



ANALYSIS OF QUALITY PLANNING GAPS IN THE IMPLEMENTATION OF MADRASAH REFORM IN THE WEST SULAWESI

**Kamus¹, Misykat Malik Ibrahim², St. Syamsudduha³,
Hasyim Haddade⁴, & Saadal Jannah⁵**

¹Sekolah Tinggi Agama Islam Negeri Majene, Indonesia

^{2,3,4}Universitas Islam Negeri Alauddin Makassar, Indonesia

⁵Sekolah Tinggi Ilmu Bahasa Arab (STIBA) Makassar, Indonesia

Correspondence Email: Kamusmustamin@stainmajene.ac.id

ABSTRACT

The Madrasah Reform program is a strategic policy initiative aimed at strengthening governance, improving learning quality, and nurturing a quality culture in madrasahs through evidence-based management. This study analyzes gaps in madrasah quality planning during the implementation of Madrasah Reform in West Sulawesi Province using the Discrepancy Evaluation Model (DEM). Data were gathered through interviews and document-based verification involving key stakeholders at provincial and district levels as well as selected madrasahs representing categories A, C, and non-accredited (TT). The results indicate that three out of four madrasahs exhibit design-stage discrepancies, mainly reflected in weak EDM-RKAM synchronization, limited operational SOPs, insufficient AKG-based CPD, and underdeveloped quality indicators. MAN 1 Majene shows minimal discrepancy, whereas the TT madrasahs demonstrate major to critical discrepancies. Overall, the findings suggest that Madrasah Reform outcomes depend heavily on design accuracy, managerial readiness, human resource capacity, and internal quality culture. Theoretically, this study extends the use of DEM as an evaluative lens for diagnosing early-stage policy implementation gaps in Islamic education governance.

Keywords: Madrasah reform; Discrepancy evaluation model; Quality planning; EDM; Madrasah management

1. INTRODUCTION

At the central level, Madrasah Reform (MR) is widely regarded as successful in promoting more accountable and data-driven governance. In West Sulawesi, however, implementation remains uneven due to incomplete quality-planning documents and limitations in human resources and infrastructure, indicating that central-level policy success does not automatically translate into local effectiveness. Initiated by Indonesia's Ministry of Religious Affairs in 2020, MR represents a comprehensive transformation agenda to enhance Islamic education quality in the digital era.

MR seeks to improve institutional governance and performance through four pillars: strengthening data-based planning, digitalizing budgeting via the e-RKAM application, enhancing teacher professionalism and competence, and reinforcing accountability and transparency in madrasah management. Operationally, madrasahs are required to develop an *Evaluasi Diri Madrasah* (EDM) as the basis for needs analysis, priority setting, and quality strategies, while budgeting through e-RKAM must be aligned with EDM recommendations. Teacher professionalism is strengthened through continuous professional development (CPD) integrated with competency assessment (AKG) and academic supervision, supported by SOPs to ensure quality assurance and accountability.

In practice, weak design readiness undermines MR implementation and its governance outcomes. This concern is reflected in accreditation dynamics: BAN-PDM's IA2024 orientation emphasizes planning and governance evidence—EDM, RKAM, academic SOPs, supervision instruments, and CPD documentation—yet several madrasahs in West Sulawesi have reportedly been downgraded or categorized as TT (*Tidak Terakreditasi*) due to missing or non-functional core documents. Assessors also note that internal evaluation systems are often inconsistent, with planning documents treated as administrative requirements rather than quality-management tools.

From an evaluation perspective, these conditions align with Provus' Discrepancy Evaluation Model, which posits that a mismatch between standards and actual performance at the design stage tends to generate subsequent failures across installation, process, and outcomes. Moreover, local challenges such as limited PNS teachers, high dependence on overloaded honorary teachers, constrained training opportunities, and low yayasan involvement weaken governance support for principals. These technical and structural constraints, combined with weak quality culture and fragile monitoring–evaluation—often intensified by rural geography and limited internet access—can stagnate MR implementation, consistent with governance and evaluation perspectives that stress the importance of early-stage capacity and readiness.

