



Evaluating the performance of madrasah ibtidaiyah principals as a strategic mechanism for enhancing educational quality assurance in Indonesia

Moh Masnun¹, Patimah², Onwardono Rit Riyanto³ Naya Hidayani Haq⁴

^{1,2,3,4} Universitas Islam Negeri Siber Syekh Nurjati Cirebon, Indonesia

*Corresponding author: mohmasnun@uinssc.ac.id

Article Info

How to cite this article:

Masnun, M., Patimah, Riyanto, O. R., Haq, N. H. (2026). Evaluating the performance of madrasah ibtidaiyah principals as a strategic mechanism for enhancing educational quality assurance in Indonesia. *AL-TARBIYAH: Jurnal Pendidikan (The Educational Journal)*, 36(1), 75-89.

doi:

<https://doi.org/10.24235/w2sxx305>

5

Article history:

Received: April 16th, 2026

Accepted: July 14th, 2026

Published: July, 2026

Copyright © 2026

Al-Tarbiyah: Jurnal

Pendidikan (The Educational

Journal), under the Creative

Commons Attribution 4.0

International License.



Abstract

Madrasah principal leadership is a crucial factor in improving school effectiveness, learning quality, and teacher professionalism. Although international literature emphasizes the importance of school leadership evaluation, empirical studies on madrasah principal performance evaluation as a quality assurance mechanism in Islamic education are still very limited. This study aims to fill this gap by analyzing the practice of madrasah principal performance evaluation in three Islamic elementary schools (Madrasah Ibtidaiyah) in Cirebon Regency, Indonesia. A qualitative approach with a multiple case study design was used to understand the actual practice of leadership evaluation within the quality assurance system. Data were collected through in-depth interviews, observations, and document analysis, involving madrasah principals, supervisors, teachers, and madrasah administrators. Thematic analysis revealed that madrasah principal performance evaluation serves as a dual accountability mechanism (internal and external), integrated within the evaluation management cycle that includes planning, implementation, reflection, and follow-up. Evaluation also encourages strengthening teacher professionalism through collaborative development programs and instructional leadership. However, the effectiveness of implementation is still influenced by limited supervisor capacity and cultural resistance to evaluation. These findings expand educational accountability theory and the Quality Assurance Framework, while also providing implications for policy development, leadership practices, and strengthening quality assurance systems in madrasahs. This study confirms that madrasah principal performance evaluation is not merely an administrative mechanism, but a strategic instrument for continuous improvement in educational quality.

Keywords: Madrasah Principal Performance Evaluation; Educational Quality Assurance; Educational Leadership

INTRODUCTION

Improving educational quality remains a major priority in education reform worldwide, including in Islamic educational institutions such as madrasahs (Firmansyah et al., 2024). School leadership is widely recognized as a critical factor influencing school effectiveness, teacher performance, and student achievement (Thien & Hallinger, 2026). Principals are expected not only to manage schools administratively but also to foster positive school cultures, improve instructional quality, and implement educational policies effectively (Day et

al., 2023; van Nieuwerburgh et al., 2020). Consequently, principal performance evaluation has become an important strategy for monitoring leadership effectiveness and supporting continuous school improvement.

Growing accountability demands have further increased expectations for school leaders. A systematic review of Scopus-indexed studies published between 2001 and 2025 shows that accountability is one of the dominant themes in contemporary school leadership research (Mu'alimin et al., 2026; Purwanto & Yuliana, 2025). As a result, principal performance evaluation has evolved beyond an administrative requirement into a management tool that supports professional development and organizational improvement (Subandi et al., 2021; Yuliana et al., 2025).

These challenges are particularly relevant to Indonesian madrasahs, where principals must balance Islamic educational values with national education standards (Purwanto et al., 2024). Previous studies have shown that effective madrasah leadership contributes to teacher professionalism, a positive school climate, and institutional performance (Jaelani et al., 2019; Yuliana et al., 2025). However, most studies have focused on leadership styles or instructional leadership rather than examining principal performance evaluation as a strategic mechanism for educational quality assurance.

Research on educational quality management emphasizes that systematic leadership evaluation is essential for continuous organizational improvement (Habibi et al., 2025; *OECD*, 2025). Nevertheless, empirical evidence on how principal performance evaluation is implemented within Islamic elementary schools remains limited. Existing studies on madrasah reform in Southeast Asia primarily address curriculum development, educational modernization, and the integration of Islamic and national educational values, leaving leadership evaluation largely unexplored (Nkurunziza et al., 2025).

