

# Performance-Based Teacher Evaluation in Indonesian Public and Private Madrasah Tsanawiyah: A Qualitative Comparative Case Study

Nur Afif<sup>1</sup>, Akhmad Shunhaji<sup>2</sup>, Khasnah Syaidah<sup>3</sup>

<sup>1</sup> Universitas PTIQ Jakarta, Jakarta, Indonesia; nurafif@ptiq.ac.id

<sup>2</sup> Universitas PTIQ Jakarta, Jakarta, Indonesia; akhmadshunhaji@ptiq.ac.id

<sup>3</sup> Universitas PTIQ Jakarta, Jakarta, Indonesia; saidahasna@ptiq.ac.id

---

## ARTICLE INFO

### *Keywords:*

teacher appraisal;  
performance evaluation;  
madrasah governance;  
instructional supervision;  
professional development

---

### *Article history:*

Received 2025-12-19

Revised 2026-01-11

Accepted 2026-03-30

---

## ABSTRACT

Teacher performance evaluation is central to educational quality assurance; however, within Islamic educational institutions it must also address moral and spiritual dimensions. This study examines how teacher performance evaluation is conceptualized and implemented in public and private Madrasah Tsanawiyah, and how institutional governance shapes its developmental impact. Adopting a qualitative comparative case study design, the research was conducted in two Islamic junior secondary schools in South Jakarta, one public and one private madrasah. Data were generated through in-depth interviews with principals, teachers, and community representatives, focus group discussions, classroom and supervision observations, and document analysis. Using cross-case thematic analysis and pattern matching, the study reveals contrasting evaluation logics: a regulation-driven, accountability-oriented model in public madrasahs and a flexible, value-integrative model in private madrasahs. While standardized systems enhance administrative consistency, value-based and dialogical evaluation practices demonstrate stronger alignment with professional development, instructional improvement, and teachers' role as moral exemplars. The findings suggest that evaluation systems in Islamic education are most effective when framed as formative and developmental rather than merely administrative. This study contributes to the literature by offering a governance-sensitive evaluation framework that integrates professional standards with Islamic ethical values, advancing a more holistic model of quality assurance in Islamic secondary education.

*This is an open access article under the [CC BY-NC-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/) license.*



## Corresponding Author:

Nur Afif

Universitas PTIQ Jakarta, Jakarta, Indonesia; nurafif@ptiq.ac.id

---

## 1. INTRODUCTION

Performance-based teacher evaluation is widely recognized as a central mechanism for improving educational quality, professional accountability, and instructional effectiveness (Parcerisa et al., 2022). In this study, performance-based teacher evaluation is operationally defined as a systematic process that assesses teachers' pedagogical practice, professional competence, social interaction, and personal conduct through observable performance indicators, supervision instruments, feedback, and documented follow-up actions, rather than merely administrative compliance (Woods et al., 2025). Beyond accountability, contemporary evaluation theory emphasizes its formative function as a tool for continuous professional

development and instructional improvement (Apelehin et al., 2025; Meylani, 2024; Samundeeswari et al., 2024).

Within Islamic educational institutions, teacher evaluation carries an expanded mandate. In addition to instructional effectiveness, evaluation is expected to reflect teachers' moral integrity, social responsibility, and commitment to Islamic values that underpin character education. Ideally, evaluation systems in madrasahs should simultaneously strengthen teaching quality and reinforce teachers' roles as moral exemplars (*uswah hasanah*). However, empirical studies consistently report that teacher evaluation in Islamic schools remains predominantly administrative, focusing on documentation, attendance, and reporting requirements, rather than reflective supervision and professional growth (Gaziel, 2008; Hanifatulqolbi et al., 2019). This creates a persistent gap between the normative goals of performance-based evaluation and its bureaucratic implementation.

