

Transformation of PAI Learning Evaluation from a Cognitive Approach toward Holistic Assessment

Hairul Amin¹, Mas'ud Zein², Irfan Abdurahman,³Siti Rahmah⁴, Sri Wahyuni⁵, Mahrina Olpa⁶

¹⁻²Universitas Islam Negeri Sultan Syarif Kasim Riau

³⁻⁵STAI Pelabuhan Ratu Sukabumi

*22590114673@students.uin-suska.ac.id

ARTICLE INFO

Article history

Received July 24, 2026

Revised August, 25, 2026

Accepted September 15, 2026

Keywords: Learning evaluation, holistic assessment, Islamic Religious Education, cognitive approach

ABSTRACT

Assessment in Islamic Religious Education (PAI) in Indonesia remains dominated by written tests, resulting in limited evaluation of students' affective and psychomotor development. This study aims to examine the shift from cognitive-oriented assessment toward a holistic approach and identify the gap between teachers' perceptions and practices. An explanatory sequential mixed-methods design was conducted with 150 PAI teachers from 20 secondary schools. Quantitative data were collected through a survey measuring the perceived importance and implementation of assessment across cognitive, affective, and psychomotor domains. Teachers were classified as recognizing holistic assessment when they rated it as important or very important, while systematic implementation required consistent assessment of all three domains based on predefined indicators. Qualitative data from interviews, classroom observations, focus group discussions, and assessment document analysis were used to explain the quantitative findings and integrated through a comparison of convergent and divergent results. The findings show that 81% of teachers considered holistic assessment important, but only 34% implemented it systematically. Document analysis found that cognitive assessment accounted for 72% of practices, compared with 18% for affective and 10% for psychomotor assessment. The results reveal a substantial implementation gap associated with limited instructional time, unclear guidelines, administrative demands, and inadequate infrastructure.

1. INTRODUCTION

Learning evaluation is a fundamental component of education, serving not only to measure student learning achievement but also to provide information for continuous improvement (Lubis & Pratami, 2026). In the context of Islamic Religious Education (PAI), evaluation holds particular significance because the objectives of PAI encompass integrated cognitive, affective, and psychomotor aspects, with emphasis on the formation of noble character and spiritual awareness (Musanna et al., 2025). However, the reality of PAI evaluation practices in the field reveals a misalignment between ideal goals and actual implementation. Various empirical studies indicate that PAI assessment practices in Indonesia remain heavily dominated by cognitive approaches emphasizing measurement through written tests (Jayanti et al., 2026). Consequently, the affective dimension encompassing values, attitudes, and intrinsic motivation and the psychomotor dimension related to practical worship skills and the application of Islamic values in daily life receive insufficient attention in the evaluation process (Faelasup & Astuti, 2025). This imbalance leads PAI learning to become more oriented toward mere knowledge transfer rather than the character and spiritual development of students, which is the true essence of PAI (Fikri et al., 2025). In response to this issue, a holistic approach to PAI learning assessment has become an urgent imperative.

Holistic assessment is defined as comprehensive evaluation that integrates balanced measurement of cognitive, affective, and psychomotor dimensions, employing various authentic and meaningful assessment techniques (Ridho et al., 2025). Recent research shows that holistic assessment not only improves the validity of learning outcome measurement but also promotes deeper learning, enhances students' intrinsic motivation, and creates a more inclusive learning environment (Barorah et al., 2025).

The transformation from a cognitive approach to holistic assessment faces various structural and cultural challenges within Indonesia's education system. Limited instructional time, teachers' insufficient understanding of alternative assessment techniques, unsupportive infrastructure, and the lack of valid and reliable assessment instruments and rubrics constitute major obstacles (Musanna et al., 2025). Therefore, this research aims to thoroughly analyze the transformation of PAI learning evaluation from a cognitive paradigm toward holistic assessment, identify implementation challenges, and formulate strategic recommendations to realize meaningful and sustainable PAI assessment. Learning evaluation in PAI must be understood within a holistic context involving the comprehensive development of human beings intellectual, moral, spiritual, and social dimensions (Liu et al., 2024). From an Islamic perspective, evaluation is not merely about measuring mastery of religious content but also about assessing behavioral change, moral development, and the internalization of Islamic values in students' daily lives (Qolbi, 2025). The concept of evaluation in classical Islamic tradition, as illustrated through the Prophet Muhammad's evaluative practices with his companions, demonstrates a comprehensive and balanced approach to assessing knowledge, attitudes, and behavior (Liu et al., 2024). Within the context of Indonesia's modern curriculum, PAI has clear competency standards encompassing the aspects of aqidah (faith), akhlak (morals), ibadah (worship practices), and muamalah (social ethics) (Jayanti et al., 2026). However, the translation of these competency standards into assessment practices often undergoes reduction, with the primary focus centering on cognitive aspects alone, leaving the affective and psychomotor dimensions which should be equally important neglected (Faelasup & Astuti, 2025). Holistic assessment emerges as a response to the limitations of traditional assessment that only measures low-level cognitive aspects such as memorization (Ristiliana et al., 2025). The holistic approach integrates multiple intelligences frameworks and recognizes that learning encompasses interconnected intellectual, emotional, social, and kinesthetic dimensions (AlAfnan, 2025). Empirical research demonstrates that holistic assessment employing various methods such as observation, portfolios, project-based assessment, and self-assessment can produce more accurate and comprehensive student learning profiles compared to single-test-based assessment (Ristiliana et al., 2025).

