

EVALUATING THE IMPLEMENTATION OF THE MERDEKA CURRICULUM IN ISLAMIC RELIGIOUS EDUCATION AT A PESANTREN-BASED JUNIOR HIGH SCHOOL: A CIPP CASE STUDY

***¹Ahmad Yaim, ²Ahmad Zainuri, ³Fitri Ofiyanti, ⁴Indah Wigati, ⁵Maryamah**

^{*1,2,3,4,5}UIN Raden Fatah Palembang

Email: ^{*1,2,3,4,5}ahmadyaim@gmail.com

Abstract

The implementation of Indonesia's Merdeka Curriculum requires schools to translate curricular flexibility into context-responsive teaching, assessment, and institutional support. This study evaluates the implementation of the Merdeka Curriculum in Islamic Religious Education (PAI) at a pesantren-based junior high school in Lempuing Jaya, South Sumatra, Indonesia, using the Context, Input, Process, and Product (CIPP) evaluation model. An evaluative qualitative case-study design was employed. Data were generated through participatory classroom and project observations, semi-structured interviews with school and curriculum leaders, PAI teachers, a supervising inspector, and students, and analysis of curriculum and assessment documents. Data were analyzed using an interactive qualitative analysis procedure involving condensation, display, and conclusion drawing, with source and method triangulation used to strengthen credibility. The findings reveal a marked implementation asymmetry. Contextually, the school's religious ecology and institutional mission strongly support PAI and character development, but the school's operational curriculum documents remain insufficiently contextualized. At the input level, digital infrastructure and learning resources are relatively strong, whereas teacher pedagogical readiness, diagnostic-assessment documentation, project guidance, and operational quality-assurance mechanisms remain limited. At the process level, teachers have adapted learning through digital tools, block scheduling, and contextual problem-based activities, but differentiated instruction, formative assessment, reflective practice, and the reflective phase of the Pancasila Student Profile Strengthening Project are inconsistently implemented. Product evidence suggests promising development of students' religious character and digital participation, but the reported 85% achievement is an institutional estimate rather than an independently measured outcome, while documentation of learning attainment, stakeholder satisfaction, and differentiated student products remains incomplete. The study argues that the principal implementation challenge is not infrastructure but the alignment of teacher capability, assessment evidence, curriculum governance, and reflective improvement. A staged improvement strategy is proposed through peer coaching, contextualization of the school operational curriculum, diagnostic assessment capacity building, integration between school and boarding-school programs, and stronger evidence-based quality assurance.

Keywords: Merdeka Curriculum, Islamic Religious Education, CIPP evaluation, pesantren, differentiated instruction, curriculum implementation, qualitative case study

Abstrak

Implementasi Kurikulum Merdeka di Indonesia menuntut sekolah untuk menerjemahkan fleksibilitas kurikulum menjadi pembelajaran, asesmen, dan dukungan institusional yang responsif terhadap konteks. Penelitian ini mengevaluasi implementasi Kurikulum Merdeka pada mata pelajaran Pendidikan Agama Islam (PAI) di sebuah sekolah menengah pertama berbasis pesantren di Lempuing Jaya, Sumatera Selatan, Indonesia, dengan menggunakan model evaluasi Context, Input, Process, and Product (CIPP). Penelitian ini menggunakan desain studi kasus kualitatif evaluatif. Data diperoleh melalui observasi partisipatif di kelas dan kegiatan proyek, wawancara terstruktur dengan pimpinan sekolah dan kurikulum, guru PAI, pengawas sekolah, serta siswa, serta analisis dokumen kurikulum dan asesmen. Data dianalisis menggunakan prosedur analisis kualitatif interaktif yang meliputi kondensasi data, penyajian data, dan penarikan kesimpulan, dengan triangulasi sumber dan metode yang digunakan untuk memperkuat kredibilitas. Temuan penelitian mengungkapkan adanya asimetri implementasi yang signifikan. Secara kontekstual, ekologi keagamaan dan misi institusional sekolah sangat mendukung PAI dan pengembangan karakter, namun dokumen kurikulum operasional sekolah masih belum cukup terkontekstualisasi. Pada tingkat input, infrastruktur digital dan sumber belajar relatif kuat, sedangkan kesiapan pedagogis guru, dokumentasi asesmen diagnostik, panduan proyek, serta mekanisme penjaminan mutu operasional masih terbatas. Pada tingkat process, guru telah mengadaptasi pembelajaran melalui perangkat digital, penjadwalan blok, dan aktivitas berbasis masalah yang kontekstual, tetapi pembelajaran terdiferensiasi, asesmen formatif, praktik reflektif, serta tahap refleksi pada Proyek Penguatan Profil Pelajar Pancasila belum diimplementasikan secara konsisten. Bukti product menunjukkan perkembangan karakter keagamaan dan partisipasi digital siswa yang menjanjikan, namun pencapaian 85% yang dilaporkan merupakan estimasi institusional dan bukan hasil pengukuran independen, sementara dokumentasi pencapaian pembelajaran, kepuasan pemangku kepentingan, dan produk siswa yang terdiferensiasi masih belum lengkap. Penelitian ini berargumen bahwa tantangan utama implementasi bukanlah infrastruktur, melainkan penyelarasan kapabilitas guru, bukti asesmen, tata kelola kurikulum, dan perbaikan berbasis refleksi. Strategi perbaikan bertahap diusulkan melalui peer coaching, kontekstualisasi Kurikulum Operasional Satuan Pendidikan (KOSP), peningkatan kapasitas asesmen diagnostik, integrasi antara program sekolah dan pesantren, serta penjaminan mutu berbasis bukti yang lebih kuat.