The literature on teacher evaluation can be organized into five interrelated strands. First, studies distinguish between evaluation purposes, highlighting tensions between accountability-oriented, summative appraisal and development-oriented, formative evaluation (OECD, 2018; Ramdane & Souad, 2022). Second, research on evaluation instruments and supervision models emphasizes the role of classroom observation, performance rubrics, and instructional supervision in shaping teaching quality (Okafor et al., 2024; Sugiar et al., 2024). Third, feedback mechanisms emerge as a critical determinant of effectiveness; dialogical, goal-oriented feedback has been shown to significantly improve subsequent teacher performance, yet remains weakly implemented in many systems (Hunter & Springer, 2022; Shang et al., 2025). Fourth, in Islamic education, scholars note the limited integration of character and value-based assessment, as standardized quality assurance models tend to privilege technical indicators over moral and spiritual dimensions (Amzat, 2022; Freeh & Hussein, 2021). Fifth, comparative education research demonstrates that public-private governance differences, including levels of autonomy, accountability, and resource control—significantly shape evaluation practices and professional learning cultures (Ather et al., 2014; OECD, 2018).

Despite these advances, few studies empirically compare teacher performance evaluation practices between public and private Madrasah Tsanawiyah, particularly in urban Indonesian contexts such as Jakarta, using multi-source qualitative data. This gap is significant because public and private madrasahs operate under distinct governance logics. Public madrasahs are embedded in centralized regulatory frameworks that emphasize standardization and accountability, while private madrasahs enjoy greater institutional autonomy, enabling contextual adaptation but often facing weaker external oversight. Understanding how these governance differences shape evaluation instruments, feedback processes, and professional follow-up is essential for designing more effective and context-sensitive evaluation systems.

Against this background, this study addresses the following research questions:

1. How are performance-based teacher evaluations implemented in public and private Madrasah Tsanawiyah in South Jakarta?
2. What governance factors influence the choice of evaluation instruments, feedback mechanisms, and follow-up practices?
3. What effects do key stakeholders perceive evaluation practices to have on instructional quality and character education?

This study contributes theoretically by extending teacher evaluation literature through a governance-sensitive and value-integrative framework that bridges mainstream evaluation theory and Islamic educational principles. Practically, it offers evidence-based insights for policymakers, madrasah leaders, and educational foundations on how to balance accountability, professional development, and Islamic values in performance-based teacher evaluation systems.

## 2. METHODS

This study employed a qualitative comparative multiple case study design, focusing on teacher performance evaluation practices in two Islamic junior secondary schools (Madrasah Tsanawiyah). This design was selected to enable in-depth, context-sensitive comparison across contrasting institutional

governance models while preserving the naturalistic character of educational practices. Rather than seeking statistical generalization, the study aims for analytic transferability, allowing findings to inform similar Islamic educational contexts through theoretical and conceptual replication (Mujiansyah, 2024).

Two cases were considered sufficient because they represent maximum variation in governance structure within the same socio-cultural and regulatory environment: a public madrasah operating under centralized state regulation and a private madrasah managed by a religious foundation. This contrast provides a robust basis for identifying how governance logic, institutional autonomy, and accountability structures shape teacher performance evaluation practices.

The research was conducted from June to August 2025 at MTsN 2 Jakarta (public madrasah) and MTsS Darunnajah Ulujami (private madrasah), both located in South Jakarta, Indonesia. MTsN 2 Jakarta is a state-administered institution under the Ministry of Religious Affairs, serving approximately 900 students with a standardized Teacher Performance Appraisal (PKG) conducted annually. MTsS Darunnajah Ulujami is a privately managed madrasah linked to a pesantren-based foundation, enrolling approximately 750 students and implementing internally designed evaluation and supervision procedures conducted each semester.

Participants were selected using purposive sampling based on their direct involvement in teacher evaluation processes. The sample comprised 2 principals, 16 teachers, and 4 school committee/community representatives. Teacher participants were selected based on the following criteria: (1) a minimum of three years of teaching experience, (2) active participation in recent performance evaluations, and (3) representation of different subject areas (religious and general subjects). Principals were included due to their formal supervisory authority, while community representatives were selected for their advisory roles and oversight functions. Participants with less than one year of institutional involvement or no exposure to evaluation activities were excluded.

