests them because a review comment can legitimately force the student back to the problem statement, disciplinary analysis, theoretical rationale, or model architecture.

Table 5. Critical comparison between the present findings and recent research

Recent research domain	What recent studies establish	What the present study shows differently	Critical contribution
PjBL syntheses (Farshad & Fortin, 2026; Yusri et al., 2024; Zhang & Ma, 2023)	PjBL is generally associated with positive outcomes, but implementation is heterogeneous and meta-analytic quality can be weak.	A project assignment alone did not ensure coherent design; evidence requirements, gates, and assessment architecture were decisive needs.	Shifts attention from whether PjBL works on average to how a postgraduate design project is engineered and audited.
Research-based learning (Matos et al., 2023; Thiem et al., 2023; Wessels et al., 2021)	Research participation can strengthen research competences and methodology learning.	Research evidence is embedded in every major design decision rather than treated as a separate research task.	Reframes research competence as professional decision infrastructure.
History inquiry/PjBL (Matitaputty et al., 2024; Rahmawati & Hadi, 2026; Sariyatun et al., 2025; Wilke et al., 2023)	Source inquiry and project work can strengthen historical thinking and disciplinary learning.	The target is postgraduate competence to design such learning, including source criteria, rubrics, and quality gates.	Moves from learner historical outcomes to teacher-designer historical pedagogy.
Instructional-design competence (Blömeke et al., 2022; Yalçın et al., 2021)	Professional competence is integrated and decision-oriented.	Competence includes auditable traceability and SKI-specific historical/value constraints.	Adds disciplinary and evidentiary accountability to generic design competence.

Recent research domain	What recent studies establish	What the present study shows differently	Critical contribution
Feedback literacy (Carless, 2023; De Kleijn, 2023; Winstone et al., 2022)	Effective feedback depends on judgement and learner action.	Feedback uptake is documented through response decisions, revision logs, and version comparison.	Turns feedback literacy into visible evidence of design competence.
Contextual Islamic education (Parhan et al., 2024)	Islamic learning can connect knowledge, context, and virtue.	Ibrah is formulated only after historical evidence and interpretation have been critically established.	Protects value reflection from premature moralising through the rule historical validity before ibrah.

Design traceability, expert feasibility, and the boundary of the claim

The 97.75% expert-feasibility result is meaningful primarily because it evaluates a traceable design rather than a detached list of features. Forward traceability connects needs to DR-01-DR-11 and then to D1-D5, syntax, artefacts, interaction, assessment, and intended outcomes. Backward traceability allows every activity, instrument, or rule to be challenged: a component without a documented need or mechanism is weakly justified, while a priority need without a design response indicates incompleteness. This design logic is consistent with recent design studies in history education that iteratively connect problems, disciplinary mechanisms, and instructional forms (Wilke et al., 2023), but the present work formalises the links as an explicit audit trail across the whole postgraduate course.

The near-identical ratings of the two expert groups are also informative. Research/design specialists judged the architecture coherent and operational, whereas Islamic-education/SKI specialists judged the historical, pedagogical, ethical, and value dimensions appropriate. Agreement across perspectives reduces the risk that technical completeness was achieved by weakening disciplinary identity, or that disciplinary richness was achieved without an implementable architecture. Nevertheless, expert agreement remains documentary and conceptual evidence. It cannot establish usability, lecturer fidelity, inter-rater reliability, equitable participation, or student learning effects. The absence of low ratings therefore supports progression to empirical testing, not an effectiveness claim.

Limitations and a staged agenda for subsequent research

The study has four main limitations. First, previous course data were used diagnostically and were not outcomes of the developed model. Second, the

purposive single-programme context limits claims about transferability. Third, four experts evaluated documents and design logic rather than live lecturer-student interaction. Fourth, D1-D5 rubrics, historical-reasoning criteria, quality gates, support tools, and fidelity instruments have not yet been tested for reliability, construct behaviour, workload, accessibility, or bias. These limitations require a progressive research programme in which each phase answers a different validity question rather than moving directly from expert approval to an effectiveness trial.