Novelty and contribution. While MR has been discussed largely in terms of program rollout and outcomes, this study explicitly shifts the analytical focus to the design stage as the earliest detectable point of failure. By applying the Discrepancy Evaluation Model to map gaps in EDM quality, EDM–RKAM alignment, SOP functionality, and assessment-driven CPD, the study offers a systematic diagnostic lens for identifying where planning standards break down in practice. This approach provides an empirical basis for differentiated, needs-based support interventions—particularly strengthening EDM and e-RKAM literacy, academic supervision, and yayasan governance—to improve the effectiveness of Madrasah Reform in resource-constrained private madrasahs.

2. METHODS

This study employs Provus's Discrepancy Evaluation Model (DEM), which evaluates program success by examining the gap between Standards (S) and Actual Performance (P). The analysis focuses on the Design Stage, assessing whether madrasah planning instruments meet the core standards required by the Madrasah Reform program—particularly the quality of EDM, the

consistency of EDM–RKAM alignment, the functionality of SOPs, assessment-driven CPD (AKG and academic supervision), and the presence of measurable quality indicators. Following Provus’s logic, discrepancies at the design stage are treated as early signals of likely weaknesses in later stages of implementation (installation and process).

The choice of a design-stage focus is theoretically grounded. In Stufflebeam’s CIPP model, Context and Input are positioned as prerequisites for effective Process and Product outcomes; similarly, theory-driven evaluation emphasizes the need to test whether program assumptions and structures are coherent before judging results. In the Madrasah Reform context, official technical guidelines (e.g., EDM–eRKAM) explicitly require budgeting decisions in RKAM/e-RKAM to follow EDM-based needs analysis and priorities, making EDM–RKAM misalignment a direct indicator of design discrepancy.

Empirically, this study adopts a qualitative evaluative approach combining document verification and semi-structured interviews with key stakeholders (madrasah education officials, supervisors, BAN-PDM assessors, and principals) across selected madrasahs representing accreditation categories A, C, and TT in West Sulawesi. Ethical procedures were observed: participation was voluntary, informed consent was obtained, and institutional identities were anonymized to ensure confidentiality.

Theorist / Model	Core Concept	Relevance to the Design Stage
Provus (DEM, 1971)	Evaluation compares Standards (S) and Performance (P).	Assesses gaps in madrasah design documents (EDM, RKAM, SOP).
Stufflebeam (CIPP, 2003)	Context–Input determines the quality of Process–Product.	Quality design determines readiness for Madrasah Reform implementation.
Chen (Theory-Driven Evaluation, 1990)	Evaluation must test design assumptions.	Invalid EDM = failed design assumptions.
Scriven (Goal-Free Evaluation, 1991)	Focuses on evidence, not program claims.	Tests the absence of key documents (EDM, SOP, CPD).
Patton (Utilization-Focused, 2008)	Evaluation must be useful for decision makers.	Discrepancy findings are used to formulate recommendations for Madrasah Reform.
JUKNIS EDM–eRKAM 2021	EDM is the basis for data-driven quality planning.	Serves as the standard for Design Stage evaluation.
PCU–DCU 2023	Document validity determines feasibility of implementation.	Madrasahs without EDM are categorized as having critical discrepancies.
IA2024 BAN-PDM	Evidence-based governance standards.	Incomplete documents = non-compliance with national standards.

Data for this study were obtained through in-depth interviews designed to capture a comprehensive understanding of Madrasah Reform implementation across different levels of bureaucracy and the madrasah ecosystem. Interviews with the Head of the Madrasah Education Division at the West Sulawesi Regional Office of the Ministry of Religious Affairs provided a macro-level overview of policy direction, the general condition of madrasah quality, and differences in readiness across districts. Meanwhile, interviews with the Heads of the Madrasah Education Sections (Kasi Penmad) in Majene and Polewali Mandar districts revealed more detailed operational dynamics, including how Madrasah Reform facilitation is conducted, administrative

challenges that arise in completing EDM–eRKAM, and human resource constraints at the madrasah level. Their perspectives are crucial because they occupy an intermediary position between policymakers and technical implementers, enabling them to portray implementation gaps from a structural standpoint.

To obtain empirically grounded, observation-based insights from the field, this study also involved madrasah supervisors and BAN-PDM assessors, two actors who hold formal mandates to assess madrasah performance and governance based on national instruments. Supervisors contributed data on learning culture, supervision practices, and madrasah leadership, while BAN-PDM assessors provided objective information on document conditions, infrastructure, teaching practices, and compliance with IA2024 indicators. Interviews with principals of madrasahs in categories A, C, and TT further enriched the analysis by providing internal contextual data on how quality planning is carried out, how EDM is compiled, how RKAM is implemented, and how managerial challenges are addressed. By integrating these cross-level interviews, the study gained depth of data and enabled robust source triangulation, an approach recommended in qualitative methods by Creswell (2014) and Miles, Huberman, and Saldaña (2018) to enhance the credibility and validity of findings.