Kata Kunci: Kurikulum Merdeka, Pendidikan Agama Islam, Evaluasi CIPP, Pesantren, Pembelajaran Terdiferensiasi, Implementasi Kurikulum, Studi Kasus Kualitatif

INTRODUCTION

Curriculum reform is consequential not because a new policy document is adopted, but because policy intentions are translated into everyday instructional practice. Indonesia's Merdeka Curriculum was designed to provide greater curricular flexibility, emphasize essential competencies, and strengthen learner development. Its implementation nevertheless depends on the capacity of schools and teachers to interpret curricular goals, assess learner needs, design responsive instruction, and use evidence for continuous improvement. Recent studies of Merdeka Curriculum implementation similarly report that teachers often recognize the value of student-centered and differentiated learning while experiencing difficulties related to professional readiness, assessment, learning-

module development, and institutional support. (Putri & Pranata, 2024; Al Arsyadhi et al., 2024; Hadi & Wahyudin, 2024)

The policy context has also evolved. The current national regulatory framework is anchored in Ministerial Regulation No. 12 of 2024 concerning the curriculum for early childhood, primary, and secondary education and was amended by Ministerial Regulation No. 13 of 2025. The 2025 amendment did not introduce a wholly new curriculum; rather, it strengthened the existing direction through, among other changes, a deeper-learning approach and revised arrangements for curricular and co-curricular learning. This policy development makes evaluation particularly important because implementation quality must be interpreted against the regulatory framework that is actually in force. (Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia, 2024; Ministry of Primary and Secondary Education of the Republic of Indonesia, 2025)

Islamic Religious Education (PAI) presents a distinctive implementation context. PAI is expected to support not only disciplinary knowledge but also ethical, spiritual, and character development. In a pesantren-based school, these aims intersect with an institutional ecology that extends beyond the formal classroom into boarding routines, religious practice, teacher–student relationships, and collaboration with pesantren educators. The resulting environment can provide a strong cultural foundation for religious and character formation, but it can also create distinctive implementation demands, including coordination between formal schooling and residential education, management of student fatigue, and adaptation of learning to the rhythms of pesantren life. (Khair et al., 2026; Yani et al., 2025)

The case examined in this study is SMP As-Shiddiqiyah Lempuing Jaya, a pesantren-based junior high school that had implemented the Merdeka Curriculum and served as a school with an early implementation role. The source study identified an apparent contrast between relatively strong physical and digital resources and continuing weaknesses in teacher readiness, differentiated instruction, diagnostic assessment, curriculum documentation, formative assessment, and project implementation. Such a configuration is analytically important because it challenges a common assumption that implementation problems are primarily caused by inadequate facilities.

Evaluation is therefore required to move beyond a simple description of whether the curriculum is being implemented. The CIPP model is particularly appropriate because it treats evaluation as a decision-oriented process for identifying needs and conditions (Context), examining the adequacy of strategies and resources (Input), monitoring implementation (Process), and judging outcomes for improvement and accountability (Product). Stufflebeam and Zhang’s formulation emphasizes that CIPP is not limited to proving whether planned activities were completed; it is intended to generate information for improvement throughout a program. (Stufflebeam & Zhang, 2017)

Although studies of the Merdeka Curriculum have expanded rapidly, fewer studies provide an integrated CIPP evaluation of PAI implementation in a pesantren-based junior

high school while simultaneously examining curriculum governance, diagnostic assessment, differentiated learning, digital resources, project-based learning, and sustainability. Recent research in PAI contexts confirms the continuing need to understand how national curriculum reform is enacted within specific institutional settings, but the evidence remains context dependent. (Batubara et al., 2025; Khair et al., 2026; Yani et al., 2025)

This study addresses that gap by asking: (1) How does the context of the pesantren-based school support or constrain PAI implementation under the Merdeka Curriculum? (2) Are the human, documentary, infrastructural, and financial inputs adequate for implementation? (3) How is the curriculum enacted in classroom and project processes? and (4) What products and improvement needs emerge from implementation? The contribution of the study is not to claim generalizability from a single school, but to provide a theoretically structured account of how implementation gaps can persist even where infrastructure is comparatively strong.