To address this gap, this study investigates the performance evaluation of Madrasah Ibtidaiyah principals as a strategic mechanism for strengthening educational quality assurance in Indonesia. Unlike previous studies that emphasize leadership characteristics or instructional practices, this research examines how performance evaluation supports quality assurance, teacher performance, and school improvement within Islamic elementary education. The findings are expected to contribute to the educational leadership literature while providing practical recommendations for strengthening quality assurance policies and leadership development in Indonesian madrasahs.

RESEARCH METHOD

Research Design

This study employed a qualitative approach using a multiple case study design. This design was selected to gain an in-depth understanding of how madrasah principal performance evaluation is implemented as part of educational quality assurance within real-life institutional settings. Multiple case studies enable researchers to examine a phenomenon across different organizational contexts through cross-case analysis, thereby identifying both common patterns and contextual differences (Yin, 2011). Accordingly, this study explored not only the implementation of principal performance evaluation but also its contribution to strengthening educational quality in madrasahs.

Research Locations and Participants

This study was conducted at three Islamic elementary schools (Madrasah Ibtidaiyah) in Cirebon Regency, West Java, Indonesia: MIN 1 Cirebon, MIN 9 Cirebon, and MI Tarbiyatul Wildan. These schools were selected purposively because they had implemented principal performance evaluation as part of their internal quality management practices.

A total of 12 participants were selected using purposive sampling based on their direct involvement in principal performance evaluation and school quality assurance. Participants consisted of three principals, three madrasah supervisors, and six teachers, each having at least five years of professional experience.

Table 1. Profile of Research Participants

Code	Role	School
P1	Principal	MIN 1 Cirebon
P2	Principal	MIN 9 Cirebon
P3	Principal	MI Tarbiyatul Wildan
S1	Madrasah Supervisor	MIN 1 Cirebon
S2	Madrasah Supervisor	MIN 9 Cirebon
S3	Madrasah Supervisor	MI Tarbiyatul Wildan
T1	Teacher	MIN 1 Cirebon
T2	Teacher	MIN 1 Cirebon
T3	Teacher	MIN 9 Cirebon
T4	Teacher	MIN 9 Cirebon
T5	Teacher	MI Tarbiyatul Wildan
T6	Teacher	MI Tarbiyatul Wildan

Data Collection Techniques

Data were collected between August and November 2025 using three complementary techniques: semi-structured interviews, non-participant observations, and document analysis. Semi-structured interviews explored participants' experiences and perceptions regarding principal performance evaluation and its role in educational quality assurance. Each interview lasted approximately 45-90 minutes, was audio-recorded with participants' consent, and subsequently transcribed verbatim.

Non-participant observations focused on principals' leadership practices, instructional supervision, professional interactions with teachers, and quality assurance activities. Documentary evidence included principal performance evaluation reports, academic supervision reports, madrasah work plans, self-evaluation reports, and internal quality assurance documents. Combining these three data sources enabled methodological triangulation and strengthened the credibility of the findings.

Data Analysis Techniques

Data were analyzed using the six-phase thematic analysis framework proposed by (Braun & Clarke, 2019). The analysis involved: (1) familiarization with the data through repeated reading of interview transcripts and field notes; (2) generation of initial codes; (3) searching for potential themes; (4) reviewing and refining themes across cases; (5) defining and naming themes; and (6) producing thematic interpretations that explained the implementation of principal performance evaluation across the three madrasahs. Cross-case comparison was subsequently conducted to identify similarities and differences in quality assurance practices among the participating schools.

To enhance the trustworthiness of the study, several validation strategies were employed. Methodological triangulation was achieved by comparing findings from interviews, observations, and documentary evidence. Member checking was conducted by inviting participants to review the researchers' interpretations of interview data. Peer debriefing with fellow qualitative researchers was undertaken to ensure analytical consistency and reduce subjective interpretation. These procedures enhanced the credibility, dependability, and confirmability of the findings.

FINDINGS & DISCUSSION

Findings

Thematic analysis of interview, observation, and document data from the three participating madrasahs yielded four themes: (1) performance evaluation as an institutional accountability practice, (2) mechanisms for implementing madrasah principal performance evaluation, (3) performance evaluation as a driver of teacher professionalism, and (4) structural and cultural challenges in implementation. Coding proceeded through open coding, thematic grouping, and constant comparison across the three cases until no new themes emerged from the data. Twelve participants contributed across the three sites: three principals (P1-P3), three madrasah supervisors (S1-S3), and six teachers (T1-T6), together with a committee chair (C1) who provided supplementary corroboration at MI Tarbiyatul Wildan. Findings are presented theme by theme, moving from the public madrasahs (MIN 1 and MIN 9 Cirebon) to the private madrasah (MI Tarbiyatul Wildan) in each case, to make cross-case variation visible.