To provide a clearer picture of the data sources used in this study, a concise mapping of the profiles of the informants involved is necessary. The diversity of their positions and roles—from structural officials, supervisors, and assessors to madrasah principals—allows the researchers to obtain a comprehensive perspective on Madrasah Reform implementation at various levels. This mapping also constitutes an important basis for data triangulation in order to ensure the depth and validity of the study’s findings. Accordingly, the following table presents the profiles of the informants who served as primary sources in the discrepancy analysis at the Design Stage.

No	Informant	Position / Role	Level of Involvement in Madrasah Reform	Data Contribution
1	Head of Penmad, West Sulawesi	Head of Madrasah Education Division	Provincial policy, regional monitoring	Policy direction for Madrasah Reform, regional constraints, quality disparities
2	Head of Penmad, Majene	District Madrasah Supervisor	Technical assistance, supervision, facilitation of Madrasah Reform	Realities of Madrasah Reform implementation and madrasah readiness
3	Head of Penmad, Polewali Mandar	District Madrasah Supervisor	Coordination of Madrasah Reform and supervision of private madrasahs	Documentation challenges, human resources, and foundation-level ecosystem issues
4	BAN-PDM Assessor	Academic and managerial supervisor	Supervision of teaching and governance	Supervision findings, quality culture, and learning climate
5	Asesor BAN-PDM	Accreditation evaluator	Visitasi dokumen dan observasi pembelajaran	Validitas dokumen, discrepancy, indikator IA2024
6	Principal of MAN 1 Majene	Principal of A-accredited madrasah	Madrasah Reform implementer with strong readiness	Model of good practice in planning and implementing Madrasah Reform

7	Principal of MAS Hajaratur Akbar	Principal of C-accredited madrasah	Madrasah implementer with moderate capacity	Reform implementation with moderate capacity	Design gaps and difficulties integrating EDM and RKAM
8	Principal of MAS Al-Muawanah	Principal of TT madrasah	Low capacity, limited documentation	limited	Structural weaknesses and limitations in planning

Observation in this study was conducted by examining, in situ, a range of official documents that serve as quality standards for madrasah administration, including the IA2024 accreditation documents and Form 4.3 completed by BAN-PDM assessors. These documents provide an objective picture of the factual conditions of the madrasahs, covering aspects of governance, teaching and learning, infrastructure, and fulfillment of national education standard indicators. Form 4.3, in particular, contains the assessors' in-depth evaluations of the 14 IA2024 items, making it a rich source of data for identifying discrepancies between established standards and the actual performance of madrasahs. Through document observation, this study revealed variations in madrasah readiness, ranging from A-accredited madrasahs with complete documentation to TT madrasahs that possess very few basic planning documents.

The documentation technique involved collecting and analyzing various core guidelines and regulations related to the Madrasah Reform program. These included the 2021 EDM–eRKAM Technical Guidelines, which underscore the importance of EDM as the basis for data-driven RKAM preparation; the 2023 PCU–DCU Technical Guidelines, which regulate the governance of Madrasah Reform implementation and its monitoring and evaluation mechanisms; the TIP–TIK Technical Guidelines, which lay out the framework for madrasah digitalization; and the 2020–2024 Madrasah Reform Guidelines, which articulate the national policy direction for madrasah reform. In addition, the 2024 Madrasah Assessment Guidelines were analyzed to determine how learning evaluation and learning outcomes are positioned as an integral part of quality planning. All of these documents served as primary references in determining the Standards (S), which were then compared with the Actual Performance (P) of the madrasahs using the DEM approach.

The integration of document observation and regulatory analysis enabled the researchers to construct a comprehensive picture of the extent to which madrasahs have fulfilled the design requirements of Madrasah Reform. Furthermore, the use of policy documents as data helps ensure that the evaluation is not merely descriptive, but firmly grounded in the official regulations in force. Thus, the techniques of observation and documentation play a crucial role in uncovering gaps in madrasah quality planning and in strengthening the validity of discrepancy analysis at the Design Stage.