The Merdeka Curriculum is treated in this study as an implementation system rather than merely a curriculum document. Implementation requires alignment among policy expectations, school-level curriculum design, teacher competence, learner assessment, instructional practices, and mechanisms for reflection and improvement. Differentiated learning is particularly important because it requires teachers to use information about learner readiness and needs to adapt content, process, and/or products rather than reproduce a uniform instructional sequence. Recent empirical work likewise identifies diagnostic assessment and continuous assessment as important supports for differentiated learning. (Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia, 2022; Nazilah, 2024; Rahmadewi et al., 2025)

CIPP consists of four interconnected evaluation domains. Context evaluation examines needs, problems, institutional goals, and environmental conditions. Input evaluation examines strategies, resources, human capacity, and implementation arrangements. Process evaluation monitors implementation and identifies deviations, barriers, and opportunities for improvement. Product evaluation examines intended and unintended outcomes and provides evidence for continuation, modification, or termination. The model is therefore well suited to curriculum evaluation because it connects implementation conditions with improvement decisions rather than treating outcomes as the only evidence of program quality. (Stufflebeam & Zhang, 2017)

CIPP DOMAIN	ANALYTICAL FOCUS	KEY EVIDENCE
IN THIS STUDY		
CONTEXT	Alignment of PAI goals, learner needs, pesantren ecology, institutional policy, and current curriculum policy	Vision–mission, learner needs, KOSP, policy documents, interviews

INPUT	Teacher readiness, curriculum documents, diagnostic assessment documents, capacity, infrastructure, project resources, financing	Training records, KOSP, modules, assessment documents, facilities, ARKAS/project records
PROCESS	Classroom enactment, differentiated instruction, assessment, digital learning, P5/co-curricular implementation, supervision	Observations, teacher/student accounts, LKPD, rubrics, reflection records
PRODUCT	Learning and character outcomes, student products, documentation of attainment, sustainability	Portfolios, project products, assessment records, observations, follow-up plans

METHOD

Research design and setting

The study employed an evaluative qualitative case-study design guided by the CIPP framework. The case was SMP As-Shiddiqiyah Lempuing Jaya, a junior high school operating within a pesantren environment in Lempuing Jaya, South Sumatra, Indonesia. The design was selected because the research sought to understand how the Merdeka Curriculum was enacted within a bounded institutional setting and to identify actionable gaps across contextual, resource, implementation, and outcome dimensions. (Yin, 2018)

Participants and data sources

Participants were selected purposively because they occupied roles directly connected to curriculum implementation. The source manuscript identifies the principal, vice principal for curriculum, PAI teachers, a supervising inspector, and students in Grades VII and VIII as key information sources. Documentary sources included the School Operational Curriculum (KOSP), teaching modules, project documentation, assessment records/grade ledgers, and academic-supervision documents. Because the source manuscript does not report the exact number of participants in each category, this revision does not invent participant counts; the final submission should add the verified numbers and participant characteristics if available. The use of purposive role-based selection is consistent with qualitative case-study logic, although the final manuscript should report the verified number of participants and their characteristics (Yin, 2018).

Data collection

Three complementary techniques were used: participatory observation, semi-structured interviews, and document analysis. Observation covered classroom learning and the implementation of the project activities. Interviews explored curriculum understanding,

teacher adaptation, assessment practices, resource constraints, and perceptions of outcomes. Document analysis examined the extent to which reported practices were supported by formal curriculum, assessment, project, and supervision records. This combination enabled discrepancies between reported practice and documentary evidence to be examined rather than resolved in favor of a single source.