Specifically in the PAI context, holistic assessment has been developed as a conceptual framework integrating the three learning domains cognitive, affective, and psychomotor through authentic and meaningful assessment instruments (Mukmin & Nuraini, 2024). This framework is grounded in constructivist theory, emphasizing that learning is an active process in which students construct knowledge through experience, reflection, and social interaction (Arega & Hunde, 2025). Although holistic assessment possesses a strong theoretical foundation and proven empirical potential, its implementation in the field faces significant challenges. First, at the teacher level, many PAI educators still have limited understanding of alternative assessment techniques and how to develop valid assessment instruments for affective and psychomotor domains (Musanna et al., 2025). Second, school infrastructure and resources are often inadequate to support the implementation of holistic assessment, which requires time, equipment, and sophisticated data management systems. Third, external accountability pressures and a focus on national examination results lead schools to prioritize cognitive assessment while neglecting other dimensions (Arlinda, 2024). From an international perspective, the shift toward holistic assessment has become a central focus of educational reform. Various countries have adopted competency-based curricula emphasizing

holistic assessment, including Rwanda, Uganda, and various institutions in Latin America (Ndayisenga, 2025). Recent developments show that holistic assessment increasingly integrates digital technology and ongoing formative assessment (Ginanjari et al., 2025). Specifically in Islamic education, recent research has identified innovative models such as the Integrative-NTA Model (Values, Technology, Authentic Assessment), which combines value-based assessment, technology integration, and authentic assessment (Uri & Saputra, 2025). However, the implementation of holistic assessment in the field still faces substantive challenges related to teacher competence, infrastructure, and policy alignment, even though awareness of the importance of a holistic approach has been built among educators (Mistiawati et al., 2025).

Table 1. Summary of Prior Literature and Specific Limitations Addressed by Present Study

Study	Focus	Context	Main finding	Limitation addressed by this study
Mukmin & Nuraini (2024)	PAI holistic assessment framework	PAI education	Verify from original paper	Conceptual focus
Musanna et al. (2025)	Assessment implementation barriers	Indonesian education/PAI	Teacher or institutional barriers	Does not combine perceptions and observed practice
Ristiliana et al. (2025)	Multiple assessment techniques	Verify context	Benefits of diversified assessment	Limited evidence on implementation conditions
Uri & Saputra (2025)	Integrative-NTA model	Islamic education	Integration of values, technology, and authentic assessment	Model-oriented rather than implementation-focused

Despite the widely recognized importance of holistic assessment in PAI, significant research gaps exist in the literature: (1) the majority of studies are conceptual and normative, while empirical evidence on practical implementation remains limited; (2) the gap between ideal perceptions and actual practices has not been fully mapped and deeply explained; (3) successful local adaptation strategies within resource-constrained contexts have not been systematically documented; and (4) holistic perspectives integrating multi-level analysis (teacher, school, system) are still rarely found in Indonesian PAI literature. This research aims to: (1) thoroughly analyze the transformation of PAI learning evaluation from a cognitive paradigm toward holistic assessment; (2) identify the gap between ideal perceptions and practical implementation; (3) uncover implementation barriers at various levels; (4) document and analyze successful local adaptation strategies; and (5) formulate strategic recommendations that are context-sensitive and practically implementable. The novelty of this research lies in: (a) the use of a mixed-methods explanatory sequential design that systematically integrates quantitative and qualitative data; (b) a multilevel approach analyzing barriers and strategies at the individual teacher level, institutional school level, and systemic level; (c) a focus on PAI learning in the Indonesian context, which has not been extensively explored empirically; and (d) the identification and analysis of best practices from teachers who have successfully implemented holistic assessment within resource limitations.