To clarify the analytical mechanism in this study, the following table maps the types of data and their respective functions in the evaluation process using the Discrepancy Evaluation Model. The table shows how each type of data is used to assess the gaps between Madrasah Reform standards and the actual performance of madrasahs in the field.

Type of Data	Data Source	Form of Data	Function in DEM Analysis
Interviews	Head of Penmad, Head of Penmad Majene, Head of Penmad Polman, Supervisors, BAN-PDM Assessors, Principals (A–C–TT)	Narrative transcripts	– Explore informants’ understanding and experiences of Madrasah Reform implementation. – Identify actual performance to be compared with standards. – Validate discrepancy findings through the perspectives of implementers and ev
Observation of Accreditation Documents	BAN-PDM (IA2024), Form 4.3 Assessors	Scores, assessor notes, visit results	– Assess the alignment of madrasah quality documents with national instruments. – Identify structural weaknesses, governance issues, and teaching practices. – Provide an objective basis for a
Madrasah Reform Policy Documentation	EDM–eRKAM JUKNIS 2021, PCU–DCU JUKNIS 2023, TIP–TIK JUKNIS	Regulatory documents	– Establish the Standards (S) used for comparison with madrasah performance. – Determine mandatory indicators for the Design Stage of Madrasah Reform. – Serve as the benchmark for evaluating madrasah compliance with national policies.
Internal Academic Documentation of Madrasahs	EDM, RKAM, SOPs, Work Programs, KOM, Academic Calendar	Quality planning documents	– Identify the design readiness of madrasahs. – Assess the synchronization between EDM, RKAM, and SOPs. – Determine discrepancy categories (minimal–moderate–major–critical).
2024 Madrasah Assessment Guidelines	Directorate of KSKK Madrasah	Technical assessment guidelines	– Assess the use of assessment data as a basis for decision-making. – Analyze the alignment of learning practices with assessment standards.
Additional Field Data	Photographs of activities, administrative files, supervision notes	Supporting data	– Provide triangulation to strengthen the validity of findings. – Confirm the consistency between informants’ claims and empirical evidence.

This mapping illustrates that the study employs a combination of policy data, field data, and documentary data, allowing discrepancy analysis to be conducted comprehensively. Understanding the function of each type of data enables the researchers to assess quality planning gaps systematically within the DEM framework.

To ensure that the analytical process proceeds systematically and in accordance with the Discrepancy Evaluation Model (DEM) framework, this study employed an analytical sequence consisting of integrated stages, beginning with data reduction, followed by the identification of standards, comparison of performance, triangulation, and finally the drawing of conclusions regarding discrepancy themes. This sequence was designed so that each component of the Madrasah Reform design could be evaluated on the basis of empirical evidence, while also ensuring that the evaluation results rest on a robust methodological foundation.

Each stage of analysis—from the identification of standards, comparison of performance, and triangulation to thematic analysis—contributes to a deeper understanding of design gaps in the implementation of Madrasah Reform. Through this approach, the study not only assesses the

alignment of planning documents, but also interprets how these gaps emerge and what their implications are for the overall effectiveness of the program.

The methods section describes the methodology or approach you used to conduct your research. It outlines the specific steps, techniques, or tools you employed to collect data, perform experiments, or analyze information. This section should provide sufficient detail for others to understand and potentially replicate your study. It is important to include information about the study design, data collection procedures, any statistical analyses performed, and ethical considerations, if applicable. Your methods would amount to around 500-1000 words.

3. RESULTS AND DISCUSSION

Quality Planning Standards under the Madrasah Reform Program

Quality planning standards in the Madrasah Reform (MR) program refer to the minimum set of guidelines and indicators that madrasahs must meet to ensure systematic, measurable, and data-based governance, budgeting, and continuous improvement. These standards are anchored in key policy references, including the 2021 EDM–eRKAM Technical Guidelines, the 2023 PCU–DCU Guidelines, MR program guidance (2020–2024), and the 2024 BAN-PDM Accreditation Instrument (IA2024), which collectively define what constitutes acceptable “design readiness” in madrasah quality management.