Data were collected using four complementary qualitative techniques to ensure depth and triangulation. Semi-structured interviews were conducted with all participants using an interview guide developed from relevant literature on teacher evaluation, supervision, and Islamic education management. Sample questions included: "How is teacher performance evaluated in this madrasah?" and "How is evaluation feedback used for professional development?" Interviews were conducted in Bahasa Indonesia, lasted 45–75 minutes, audio-recorded with consent, and transcribed verbatim.

Participant observation was conducted over an eight-week period, encompassing 12 observation sessions. Observations focused on classroom teaching, academic supervision meetings, PKG evaluation discussions, and post-evaluation feedback sessions. An observation protocol was used to record evaluative interactions, feedback practices, and teacher responses. Focus group discussions (FGDs) were conducted twice (one per madrasah), each involving 6–8 teachers. FGDs lasted approximately 90 minutes and were facilitated by the researcher using guided prompts to explore shared perceptions of evaluation effectiveness, fairness, and developmental impact. Document analysis included institutional evaluation guidelines, PKG instruments, supervision notes, teacher performance reports, and professional development plans. Documents were selected based on relevance to evaluation processes during the previous two academic years.

Data analysis followed an interactive thematic analysis model, involving data reduction, data display, and conclusion drawing conducted iteratively throughout the study (Miles et al., 2014). Coding was conducted inductively and deductively, informed by evaluation theory and emerging field data. A cross-case thematic analysis and pattern-matching strategy was employed to identify convergent and divergent evaluation practices across the two cases.

To ensure trustworthiness, the study employed methodological and source triangulation. Interview data were cross-checked with observational records and documentary evidence, as well as across participant groups. Member checking was conducted with selected participants to validate interpretive accuracy. These strategies enhanced the credibility, dependability, and confirmability of the findings.

### 3. FINDINGS AND DISCUSSION

#### 3.1 Findings

##### 3.1.1. Similarities and Differences in the Implementation of Performance-Based Teacher Evaluation

Data from interviews, observations, and documents indicate that both public (MTs Negeri) and private (MTs Swasta) madrasahs conceptualize teacher performance evaluation as a structured process aimed at improving instructional quality and teacher professionalism. In both settings, evaluation formally covers pedagogical, professional, social, and personal competencies. A principal at the public madrasah stated:

*“Penilaian kinerja guru kami dasarkan pada empat kompetensi utama sesuai regulasi, terutama bagaimana guru merencanakan dan melaksanakan pembelajaran di kelas”* (Principal, MTsN).

Similarly, teachers in the private madrasah described evaluation as a routine mechanism linked to instructional improvement:

*“Evaluasi di sini tidak hanya menilai administrasi, tapi bagaimana cara mengajar dan sikap guru sehari-hari”* (Teacher A, MTsS).

Despite these shared objectives, substantial differences were observed in implementation. In the public madrasah, evaluation relies on standardized instruments mandated by the Ministry of Religious Affairs, particularly the Teacher Performance Appraisal (PKG), conducted annually and integrated with formal academic supervision. Document analysis of PKG forms shows detailed scoring rubrics, fixed performance indicators, and standardized reporting formats. Classroom observations were followed by brief, formal feedback sessions focused on compliance with indicators.

In contrast, the private madrasah applies a more flexible and context-sensitive evaluation model shaped by the foundation’s vision and pesantren culture. Evaluation instruments are internally developed and include indicators of religious commitment, discipline, and contribution to institutional values. Observation data show that feedback is often delivered informally through mentoring conversations. As one senior teacher noted:

*“Biasanya setelah supervisi, kami langsung diskusi santai dengan kepala sekolah tentang kekuatan dan kekurangan mengajar”* (Teacher B, MTsS).

These findings demonstrate that differences are not merely technical but reflect distinct managerial and cultural orientations.