Performance Evaluation as an Institutional Accountability Practice

Across the three participating madrasahs, principal performance evaluation was consistently described as serving two complementary accountability functions: internal accountability to the school community and external accountability to the Ministry of Religious Affairs.

At MIN 1 Cirebon, P1 emphasized that performance evaluation was viewed as a tool for continuous improvement rather than merely an administrative requirement:

“Performance evaluation is not something we fear. It is our tool to check whether the madrasah is on the right track. If there is any deviation, we know what needs to be corrected. Without evaluation, we have no mirror.” (P1, Interview, September 2025).

This statement was supported by observations of monthly evaluation meetings, during which the principal presented school achievements and areas requiring improvement to all teachers. Document analysis of the 2025 Madrasah Self-Evaluation (EDM) further showed that the results of the principal's performance evaluation were used as the basis for preparing the following year's Madrasah Work Plan (RKM).

A similar pattern was observed at MIN 9 Cirebon. S2 described the supervisor's role as ensuring that school programs were implemented according to established quality standards:

"As supervisors, we do not come to judge. We verify whether planned programs are being implemented and whether quality standards have been achieved. The evaluation results become part of our official report to the Ministry of Religious Affairs." (S2, Interview, October 2025).

This finding was reinforced by T3, who explained *"After the principal's performance evaluation, teacher development activities became more structured and had a clearer direction."* (T3, Interview, September 2025).

At MI Tarbiyatul Wildan, participants also described performance evaluation as an accountability mechanism, although documentation procedures were less formal than those found in the two public madrasahs. P3 stated:

"We are not as well documented as public madrasahs, but our commitment to accountability is there. We report our performance to both the committee and the supervisor, although not everything is formally documented." (P3, Interview, October 2025).

This statement was supported by C1, Chair of the School Committee, who noted:

"The principal always presents activity reports during committee meetings. Transparency is already in place, although the documentation system still needs improvement." (C1, Interview, October 2025).

Observation records and institutional documents from the three schools consistently showed that principal performance evaluation was integrated into routine school management practices, including planning, monitoring, reporting, and school improvement activities. The three madrasahs differed primarily in the level of procedural formalization and documentation.

Mechanisms for Implementing Madrasah Principal Performance Evaluation

The findings showed that principal performance evaluation in all three madrasahs followed four sequential stages: planning, assessment, evaluation of results, and follow-up. These stages were implemented using the Madrasah Principal Performance Assessment (PKKM) instrument issued by the Ministry of Religious Affairs of the Republic of Indonesia. Although the same evaluation framework was applied, differences were observed in the level of implementation and documentation across the three schools. A summary of these differences is presented in Table 1.

Planning

Planning practices varied across the participating schools.

At MIN 1 Cirebon, evaluation activities were formally planned at the beginning of the academic year. P1 explained:

“At the beginning of each year, the supervisor and I agree on an evaluation schedule consisting of two assessment cycles. I also prepare my performance portfolio from the beginning so that all supporting evidence is available during the assessment.” (P1, Interview, September 2025).

This account was confirmed by S1 and verified through the 2025 Supervision Schedule Agreement, which documented the evaluation schedule and assessment components.

At MIN 9 Cirebon, planning was carried out using a participatory approach. P2 explained that the evaluation focus was aligned with the school's annual development priorities, while T3 confirmed that senior teachers actively participated in planning meetings.

By contrast, planning at MI Tarbiyatul Wildan remained more reactive. P3 explained: *“Evaluation activities are usually planned only after the supervisor confirms the visit schedule.” (P3, Interview, October 2025).* This statement was supported by C1, and document analysis found no formal annual evaluation schedule for 2025.

Assessment

Assessment activities combined external evaluation conducted by madrasah supervisors using the PKKM instrument with internal self-evaluation through the Madrasah Self-Evaluation (EDM). At MIN 1 Cirebon, S1 conducted classroom observations, reviewed institutional documents, and interviewed teachers across four PKKM dimensions: personality and social competence, instructional leadership, school development, and resource management. T1 confirmed: *“The supervisor also asked us about the principal's daily leadership practices during the assessment process.” (T1, Interview, September 2025).* Research observations also documented classroom visits and document verification during the assessment process.