2. METHODS

This research adopts a mixed-methods design with an explanatory sequential approach, combining quantitative and qualitative data to provide an in-depth understanding of the phenomenon of transforming PAI learning evaluation (Battista et al., 2024). The explanatory sequential design was chosen because it enables researchers to first collect and analyze quantitative data to identify

general trends and patterns, and subsequently deepen that understanding through qualitative data (Hutson & He, 2024). This approach aligns with the complexity of learning assessment issues, which involve technical, pedagogical, and contextual dimensions that cannot be fully comprehended through a single methodology alone (Onwuegbuzie et al., 2023). Philosophically, this research is grounded in the pragmatic paradigm, which acknowledges that the social reality of education is complex and multi-perspectival, thus requiring the integration of diverse perspectives for comprehensive understanding (Soicher et al., 2024). In this context, quantitative data provide an overview of the distribution and intensity of phenomena at the population level, while qualitative data reveal the meanings, interpretations, and contexts surrounding assessment practices in the field (Fitrah, 2025). This study involves a population of PAI teachers, school principals, and students from secondary education units (junior and senior high schools) across several provinces in Indonesia, representing variations in geographical, socioeconomic, and administrative contexts. For the quantitative phase, stratified random sampling was employed, taking into account the proportions of public and private schools, as well as regional development levels (urban, suburban, rural). The quantitative sample comprises 150 PAI teachers distributed across 20 schools, with the following inclusion criteria: (a) having taught PAI for at least 3 years, (b) being directly involved in the learning assessment process, and (c) being willing to participate in the research.

Table 2. Presents the target population, accessible population, sampling techniques, and final sample involved in the quantitative and qualitative phases of the study.

Participant group	Target population	Accessible population or sampling frame	Sampling technique	Sample
Secondary schools	All public and private junior and senior high schools in the selected provinces	Official list of eligible schools in the selected provinces	Stratified random sampling by province, school status, and location	20 schools
PAI teachers	All PAI teachers working in eligible secondary schools in the selected provinces	Eligible PAI teachers in the 20 selected schools	Stratified/random sampling for the quantitative phase	150 teachers
School principals	Principals of the 20 selected schools	Principals of participating schools	Purposive sampling based on school characteristics and implementation patterns	10 principals
Students	Students enrolled in PAI classes in the 20 selected schools	Students in schools where qualitative activities were conducted	Purposive sampling for focus group discussions	20 students

Participant group	Target population	Accessible population or sampling frame	Sampling technique	Sample
Education policymakers	Officials responsible for curriculum or assessment in the relevant districts or education offices	Officials from districts containing participating schools	Purposive sampling based on policy responsibilities	5 policymakers

For the qualitative phase, purposive sampling was conducted to select participants possessing in-depth knowledge of PAI assessment practices and diverse implementation experiences (Love et al., 2023). The qualitative sample consists of 30 PAI teachers (selected from the quantitative sample based on their characteristics and variation in experience), 10 school principals, 20 students (for focus group discussions), and 5 education policymakers at the district/education office level. This sampling strategy ensures depth and triangulation of perspectives from various stakeholders (Zhou et al., 2023). Data Collection Instruments for quantitative data, the instruments include: Questionnaire Survey: A structured questionnaire was developed using a 5-point Likert scale to measure the following dimensions: (1) Teachers' perceptions of the importance of holistic assessment (5 items), (2) Practical implementation of cognitive, affective, and psychomotor assessment (12 items), (3) Barriers to implementing holistic assessment (8 items), (4) Teacher competence in developing assessment instruments (7 items), and (5) Institutional support and resources (6 items). The questionnaire was developed through an iterative process involving: (a) an extensive literature review, (b) focus group discussions with 5 experts in Islamic education and educational assessment, (c) content validity testing using Lawshe's CVR (Content Validity Ratio), and (d) reliability testing using Cronbach's Alpha. The overall Cronbach's Alpha value was 0.87, indicating good internal consistency (Corr et al., 2022). Document Analysis of Assessment Artifacts: A total of 100 samples of student assessment documents (report cards, assessment sheets, portfolios) from 20 schools were analyzed to identify: the proportion of assessment types used (written tests, oral tests, observation, portfolios), the learning domains assessed (cognitive, affective, psychomotor), and the quality of assessment rubrics (Matondang et al., 2026).