Data analysis

Data analysis followed the interactive qualitative procedure described by Miles, Huberman, and Saldaña: data condensation, data display, and conclusion drawing/verification. The CIPP dimensions provided an initial analytical structure, while subthemes were developed from the empirical material. Particular attention was paid to convergence and divergence across interviews, observations, and documents. For example, claims that diagnostic assessment had been conducted were compared with the availability of documentary evidence; claims about learning outcomes were compared with assessment and portfolio records. (Miles et al., 2014)

Trustworthiness

Credibility was strengthened through source triangulation and technique triangulation, supplemented by peer discussion. The analysis therefore distinguishes between reported perceptions, observed practices, and documented evidence. This is particularly important in program evaluation because administrative compliance and substantive implementation may not be identical. (Heale & Forbes, 2013)

Ethical considerations

The source manuscript does not report an ethics approval number, informed-consent procedure, or institutional review process. These details should be inserted from the authors' actual records before submission to an international journal. No approval number or consent procedure is fabricated in this revision.

RESULTS AND DISCUSSION

Context: strong institutional fit but incomplete operationalization

The context evaluation indicates a strong normative fit between PAI, the school's religious mission, and the Merdeka Curriculum's emphasis on competency and character development. The school integrates formal PAI learning with religious routines such as congregational prayer, Qur'anic recitation, tahfiz/tahsin activities, and character formation. This institutional ecology provides an important contextual resource because students encounter religious learning not only as classroom content but also as lived practice. (Khair et al., 2026; Yani et al., 2025)

However, contextual alignment at the level of institutional values does not automatically produce contextualized curriculum governance. Documentary evidence indicated that the KOSP had not been sufficiently adapted to the distinctive characteristics of the pesantren-based school, and school-level policy documents lacked sufficiently explicit standard

operating procedures and role allocation. The central contextual finding is therefore a mismatch between strong institutional purpose and weaker operational translation.

A second contextual issue concerns learner diversity. Teachers reported conducting initial diagnostic activities involving Qur’anic reading, oral or written understanding of religious concepts, and behavioral observation. Yet the document review did not identify a complete, formal record of the initial PAI diagnostic assessment. The evidence should therefore be interpreted carefully: the problem is not necessarily that no diagnostic activity occurred, but that the diagnostic process was insufficiently documented to function as a reliable planning baseline. This distinction is methodologically important and resolves an apparent contradiction in the original manuscript.

Input: infrastructure is stronger than pedagogical readiness

The input evaluation reveals a pronounced asymmetry. The school possesses substantial physical and digital resources, including classrooms, a library, Qur’anic and PAI learning materials, LCD projectors, internet access, and student Chromebooks. These resources provide a favorable infrastructure for student-centered and technology-supported learning.

The principal weakness lies in human and organizational capacity. Teachers had limited exposure to sustained professional development on the substantive demands of the Merdeka Curriculum. Training was described as insufficiently intensive to support the design of differentiated teaching modules and diagnostic assessment instruments. As a result, teaching materials and student worksheets remained relatively uniform despite heterogeneous learner readiness. (Putri & Pranata, 2024; Al Arsyadhi et al., 2024; Mahesa et al., 2025)

Curriculum governance also showed weaknesses. The KOSP was reported to be insufficiently contextualized, internal MGMP activity was described as predominantly administrative, and an operational instrument for systematic review of teaching modules was not available. In addition, the project module for the Pancasila Student Profile Strengthening Project was not sufficiently operationalized, while school budget planning did not consistently provide dedicated resources for practical PAI activities and student project products. The input finding is therefore clear: the principal constraint is not a lack of basic infrastructure but incomplete alignment among teacher capability, documentation, quality assurance, and resource allocation.

Process: adaptive practice alongside inconsistent differentiation and reflection

Classroom implementation was broadly consistent with the intended teaching scenarios, and digital tools—particularly Chromebooks and Google Classroom—were reported to increase student engagement with digital learning. However, the original manuscript’s estimate of approximately 80% alignment between implementation and teaching modules is not supported by a reported measurement procedure. It is therefore treated here as an observational or institutional estimate rather than a quantitative implementation score.

Time management emerged as a recurrent process challenge, especially during group discussion and whole-class presentations. School leaders responded by modifying scheduling arrangements and using a block-based schedule to provide longer instructional periods. Some memorization-oriented activities were shifted to the boarding-school environment so that school hours could prioritize discussion, problem solving, and applied learning. This represents an important form of contextual adaptation rather than simple noncompliance with the curriculum.

Student fatigue was another contextual constraint. The residential schedule sometimes extended into the evening, reducing concentration and stamina during morning classes. Teachers responded by moving from predominantly lecture-based instruction toward problem-based and contextually relevant activities connected to students' experiences in the boarding environment. The adaptation demonstrates the value of teacher agency, but it also indicates that curriculum implementation must account for the interaction between formal school schedules and residential education.