At MIN 9 Cirebon, assessment focused primarily on instructional leadership through direct observation of academic supervision activities. P2 explained: *“The supervisor evaluates my performance while observing how I supervise teachers during classroom instruction.” (P2, Interview, September 2025).* This procedure was confirmed by T4 and by classroom observations conducted during the study.

At MI Tarbiyatul Wildan, assessment relied mainly on document verification, with limited classroom observation. Analysis of the 2025 PKKM assessment form showed that only two of the six observation-based assessment components had been completed.

Evaluation of Results

The evaluation of results also differed among the three schools.

At MIN 1 Cirebon, evaluation results were discussed through structured feedback sessions involving the principal and supervisor. S1 described the process as follows: *“This is not merely about announcing scores; it is a professional dialogue aimed at improving leadership performance.” (S1, Interview, October 2025).*

Document analysis of the 2025 Madrasah Principal Performance Assessment (PKKM) Report showed that the report included performance scores, performance-gap analysis, and recommendations for improvement.

At MIN 9 Cirebon, P2 shared the evaluation results with the school's core management team. *"I discuss the evaluation results with the management team because improving school performance is a shared responsibility."* (P2, Interview, September 2025). T3 confirmed that this collaborative discussion encouraged broader participation in school improvement.

At MI Tarbiyatul Wildan, evaluation results were primarily communicated through final performance scores. P3 acknowledged: *"Structured discussions about areas that need improvement have not yet become a regular practice."* (P3, Interview, October 2025). Similarly, T5 reported that teachers rarely received detailed information regarding evaluation results or subsequent improvement plans.

Follow-up

Follow-up activities demonstrated noticeable differences among the three schools.

At MIN 1 Cirebon, evaluation recommendations were translated into a formal Follow-up Action Plan (RTL) integrated within the Madrasah Work Plan (RKM). Document analysis of the 2025 RKM indicated that most priority programs referred directly to recommendations contained in the previous PKKM Report. S1 explained: *"Evaluation should continue through concrete improvement programs rather than end with the assessment itself."* (S1, Interview, October 2025).

At MIN 9 Cirebon, follow-up activities were implemented through a biweekly professional learning community. P2 explained that this initiative addressed recommendations related to collaborative leadership, while T4 confirmed that the meetings were conducted regularly. At MI Tarbiyatul Wildan, follow-up activities remained limited. Document analysis of the 2025 PKKM Report identified documented implementation for only one of four recommended actions. C1 emphasized: *"A more transparent mechanism is needed to monitor the implementation of evaluation recommendations."* (C1, Interview, October 2025).

Overall, interview, observation, and document data showed variations among the three madrasahs in the implementation of each evaluation stage, particularly regarding planning procedures, assessment practices, documentation, stakeholder involvement, and follow-up activities. These similarities and differences are summarized in Table 1.

Table 1.

Cross-Case Comparison of Madrasah Principal Performance Evaluation Mechanisms			
Evaluation Stage	MIN 1 Cirebon	MIN 9 Cirebon	MI Tarbiyatul Wildan
Planning	Structured and scheduled, based on formal agreement documents	Needs-based, developed through a participatory approach	Reactive, not yet systematically scheduled
Assessment	Combines field observation and document analysis across multiple dimensions	Emphasizes authentic observation integrated with academic supervision	Predominantly document-based, with limited observation
Evaluation	of Structured feedback	Collegial discussion	One-directional, limited

Evaluation Stage	MIN 1 Cirebon	MIN 9 Cirebon	MI Tarbiyatul Wildan
Results	with detailed performance gap analysis	involving the core management team	to reporting final scores
Follow-up	Formalized and integrated into the Madrasah Work Plan (RKM)	Implemented through an active professional learning community	Largely informal; few recommendations documented as implemented

Performance Evaluation as a Driver of Teacher Professionalism

Participants across the three participating madrasahs consistently associated principal performance evaluation with teacher professional development. They explained that evaluation results were used to identify areas requiring improvement and to determine subsequent professional development activities.

At MIN 1 Cirebon, P1 explained that the findings of principal performance evaluation were translated into concrete teacher development programs:

"The results of the principal's performance evaluation do not stop with the assessment. They become the basis for determining teacher development programs, such as academic supervision training, lesson study, and other professional learning activities." (P1, Interview, September 2025).

This statement was supported by S1, who observed:

"Principals who undergo regular and comprehensive performance evaluations tend to be more proactive in developing teachers because they clearly understand the areas that require improvement." (S1, Interview, October 2025).