Table 6. Specific staged agenda for future empirical research

Phase	Suggested design and sample	Primary evidence	Decision before progression
1. Readability/ usability	Cognitive walkthrough with 6-10 lecturers and 12-20 master's students from at least two cohorts.	Task completion, misunderstood terms, navigation errors, think-aloud data, time-on-task.	Revise any component that produces recurrent interpretation or navigation failure.
2. Limited semester pilot	Two classes or approximately 30-50 students using the full 16-meeting sequence.	Fidelity logs, quality-gate pass/rework rates, artefact progression, scaffolding frequency, workload, access barriers, acceptance.	Establish feasible time, support load, and minimum implementation conditions.
3. Measurement study	Multiple trained raters scoring a common bank of student artefacts.	Inter-rater agreement/ICC or weighted kappa, moderation drift, rubric sensitivity, evidence of score structure and fairness.	Retain only criteria that can be interpreted consistently and without systematic bias.
4. Effectiveness study	Matched-cohort or quasi-experimental design across 4-6 course sections, with pre/post or repeated D1-D5 performance evidence.	D1-D5 gains, artefact quality, effect sizes with uncertainty, fidelity, prior experience, and feedback uptake as moderators/mediators.	Make effectiveness claims only if outcome improvement co-occurs with adequate fidelity and reliable measurement.
5. Transfer and adaptation	Replication in at least three institutions and across school, madrasah, and higher-education design targets.	Context-mechanism-component comparisons, low-tech feasibility, local adaptations, equity of access.	Distinguish invariant core principles from adaptable supports, schedules, and media.

This sequence makes future research falsifiable and cumulative. A weak usability result would require redesign before a pilot; poor inter-rater agreement would prevent strong claims about competence gains; and effectiveness estimates would be interpreted together with fidelity and access rather than as decontextualised averages. Cross-programme replication would then determine which elements constitute the model's core theory and which should remain context-sensitive adaptations.

Type your text here using Cambria 12pt font, 1.15 spacing. The discussion section provides the author with the interpretation of the results. In this section, the author should compare the findings with existing literature, analyze whether the research results align with previously proposed theories or hypotheses, and explore the implications of the findings. The author should also explain the limitations of the study and how these limitations may impact the results or generalizability of the findings. Furthermore, the discussion should focus on the relevance of the findings to both theory and practical applications, and provide suggestions for further research. This section provides an opportunity to delve deeper into the meaning of the research results and provide a broader perspective on the topic under study.

CONCLUSION

This study produced a conceptually validated Research-Based PjBL-SKI Model that prepares master's students in Islamic Religious Education to investigate instructional needs and use evidence to design Islamic Cultural History instruction. The integrated needs analysis showed that written curriculum expectations, professional teaching experience, and even a project assignment were insufficient when research requirements, milestones, historical reasoning criteria, aligned assessment, quality gates, feedback uptake, and documented revision were not systematically connected. These findings generated eleven design requirements and five competence dimensions covering evidence-based diagnosis, curriculum and SKI analysis, theory-research synthesis, coherent model construction, and the assessment, defence, and revision of design decisions.

The model translates those requirements into a nine-stage course architecture: alignment, authentic diagnosis, SKI-context analysis, theory-evidence synthesis, project planning, concept selection, prototype construction, academic review-revision, and reflective communication. Its distinctive contribution is the auditable relationship among problem evidence, theory, disciplinary reasoning, design decisions, artefacts, assessment, and revision. The principle of historical validity before ibrah ensures that value reflection follows rather than replaces responsible historical interpretation. Four experts awarded 1,955 of 2,000 possible

points, a mean of 4.89, and 97.75% conceptual feasibility; their recommendations concerned implementation readiness rather than reconstruction of the model's core logic.

The study therefore offers a defensible design hypothesis rather than a proven intervention. The contribution should be judged by the clarity of its design logic, disciplinary specificity, and traceability, not by untested claims of effectiveness. Empirical progression should begin with readability and usability, continue through a limited semester pilot and measurement-reliability study, and only then proceed to multi-section effectiveness and cross-institutional transfer studies. This staged agenda is necessary to determine whether the model can be implemented with adequate fidelity, manageable workload, reliable assessment, equitable access, and measurable improvement in postgraduate instructional-design competence.