In operational terms, MR design readiness is assessed through five core pillars. First, a valid and analytical EDM is required as the foundation for needs diagnosis and priority setting. Second, EDM–RKAM synchronization through the e-RKAM system ensures that budgeting follows EDM-based priorities and supports annual quality targets with measurable indicators. Third, the availability of complete managerial and academic SOPs is essential to standardize governance and reduce implementation deviations; IA2024 explicitly treats SOPs as compulsory governance evidence. Fourth, teacher and principal development must be implemented through CPD that is driven by AKG results and academic supervision, reflecting an evidence-based approach to professional improvement. Fifth, madrasahs are expected to formulate measurable quality indicators, aligned with SMART criteria, to enable monitoring and evaluation across governance, learning, and quality culture.

Together, these five pillars constitute the benchmark standards for the Design Stage within the Discrepancy Evaluation Model (DEM). The degree to which each pillar is fulfilled is compared with actual madrasah conditions to determine discrepancy levels.

Findings on Design Discrepancies

Design Stage analysis using DEM shows marked variation in design readiness across the four sample madrasahs. Overall, only one madrasah meets most MR design standards, while the remaining three demonstrate moderate to critical gaps. The most recurrent discrepancies are: (1) shallow or non-operational EDM, (2) weak EDM–RKAM alignment, (3) incomplete or non-functional SOPs, (4) CPD not based on AKG/supervision, and (5) weak or absent quality indicators. These patterns suggest that discrepancy levels are not merely “document gaps” but indicators of wider institutional capacity constraints.

1) MAN 1 Majene – Minimal Discrepancy

MAN 1 Majene demonstrates strong design readiness. Its EDM is data-based and analytical, priorities are logically formulated, and budgeting is aligned through e-RKAM. The madrasah maintains complete SOPs and implements CPD documented and linked to AKG and supervision. This alignment indicates minimal discrepancy and provides a solid base for MR implementation.

2) MAS Hajaratur Akbar – Moderate Discrepancy

MAS Hajaratur Akbar has EDM and RKAM, but the EDM analysis remains descriptive and priorities are not sufficiently diagnostic. RKAM is only partially linked to EDM, SOPs are incomplete or inconsistently applied, CPD activities exist but are not systematically based on AKG/supervision, and quality indicators are weak. This reflects moderate discrepancy—design components exist but lack coherence and functionality.

3) MAS Al-Muawanah – Major Discrepancy

At MAS Al-Muawanah, EDM exists but is non-operational and does not guide needs analysis or priority setting. RKAM is disconnected from EDM, SOPs are largely absent, and there is no evidence of assessment-driven CPD or quality indicators. This constitutes major discrepancy, indicating substantial misalignment likely to constrain MR implementation.

4) MA An-Nur Bakengkeng – Critical Discrepancy

MA An-Nur Bakengkeng shows the most severe gaps: EDM, RKAM, SOPs, CPD evidence, and quality indicators are absent. This critical discrepancy implies the madrasah does not meet the minimum design prerequisites to enter the reform process meaningfully.

Madrasah	EDM	RKAM	SOP	CPD	Quality Indicators	Discrepancy Category
MAN 1 Majene	✓ Valid & analytical	✓ Synchronized	✓ Complete	✓ AKG-based	✓ SMART	Minimal
MAS Hajaratur Akbar	✓ Exists, limited depth	✗ Poorly synchronized	✓ Partial	✗ Not AKG-based	✗ Weak	Moderate
MAS Al-Muawanah	✓ Exists, non-operational	✗ Not synchronized	✗ Absent	✗ Absent	✗ Absent	Major
MA An-Nur Bakengkeng	✗ Absent	✗ Absent	✗ Absent	✗ Absent	✗ Absent	Critical

From this summary it can be concluded that the readiness of madrasah quality planning design in West Sulawesi Province displays very sharp disparities. MAN 1 Majene stands out as the only madrasah that meets almost all Madrasah Reform standards, thereby falling into the minimal discrepancy category and possessing a solid foundation for subsequent implementation stages. In contrast, C-category madrasahs and particularly TT-status madrasahs exhibit fundamental misalignments in nearly all design components, ranging from the absence of EDM and the lack of

data-based RKAM to the absence of SOPs, CPD, and quality indicators. This pattern confirms that many of the obstacles to MR implementation do not originate at the implementation stage, but are rooted in structural weaknesses in the initial design. Consequently, the success of Madrasah Reform depends heavily on strengthening planning capacity and governance at the design stage before interventions at the installation or process stages can be effectively undertaken.