Process evidence nevertheless revealed persistent weaknesses in differentiated instruction and formative assessment. Student worksheets were often insufficiently differentiated, portfolio rubrics did not fully specify non-cognitive performance evidence, and teacher reflection records were not consistently completed. Classroom observation also indicated limited two-way interaction in some Grade VII learning situations. (Kahmann et al., 2022; Roqobih et al., 2024)

The project component showed a similar pattern. Students were active in the introductory and action stages of a project on diversity and tolerance, including poster production, but contextualization, external partnership, and structured reflection were weaker. The absence of a systematic reflection instrument reduced the opportunity to convert project activity into deeper learning and sustained character development. This is a critical process finding because project-based learning should not be judged primarily by the visibility of the final product. (Yani et al., 2025)

Product: promising character development but weak outcome documentation and sustainability

Product evidence indicates positive developments in students' religious character, participation, and digital presentation skills. School leaders and observations suggested improvements in students' confidence in presenting religious ideas, use of digital tools, tolerance, respect, and religious habituation. The pesantren environment appears to provide a strong ecological support for these outcomes.

The source manuscript reports an approximately 85% level of achievement of essential curricular targets according to the principal. This figure is retained only as a reported institutional estimate. It should not be presented as an independently verified effectiveness percentage because the manuscript does not provide a calculation formula, denominator, measurement instrument, or independent assessment procedure. This is a major methodological correction for international publication.

The product evaluation also identifies important evidence gaps. Documentation of phase-level learning attainment was incomplete; stakeholder satisfaction data from students and parents were unavailable; portfolios and project products were relatively uniform; and a concrete post-project follow-up plan was insufficiently documented. Thus, the product component presents a mixed conclusion: substantive signs of positive character and digital engagement coexist with limited evidence for systematic outcome measurement and sustainability.

CIPP COMPONENT	STRENGTHS IDENTIFIED	CRITICAL GAPS	EVALUATION JUDGEMENT
CONTEXT	Strong alignment with pesantren mission; PAI integrated with religious life; learner needs recognized	KOSP insufficiently contextualized; SOP/roles not fully operational; diagnostic evidence incompletely documented	Strong contextual potential, moderate operational readiness
INPUT	Digital infrastructure, learning resources, supervision instruments	Limited teacher training; weak diagnostic-assessment capacity; weak project documentation; limited dedicated project/practical funding	Infrastructure-ready but capability/governance constrained
PROCESS	Digital learning; adaptive scheduling; contextual PBL; supportive supervision	Inconsistent differentiation, formative assessment, reflection, and project contextualization	Adaptive but uneven implementation
PRODUCT	Religious character, participation, digital presentation, tolerance	Outcome documentation incomplete; products relatively uniform; weak stakeholder	Promising outcomes with limited accountability evidence

DISCUSSION

The implementation problem is an alignment problem

The most important finding is the asymmetry between material readiness and pedagogical-institutional readiness. The school possesses the infrastructure needed to support contemporary learning, but infrastructure alone does not produce differentiated instruction or evidence-based assessment. This finding is consistent with recent Merdeka Curriculum research showing that teacher adaptation is shaped by professional knowledge, assessment literacy, technological competence, and institutional support rather than by policy adoption alone. (Putri & Pranata, 2024; Al Arsyadhi et al., 2024; Mahesa et al., 2025)

From a CIPP perspective, the school's central problem is therefore not a single deficient component. It is a coupling problem across components: context establishes ambitious goals; input provides substantial infrastructure but insufficient pedagogical capacity; process contains adaptive practices but inconsistent differentiation and reflection; and product shows encouraging character-related evidence without a sufficiently robust measurement and sustainability system. CIPP is particularly useful here because it makes these interdependencies visible. (Stufflebeam & Zhang, 2017)

Diagnostic assessment is the missing bridge between curriculum flexibility and differentiation

The discrepancy between reported diagnostic activities and missing documentary evidence is one of the most consequential findings. Differentiated learning requires information about students' starting points, but such information must be sufficiently structured to inform instructional decisions. Official implementation guidance has emphasized that educators can begin the transition by conducting initial assessment and becoming increasingly responsive to differences in learner needs. Recent empirical work similarly identifies continuous assessment and readiness mapping as important mechanisms for differentiated learning. (Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia, 2022; Nazilah, 2024; Rahmadewi et al., 2025; Roqobih et al., 2024)

The implication is not that every diagnostic activity must become administratively burdensome. Rather, schools need a minimum evidence architecture: what was assessed, when it was assessed, what patterns were found, and how those findings changed content, process, grouping, scaffolding, or product options. Without this chain, differentiated learning can become a rhetorical commitment without traceable instructional consequences.