For qualitative data, the instruments include: Semi-Structured Interviews: An interview guide was developed containing open-ended questions exploring: (1) Teachers' experiences and perceptions of PAI assessment, (2) The practical process of developing and implementing assessment instruments, (3) Challenges and adaptation strategies within the school context, and (4) Perspectives on the transformation toward holistic assessment. Interviews with PAI teachers lasted 45–60 minutes, with school principals lasting 30–40 minutes, and with policymakers lasting 40–50 minutes. All interviews were audio-recorded and transcribed verbatim (Pham & Le, 2024). Classroom Observations: Non-participant observations were conducted during 25 PAI learning and evaluation sessions using a structured observation guide that recorded: (1) The types of assessment activities employed (tests, observation, discussion, projects, etc.), (2) Teacher-student interactions during the assessment process, (3) The use of assessment instruments (rubrics, forms, etc.), and (4) Feedback and student responses to assessments. Each observation lasted for one full teaching period (40–45 minutes), accompanied by detailed field notes (Suranto et al., 2024).

3. RESULTS AND DISCUSSION

Quantitative Analysis Results

Of the 150 PAI teachers who responded to the survey, the demographic distribution revealed that 58% were female, with an average teaching experience of 9.4 years ($SD = 6.2$), and 72% held pedagogical certification. The majority (67%) taught in public schools, while 33% taught in private schools. In terms of geographic location, 40% were based in urban areas, 35% in suburban areas, and 25% in rural areas. Regarding educational attainment, 85% held a bachelor's degree (S1), and 15% held a master's degree (S2). This profile indicates that the sample is representative in terms of variations in experience and educational contexts (Corr et al., 2022). The analysis of assessment documents from the 20 schools revealed an unbalanced composition in the types of PAI assessments employed. Written-test-based assessment predominated, accounting for 68% of the total assessments, followed by oral tests (15%), observation (10%), and alternative assessment methods (portfolios, project-based assessment, self-assessment) at only 7%. This finding aligns with the literature, which indicates the predominance of cognitive assessment in PAI practices in Indonesia (Jayanti et al., 2026). With regard to the learning domains assessed, the analysis of assessment rubrics showed that 72% of rubrics focused on the cognitive dimension, 18% addressed the affective aspect but without standardized instruments, and only 10% encompassed the psychomotor aspect. Upon further analysis of the affective and psychomotor rubrics, it was found that the majority (65%) lacked clear descriptions of performance levels, rendering the assessments subjective and unreliable. This indicates a significant gap between the holistic objectives of PAI and the implementation of assessment practices in the field (Musanna et al., 2025). The analysis of the assessment rubrics further demonstrates a considerable disparity in the distribution of the assessed learning domains.

Table 3. Distribution of Assessed Learning Domains in PAI Evaluation Practices

Assessment Domain	Percentage
Cognitive	72%
Affective	18%
Psychomotor	10%

Although affective and psychomotor aspects appeared in some assessment rubrics, the quality of the instruments employed remained low. Specifically, 65% of affective and psychomotor rubrics lacked clear and measurable performance indicators. Consequently, the assessment process relied heavily on teacher subjectivity, raising questions about the reliability and validity of the assessment outcomes. These findings indicate a disparity between the objectives of Islamic education which emphasize the balance of cognitive, affective, and psychomotor aspects and the reality of learning evaluation practices in the field (Musanna et al., 2025). A teacher perception survey using a 1–5 scale revealed that 81% of teachers ($n = 121$) agreed or strongly agreed that "holistic assessment is important for meaningful PAI" ($M = 4.2$, $SD = 0.8$). However, when asked about practical implementation, only 34% ($n = 51$) reported systematically integrating affective and psychomotor assessments into their instruction ($M = 2.1$, $SD = 1.3$), indicating a significant gap between ideal perceptions and actual practices (Al-Hakim et al., 2025). To further examine these gaps, an independent samples t-test was conducted to compare teacher competence in developing affective–psychomotor assessment instruments between those who had received training and those who had not. The results revealed a statistically significant difference, with trained teachers scoring higher ($M = 3.4$, $SD = 0.9$) than their untrained counterparts ($M = 2.2$, $SD = 1.1$), $t(148) = 7.12$, $p < 0.001$, Cohen's $d = 1.15$ (large effect). This indicates that professional development is a crucial factor in enhancing teacher competence (Mistiawati et al., 2025).