Analytical Discussion

The findings reinforce the central proposition of DEM: design-stage misalignment is an early predictor of program failure before implementation advances. In the MR context, “data-based planning” cannot be realized if the foundational design instruments—EDM, EDM–RKAM alignment, SOPs, CPD, and indicators—are incomplete or non-functional. The discrepancy pattern across cases suggests that design readiness reflects deeper institutional capacity (data literacy, managerial routines, and governance support), not merely compliance with documentation.

The clearest systemic fault line is EDM quality and EDM–RKAM coherence. Where EDM is descriptive or absent, budgeting tends to follow administrative habits rather than evidence-based priorities, weakening the logic of MR as a reform program. Similarly, the absence of SOPs and AKG-based CPD indicates limited internal quality control and weak professional improvement mechanisms, which IA2024 treats as core governance evidence. These constraints are more pronounced in TT-status madrasahs, where limited human resources, weak foundation support, and fragile supervision ecosystems restrict the ability to institutionalize quality routines.

Leadership and managerial capacity appear to differentiate discrepancy levels: MAN 1 Majene demonstrates the capability to maintain continuous documentation and quality routines, while TT madrasahs lack the minimal infrastructure and organizational support to sustain planning standards. This implies that MR implementation is highly sensitive to local readiness conditions and supporting ecosystems (supervision, foundation governance, and capacity-building access).

Policy-wise, these results highlight a national implementation gap: MR progress reported at the central level may mask uneven readiness across regions and especially among small private madrasahs. Therefore, national MR strategies need a stronger differentiated support architecture—where TT and low-capacity madrasahs receive structured design-stage assistance (EDM clinics, e-RKAM mentoring, SOP packages, and supervision strengthening), while MR-ready madrasahs are accelerated as models and peer-learning hubs. In short, strengthening design readiness should be positioned as a national priority mechanism to ensure MR equity and scalability across diverse local contexts.

Discrepancy Scores at the Design Stage

Madrasah	Discrepancy Score	Visual Level
MAN 1 Majene	1 (Minimal)	
MAS Hajaratur Akbar	2 (Moderate)	
MAS Al-Muawanah	3 (Major)	
MA An-Nur Bakengkeng	4 (Critical)	

Interpretation of the Design Stage Discrepancy Table

The table above displays the degree of design discrepancy in the four sample madrasahs, with scores ranging from 1 to 4. Higher scores indicate greater misalignment between Madrasah Reform standards and actual madrasah performance. A salient pattern emerges: the lower a madrasah's institutional capacity and the weaker its planning documents, the higher the discrepancy level.

MAN 1 Majene, with a discrepancy score of 1 (Minimal), is the only madrasah that demonstrates strong alignment between MR planning standards and factual conditions. This score is consistent with findings that MAN 1 has a valid EDM, synchronized RKAM, complete SOPs, AKG-based CPD, and SMART quality indicators. These conditions indicate a robust managerial foundation for effective MR implementation.

In contrast, MAS Hajaratur Akbar falls at level 2 (Moderate), indicating that although some design components are in place, their quality and consistency are not yet optimal. This madrasah still shows misalignments in EDM analysis, RKAM synchronization, and the measurability of quality indicators. A moderate discrepancy level suggests that the madrasah has potential for development but requires intensive facilitation to fully meet MR standards.

Meanwhile, MAS Al-Muawanah, with a score of 3 (Major), exhibits substantial gaps across most design components. Its non-operational EDM, absence of SOPs, and lack of assessment-based CPD indicate weak governance structures. This discrepancy level signals that MR implementation in this madrasah will be severely constrained without structural intervention and systematic guidance.

The most concerning position is held by MA An-Nur Bakengkeng, which earns a score of 4 (Critical), signifying pervasive misalignment across all design components. The absence of EDM, RKAM, SOPs, CPD, and quality indicators shows that this madrasah does not yet possess the basic foundation necessary to enter the reform process. Within the DEM framework, this is referred to as a critical design discrepancy, a situation in which an institution fails to meet the minimum requirements for implementing any program, including Madrasah Reform.

Overall, the table clarifies that design gaps not only vary across madrasahs but are also directly associated with governance capacity, quality culture, and documentary readiness. Madrasahs with low discrepancy scores have a higher probability of successful MR implementation, whereas those with high scores require deeper, sustained need-based interventions.