Pesantren should be treated as an implementation resource, not merely a background setting

The case demonstrates that the pesantren environment can function as a curricular resource. Religious routines and boarding activities create opportunities to connect formal PAI learning with lived practice. The decision to relocate some memorization activities to the boarding environment and use school hours for discussion and applied learning illustrates how curriculum can be reorganized around the actual ecology of the institution.

However, the same ecology creates constraints, including student fatigue and the need to coordinate formal school and boarding programs. The implication for curriculum governance is that a pesantren-based school should not simply copy a generic KOSP. Its operational curriculum should explicitly map the relationship between school-based instruction, boarding activities, religious habituation, assessment, and student well-being.

The 85% achievement claim illustrates the difference between perception and evidence

International readers will likely challenge the reported 85% achievement figure because the manuscript does not explain how the percentage was derived. The figure may be useful as evidence of institutional perception, but it should not be treated as an outcome measure. This distinction strengthens rather than weakens the study because evaluation research depends on transparent evidence claims. (Stufflebeam & Zhang, 2017)

A stronger evaluation system would define measurable indicators for each intended outcome, specify data sources, and report how evidence was aggregated. For example, character-related outcomes could be assessed through structured observation or validated instruments; learning attainment could be linked to explicit criteria; project outcomes could be evaluated with analytic rubrics; and stakeholder perceptions could be collected through transparent satisfaction measures. The present study identifies the absence of such a system as an important improvement need.

From administrative supervision to instructional improvement

The school's shift toward interactive academic supervision and peer discussion represents one of the most promising improvement mechanisms in the case. This approach is compatible with the improvement orientation of CIPP: evaluation should generate actionable information and support decisions rather than merely certify compliance. (Stufflebeam & Zhang, 2017)

Peer coaching could be organized around specific evidence problems identified in the case: constructing diagnostic assessments, converting diagnostic findings into differentiated modules, developing analytic rubrics, designing structured project reflection, and interpreting assessment evidence. Such coaching would be more likely to produce change than generic training because it targets the precise points at which implementation breaks down. (Kahmann et al., 2022; Kraft et al., 2018)

Implications for Practice and Policy

The findings have three practical implications. First, school leaders should treat diagnostic assessment as an instructional infrastructure rather than an additional administrative requirement. A concise baseline should identify what students already know, where learning difficulties occur, and how those findings will alter grouping, scaffolding, content, process, or product. This is consistent with the logic of differentiated instruction and assessment for learning (Kahmann et al., 2022; Nazilah, 2024; Roqobih et al., 2024).

Second, teacher development should move from one-off curriculum socialization toward sustained, practice-embedded coaching. Evidence from teacher-coaching research indicates that coaching can improve instructional practice, while professional development focused specifically on differentiated instruction is more useful when it is connected to teachers’ subject-specific work (Kahmann et al., 2022; Kraft et al., 2018). In this case, coaching should focus on the concrete implementation bottlenecks identified by the evaluation: diagnostic tools, differentiated modules, formative assessment, analytic rubrics, project reflection, and interpretation of student evidence.

Third, curriculum governance in pesantren-based schools should explicitly connect the school operational curriculum with the boarding-school ecology. Rather than treating the pesantren as an external setting, the KOSP should identify how classroom learning, religious routines, Qur’anic programs, student well-being, assessment, scheduling, and responsibility-sharing interact. This would allow institutional strengths to become a deliberate implementation resource while making potential constraints, such as student fatigue and fragmented responsibilities, visible within the quality-assurance system (Khair et al., 2026; Yani et al., 2025).

PRIORITY	RECOMMENDED ACTION	RESPONSIBLE ACTORS	SUCCESS EVIDENCE
1. DIAGNOSTIC ASSESSMENT	Develop a common but flexible PAI diagnostic-assessment package and record results at the start of each learning cycle.	PAI teachers; curriculum team	Documented baseline and evidence of instructional adaptation
2. TEACHER CAPABILITY	Establish sustained peer coaching focused on modules, differentiation, assessment, and reflection rather than one-off socialization.	Principal; curriculum coordinator; senior teachers	Revised modules and observed improvement in differentiated practice

