JME: Journal of Management in Education
Vol. 02, No. 02 (2026)
Available online at: https://ejournal.narasikhatulistiwa.org/index.php/jme/index

MANAGEMENT OF KITAB KUNING LEARNING QUALITY EVALUATION IN DEVELOPING A QUALITY CULTURE IN TRADITIONAL ISLAMIC BOARDING SCHOOLS

Fingki Himami Hafshawati[1], Siti Widiyatus Sholehah[2], Nurul Fadila[3], Wahyu Ningtiyas[4], Luluk Khoiriyah[5]

[1,2,3,4,5] Universitas Nurul Jadid, Indonesia

himamihafshawati@gmail.com[1], sitiwidiatussholeha@gmail.com[2], nurulfadilaf3@gmail.com[3], wahyuningtiyas050@gmail.com[4], lulukkhoiriyah449@gmail.com[5]

Received: 26 September 2026 ; Revised: 9 October 2026; Publish: 14 November 2025

ABSTRAK

Manajemen evaluasi mutu pembelajaran kitab kuning merupakan salah satu komponen penting dalam peningkatan kualitas pendidikan di pondok pesantren salaf. Namun, hubungan antara pengelolaan evaluasi dan pembentukan budaya mutu kelembagaan masih memerlukan kajian lebih mendalam, khususnya dalam lingkungan pendidikan yang memiliki karakteristik pedagogis tradisional dan otoritas keilmuan yang kuat. Penelitian ini bertujuan menganalisis manajemen evaluasi pembelajaran kitab kuning dan kontribusi potensialnya terhadap pembentukan budaya mutu di pondok pesantren salaf. Kajian ini menggunakan kerangka analisis kualitatif dengan fokus pada perencanaan evaluasi, pelaksanaan penilaian, pemanfaatan hasil evaluasi, dan perbaikan pembelajaran secara berkelanjutan. Hasil analisis mengidentifikasi tiga dimensi yang saling berkaitan, yaitu keterpaduan proses manajemen evaluasi, hubungan pemanfaatan hasil penilaian dengan efektivitas pembelajaran, serta kondisi organisasional yang mendukung pembentukan budaya mutu. Temuan menunjukkan bahwa manajemen evaluasi perlu diarahkan tidak hanya untuk mengukur capaian belajar santri, tetapi juga untuk mendukung pengambilan keputusan pembelajaran berdasarkan bukti dan pelaksanaan perbaikan secara berkelanjutan. Pembentukan budaya mutu juga memerlukan refleksi kolektif, keterlibatan pengajar, komitmen pimpinan, dan konsistensi tindak lanjut evaluasi. Penelitian ini menyimpulkan bahwa pengelolaan evaluasi yang sistematis dapat menjadi landasan konseptual bagi penguatan budaya mutu pembelajaran kitab kuning tanpa mengabaikan tradisi pedagogis dan nilai keilmuan pesantren salaf. Kesimpulan tersebut masih bersifat konseptual dan memerlukan penguatan melalui bukti empiris yang terdokumentasi.

Kata kunci: Kitab Kuning; Manajemen Evaluasi Pembelajaran; Budaya Mutu; Pondok Pesantren Salaf; Peningkatan Mutu Berkelanjutan.

ABSTRACT

The management of kitab kuning *learning evaluation is an important component of educational quality improvement in traditional Islamic boarding schools (*pesantren salaf*). However, the relationship between evaluation management and institutional quality culture requires further examination, particularly within educational environments characterized by traditional pedagogical practices and*

established scholarly authority. This study aims to analyze the management of kitab kuning *learning evaluation and its potential contribution to developing a quality culture in traditional Islamic boarding schools. The study adopts a qualitative analytical framework focusing on evaluation planning, assessment implementation, utilization of evaluation findings, and continuous instructional improvement. The analysis identifies three interconnected dimensions: the integration of evaluation management processes, the relationship between assessment utilization and instructional effectiveness, and the organizational conditions supporting quality culture development. The findings indicate that evaluation management should extend beyond measuring students' learning achievement toward supporting evidence-informed instructional decisions and sustained improvement practices. Furthermore, quality culture development requires collective reflection, teacher participation, leadership commitment, and consistency in implementing instructional follow-up. The study concludes that systematic evaluation management offers a contextual foundation for integrating quality improvement into traditional pesantren education while preserving its distinctive pedagogical and scholarly traditions. These conclusions remain conceptual and require confirmation through documented empirical evidence.*

Keywords: *Kitab Kuning; Learning Evaluation Management; Quality Culture; Traditional Islamic Boarding Schools; Continuous Quality Improvement.*

INTRODUCTION

Educational institutions are increasingly expected to provide high-quality learning experiences, effective administrative services, and accountable institutional governance. Educational quality is no longer evaluated exclusively through students' academic achievement but also through the consistency of instructional processes, institutional leadership, stakeholder satisfaction, and the capacity to respond to changing educational needs. These expectations require schools and universities to establish systematic quality management mechanisms that connect institutional planning, service delivery, performance monitoring, and continuous improvement. Quality management systems have consequently gained considerable attention as organizational approaches for strengthening educational effectiveness and accountability. Research has demonstrated that the application of quality management models can support institutional self-assessment, improve organizational processes, and identify opportunities for corrective action (Girmanová et al., 2022). Similarly, the implementation of standardized quality management practices has been associated with improvements in educational governance, organizational coordination, and service consistency (Aquirera & Radiana, 2025).