ACKNOWLEDGMENT

The authors acknowledge UIN Jurai Siwo Lampung and its Institute for Research and Community Service for institutional support under Assignment Letter No. B-0051a/Un.36/L.1/PP.00.9/02/2026. Appreciation is also extended to the participating students, lecturers, teachers, programme management, and expert validators whose documents, accounts, and appraisals informed the development of the conceptual model.

REFERENCES

Adeoye-Olatunde, O. A., & Olenik, N. L. (2021). Research and scholarly methods: Semi-structured interviews. *JACCP: JOURNAL OF THE AMERICAN COLLEGE OF CLINICAL PHARMACY*, *4*(10), 1358–1367. https://doi.org/10.1002/jac5.1441

Blömeke, S., Jentsch, A., Ross, N., Kaiser, G., & König, J. (2022). Opening up the black box: Teacher competence, instructional quality, and students' learning progress. *Learning and Instruction*, *79*, 101600. https://doi.org/10.1016/j.learninstruc.2022.101600

Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, *18*(3), 328–352. https://doi.org/10.1080/14780887.2020.1769238

Carless, D. (2023). Teacher feedback literacy, feedback regimes and iterative change: Towards enhanced value in feedback processes. *Higher Education*

Research & Development, *42*(8), 1890–1904. https://doi.org/10.1080/07294360.2023.2203472

Claravall, E. B., & Irey, R. (2022). Fostering historical thinking: The use of document based instruction for students with learning differences. *The Journal of Social Studies Research*, *46*(3), 249–264. https://doi.org/10.1016/j.jssr.2021.08.001

De Kleijn, R. A. M. (2023). Supporting student and teacher feedback literacy: An instructional model for student feedback processes. *Assessment & Evaluation in Higher Education*, *48*(2), 186–200. https://doi.org/10.1080/02602938.2021.1967283

Farshad, S., & Fortin, C. (2026). An umbrella review of meta-analyses on project-based learning: Effects on academic achievement, higher-order thinking, and 21st-century skills. *Educational Research Review*, *52*, 100809. https://doi.org/10.1016/j.edurev.2026.100809

Hailikari, T., Virtanen, V., Vesalainen, M., & Postareff, L. (2022). Student perspectives on how different elements of constructive alignment support active learning. *Active Learning in Higher Education*, *23*(3), 217–231. https://doi.org/10.1177/1469787421989160

Lionar, U., Mulyana, A., Sapriya, Winarti, M., Yefterson, R. B., Syafrina, Y., Naldi, H., & Yuhardi. (2025). Investigating Local Historical Figures: The Implementation of the PjBL Model to Increase Historical Literacy. *Paramita: Historical Studies Journal*, *35*(1). https://doi.org/10.15294/paramita.v35i1.23411

Loughlin, C., Lygo-Baker, S., & Lindberg-Sand, Å. (2021). Reclaiming constructive alignment. *European Journal of Higher Education*, *11*(2), 119–136. https://doi.org/10.1080/21568235.2020.1816197

Love, H. R., & Corr, C. (2022). Integrating Without Quantitizing: Two Examples of Deductive Analysis Strategies Within Qualitatively Driven Mixed Methods Research. *Journal of Mixed Methods Research*, *16*(1), 64–87. https://doi.org/10.1177/1558689821989833

Matitaputty, J. K., Saputra, N., Judijanto, L., Susanto, N., Hanif, M., Sopacua, J., & Fadli, M. R. (2024). PjBL-based digital history model to improve historical concept skills and historical consciousness. *Journal of Education and*

Learning (EduLearn), *18*(2), 430–440. https://doi.org/10.11591/edulearn.v18i2.21152

Matos, J. F., Piedade, J., Freitas, A., Pedro, N., Dorotea, N., Pedro, A., & Galego, C. (2023). Teaching and Learning Research Methodologies in Education: A Systematic Literature Review. *Education Sciences*, *13*(2), 173. https://doi.org/10.3390/educsci13020173

Morgan, D. L., & Hoffman, K. (2021). Searching for qualitatively driven mixed methods research: A citation analysis. *Quality & Quantity*, *55*(2), 731–740. https://doi.org/10.1007/s11135-020-01025-2

Parhan, M., Syahidin, S., Somad, M. A., Abdulah, M., & Nugraha, R. H. (2024). Developing a Contextual Learning Model in Islamic Education to Improve Applicable Knowledge and Foster Knowledge-Based Virtues. *Jurnal Pendidikan Islam*, *10*(1), 75–86. https://doi.org/10.15575/jpi.v10i1.35205