Research observations further showed that monthly in-house training and lesson study activities were routinely conducted following each evaluation cycle. Document analysis of the 2025 Academic Supervision Report and the Madrasah Work Plan indicated that these activities were planned as follow-up actions to recommendations contained in the Madrasah Principal Performance Assessment report.

Teachers also confirmed the implementation of these professional development activities. T1 explained:

"After the principal's performance evaluation, we established a teacher working group involving teachers from different grade levels to improve formative assessment practices." (T1, Interview, September 2025).

Similarly, T2 reported:

"Each evaluation cycle is usually followed by a new teacher development program. It is not just a recommendation to improve, but there are concrete activities that we carry out together." (T2, Interview, September 2025).

A comparable pattern was found at MIN 9 Cirebon. P2 explained that evaluation results were used to strengthen collaborative learning among teachers through a professional learning community. This initiative was confirmed by T4, who stated:

"Since the principal's evaluation, we have held regular professional learning community meetings every two weeks. Teachers now discuss classroom practices together instead of working individually." (T4, Interview, October 2025).

Observation data also showed that teachers regularly participated in collaborative discussions on instructional strategies and classroom practices during these meetings.

At MI Tarbiyatul Wildan, participants reported that teacher professional development had become more visible following the most recent principal evaluation, although activities remained less structured than in the two public madrasahs. T6 explained:

“After the latest principal evaluation, the principal became more active in inviting teachers to discuss learning problems and instructional improvement. There is a stronger enthusiasm to improve together.” (T6, Interview, October 2025).

Document analysis further showed differences in the level of institutional documentation across the three schools. At MIN 1 Cirebon and MIN 9 Cirebon, teacher professional development activities documented in the 2024–2025 Academic Supervision Reports, Madrasah Work Plans and PKKM Reports were explicitly linked to recommendations arising from principal performance evaluations. In contrast, at MI Tarbiyatul Wildan, although teacher development activities were documented, no formal documents explicitly connected these activities with recommendations contained in the principal performance evaluation reports.

Structural and Cultural Challenges in the Implementation of Principal Performance Evaluation

Participants identified two main challenges affecting the implementation of principal performance evaluation across the three madrasahs: structural constraints related to supervisory capacity and cultural resistance toward evaluation.

Structural Constraints

The most frequently reported structural challenge was the limited number of madrasah supervisors relative to the number of schools under their responsibility.

At MIN 9 Cirebon, P2 explained: *“Evaluation schedules are sometimes delayed because one supervisor is responsible for supervising more than twenty madrasahs.”* (P2, Interview, October 2025). This statement was supported by S1, who noted: *“Ideally, each madrasah should receive four supervisory visits each year. However, because of workload and the large number of schools, we are often only able to conduct one or two visits.”* (S1, Interview, October 2025).

Document analysis of the 2025 Supervisory Activity Report also showed differences between the planned and actual frequency of supervisory visits during the academic year.

Cultural Challenges

At MI Tarbiyatul Wildan, P3 explained that some school members still perceived principal performance evaluation as a threatening process rather than an opportunity for improvement: *“Some members of the madrasah still view the principal's performance evaluation as something threatening. They worry that poor results could negatively affect the school.”* (P3, Interview, October 2025). This perception was also described by C1, Chair of the School Committee: *“Some teachers are concerned that the evaluation may lead to unfavorable consequences for the madrasah, although the committee itself fully supports the evaluation process.”* (C1, Interview, October 2025).

Meanwhile, T6 expressed a different perspective: *“For me, the evaluation shows that the school receives attention and guidance. However, not everyone shares the same opinion.”* (T6, Interview, October 2025).

In contrast, participants at MIN 1 Cirebon and MIN 9 Cirebon generally described principal performance evaluation as a routine component of school quality management. Interviews, observations, and document analysis showed that evaluation activities at both schools were regularly integrated into planning, supervision, reporting, and follow-up processes.

Participants also described differences in institutional capacity between public and private madrasahs. S3 explained:

“Public madrasahs generally have stronger financial and human-resource support, making it easier to implement the PKKM system consistently. Private madrasahs often face greater resource limitations.” (S3, Interview, October 2025).

Document analysis similarly showed that the two public madrasahs maintained more comprehensive evaluation documentation than the private madrasah, particularly regarding annual planning, supervision records, and follow-up documentation.