Despite these developments, educational institutions continue to experience challenges in maintaining consistent service quality and translating quality assurance procedures into meaningful institutional improvements. Administrative fragmentation, unclear operational responsibilities, inconsistent documentation, and limited utilization of evaluation results may weaken the effectiveness of educational management. These challenges indicate that quality assurance cannot be achieved merely through the existence of institutional policies or formal accreditation documents. Instead, quality

management requires coordinated processes, effective leadership, employee involvement, evidence-based decision-making, and sustained corrective action. ISO 9001:2008 historically provided a standardized framework for addressing these concerns through process-oriented management, documentation, stakeholder satisfaction, and continual improvement. Although this edition has been withdrawn, its principles remain relevant for understanding the historical development of organizational quality management systems. Subsequent standards have retained the emphasis on process consistency and improvement while introducing broader expectations concerning organizational context, leadership, and risk management (International Organization for Standardization [ISO], 2008, 2015, 2026).

Empirical research within educational institutions provides evidence of how ISO-based quality management systems operate in organizational settings. A study involving 83 educational institutions and 2,132 teachers and management personnel in Spain examined the implementation of ISO 9001 and identified relatively strong perceived effects on management and quality systems, although the effects differed across organizational dimensions, including communication, learning processes, institutional climate, and satisfaction (Rodríguez-Mantilla et al., 2024). Research at a private university in Slovakia similarly documented the application of ISO 9001 and the Common Assessment Framework to support institutional evaluation, identify performance limitations, and formulate improvement opportunities (Girmanová et al., 2022). Within Indonesia, a literature-based investigation also highlighted the relevance of ISO 9001 for improving organizational arrangements, documentation practices, and stakeholder-oriented educational services (Aquirera & Radiana, 2025). These studies illustrate that the adoption of quality management systems may strengthen institutional procedures, but the effectiveness of implementation depends on organizational conditions and the extent to which quality requirements are translated into actual management practices.

Existing literature on ISO-based educational quality management can be classified into three major orientations. The first examines the application of ISO 9001 and related quality management models in educational institutions, particularly their contributions to organizational procedures, institutional assessment, and administrative effectiveness (Girmanová et al., 2022; Aquirera & Radiana, 2025). The second investigates institutional differences in the outcomes of implementing ISO-based quality management systems, including variations in leadership, management practices, organizational climate, and stakeholder satisfaction (Rodríguez-Mantilla et al., 2024). The third addresses the development of educational organization management standards through ISO 21001, which provides more specific guidance for educational services, learner needs, stakeholder engagement, and continuous improvement (ISO, 2025). These scholarly orientations indicate a gradual development from general organizational quality management toward educationally contextualized quality assurance. Historically, ISO 9001:2008 emphasized eight quality management principles, including customer focus, leadership, involvement of people, process approach, systems approach, continual improvement, factual decision-making, and mutually beneficial supplier relationships.

Later revisions reorganized these principles, reflecting the development of quality management concepts and their application across organizational settings (ISO, 2008, 2015).

The research gap addressed in this study concerns the need for a more integrated conceptual explanation of the nature, objectives, benefits, and historical principles of ISO 9001:2008 within educational management. Previous studies have examined the implementation of quality management models, their organizational effects, and the practical implications of ISO standards for educational institutions (Girmanová et al., 2022; Rodríguez-Mantilla et al., 2024; Aquirera & Radiana, 2025). However, these perspectives do not fully resolve the conceptual distinction between quality management as a procedural compliance system and quality management as an organizational mechanism for continuous educational improvement. Furthermore, the progression from ISO 9001:2008 to subsequent standards requires careful interpretation to distinguish enduring management principles from requirements that have been revised or replaced. The central issue is therefore not whether ISO 9001:2008 remains the current certification standard, but how its foundational principles can be critically understood within the development of educational quality management. This distinction is particularly important to prevent historical quality management concepts from being presented as unchanged contemporary certification requirements.

The novelty of this study lies in developing an integrated historical-conceptual interpretation of ISO 9001:2008 quality management principles and their relevance to educational institutions within the evolving landscape of international quality assurance standards. Rather than treating ISO 9001:2008 solely as an administrative certification instrument, the study examines its foundational principles as interconnected managerial elements involving leadership, stakeholder orientation, process control, personnel participation, evidence-based evaluation, and continuous improvement. The analysis also considers the conceptual continuity and differences between ISO 9001:2008 and later quality management standards, including ISO 9001:2015, ISO 9001:2026, and the education-specific ISO 21001:2025 (ISO, 2008, 2015, 2025, 2026). Accordingly, this study aims to analyze the nature, objectives, institutional benefits, and eight historical quality management principles of ISO 9001:2008, as well as their conceptual contribution to improving educational management. The study is expected to provide a theoretical foundation for understanding how standardized quality management principles can inform institutional governance, educational service consistency, and continuous improvement without conflating historical standards with currently applicable requirements.

METODE PENELITIAN

This study employed a qualitative approach to examine the management of kitab kuning learning evaluation in developing a quality culture within traditional Islamic

boarding schools (pesantren salaf). The research was designed to explore how educational evaluation is planned, implemented, and utilized to support instructional improvement while maintaining the distinctive pedagogical traditions of Islamic boarding schools. The primary focus included evaluation planning, assessment implementation, utilization of evaluation findings, and institutional practices supporting continuous quality improvement. A qualitative approach was considered appropriate because the study sought to understand educational management processes, institutional experiences, and the meanings attributed to evaluation practices rather than measure causal relationships between predetermined variables. Qualitative case study research in educational settings commonly uses interviews, observation, and documentary evidence to examine institutional processes within their organizational contexts (Khairiyah et al., 2026).