##### 3.1.2. Institutional Policy and Its Influence on Evaluation Practices

Institutional policy emerged as a decisive factor shaping evaluation practices. In MTs Negeri, evaluation procedures are tightly regulated by national policies, requiring uniform instruments, formal documentation, and scheduled supervision cycles. A principal explained:

*“Kami tidak bisa keluar dari juknis PKG karena itu sudah ditetapkan kementerian”* (Principal, MTsN).

In MTs Swasta, greater institutional autonomy allows the foundation to design evaluation policies aligned with its educational philosophy. Internal guidelines emphasize moral conduct, religious discipline, and collegial supervision. However, document review reveals variation in evaluation frequency and follow-up procedures across departments, indicating weaker standardization.

#### 3.2. Discussion

##### 3.2.1. Evaluation Purposes: Accountability versus Professional Development

The findings demonstrate that teacher performance evaluation in both public and private madrasahs operates within a dual evaluative logic, simultaneously serving accountability and professional development functions. However, these functions are not equally institutionalized. In

public madrasahs, evaluation practices are predominantly structured around standardized instruments and bureaucratic procedures, positioning accountability as the dominant evaluative orientation. This configuration reflects a compliance-driven logic commonly observed in public education systems, where uniformity and administrative control tend to supersede pedagogical reflexivity, as noted in comparative education studies (Ather et al., 2014). Consequently, evaluation outcomes function more as formal performance records than as catalysts for instructional transformation.

In contrast, private madrasahs exhibit a stronger developmental orientation, with evaluation embedded in mentoring relationships and reflective supervisory dialogue. Evidence from sustained mentoring interactions and post-evaluation lesson revisions suggests that formative, dialogical feedback fosters professional trust and enhances teachers' willingness to critically reassess their instructional practices. This finding aligns with scholarship emphasizing the effectiveness of formative evaluation as a mechanism for continuous professional learning rather than mere performance judgment (Kramer et al., 2025; Li & Gu, 2024). The divergence between the two contexts underscores how institutional governance shapes the pedagogical meaning and developmental potential of teacher evaluation.

### 3.2.2. Feedback Mechanisms and Their Effects on Teaching Quality

The study offers empirical reinforcement for scholarship highlighting the pivotal role of feedback quality within teacher evaluation systems (Al Maktoum & Al Kaabi, 2024; Hunter & Springer, 2022; Kraft & Christian, 2022). Cross-case observations reveal marked differences in the nature and pedagogical function of feedback across institutional contexts. In MTs Negeri, feedback practices tend to be concise, indicator-driven, and primarily oriented toward documentation and procedural compliance. While such feedback fulfills administrative requirements, it provides limited scaffolding for reflective practice or instructional refinement.

By contrast, feedback in MTs Swasta is predominantly dialogical and reflective, emphasizing shared meaning-making, goal clarification, and instructional problem-solving. Teachers who received sustained, goal-oriented feedback demonstrated tangible pedagogical improvements, including the revision of lesson plans, the adoption of more varied instructional strategies, and heightened student engagement during classroom interactions. These findings substantiate theoretical claims that feedback exerts its strongest impact when it is formative, specific, and future-oriented, functioning not merely as an evaluative judgment but as a mechanism for professional learning. The results further suggest that feedback quality mediates the extent to which evaluation systems translate into meaningful instructional change.

### 3.2.3. Evaluation, Islamic Values, and Character Education

A distinctive contribution of this study lies in its identification of divergent modes of integrating Islamic values within teacher evaluation frameworks. In MTs Swasta, evaluative criteria explicitly incorporate dimensions such as integrity (*amanah*), discipline (*iltizam*), and religious commitment, thereby reinforcing teachers' roles as *uswah hasanah* within the educational community. Teachers assessed along these moral-spiritual dimensions demonstrated consistent value modeling in classroom conduct, collegial interactions, and student engagement. This finding lends empirical support to theories of holistic Islamic education that conceptualize teacher professionalism as an integrated synthesis of intellectual competence and moral exemplarity (Damni, 2023; Surbakti et al., 2024).