Theoretical and Practical Implications for Madrasah Reform Implementation

Theoretical implications: The study supports DEM's claim that design discrepancy generates downstream constraints in installation and process. It also aligns with evaluation perspectives emphasizing that early-stage context/input readiness is foundational for later outcomes.

Practical implications: Improvements should begin from design-stage capacity building: EDM literacy, EDM–RKAM integration, SOP institutionalization, assessment-driven CPD, and measurable indicators—especially for TT and low-capacity private madrasahs.

Limitations and Critical Interpretation of the Study

This study is limited to four purposively selected madrasahs across categories A, C, and TT, and thus does not represent the full population of West Sulawesi madrasahs. Document scarcity in TT madrasahs required greater reliance on interviews and assessor observations, although triangulation was used to reduce bias. Contextual factors (foundation support, geography, internet access, and staffing patterns) also shape design readiness but are not always fully captured in formal MR documents.

Policy Recommendations

Based on the Design Stage evaluation using the Discrepancy Evaluation Model (DEM), it was found that differences in the quality of quality-planning designs across madrasahs are influenced by institutional capacity, human resource readiness, managerial documentation, and the support of madrasah supervisors. Therefore, policy recommendations are needed that address not only technical aspects but also structural contexts and the capabilities of individual madrasahs. The following recommendations are formulated in accordance with the levels of need and the disparities in madrasah readiness in West Sulawesi Province.

1. **Strengthening Quality Planning Competence Based on EDM–eRKAM.** Provincial and district offices of the Ministry of Religious Affairs need to enhance technical assistance for preparing valid EDMs through intensive training, document preparation clinics, and ongoing supervision. Such assistance should enable principals and planning teams to transform education report cards, madrasah assessment results, and EMIS data into sharp needs analyses. The automatic integration of EDM into RKAM via e-RKAM must be enforced consistently so that every budget allocation is genuinely evidence-based rather than driven by previous administrative habits.
2. **Differentiated Interventions Based on Discrepancy Levels.** Reform cannot be implemented in a uniform manner. Madrasahs with minimal discrepancy require acceleration programs, whereas those with moderate discrepancy need strengthened supervision and document consolidation. Madrasahs with major and critical discrepancies require special intervention patterns, such as intensive coaching clinics, placement of dedicated madrasah operators, and technical guidance for preparing basic documents. This need-based intervention scheme is aligned with national education quality improvement programs.
3. **Strengthening the Role of Madrasah Supervisors and BAN-PDM Assessors.** Madrasah supervisors should be given additional mandates as quality planning facilitators, not merely as evaluators of teaching and learning. They can guide madrasahs in preparing SOPs, designing teacher CPD, and ensuring consistency between planning documents and day-to-day implementation. BAN-PDM assessors should be involved in local technical assistance activities, particularly for TT-status madrasahs, to ensure coherence between accreditation standards and MR guidance.
4. **Preparation and Standardization of Managerial and Academic SOPs.** Madrasahs require a minimum package of SOPs that can be used as a basic governance standard. Provincial offices may develop official SOP templates covering planning, supervision,

assessment, student services, and administration. Providing standard SOPs will reduce disparities among madrasahs and support low-capacity institutions.

5. **Revitalizing AKG- and Supervision-Based CPD.** CPD programs for teachers and principals must be directed by AKG results, instructional needs, and recommendations from academic supervision. Districts should establish an annual CPD needs mapping system so that each teacher receives competency development aligned with identified weaknesses, rather than participating only in generic training of unmeasured impact. Implementing data-based CPD will strengthen teaching quality and serve as an indicator of continuous quality improvement.
6. **Strengthening the Role of Madrasah Foundations.** Many discrepancies in private madrasahs stem from weak foundation roles in supporting governance and quality management. Foundation capacity-building should therefore become a policy priority. The Ministry of Religious Affairs can introduce capacity-building programs for foundations, including training in document management, funding, and internal supervision. Strong foundations constitute an essential pillar for MR success.
7. **Affirmative Policies for TT-Status Madrasahs.** Non-accredited (TT) madrasahs require the most intensive support, including a "MR rescue" program, provision of administrative operators, facilitation of EDM and RKAM preparation, and strict quarterly monitoring. Such affirmative interventions not only improve design readiness but also pave the way for TT madrasahs to achieve at least a B accreditation rating in the next cycle.