The data collection approach described in the original manuscript was based on interviews. Accordingly, the research focused on information concerning the management of kitab kuning learning evaluation, including the formulation of assessment objectives, implementation of learning assessments, identification of students' competencies, and follow-up actions undertaken by educational personnel. Relevant perspectives included those of pesantren administrators and teachers directly involved in the learning and evaluation process. The interview inquiry was organized around the relationship between traditional instructional practices, evaluation procedures, and institutional quality improvement. Where available, institutional documents such as assessment schedules, student learning records, evaluation instruments, and follow-up reports could provide supporting evidence for interpreting interview findings. The methodological value of combining interview accounts with institutional documents is demonstrated in qualitative educational case studies that examine organizational practices through multiple sources of evidence (Khairiyah et al., 2026). However, the original manuscript does not specify the research site, number of informants, interview dates, or available documentary records; these details must be established from the actual research records before the method can be reported as fully implemented.

Data analysis was structured around qualitative thematic analysis to identify patterns in the management of learning evaluation and their relationship with institutional quality culture. The analytical procedure involved organizing interview material, reading the data repeatedly, identifying relevant meaning units, developing initial codes, grouping related codes into themes, and interpreting the relationships among evaluation planning, assessment implementation, utilization of results, and continuous improvement practices. The analysis was guided by the requirement that themes represent meaningful patterns supported by the data rather than merely reproduce the headings of interview questions. Methodological coherence, transparency in theme development, and researcher reflexivity are essential considerations in qualitative thematic analysis (Braun & Clarke, 2023). The credibility of the findings would be strengthened by examining consistency across informant accounts, comparing interview information with available documents, and maintaining clear links between

empirical evidence and analytical interpretations. Any claims concerning the formation of a quality culture must be supported by evidence of recurring evaluation practices, collective engagement, and documented follow-up activities rather than inferred solely from the existence of formal assessment procedures.

RESULT

Management of Kitab Kuning Learning Evaluation

The analysis of the research manuscript identified four interconnected components of *kitab kuning* learning evaluation management: evaluation planning, assessment implementation, analysis of assessment results, and instructional follow-up. These components describe the management sequence through which learning objectives are connected to the assessment of students' competencies. The evaluation framework focuses on students' ability to read classical Arabic texts, understand their linguistic structures, explain their meanings, and demonstrate comprehension of the religious knowledge being studied.

The management sequence places evaluation planning as the initial component, followed by the implementation of assessments appropriate to traditional pesantren instruction. Evaluation results are subsequently examined to identify students' learning achievements and areas requiring additional support. The final component concerns the use of assessment findings to determine instructional responses, including material reinforcement, additional learning assistance, or adjustments to teaching approaches. These four components form an analytical framework for investigating how evaluation is managed in *pesantren salaf*; their actual implementation and continuity at the research site remain subject to verification through field evidence.

The available documentary evidence establishes that the original manuscript identifies evaluation planning, assessment implementation, result analysis, and follow-up as the principal aspects to be investigated. It also identifies interviews as the intended source of information concerning these practices. However, the manuscript does not provide interview transcripts, attributable statements from pesantren administrators or teachers, observation records, or institutional evaluation documents demonstrating how these processes were implemented. Therefore, the evidence currently supports the identification of evaluation-management components but does not establish their actual performance at a particular pesantren.

Table 1. Analytical Matrix of Kitab Kuning Learning Evaluation Management

EVALUATION COMPONENT	ASPECT IDENTIFIED IN THE MANUSCRIPT	EVIDENCE REQUIRED FOR EMPIRICAL CONFIRMATION
EVALUATION PLANNING	Determination of learning objectives, assessed competencies, and evaluation procedures	Interview accounts, assessment plans, learning objectives
ASSESSMENT	Assessment of students' reading,	Teacher interviews, assessment

IMPLEMENTATION	linguistic understanding, and textual comprehension	instruments, examination records
ANALYSIS OF RESULTS	Identification of achievement levels and learning difficulties	Assessment results, teacher analysis records
INSTRUCTIONAL FOLLOW-UP	Reinforcement, additional guidance, and adjustment of instructional practices	Follow-up records, documented instructional changes, interview evidence

Source: Authors' analytical organization of the original research manuscript.

The findings at this stage indicate that kitab kuning learning evaluation management can be examined through four sequential but interconnected components. The manuscript provides a basis for categorizing these components, while the relationship between assessment results and actual instructional improvements cannot yet be established. The empirical status of each component must therefore be determined from documented field data before conclusions about institutional implementation are drawn.

Effectiveness of Evaluation Management in Improving Learning Quality

The analysis identified the utilization of evaluation results as a central dimension in assessing the effectiveness of *kitab kuning* learning evaluation management. Within the research framework, effectiveness is associated with the extent to which evaluation activities contribute to identifying students' learning needs and informing subsequent instructional decisions. The manuscript distinguishes between the implementation of assessment procedures and the use of assessment findings to improve learning. This distinction positions evaluation as a management process that extends beyond the measurement of students' academic achievement.