By contrast, public madrasahs tend to acknowledge moral dimensions at a normative level while operationalizing evaluation primarily through technical and procedural indicators. Such an approach risks marginalizing character formation within evaluative practice, reducing ethical commitments to implicit expectations rather than measurable responsibilities. This pattern resonates with broader international critiques of technocratic evaluation regimes, which caution that excessive standardization may obscure context-sensitive and value-laden dimensions of teaching (OECD, 2018). The contrast

underscores how governance structures shape not only evaluation mechanisms but also the ethical horizons of teacher professionalism in Islamic education.

### 3.3. Strengthening Performance-Based Evaluation Systems

#### 3.3.1. Toward an Integrative Evaluation Model

The findings suggest that an effective performance-based evaluation system in Islamic education requires an integrative architecture that aligns national professional standards with value-based indicators and continuous development mechanisms. Evaluation should not be confined to periodic appraisal cycles but reconceptualized as an ongoing developmental process sustained through mentoring, structured feedback, and systematic professional follow-up. Such a shift reframes evaluation from a compliance-driven activity into a formative infrastructure for instructional and moral growth.

Leadership emerges as a decisive mediating factor in this process. Principals who function simultaneously as instructional leaders and moral exemplars are more capable of translating evaluative data into meaningful professional learning opportunities. Consistent with Murtado & Kurniawan (2025) findings, leadership practices grounded in trust, dialogical supervision, and ethical commitment strengthen teachers' engagement with evaluation outcomes. This underscores that the effectiveness of evaluation systems in Islamic education is contingent not only on technical design but also on leadership capacity to integrate pedagogical improvement with Islamic ethical values.

#### 3.3.2. Policy Implications for Continuous Improvement

From a policy perspective, teacher performance evaluation should be institutionally embedded within continuous improvement cycles and explicitly linked to strategic development planning. Positioning evaluation as an integral component of organizational learning ensures that appraisal outcomes inform systematic interventions such as mentoring, in-house training, and professional learning communities. Conceptualizing evaluation as *professional muhasabah*, a reflective and ethically grounded self-assessment, fosters intrinsic motivation, moral accountability, and sustained professional growth among teachers.

This policy orientation aligns with quality management theories that emphasize iterative cycles of planning, implementation, evaluation, and refinement as drivers of institutional effectiveness. By integrating technical standards with reflective and value-based practices, evaluation systems can function not merely as instruments of control but as catalysts for pedagogical and ethical transformation. Consequently, such an approach offers a strategic pathway for strengthening teacher professionalism while simultaneously enhancing the holistic quality of Islamic education, encompassing intellectual, moral, and spiritual development.

**Table 1.** Cross-Case Summary of Teacher Performance Evaluation Practices (Narrative Table)

Dimension	MTs Negeri (Public)	MTs Swasta (Private)
Evaluation instrument	National PKG, standardized rubrics	Internally developed, adaptive
Frequency	Annual, scheduled	Semester-based, flexible
Evaluators	Principal, official supervisors	Principal, senior teachers
Feedback mode	Formal, written, indicator-based	Dialogical, mentoring-based
Follow-up PD	Formal training, workshops	Mentoring, in-house learning
Value/character indicators	Implicit, normative	Explicit, integrated

Source: Comparative analysis of qualitative field data (2025)

Table 1 reveals clear cross-case differences in teacher performance evaluation practices between public and private madrasahs that reflect contrasting governance and pedagogical orientations. MTs Negeri applies nationally standardized PKG instruments through scheduled annual evaluations, positioning performance appraisal primarily as a mechanism of administrative accountability and bureaucratic quality assurance. Feedback in this context is predominantly formal, written, and

indicator-based, with follow-up professional development largely limited to structured training and workshops. In contrast, MTs Swasta employs internally developed and adaptive instruments implemented on a more flexible, semester-based schedule, involving principals and senior teachers as evaluators. Evaluation in private madrasahs emphasizes dialogical, mentoring-based feedback and in-house professional learning, framing appraisal as a continuous developmental process. The most salient distinction lies in the integration of value and character indicators: while public madrasahs address moral dimensions implicitly and normatively, private madrasahs explicitly embed Islamic values into evaluation criteria. Collectively, these differences suggest that the effectiveness of teacher evaluation systems depends not only on standardization and frequency but also on feedback quality, developmental follow-up, and the meaningful integration of value-based dimensions within Islamic educational contexts.