Rahmawati, B. F., & Hadi, M. S. (2026). The Effectiveness of Project-Based Learning Integrated with Historical Methods in Improving Students' Historical Thinking Skills. *Fajar Historia: Jurnal Ilmu Sejarah Dan Pendidikan*, *10*(2), 636–646. https://doi.org/10.29408/fhs.v10i2.36001

Rehman, N., Huang, X., Mahmood, A., AlGerafi, M. A. M., & Javed, S. (2024). Project-based learning as a catalyst for 21st-Century skills and student engagement in the math classroom. *Heliyon*, *10*(23), e39988. https://doi.org/10.1016/j.heliyon.2024.e39988

Roro Wilis & Dyah Kumalasari. (2025). The Jogja Library Center Tour: Literacy-Based Inquiry Model to Improve Historical Thinking Skill. *Jurnal Pedagogi Dan Pembelajaran*, *8*(3), 594–604. https://doi.org/10.23887/jp2.v8i3.100800

Sariyatun, S., Sutimin, L. A., & Abidin, N. F. (2025). Enhancing Historical Thinking and HOTS through Digital Learning Materials: An Inquiry-Based Approach Using Primary Sources for High School Students. *Journal of Curriculum Studies Research*, *7*(2), 228–248. https://doi.org/10.46303/jcsr.2025.19

Theobald, M. (2021). Self-regulated learning training programs enhance university students' academic performance, self-regulated learning strategies, and

motivation: A meta-analysis. *Contemporary Educational Psychology*, *66*, 101976. https://doi.org/10.1016/j.cedpsych.2021.101976

Thiem, J., Preetz, R., & Haberstroh, S. (2023). How research-based learning affects students' self-rated research competences: Evidence from a longitudinal study across disciplines. *Studies in Higher Education*, *48*(7), 1039–1051. https://doi.org/10.1080/03075079.2023.2181326

Universitas Riau Indonesia, Isjoni, I., Sumarno, S., & Universitas Riau Indonesia. (2025). Development of a research-based learning model to enhance historical thinking skills. *Perspectives of Science and Education*, *75*(3), 201–217. https://doi.org/10.32744/pse.2025.3.13

Van Doorsselaere, J. (2025). Meaningful learning beyond the textbook: A case study of student experiences during an authentic historical inquiry on local heritage. *History Education Research Journal*, *22*(1). https://doi.org/10.14324/HERJ.22.1.09

Wessels, I., Rueß, J., Gess, C., Deicke, W., & Ziegler, M. (2021). Is research-based learning effective? Evidence from a pre–post analysis in the social sciences. *Studies in Higher Education*, *46*(12), 2595–2609. https://doi.org/10.1080/03075079.2020.1739014

Wilke, M., Depaepe, F., & Van Nieuwenhuyse, K. (2023). Fostering Secondary Students' Historical Thinking: A Design Study in Flemish History Education. *Journal of Formative Design in Learning*, *7*(1), 61–81. https://doi.org/10.1007/s41686-023-00074-8

Winstone, N. E., Balloo, K., & Carless, D. (2022). Discipline-specific feedback literacies: A framework for curriculum design. *Higher Education*, *83*(1), 57–77. https://doi.org/10.1007/s10734-020-00632-0

Yalçın, Y., Ursavaş, Ö. F., & Klein, J. D. (2021). Measuring instructional design competencies of future professionals: Construct validity of the ibstpi® standards. *Educational Technology Research and Development*, *69*(3), 1701–1727. https://doi.org/10.1007/s11423-021-10006-7

Yusri, R., Yusof, A. M., & Sharina, A. (2024). A systematic literature review of project-based learning: Research trends, methods, elements, and

frameworks. *International Journal of Evaluation and Research in Education (IJERE)*, *13*(5), 3345. https://doi.org/10.11591/ijere.v13i5.27875

Zhang, L., & Ma, Y. (2023). A study of the impact of project-based learning on student learning effects: A meta-analysis study. *Frontiers in Psychology*, *14*, 1202728. https://doi.org/10.3389/fpsyg.2023.1202728