The evaluation management framework connects instructional effectiveness with four operational dimensions: consistency in assessment implementation, analysis of students' learning outcomes, utilization of evaluation results, and monitoring of instructional follow-up. Assessment findings are expected to provide information about students' difficulties in reading, translating, and understanding classical Islamic texts. Such information may inform instructional adjustments, additional guidance, and reinforcement of learning materials. However, the actual effectiveness of these practices depends on whether evaluation results are translated into observable instructional changes rather than remaining as administrative records or final achievement scores.

The available documentary evidence shows that the original manuscript explicitly identifies the utilization of assessment results, instructional follow-up, and monitoring of learning improvements as important dimensions of evaluation management. It also proposes obtaining information from pesantren administrators and teachers concerning assessment implementation, decision-making, and changes in instructional practices. Nevertheless, the document does not include interview excerpts, assessment records, comparative student achievement data, or documented instructional changes. Consequently, the available evidence establishes the criteria for examining evaluation

effectiveness but does not demonstrate that learning quality has improved at the research site.

Table 2. Analytical Matrix of Evaluation Management and Learning Quality

EVALUATION DIMENSION	ANALYTICAL FOCUS	EVIDENCE REQUIRED TO ESTABLISH EFFECTIVENESS
ASSESSMENT CONSISTENCY	Regularity and suitability of learning assessments	Assessment schedules, teacher accounts, examination records
ANALYSIS OF LEARNING OUTCOMES	Identification of students' competencies and learning difficulties	Student assessment results, teacher evaluation notes
UTILIZATION OF EVALUATION RESULTS	Use of assessment information in instructional decision-making	Meeting records, teacher interviews, instructional plans
INSTRUCTIONAL FOLLOW-UP	Implementation of additional guidance or teaching adjustments	Remedial records, changes in teaching procedures
MONITORING OF IMPROVEMENT	Examination of learning developments following corrective actions	Follow-up evaluations, comparative learning records

Source: Authors' analytical organization of the original research manuscript.

These findings indicate that evaluation management effectiveness should be understood through the relationship between assessment implementation, evidence-based instructional decisions, and documented improvement efforts. The central analytical distinction is between the existence of evaluation activities and their demonstrated contribution to instructional quality. Although the manuscript provides a framework for examining this relationship, empirical conclusions regarding effectiveness require evidence of how evaluation findings were utilized and what changes followed their implementation.

The Role of Learning Evaluation in Developing a Quality Culture

The analysis identified the development of institutional quality culture as a broader dimension of *kitab kuning* learning evaluation management. In the original research framework, quality culture is associated with regular evaluation practices, openness to assessment findings, collective reflection, teacher involvement, and consistency in implementing instructional improvements. These dimensions describe how evaluation may extend beyond individual classroom assessment and become part of the institution's shared approach to educational quality. Quality culture is therefore examined not merely through formal procedures but through the continuity of quality-oriented practices among educational personnel.

The proposed relationship between evaluation management and quality culture involves a sequence of institutional activities. Evaluation findings provide information about instructional performance and students' learning needs. These findings may subsequently become the basis for discussions among teachers and pesantren administrators, decisions concerning instructional improvements, and monitoring of the actions undertaken. When these processes are consistently integrated into everyday

educational management, they may contribute to collective responsibility and continuous improvement. Within *pesantren salaf*, the development of such practices must also account for the authority of pesantren leaders, the pedagogical relationships between teachers and students, and the preservation of classical Islamic scholarly traditions.

The documentary evidence indicates that the manuscript identifies several elements relevant to quality culture, including evaluation regularity, reflective practices, teacher participation, institutional commitment, and continuous instructional improvement. It also emphasizes the importance of examining how pesantren leaders and teachers understand and implement evaluation within their educational environment. However, the document does not contain interview evidence demonstrating shared institutional commitments, records of collective evaluation meetings, or documented examples of sustained instructional improvement. Accordingly, the available evidence supports an analytical framework for investigating quality culture but does not establish that such a culture has already developed within a particular institution.

Table 3. Analytical Matrix of Learning Evaluation and Quality Culture Development

QUALITY CULTURE DIMENSION	INSTITUTIONAL PRACTICE TO BE EXAMINED	EVIDENCE REQUIRED FOR EMPIRICAL CONFIRMATION
EVALUATION REGULARITY	Consistent implementation of learning assessments	Evaluation schedules, assessment records
REFLECTIVE PRACTICES	Collective examination of instructional strengths and weaknesses	Teacher interviews, evaluation meeting records
TEACHER PARTICIPATION	Involvement of teachers in assessment and improvement decisions	Interview accounts, meeting attendance, decision records
LEADERSHIP COMMITMENT	Institutional support for evaluation and instructional improvement	Pesantren leadership interviews, institutional policies
CONTINUOUS IMPROVEMENT	Consistent implementation and monitoring of corrective actions	Follow-up reports, documented instructional adjustments
PRESERVATION OF EDUCATIONAL TRADITIONS	Alignment of quality improvement practices with pesantren pedagogy	Teacher and leader accounts, instructional documentation

Source: Authors' analytical organization of the original research manuscript.

The findings suggest that the contribution of learning evaluation to quality culture should be assessed through the institutional use of evaluation results, collective participation, and the continuity of improvement practices. Evaluation regularity alone is insufficient to demonstrate the existence of a quality culture unless accompanied by evidence of reflective decision-making and sustained corrective action. The relationship between evaluation management and quality culture therefore remains an analytical proposition requiring confirmation through institutional experiences and documented educational practices.