#### 4. CONCLUSION

This study examined how performance-based teacher evaluation is implemented in public and private Madrasah Tsanawiyah and how institutional governance shapes its implications for instructional quality and Islamic educational values. The findings show that public madrasahs employ a standardized, regulation-driven evaluation system emphasizing accountability, compliance, and formal quality assurance, whereas private madrasahs adopt a more flexible, value-oriented approach that integrates pedagogical performance with moral and spiritual dimensions rooted in institutional culture. While the public model strengthens comparability and administrative consistency, it tends to constrain formative reflection and professional learning; conversely, the private model fosters developmental supervision and character-based professionalism but faces challenges related to standardization and external accountability. Across both contexts, evaluation is most effective when positioned as a continuous developmental process supported by structured feedback, leadership-driven mentoring, and systematic follow-up professional development, highlighting the decisive role of principals as instructional and moral leaders. This study is limited by its qualitative design and focus on two madrasahs in a single urban setting, which restricts statistical generalization but allows for analytic transferability. Practically, the findings suggest the need for hybrid evaluation frameworks that balance national performance standards with Islamic value-based indicators. Future research should adopt longitudinal and mixed-method designs to examine the sustained effects of evaluation on teacher pedagogical transformation and student character development, as well as explore the potential of digital performance appraisal systems in strengthening transparency, reflective supervision, and continuous improvement in Islamic educational institutions.

#### REFERENCES

- Al Maktoum, S. B., & Al Kaabi, A. M. (2024). Exploring teachers' experiences within the teacher evaluation process: A qualitative multi-case study. *Cogent Education*, 11(1), 2287931.
- Amzat, I. H. (2022). *Supporting Modern Teaching in Islamic Schools: Pedagogical Best Practice for Teachers*. Routledge. <https://doi.org/10.4324/9781003193432>
- Apelehin, A. A., Imohiosen, C. E., Ajuluchukwu, P., Abutu, D. E., Udeh, C. A., & Iguma, D. R. (2025). Assessment and evaluation for social improvement in education: Strategies for equity and fairness. *International Journal of Social Science Exceptional Research*. 2025b, 4(1), 119–125.
- Ather, H. M., Hassaan, J., & Khan, T. (2014). Comparison of performance appraisal system in public and private schools. *Pakistan Journal of Commerce and Social Sciences*, 8(1), 272–278.
- Damni, A. (2023). Conceptualizing Professionalism: Perspectives of Science and Religion Teacher Candidates in Islamic Higher Education. *Symbiotic: Journal of Biological Education and Science*, 4(2), 90–103.
- Freeh, M. B., & Hussein, M. S. (2021). The Level of Teaching Performance of Islamic Education Teachers