At this point, the authors conclude that these policy recommendations underscore that the success of Madrasah Reform can only be achieved if improvements begin at the Design Stage. Without a valid, accurate, and well-documented planning foundation, all subsequent MR stages will encounter obstacles. Therefore, madrasah reform efforts must be directed not only at implementation aspects but also at institutional capacity building, professional human resource development, and the comprehensive strengthening of supporting ecosystems.

4. CONCLUSION

This study shows that most madrasahs in West Sulawesi have not yet met the quality planning standards of the Madrasah Reform program. Only MAN 1 Majene has a planning design that is aligned with the EDM-eRKAM standards, SOPs, CPD, and quality indicators, and is therefore categorized as having minimal discrepancy. In contrast, C-category madrasahs and especially TT-status madrasahs exhibit serious discrepancies ranging from moderate to critical, as indicated by the absence of basic planning documents, weak synchronization between EDM and RKAM, and the non-functional implementation of SOPs and AKG-based CPD.

These findings reinforce Provus's theory that failure at the design stage will generate failure at the installation and implementation stages. They are also consistent with the CIPP model, which emphasizes that weaknesses in the context and input components will hinder quality attainment at the process and product levels. Accordingly, discrepancies in madrasahs in West Sulawesi are not merely administrative shortcomings, but reflect limited institutional capacity in adopting the principles of evidence-based management that lie at the core of Madrasah Reform.

Based on these findings, strengthening the managerial capacity of principals, enhancing data literacy, improving documentation systems, and establishing continuous monitoring represent urgent needs if Madrasah Reform is to be implemented effectively. Affirmative interventions for TT-status and C-category madrasahs should be designed through technical assistance for EDM–eRKAM preparation, the strengthening of academic supervision, and the systematic development of foundation governance. Without reinforcement at the design stage, the success of Madrasah Reform will remain partial and uneven across regions.

REFERENCES

- Badan Akreditasi Nasional Pendidikan Dasar dan Menengah. (2024a). Instrumen Akreditasi Madrasah (IA 2024). BAN-PDM RI.
- Badan Akreditasi Nasional Pendidikan Dasar dan Menengah. (2024b). Format Asesmen Lapangan 4.3. BAN-PDM RI.
- Chen, H.-T. (1990). *Theory-driven evaluations*. Sage.
- Creswell, J. W. (2014). *Research design: Qualitative, Quantitative, and Mixed Methods Approaches* (4th ed.). Sage.
- Direktorat Jenderal Pendidikan Islam. (2021). *Petunjuk teknis Evaluasi Diri Madrasah (EDM) dan e-RKAM*. Kementerian Agama Republik Indonesia.
- Direktorat Jenderal Pendidikan Islam. (2022). *Petunjuk teknis Bimtek EDM–eRKAM 2022*. Kementerian Agama Republik Indonesia.
- Direktorat Jenderal Pendidikan Islam. (2023). *Petunjuk teknis PCU–DCU 2023*. Kementerian Agama Republik Indonesia.
- Direktorat KSKK Madrasah. (2020). *Pedoman implementasi Madrasah Reform*. Kementerian Agama Republik Indonesia.
- Direktorat KSKK Madrasah. (2024). *Pedoman Asesmen Madrasah 2024*. Kementerian Agama Republik Indonesia.
- Fullan, M. (2007). *The new meaning of educational change* (4th ed.). Teachers College Press.
- Kementerian Agama Republik Indonesia. (2023). *Juknis TIK/TIP madrasah*. Kementerian Agama Republik Indonesia.
- Leithwood, K., Harris, A., & Hopkins, D. (2020). Seven Strong Claims about Successful School Leadership Revisited. *School Leadership & Management*, 40(1), 5–22.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2018). *Qualitative data analysis: A methods sourcebook* (4th ed.). Sage.
- OECD. (2016). *School leadership for learning: Insights from TALIS 2013*. OECD Publishing.
- Osborne, D., & Gaebler, T. (1992). *Reinventing government*. Addison-Wesley.
- Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Sage.
- Provus, M. (1971). *Discrepancy Evaluation: For Educational Program Improvement and Assessment*. McCutchan.
- Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Sage.
- UNESCO. (2017). *Data-driven school improvement: Guidelines for education systems*. UNESCO Publishing.