DISCUSSION

Management of Kitab Kuning Learning Evaluation

The analytical framework developed in the first finding indicates that the management of *kitab kuning* learning evaluation should be understood as an integrated educational process rather than a collection of separate assessment activities. Evaluation planning, assessment implementation, analysis of results, and instructional follow-up represent interconnected management functions that determine how educational institutions identify learning needs and respond to instructional problems. In traditional Islamic boarding schools, these functions acquire particular significance because educational quality cannot be adequately represented by examination results alone. Students' mastery of classical Islamic texts involves linguistic understanding, interpretative competence, conceptual comprehension, and the ability to explain religious knowledge within established scholarly traditions. Consequently, evaluation management must accommodate both the measurable dimensions of learning and the distinctive pedagogical characteristics of *pesantren salaf*.

This interpretation corresponds with research on quality management in *kitab kuning* instruction, which identifies quality planning, quality control, and quality improvement as interconnected institutional processes (Hidayah, 2023). In that study, quality planning involved determining graduate competency standards and organizing instructional arrangements, whereas quality control included assessing students' competencies and comparing their achievement with institutional expectations. Quality improvement subsequently involved reviewing instructional performance and strengthening educational management. These findings provide an important conceptual comparison because they demonstrate that evaluation is situated within a broader quality management system. However, the present analytical framework emphasizes a more specific relationship between assessment procedures and instructional follow-up. This distinction suggests that institutional quality management and learning evaluation management are closely related but should not be treated as identical concepts.

The relationship between planning and implementation is also supported by research examining the management of classical Islamic instruction at Pondok Pesantren Nurul Qodim Al-Mansuriyah. That investigation documented collaborative instructional planning, the implementation of *bandongan*, scheduled written and oral examinations, reading assessments, and daily teacher monitoring (Ja'far, 2025). These practices demonstrate how evaluation procedures may be integrated into the organization of traditional instruction rather than introduced as independent administrative requirements. Nevertheless, the existence of various assessment methods does not necessarily establish their managerial effectiveness. An institution may conduct regular examinations and maintain detailed records while experiencing difficulties in utilizing assessment findings to improve teaching. Therefore, the analytical significance of

evaluation management lies in the relationships among assessment objectives, evidence collection, interpretation of student performance, and subsequent educational decisions.

From the perspective of educational evaluation theory, the Context, Input, Process, and Product (CIPP) model offers a complementary framework for examining these relationships. Research on the *sorogan* method at Pondok Pesantren Alluqmaniyyah employed the CIPP framework to examine instructional needs, resources, implementation processes, and learning outcomes (Fuadi et al., 2025). A related study of *kitab kuning* learning at Pondok Pesantren Alfakhriyah Multidimensi identified differences between students' textual understanding and their capacity for critical reasoning and public communication, alongside limitations in facilities and technology utilization (Huseng et al., 2025). These studies reinforce the argument that evaluation should examine more than the final achievement of learners. However, the CIPP model primarily provides a structure for comprehensive program assessment, whereas evaluation management also concerns the organizational processes through which assessment information is translated into instructional action. Integrating these perspectives may therefore provide a more complete understanding of how traditional pesantren evaluate and manage learning quality.

A further consideration concerns the relationship between standardization and pedagogical flexibility. Standardized assessment criteria may improve consistency, particularly when different teachers evaluate similar competencies. However, excessive standardization may overlook the individual guidance, interpretative depth, and teacher–student interaction that characterize traditional Islamic learning. The educational value of *sorogan*, for example, partly arises from opportunities for individualized correction and direct academic engagement. Evaluation management should therefore establish clear learning expectations without imposing uniform procedures that disregard differences in instructional methods, student readiness, or scholarly traditions. This approach is consistent with the need to balance organizational quality assurance with pedagogical responsiveness in Islamic educational institutions (Hidayah, 2023; Fuadi et al., 2025).

The theoretical implication is that *kitab kuning* learning evaluation management can be conceptualized as a coordinated process connecting educational standards, assessment practices, interpretation of learning evidence, and instructional responses. Its effectiveness cannot be determined solely by the availability of assessment schedules, instruments, or formal evaluation policies. Instead, it requires examination of how these elements interact in actual institutional practice. For traditional pesantren, a meaningful evaluation system should preserve established scholarly traditions while enabling teachers and administrators to identify learning difficulties and organize appropriate improvements. Whether these relationships function consistently in the institution under investigation remains an empirical question requiring interview and documentary evidence.

Effectiveness of Evaluation Management in Improving Learning Quality

The second finding establishes an analytical distinction between the implementation of learning evaluation and its effectiveness in improving instructional quality. Evaluation may generate information about students' achievement, yet the production of assessment data does not automatically lead to educational improvement. Its managerial effectiveness depends on whether the information is interpreted accurately, communicated to relevant educational personnel, and used to determine appropriate instructional responses. This distinction is especially relevant to *kitab kuning* learning because students may demonstrate different levels of competence in reading Arabic texts, applying grammatical rules, explaining textual meanings, and interpreting religious concepts. A single assessment score may therefore provide an incomplete representation of their learning needs.

Research on *kitab kuning* learning management demonstrates the relevance of connecting assessment results with institutional decision-making. Quality control at Pondok Pesantren Al-Iman involved assessing student competencies against expected graduate standards and determining educational policies on the basis of assessment outcomes (Hidayah, 2023). Similarly, the integration of formal examinations and daily teacher monitoring in traditional Islamic instruction provides different sources of information about students' academic development (Ja'far, 2025). These studies suggest that effective evaluation requires more than periodic examinations. Nevertheless, they must be interpreted according to their respective research contexts. Their findings cannot, without additional evidence, establish that identical mechanisms operate effectively in every traditional Islamic boarding school.