- in the Intermediate Stage in the State of Iraq in the Light of Quality Assurance Standards. *Review of International Geographical Education Online*, 11(5), 3985–3997. <https://doi.org/10.48047/rigeo.11.05.280>
- Gaziel, H. (2008). Principals' performance assessment: Empirical evidence from an Israeli case study. *Educational Management Administration and Leadership*, 36(3), 337–351. <https://doi.org/10.1177/1741143208090593>
- Hanifatulqolbi, D., Ismail, I. E., Hammad, J., & Al-Hooti, M. H. (2019). Decision support system for considering the best teacher performance using MOORA method. *Journal of Physics: Conference Series*, 1193(1). <https://doi.org/10.1088/1742-6596/1193/1/012018>
- Hunter, S. B., & Springer, M. G. (2022). Critical Feedback Characteristics, Teacher Human Capital, and Early-Career Teacher Performance. *Educational Evaluation and Policy Analysis*, 44(3), 380–403. <https://doi.org/10.3102/01623737211062913>
- Kraft, M. A., & Christian, A. (2022). Can teacher evaluation systems produce high-quality feedback? An administrator training field experiment. *American Educational Research Journal*, 59(3), 500–537.
- Kramer, S. L., Posner, M. A., Browman, A. S., Lawrence, N. R., Roem, J., & Krier, K. (2025). The impacts of a standards-based grading system emphasizing formative assessment, feedback, and re-assessment: A mixed methods, cluster randomized control trial in ninth grade mathematics classrooms. *Journal of Research on Educational Effectiveness*, 18(1), 56–87.
- Li, J., & Gu, P. Y. (2024). Formative assessment for self-regulated learning: Evidence from a teacher continuing professional development programme. *System*, 125, 103414.
- Meylani, R. (2024). A comparative analysis of traditional and modern approaches to assessment and evaluation in education. *Batı Anadolu Eğitim Bilimleri Dergisi*, 15(1), 520–555.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative Data Analysis: A Methods Sourcebook* (3rd ed.). SAGE Publications.
- Mujiansyah, M. (2024). Interdisciplinary Reflections in Teaching for Educational Statistics Perspectives of Educators with Islamic Religious Education Backgrounds. *Glocal Praxis in Elementary Education*, 1(1), 32–47.
- Murtado, D., & Kurniawan, C. S. (2025). Enhancing Teacher Performance through Participatory Supervision in Indonesian Islamic Schools. *EDUKASI: Jurnal Penelitian Pendidikan Agama Dan Keagamaan*, 23(2), 306–321.
- OECD. (2018). *Preparing Our Youth for an Inclusive and Sustainable World: The OECD PISA Global Competence Framework*. Organisation for Economic Co-operation and Development.
- Okafor, R. N., Obona, E. E., Ngene, A. N., & Eji, E. E. (2024). School supervision and inspection: Enhancing educational quality, accountability, and student social welfare in Nigerian secondary schools. *Unizik Journal of Educational Research and Policy Studies*, 18(3).
- Parcerisa, L., Verger, A., & Browes, N. (2022). Teacher Autonomy in the Age of Performance-Based Accountability: A Review Based on Teaching Profession Regulatory Models (2017-2020). *Education Policy Analysis Archives*, 30(100).
- Ramdane, T., & Souad, M. (2022). Crafting a Strategy to Assess the Learning of Islamic Studies in Elementary Schools. In *Supporting Modern Teaching in Islamic Schools: Pedagogical Best Practice for Teachers* (pp. 245–258). <https://doi.org/10.4324/9781003193432-23>
- Samundeeswari, D., Angayarkanni, R., Raju, G., Rana, N., & Sharma, A. (2024). Teacher professional development: Effective strategies and evaluation methods. *Educational Administration: Theory and Practice*, 30(6), 1726–1733.
- Shang, R., Mallari, K., Au Yeong, W. Bin, Yasuhara, K., Tang, A., & Hsieh, G. (2025). Rethinking Teaching Evaluation Reports: Designing AI-transformed Student Feedback for Instructor Engagement. *Proceedings of the ACM on Human-Computer Interaction*, 9(7), 1–40.
- Sugiar, L., Sukirman, S., & Sanusi, S. (2024). Academic Supervision as a Strategy for Improving Teaching and Learning Quality. *International Journal of Educational Administration, Management, and Leadership*, 31–48.

- Surbakti, S. S. B., Harahap, R., & Hasanah, U. (2024). Future perspectives on the islamic personality model: Integrating spiritual, moral, intellectual, social, personal, and behavioral dimensions for holistic development. *Journal on Islamic Studies*, 1(1), 17–35.
- Woods, J. J., Romano, M. K., & Cripe, K. J. (2025). Evaluation of a Statewide Professional Development Model for Caregiver Coaching. *Topics in Early Childhood Special Education*, 45(3), 182–193.