Discussion

The results of this study indicate that the performance evaluation of madrasah principals in Madrasah Ibtidaiyah functions beyond an administrative procedure, but rather as a strategic mechanism in educational governance that emphasizes leadership accountability and improves the quality of school organizations. These findings demonstrate that the practice of madrasah principal performance evaluation is understood by informants as a means of organizational reflection that allows educational institutions to assess leadership effectiveness and identify areas for improvement. This phenomenon aligns with the Accountability Theory perspective, which views evaluation as an institutional mechanism to ensure that organizational leaders are accountable for their decisions, actions, and performance outcomes to various stakeholders (Day et al., 2023; Thien & Hallinger, 2026). In the educational context, accountability is not only related to regulatory compliance but also encompasses transparency, legitimacy, and improving the quality of educational institution management (Mu'alimin et al., 2026). Therefore, the madrasah principal performance evaluation found in this study reflects the accountability practices of educational organizations that involve internal and external dimensions, where madrasah principals are accountable to government authorities, the school community, and the community as stakeholders in education.

In addition to serving as an accountability mechanism, this study's findings indicate that madrasah principal performance evaluation occurs within a systematic management cycle that encompasses evaluation planning, implementation, reflection on evaluation results, and follow-up for improvement. This pattern demonstrates that performance evaluation does not stop at performance measurement alone but is part of a continuous organizational learning process. From an educational quality management perspective, this mechanism aligns with the principle of continuous improvement, which emphasizes the importance of continuous evaluation and improvement cycles in enhancing the quality of educational organizations (Gorgulu et al., 2026; Long et al., 2024).

These findings also align with the Quality Assurance Framework, which positions evaluation as an integral part of the educational quality assurance system through a process of evaluation, reflection, and strategic decision-making (*OECD*, 2025). In this context, school leadership evaluation serves not only as an administrative oversight instrument but also as a source of strategic information for improving the effectiveness of educational management and leadership (Thien & Hallinger, 2026; Vanari et al., 2025).

The integration of madrasah principal performance evaluation results into madrasah work plans, as found in the cases of MIN 1 Cirebon and MIN 9 Cirebon, demonstrates that evaluation can form the basis for evidence-based decision-making in educational organizational management. This approach allows evaluation data to be used as a source of organizational knowledge to design more systematic and targeted school quality improvement programs (Datnow & Park, 2014).

This study also found that the performance evaluation of madrasah principals has a functional relationship with the improvement of teacher professionalism. Madrasah principals who are systematically evaluated tend to be more active in developing teacher professional development programs, such as internal training, teacher learning communities, and academic supervision activities. This suggests that school leadership evaluation can encourage the strengthening of instructional leadership oriented towards improving the quality of learning. This finding is consistent with various international studies showing that school leadership is a key factor in developing teacher professional capacity and improving the quality of classroom learning (Thien & Hallinger, 2026; van Nieuwerburgh et al., 2020). From an Accountability Theory perspective, this phenomenon can be explained through the mechanism of cascading accountability, where demands for accountability on school organizational leaders will encourage the formation of professional accountability at the teacher level through more structured supervision and professional development (Coule, 2015).

Furthermore, this study shows that madrasah principal performance evaluation also plays a role in building a more reflective and collaborative school organizational culture. In some cases, evaluation results are used not only as administrative reports but also as material for collective discussions in reflection forums with teachers and the madrasah management team. This practice demonstrates that performance evaluation can function as a mechanism for collective learning within educational organizations. Educational leadership literature indicates that an effective accountability system serves not only as a performance control mechanism but also as a means of developing organizational capacity through continuous professional learning (Day et al., 2023; Eryilmaz & Strietholt, 2025). Therefore, madrasah principal performance evaluation can be understood as an instrument that not only strengthens organizational accountability but also encourages the formation of a culture of continuous improvement within the school environment (Darwis et al., 2025).

However, this study also identified a number of structural and cultural challenges that impact the effectiveness of madrasah principal performance evaluations. Structurally, the limited number and capacity of supervisors compared to the number of madrasahs limits the frequency and depth of evaluations. This situation aligns with findings that limited

supervisory capacity can hinder the effectiveness of educational accountability systems and the sustainability of the quality assurance cycle (OECD, 2025).

Furthermore, there are cultural challenges in the form of the perception that evaluations are judgmental rather than a means of professional learning. This perception highlights the importance of transforming organizational culture so that evaluation is understood as a reflective process that supports school leadership development (Hughes & Davis, 2025; Thien & Hallinger, 2026).

Cross-case findings also indicate a gap in institutional capacity between public and private madrasahs. Public madrasahs generally have stronger administrative support, documentation, and management systems, while private madrasahs face resource constraints. This confirms that the effectiveness of leadership evaluation systems is strongly influenced by the institutional capacity of educational organizations (Alzouebi et al., 2025; Habibi et al., 2025). Therefore, strengthening the evaluation system requires policy support that can increase institutional capacity, especially in private madrasah.