The effectiveness of evaluation management can be further understood through the relationship between diagnostic assessment and instructional follow-up. Evaluation findings may identify difficulties resulting from students' limited vocabulary, insufficient understanding of Arabic grammatical structures, differences in prior learning experiences, or inappropriate instructional pacing. Once such difficulties are identified, educational personnel can consider additional guidance, repetition of learning materials, adjustments in teaching methods, or differentiated support. The central managerial challenge is determining whether these responses address the actual causes of learning difficulties. In this respect, evaluation must move beyond describing achievement levels toward supporting decisions about instructional improvement.

The CIPP-based evaluation of *kitab kuning* learning in Makassar provides a useful comparative perspective. The study reported relatively strong textual comprehension while identifying the need to strengthen critical thinking and students' ability to communicate religious knowledge (Huseng et al., 2025). This distinction demonstrates that apparent achievement in one learning dimension may coexist with limitations in another. Consequently, the effectiveness of evaluation cannot be assessed only through students' ability to reproduce textual explanations. It also requires attention to the specific competencies that the institution intends to develop. Similarly, the application of the CIPP framework to *sorogan* instruction highlights the importance of examining educational resources and instructional processes alongside learning outcomes (Fuadi et

al., 2025). Together, these studies support a multidimensional approach to evaluating the quality of classical Islamic instruction.

Continuous improvement theory provides an additional perspective through the Plan–Do–Check–Act (PDCA) cycle. Within this framework, educational institutions establish learning objectives and improvement priorities during the planning stage, implement instructional activities and assessment procedures, examine evidence of learning achievement, and introduce corrective actions based on the results. The critical element is the connection between the checking and acting stages. When assessment findings are collected but not translated into instructional changes, the improvement cycle remains incomplete. A study examining quality management implementation in an Indonesian Islamic university identified a comparable problem: strong documentation and formal quality assurance arrangements did not consistently produce actionable institutional improvements (Umam & Rahman, 2025). Although the institutional setting differs from traditional pesantren, this comparison illustrates why formal compliance should not be equated with educational effectiveness.

This perspective also requires a distinction between immediate corrective action and demonstrated learning improvement. Additional teaching sessions, revised assessment instruments, or changes in instructional schedules may represent reasonable managerial responses, but their implementation does not independently establish their effectiveness. Such conclusions require evidence showing whether students' learning difficulties were reduced or their competencies developed following the intervention. For example, a change in teaching method should ideally be accompanied by subsequent assessment or systematic teacher observation. Without this information, research can describe the existence of follow-up procedures but should avoid attributing improved learning outcomes to those procedures.

The theoretical contribution of this discussion is the distinction among three interconnected levels of evaluation effectiveness: assessment implementation, utilization of assessment findings, and demonstrated instructional improvement. The first concerns whether evaluation is conducted appropriately; the second concerns whether its results inform educational decisions; and the third concerns whether those decisions produce observable improvements in learning. These levels offer a framework for evaluating the performance of educational management without assuming that procedural consistency automatically produces better outcomes. In practical terms, pesantren administrators and teachers can use this distinction to examine whether their evaluation arrangements function primarily as assessment systems or as mechanisms for continuous instructional development. Establishing actual effectiveness, however, requires evidence linking evaluation results, institutional responses, and subsequent learning developments.

The Role of Learning Evaluation in Developing a Quality Culture

The third finding extends the discussion beyond instructional evaluation toward the organizational conditions required for developing a quality culture. In this context, quality culture refers to shared educational values, collective responsibility, reflective

professional practices, and sustained commitment to improving institutional performance. Evaluation contributes to such a culture when its findings become part of recurring discussions, collaborative decisions, and continuous improvement activities. The central distinction is therefore between evaluation as a formal requirement and evaluation as an established institutional practice. Regular assessment may demonstrate administrative consistency, but the formation of a quality culture requires evidence that educational personnel collectively interpret assessment information and use it to improve their work.

This interpretation is supported by research examining the relationship between quality management practices, organizational culture, and student performance. A mixed-methods investigation involving public secondary schools in Pakistan found that school culture mediated the relationship between quality management practices and student performance (Parveen et al., 2024). The significance of this finding lies in its indication that organizational practices may influence educational outcomes through the cultural conditions in which they are implemented. Nevertheless, its quantitative findings should not be transferred directly to *pesantren salaf*, where institutional authority, educational objectives, and pedagogical traditions may differ substantially. Rather, the study offers a conceptual basis for investigating whether evaluation management becomes more influential when supported by shared values, professional participation, and organizational commitment.

Within traditional Islamic boarding schools, the development of quality culture must be interpreted in relation to the distinctive position of pesantren leadership and the continuity of Islamic scholarly traditions. The authority of the *kiai* or pesantren leadership may provide an important foundation for institutional direction, while teachers contribute practical knowledge about students' learning needs and instructional difficulties. These roles suggest that evaluation management may support quality culture through the interaction between institutional leadership and teacher participation. However, leadership commitment alone does not demonstrate the existence of a participatory quality culture. Such a conclusion would require evidence that teachers have meaningful opportunities to discuss assessment findings, identify problems, propose improvements, and participate in reviewing the outcomes of institutional decisions.