3. KOSP CONTEXTUALIZATION	Rewrite the KOSP to explicitly integrate school and pesantren goals, routines, learner needs, and quality-assurance responsibilities.	School leadership; curriculum team; pesantren management	Context-specific KOSP and operational SOPs
4. FORMATIVE ASSESSMENT	Develop simple evidence routines for feedback, reflection, and follow-up without increasing unnecessary administrative burden.	Teachers; supervisors	Regular formative records linked to instructional decisions
5. PROJECT QUALITY	Strengthen contextualization, partnership, reflection, and differentiated products in co-curricular/project activities.	Project facilitators; teachers	Structured project rubric, reflection evidence, and diverse products
6. SUSTAINABILITY	Create post-project action plans and collect student/parent feedback where appropriate.	School leadership; teachers	Documented follow-up and stakeholder feedback
7. SCHOOL-BOARDING INTEGRATION	Formalize coordination between school PAI and boarding Qur'anic/religious programs.	School and pesantren management	Integrated schedule, responsibilities, and monitoring

7. Limitations

This study has several limitations. First, it is a single-site qualitative case study; its findings are analytically transferable only to contexts with sufficiently similar

characteristics and should not be interpreted as statistically representative of pesantren-based schools in Indonesia. Second, the source manuscript does not report exact participant numbers and detailed participant characteristics; these should be added from the original field records. Third, some reported percentages, particularly the 80% implementation alignment and 85% achievement estimate, lack sufficiently transparent calculation procedures and are therefore interpreted as institutional estimates rather than independently verified quantitative indicators. Fourth, ethical approval and informed-consent procedures are not documented in the source manuscript and must be supplied before submission. Finally, the study evaluates implementation at one point in time and therefore provides limited evidence about long-term sustainability.

CONCLUSION

The CIPP evaluation shows that implementation of the Merdeka Curriculum in PAI at SMP As-Shiddiqiyah Lempuing Jaya is characterized by substantial contextual and infrastructural strengths alongside persistent pedagogical and governance gaps. The pesantren environment provides a strong foundation for religious character formation and contextual learning, while digital resources create opportunities for more active and flexible instruction. Nevertheless, these advantages are not sufficient to ensure high-quality implementation.

The principal bottleneck lies in the alignment of teacher capability, diagnostic assessment, differentiated instruction, formative assessment, curriculum governance, and evidence-based improvement. The case demonstrates that a school can possess adequate technology and still experience incomplete curriculum transformation when teachers lack sustained support and when evidence generated through assessment is not systematically connected to instructional decisions. (Kahmann et al., 2022; Roqobih et al., 2024)

The most defensible improvement pathway is therefore developmental rather than compliance-oriented: contextualize the KOSP, strengthen diagnostic assessment, provide sustained peer coaching, improve formative and project assessment, formalize school–boarding coordination, and establish a stronger evidence system for monitoring outcomes and sustainability. The contribution of this case is consequently not a claim that the Merdeka Curriculum has simply 'succeeded' or 'failed,' but an empirically grounded explanation of where implementation becomes decoupled and where institutional intervention can reconnect policy intent with classroom practice.

REFERENCES

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Al Arsyadhi, N. L., Dewi, L., & Hernawan, A. H. (2024). Evaluation of teacher readiness in implementing Kurikulum Merdeka in elementary schools. *Inovasi Kurikulum*, 21(2), 1149–1160. <https://doi.org/10.17509/jik.v21i2.64113>

- Anisa, R. D., & Fatmariza. (2026). Teachers' adaptation to the implementation of the Merdeka curriculum: A case study. *Indonesian Research Journal in Education*, 10(1), 536–552. <https://doi.org/10.22437/irje.v10i1.29255>
- Aziz, M., Napitupulu, D. S., & Tanjung, S. A. (2024). Implementation of differentiated learning in the Merdeka Belajar curriculum for elementary schools. *Journal of Elementary Educational Research*, 4(2), 127–142.
- Batubara, R., Ananda, R., & Salim. (2025). Implementation of the Merdeka Curriculum for Islamic Religious Education subject. *Edukasi Islami: Jurnal Pendidikan Islam*, 14(3), 899–914. <https://doi.org/10.30868/ei.v14i03.8177>
- Hadi, S., & Wahyudin, D. (2024). School readiness analysis in implementing differentiated learning in Kurikulum Merdeka. *Inovasi Kurikulum*, 21(2). <https://doi.org/10.17509/jik.v21i2.68994>
- Heale, R., & Forbes, D. (2013). Understanding triangulation in research. *Evidence-Based Nursing*, 16(4), 98. <https://doi.org/10.1136/eb-2013-101494>
- Kahmann, R., Droop, M., & Lazonder, A. W. (2022). Meta-analysis of professional development programs in differentiated instruction. *International Journal of Educational Research*, 114, 102072. <https://doi.org/10.1016/j.ijer.2022.102072>
- Khair, U., Harmi, H., & Sari, D. P. (2026). Curriculum reform in practice: Implementation of Kurikulum Merdeka in Islamic Religious Education at Sekolah Penggerak. *AL-ISHLAH: Jurnal Pendidikan*, 18(1), 271–281. <https://doi.org/10.35445/alishlah.v18i1.7681>
- Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, 88(4), 547–588. <https://doi.org/10.3102/0034654318759268>
- Mahesa, K., Azzahra, L., Azzahra, F., Sari, Y., & Afriantoni. (2025). An analysis of teachers' challenges in implementing the Merdeka Curriculum policy in the teaching process at SDN 158 Palembang. *AL-ADABIYAH: Jurnal Pendidikan Agama Islam*, 6(3), 253–260. <https://doi.org/10.35719/adabiyah.v6i3.1155>
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE Publications.
- Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia. (2022). *Panduan pembelajaran dan asesmen pendidikan anak usia dini, pendidikan dasar, dan pendidikan menengah*. Badan Standar, Kurikulum, dan Asesmen Pendidikan.
- Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia. (2024). *Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Nomor 12 Tahun 2024 tentang kurikulum pada pendidikan anak usia dini, jenjang pendidikan dasar, dan jenjang pendidikan menengah*. <https://peraturan.bpk.go.id/Details/281847/permendikbudriset-no-12-tahun-2024>