Theoretically, this study expands the study of educational leadership by integrating Accountability Theory and the Quality Assurance Framework to explain madrasah principal performance evaluation as a mechanism for accountability and organizational learning. From a policy perspective, these findings emphasize the need to strengthen madrasah leadership evaluation systems by increasing supervisor capacity and developing data-based evaluation mechanisms and professional reflection. Practically, performance evaluation can be utilized as a leadership reflection instrument to strengthen teacher professional development and the quality of learning. Overall, this study demonstrates that leadership evaluation integrated with organizational learning processes plays a crucial role in strengthening accountability and encouraging continuous improvement in madrasah quality.

CONCLUSION

This research demonstrates that madrasah principal performance evaluation is a strategic mechanism for strengthening the education quality assurance system. Evaluation serves not only as an administrative accountability instrument but also as an organizational learning process that drives continuous improvement through a cycle of planning, implementation, reflection, and follow-up. Evaluation results form the basis for decision-making to improve instructional leadership, teacher professionalism, and madrasah quality. Evaluation effectiveness is influenced by structural and cultural factors. Limited supervisor capacity, resource disparities between public and private madrasahs, and the perception that evaluation is judgmental remain obstacles to implementation. Therefore, strengthening evaluation policies, increasing supervisor capacity, and developing a reflective evaluation culture are necessary to improve the effectiveness of the quality assurance system. Theoretically, this research enriches the study of educational leadership by demonstrating that madrasah principal performance evaluation is a mechanism that integrates accountability, organizational learning, and professional development in the context of Islamic education. Practically, these findings provide a foundation for developing a more effective, sustainable, and evidence-based evaluation system. This study has limitations due to its qualitative approach with case studies of three elementary madrasahs in Cirebon Regency. Therefore, the

findings are contextual and not intended to be statistically generalized. Future research is recommended to use a quantitative or mixed methods approach with a broader coverage of regions and madrasa levels to examine the relationship between madrasa principal performance evaluation, teacher professionalism, and improvements in educational quality.

REFERENCES

- Alzouebi, K., Khurma, O. A., AlArabi, K., Al Mansoori, M., & Alamoush, I. (2025). School leadership and teacher stress: Unpacking the mediating role of job satisfaction and the employment status as a moderator. *Acta Psychologica*, 260, 105770. <https://doi.org/10.1016/j.actpsy.2025.105770>
- Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11(4), 589–597. <https://doi.org/10.1080/2159676X.2019.1628806>
- Coule, T. M. (2015). Nonprofit Governance and Accountability. *Nonprofit and Voluntary Sector Quarterly*, 44(1), 75–97. <https://doi.org/10.1177/0899764013503906>
- Darwis, A., Bafadal, I., Wiyono, B. B., Sultoni, & Malik, A. R. (2025). Designing an assessment tool for teacher leadership competencies in aspiring elementary school principals in Indonesia. *Sustainable Futures*, 9, 100803. <https://doi.org/10.1016/j.sfr.2025.100803>
- Datnow, A., & Park, V. (2014). *Data-driven leadership*. John Wiley & Sons.
- Day, C., Sun, J., & Grice, C. (2023). Research on successful school leadership. In *International Encyclopedia of Education (Fourth Edition)* (pp. 62–72). Elsevier. <https://doi.org/10.1016/B978-0-12-818630-5.05024-7>
- Eryilmaz, N., & Strietholt, R. (2025). School leadership matters: Evidence from TIMSS on teacher job satisfaction and students' academic outcomes. *International Journal of Educational Research Open*, 9, 100520. <https://doi.org/10.1016/j.ijedro.2025.100520>
- Firmansyah, F., Ferianto, F., & Pehlić, I. (2024). Challenges to the Implementation of Blended Learning in Islamic Religious Education Learning in Indonesia. *AL-TARBIYAH: Jurnal Pendidikan (The Educational Journal)*, 34(1), 51. <https://doi.org/10.24235/ath.v34i1.18270>
- Gorgulu, D., Comert, M., & Fayda-Kinik, F. S. (2026). Organizational commitment and ostracism in schools: The mediating role of chaotic leadership in crises. *Acta Psychologica*, 264, 106395. <https://doi.org/10.1016/j.actpsy.2026.106395>
- Habibi, A., Attar, R. W., Saudagar, F., Tomczyk, Ł., Hendra, R., & Alhazmi, A. H. (2025). Instructional leadership, self-efficacy, commitment and school climate on teachers' job satisfaction and performance: Asian context. *Asian Education and Development Studies*, 14(5), 1002–1021. <https://doi.org/10.1108/AEDS-11-2024-0262>
- Hughes, J., & Davis, K. (2025). Strategic leadership for sustainable innovation in primary education: a case study of forest school integration. *International Journal of Educational Management*, 39(7), 1756–1773. <https://doi.org/10.1108/IJEM-05-2025-0369>
- Jaelani, A., Patimah, D., Sanusi, U., & Arifuddin, A. (2019). The Implementation of Principal Managerial Competence for the Primary School Student Discipline. *Universal Journal of Educational Research*, 7(8), 1832–1838. <https://doi.org/10.13189/ujer.2019.070823>