The relationship between evaluation and organizational learning is particularly important. When educational personnel systematically examine assessment findings, they have opportunities to identify recurring learning difficulties and reconsider established instructional practices. Collective reflection can help distinguish problems arising from students' individual learning needs from those associated with instructional arrangements, resource limitations, or assessment procedures. Through repeated processes of reflection and adjustment, evaluation may become part of the institution's ordinary educational activities. Yet this process should not be assumed simply because evaluation meetings occur. The institutional significance of such meetings depends on

whether the discussion produces decisions, whether those decisions are implemented, and whether their consequences are subsequently examined.

Research on quality management in Islamic educational institutions reinforces this concern. The examination of ISO-based quality management at a state Islamic university identified a gap between the existence of formal management systems and the implementation of corrective actions (Umam & Rahman, 2025). This comparison highlights the possibility that educational institutions may possess evaluation procedures and documentation without establishing a sustainable improvement culture. For *pesantren salaf*, the implication is not that formal quality management systems must be adopted without modification, but that assessment activities should be linked to meaningful decisions and consistent follow-up. Institutional quality culture develops through repeated improvement-oriented practices rather than through the mere accumulation of evaluation documents.

At the same time, quality culture in traditional pesantren should not be conceptualized as the replacement of inherited pedagogical traditions with standardized managerial procedures. Classical Islamic learning is sustained through particular relationships of knowledge transmission, teacher authority, ethical formation, and institutional continuity. These characteristics may provide resources for developing a contextually appropriate quality culture. Teacher responsibility, collective scholarly discussion, and attention to students' intellectual development can potentially support reflective evaluation when they are integrated into institutional management. Studies of *kitab kuning* instruction demonstrate that traditional methods can coexist with organized assessment and contemporary approaches to educational evaluation (Ja'far, 2025; Fuadi et al., 2025; Huseng et al., 2025). The practical challenge is therefore not choosing between tradition and quality management, but identifying forms of evaluation that strengthen educational practice while respecting its intellectual foundations.

The proposed relationship between evaluation management and quality culture can be explained through three interconnected processes. The first is informational: evaluation produces evidence concerning students' learning and instructional performance. The second is organizational: teachers and administrators interpret the evidence, establish improvement priorities, and coordinate educational responses. The third is cultural: repeated and meaningful engagement with these activities may encourage shared expectations that educational problems should be examined and addressed collectively. This conceptual sequence extends the discussion beyond the technical execution of assessment procedures. However, the transition from organizational practices to established cultural norms cannot be demonstrated through documentary expectations alone. It requires evidence of continuity, participation, institutional commitment, and changes in everyday educational behavior.

Accordingly, the principal theoretical contribution of this study is a contextual framework that links *kitab kuning* learning evaluation management with the potential development of institutional quality culture. Evaluation planning, implementation, analysis, and follow-up provide the managerial foundation; the utilization of findings

connects evaluation to instructional improvement; and repeated collective engagement may support the institutionalization of quality-oriented practices. This framework offers a basis for examining quality culture within *pesantren salaf* without treating administrative standardization as its defining characteristic. Its empirical validity remains dependent on evidence showing how educational personnel understand evaluation, participate in improvement decisions, and sustain these practices over time. Until such evidence is available, the relationship between evaluation management and quality culture should be presented as an analytically grounded proposition rather than an established institutional outcome.

CONCLUSION

The management of kitab kuning learning evaluation provides a conceptual foundation for improving instructional quality in traditional Islamic boarding schools. The analysis identifies three interconnected dimensions: systematic evaluation management, the utilization of assessment findings for instructional improvement, and the potential contribution of evaluation practices to institutional quality culture. These dimensions indicate that the educational significance of evaluation extends beyond measuring students' mastery of classical Islamic texts. Evaluation becomes more meaningful when assessment information informs instructional decisions, addresses learning difficulties, and supports continuous improvement. However, the existence of formal assessment procedures alone is insufficient to establish their effectiveness without evidence of subsequent instructional changes.

The development of a quality culture requires the integration of evaluation practices with collective reflection, institutional commitment, and sustained instructional follow-up. Within pesantren salaf, this integration should remain responsive to traditional pedagogical methods, scholarly authority, and the distinctive educational values of Islamic boarding schools. The study contributes a contextual analytical framework connecting evaluation management, instructional improvement, and quality culture development. Its principal limitation is the absence of sufficiently documented primary field evidence to establish how these relationships operate in a particular institution. Future research should therefore examine actual evaluation practices through interviews, observations, institutional documents, and evidence of instructional improvement to determine how quality-oriented practices become embedded in everyday pesantren education.

REFERENCES

Braun, V., & Clarke, V. (2023). Toward good practice in thematic analysis: Avoiding common problems and be(com)ing a knowing researcher. *International Journal of*

Transgender Health, 24(1), 1–6. https://doi.org/10.1080/26895269.2022.2129597

Fuadi, M. N., Afifah, A. R., Azizah, Y. L., & Suwadi. (2025). Evaluasi metode sorogan dalam pembelajaran kitab kuning menggunakan model CIPP: Studi kasus di Pondok Pesantren Alluqmaniyyah Yogyakarta. *At-Tadbir: Jurnal Manajemen Pendidikan Islam, 5*(2), 102–112. https://doi.org/10.51700/attadbir.v5i2.1051

Hendrisab, H., Samsurizal, S., & Ringgit, A. (2022). Application of the Al-Miftah Al-Ulum method in learning the Kitab Kuning Islamic boarding schools. *Ahlussunnah: Journal of Islamic Education, 1*(2), 67–74. https://doi.org/10.58485/jie.v2i3.167