- Ministry of Primary and Secondary Education of the Republic of Indonesia. (2025). *Peraturan Menteri Pendidikan Dasar dan Menengah Nomor 13 Tahun 2025 tentang perubahan atas Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Nomor 12 Tahun 2024 tentang kurikulum pada pendidikan anak usia dini, jenjang pendidikan dasar, dan jenjang pendidikan menengah*. <https://peraturan.bpk.go.id/Details/322506/permendikdasmen-no-13-tahun-2025>
- Nazilah, A. S. (2024). Diagnostic assessment in differentiated learning: Supporting learners' needs for improved learning achievement. *Research Journal on Teacher Professional Development*, 2(1), 87–92. <https://doi.org/10.21580/rjtpd.v2i1.20662>
- Putri, M. T., & Pranata, O. D. (2024). Merdeka Curriculum implementation at secondary schools: Science teachers' perspective. *IJECA (International Journal of Education and Curriculum Application)*, 7(3), 331–345. <https://doi.org/10.31764/ijeca.v7i3.26282>
- Rahmadewi, S., Habib, I., & Kabalmay, T. (2025). Diagnostic assessment as the foundation for adaptive and differentiated learning transition to the Merdeka Curriculum. *El-Tsaqafah: Jurnal Jurusan PBA*, 24(2). <https://doi.org/10.20414/tsaqafah.v24i2.14168>
- Roqobih, F. D., Sudibyo, E., Subekti, H., Purnomo, A. R., & Rosdiana, L. (2024). Pendampingan implementasi assessment for learning dalam mewujudkan pembelajaran berdiferensiasi pada Kurikulum Merdeka. *Jurnal ABDINUS: Jurnal Pengabdian Nusantara*, 8(3), 755–762. <https://doi.org/10.29407/ja.v8i3.23458>
- Stufflebeam, D. L., & Zhang, G. (2017). *The CIPP evaluation model: How to evaluate for improvement and accountability*. Guilford Press.
- Trisna, N., Arafat, Y., & Furkan, N. (2024). Liberating learning: Insights into the Merdeka Curriculum implementation. *Jurnal Ilmiah Ilmu Administrasi Publik*, 14(1). <https://doi.org/10.26858/jiap.v14i1.64713>
- Untirta, S. S., Hidayat, S., Priyadi, R. A., & Dewi, R. S. (2025). Analisis implementasi Kurikulum Merdeka pada abad 21 di sekolah dasar. *Jurnal Pendidikan Dasar*, 16(1). <https://doi.org/10.21009/jpd.v16i1.51423>
- Yani, Y. I., Nasution, W. N., & Halimah, S. (2025). Implementation of the Merdeka Curriculum in Islamic Religious Education and Character Education: A multisite study at SMA Negeri 1 and 2 Sibolga. *Jurnal Pendidikan Islam*, 14(1), 35–49. <https://doi.org/10.14421/jpi.2025.141.35-49>
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE Publications.
- Yunita, Y., Zainuri, A., Ibrahim, I., Zulfi, A., & Mulyadi, M. (2023). Implementasi Kurikulum Merdeka Belajar. *Jambura Journal of Educational Management*, 4(1), 16–25. <https://doi.org/10.37411/jjem.v4i1.2122>