- Long, Y., Xia, Y., Cao, F., & Long, Y. (2024). The influence of authentic leadership on teacher performance in Guiyang public primary schools. *Studies in Educational Evaluation*, 83, 101400. <https://doi.org/10.1016/j.stueduc.2024.101400>
- Mariani, A., Wanto, D., & Yanto, M. (2022). The Principal's Role in Improving Teacher Professionalism at Madrasah Ibtidaiyah Negeri. *Munaddhomah: Jurnal Manajemen Pendidikan Islam*, 3(1), 84–94. <https://doi.org/10.31538/munaddhomah.v3i1.204>
- Mu'alimin, M., Muhith, A. M., & Hepni, H. (2026). School principals' and teachers' perspectives on school quality management in Islamic schools: insights from Indonesia. *International Journal of Educational Management*, 1–15. <https://doi.org/10.1108/IJEM-05-2025-0387>
- Nkurunziza, G., Ntayi, J. M., Kyambade, M., Ssemusu, L., Nasande, M., Obedgiu, V., & Luganda, F. I. (2025). Institutional leadership and school board role performance in developing world context: moderating role of collective action. *International Journal of Educational Management*, 39(7), 1729–1755. <https://doi.org/10.1108/IJEM-11-2024-0721>
- OECD Skills Outlook 2025. (2025). https://www.oecd.org/en/publications/oecd-skills-outlook-2025_26163cd3-en.html
- Purwanto, N. A., & Yuliana, L. (2025). Data-Driven Evaluation of Leadership Competency Development for Elementary School Principals: A Systems Engineering Approach. *ASEAN Journal of Science and Engineering*, 5(2), 447–464. <https://doi.org/10.17509/ajse.v5i2.88663>
- Purwanto, N. A., Yuliana, L., Surya, P., Suharyadi, A., Mat Nashir, I., Abullah Kamal, S. S. L., & Darmadji, A. (2024). Government Direction on the Vocational School's Performance Evaluation in Indonesia and Malaysia: Some Notes from Islamic Education's Perspective. *Millah: Journal of Religious Studies*, 331–374. <https://doi.org/10.20885/millah.vol23.iss1.art11>
- Subandi, Thoyib, M., Fauzan, A., & Jaenullah. (2021). Pesantren-based Transformational Leadership: Strategies toward International Superior Madrasah in Indonesia. *Webology*, 18(SI05), 1023–1040. <https://doi.org/10.14704/WEB/V18SI05/WEB18279>
- Thien, L. M., & Hallinger, P. (2026). Linear and nonlinear dynamics between principal instructional leadership, teacher professional learning and teacher self-efficacy: a complexity theory approach. *Journal of Educational Administration*, 64(1), 31–46. <https://doi.org/10.1108/JEA-06-2025-0248>
- van Nieuwerburgh, C., Barr, M., Munro, C., Noon, H., & Arifin, D. (2020). Experiences of aspiring school principals receiving coaching as part of a leadership development programme. *International Journal of Mentoring and Coaching in Education*, 9(3), 291–306. <https://doi.org/10.1108/IJMCE-11-2019-0107>
- Vanari, K., Reinaru, D., & Eisenschmidt, E. (2025). Data use for school self-evaluation – The process, factors, and leadership practices in Estonian schools. *Studies in Educational Evaluation*, 85, 101465. <https://doi.org/10.1016/j.stueduc.2025.101465>
- Yin, R. K. (2011). *Qualitative research from start to finish*. THE GUILFORD PRESS.
- Yuliana, L., Prasajo, L. D., Akalili, A., & Firdaus, F. M. (2025). Development of a learning leadership model for vocational high school principals to foster student well-being within the framework of Merdeka Belajar. *Jurnal Cakrawala Pendidikan*, 44(1), 1–8. <https://doi.org/10.21831/cp.v44i1.78373>