Hidayah, N. (2023). Manajemen mutu pembelajaran kitab kuning. *Jurnal Studi Islam dan Kemuhammadiyahan (JASIKA), 3*(2), 77–99. https://doi.org/10.18196/jasika.v3i2.55

Huseng, A. M., Auliyauddin, S., Rasyid, M. N. A., & Nursalam. (2025). Evaluasi pengajian kitab kuning menggunakan model CIPP di Pondok Pesantren Alfakhriyah Multidimensi Makassar. *Socius: Jurnal Penelitian Ilmu-Ilmu Sosial, 3*(1), 526–541. https://doi.org/10.5281/zenodo.15863593

Isnadia, R., Safitri, D. A., & Neni. (2026). Management of the integration of Kitab Kuning in PAI teaching materials: Optimization of the sorogan and bandongan methods in contemporary madrasahs. *MANAGIERE: Journal of Islamic Educational Management, 5*(1), 55–64. https://doi.org/10.35719/managiere.v4i2.2454

Ja'far, J. (2025). Designing, implementing, and evaluating: Management of teaching the Kitab Kuning in pesantren. *Jurnal Inovatif Manajemen Pendidikan Islam, 4*(2), 40–49. https://doi.org/10.38073/jimpi.v4i2.2611

Jannah, U. Q., & Arifin, M. (2025). Strategi kolaboratif akreditasi dan Total Quality Management dalam transformasi mutu pendidikan di pesantren. *Jurnal Penelitian dan Evaluasi Pendidikan, 12*(2). https://doi.org/10.64540/y4k50q62

Jabir, M., Nursyam, N., & Hidayat, L. (2023). Kitab Kuning learning model in modern Islamic boarding school. *Al-Hayat: Journal of Islamic Education, 7*(2), 589–600. https://doi.org/10.35723/ajie.v7i2.429

Khairiyah, U., Mariati, P., Faizah, S. N., Silvia, K. C., & Sasmita, F. E. (2026). Qualitative case study research in education: A systematic review of methodological approaches and analytical practices. *ELSE (Elementary School Education Journal), 10*(1). https://doi.org/10.30651/else.v10i1.24168

Lathifah, Z. K., Fauziah, R. S. P., Rusli, R. K., Roestamy, M., Martin, A. Y., Indra, S., & Suherman, U. (2025). Quality assurance in pesantren: Modernization, adaptability, and integration into Indonesia's education system. *Jurnal Pendidikan Islam, 11*(1), 101–114. https://doi.org/10.15575/jpi.v11i1.43951

Matnasir, M., Sujarwanto, S., & Hariyati, N. (2026). Extending Juran's Quality Trilogy in Islamic boarding schools: Empirical insights for realizing SDG 4 in Indonesia. *Journal of Current Studies in SDGs, 2*(3), Article 272. https://doi.org/10.63230/jocsis.2.4.272

Mursid, A., & Firdaus, A. (2026). Manajemen pembelajaran kitab kuning di Pondok Pesantren Sa'adatuddarein Kelurahan Tahtul Yaman Pelayangan Kota Jambi. *Jurnal Manajemen Pendidikan, 11*(2). https://doi.org/10.34125/jmp.v11i2.2108

Mustajib, Mutohar, P. M., Akhyak, & Aziz, A. (2025). Change management in improving the quality and competitiveness of Islamic boarding schools. *International Journal of Education Management and Religion, 2*(2), 162–179. https://doi.org/10.71305/ijemr.v2i2.508

Mustaqim, M., Razy, F., Mardhatillah, A. S., Syahweli, F., & Ilmi, D. (2025). Islamic education planning and its implications for the learning of Kitab Kuning in Islamic boarding schools. *Ahlussunnah: Journal of Islamic Education, 4*(3), 392–402. https://doi.org/10.58485/jie.v4i3.414

Pamungkas, T. A., Saefullah, S. R., Hadian, M. H., Hendriawan, E., & Adiningsih, N. U. (2026). Development of an Islamic values-based quality assurance model for strengthening quality culture in pesantren: A bibliometric and conceptual approach. *Edusoshum: Journal of Islamic Education and Social Humanities, 6*(2), 1565–1575. https://doi.org/10.52366/edusoshum.v6i2.606

Parveen, K., Phuc, T. Q. B., Alghamdi, A. A., Kumar, T., Aslam, S., Shafiq, M., & Saleem, A. (2024). The contribution of quality management practices to student performance: Mediated by school culture. *Heliyon, 10*(15), Article e34892. https://doi.org/10.1016/j.heliyon.2024.e34892

Sudir, Hidayatullah, M. F., Yusuf, M., & Subagya. (2025). Total Quality Management (TQM) in Islamic boarding schools: Teacher and principal perspectives. *Educational Process: International Journal, 15*, Article e2025136. https://doi.org/10.22521/edupij.2025.15.136

Umam, K., & Rahman, M. F. (2025). From compliance to improvement: Assessing ISO 9001:2015 and ISO 21001:2018 integration at a state Islamic university. *Proceedings of the Annual Conference on Islamic Education*. https://doi.org/10.4108/eai.14-10-2024.2356498

Mubah, H. Q. (2021). Manajemen pembelajaran pondok pesantren dalam mewujudkan pemahaman santri terhadap kitab kuning. *Re-JIEM (Research Journal of Islamic Education Management), 4*(2), 137–153. https://doi.org/10.19105/re-jiem.v4i2.5347