Table 4. Effect of Training on Teacher Competence

Group	N	Mean	SD	t	df	Sig. (2-tailed)	Cohen's d
Received training	82	3.40	0.90	—	—	—	—
No training	68	2.20	1.10	—	—	—	—
Overall Comparison				7.12	148	<0.001	1.15

The analysis reveals a significant difference between the two groups, $t(148) = 7.12$, $p < 0.001$. The Cohen's d value of 1.15 indicates a large effect size, meaning that training exerts a very strong influence on enhancing teacher competence. These findings indicate that the low implementation of holistic assessment is more attributable to limitations in technical competence than to a lack of conceptual awareness among teachers regarding its importance (Mistiawati et al., 2025). A frequency distribution analysis of items concerning implementation barriers shows that the most frequently cited obstacles were limited instructional time, lack of technical guidelines and assessment instruments, excessive administrative burden, limited infrastructure and technology, and insufficient conceptual understanding of holistic assessment.

Table 5. Implementation Barriers

Barrier	Percentage
Limited instructional time	89%
Lack of technical guidelines and instruments	78%
Administrative burden	72%
Limited infrastructure and technology	65%
Insufficient conceptual understanding	61%

These data indicate that the primary barriers are structural and systemic in nature, rather than stemming solely from individual teacher factors. To determine whether barriers varied according to school location, a one-way ANOVA was conducted specifically on the infrastructure barrier variable.

Table 6. Differences in Barriers Based on School Location

Source of Variation	SS	df	MS	F	Sig.
Between Groups	48.73	2	24.37	15.34	<0.001
Within Groups	233.58	147	1.59		
Total	282.31	149			

Table 7. Descriptive Statistics

School Location	Mean	SD
Urban	2.40	0.91
Suburban	3.20	1.03
Rural	4.50	0.87

The ANOVA results indicate significant differences in the level of infrastructural barriers based on school location, $F(2, 147) = 15.34$, $p < 0.001$. Teachers in rural areas reported substantially higher infrastructural barriers ($M = 4.5$) compared to those in suburban ($M = 3.2$) and urban areas ($M = 2.4$). These findings suggest that the implementation of holistic assessment is influenced not only by teacher competence but also by the availability of resources and institutional support, which differ significantly across regions (Hasniah & Sugeng, 2025).

Qualitative Analysis Results

Thematic analysis of the 30 teacher interviews generated the first theme, explaining how teachers conceptualize holistic assessment. The majority of teachers (26 out of 30) verbally acknowledged that "assessment must encompass cognitive, affective, and psychomotor aspects," reflecting awareness of the holistic paradigm. However, deeper probing revealed that this conceptualization often remained abstract and had not been translated into operational understanding (Fikri et al., 2025). One senior teacher stated: "I know assessment should be holistic, but honestly, I struggle to explain how to do it in practice. When I give an essay question about Islamic values, does that already count as affective assessment? Or do I have to observe their attitudes in class? I'm still unclear about the distinction." This finding indicates that teachers possess awareness but lack the

theoretical grounding needed to measure and distinguish between cognitive, affective, and psychomotor dimensions within the PAI context (Fadilah et al., 2026). Observations of 25 PAI learning sessions revealed a contradiction between teachers' stated beliefs about holistic assessment and what actually occurred in the classroom. Of the 25 observed sessions, 18 (72%) were dominated by teacher instruction and factual questions requiring short answers, often recorded as "daily assessments" in the form of simple checkmarks or scores. Only 7 sessions (28%) involved more meaningful assessment activities, such as reflective discussions, project-based tasks, or structured attitude observations (Matondang et al., 2026). Focus group discussions (FGDs) with 20 students from 4 schools revealed student perspectives on their assessment experiences. One student recounted: "My PAI assessments are mostly written tests and memorization. Occasionally, we have group presentations, but the final result remains a numerical grade based on what we presented. I never receive helpful feedback on how my attitudes or behaviors are developing in terms of Islamic values."

These findings suggest that although some meaningful assessment activities are conducted, the final outcomes are still reduced to quantitative, test-based evaluations, thereby preventing the full potential of alternative assessments from being maximized (Nisa, 2025). Interviews explored specific challenges faced by teachers in implementing holistic assessment. Almost all teachers (28 out of 30) mentioned that limited PAI instructional time generally 2–3 hours per week made it difficult to apply varied assessments. One teacher stated that observing every student in detail, creating complicated rubrics, and writing personal feedback for 40–50 students every week was unrealistic, particularly given that they were already full-time and still bore other administrative burdens. These administrative tasks, which include preparing report cards, inputting data into information systems, and completing various other reports, consume up to 20–30% of teachers' total working time (Musanna et al., 2025). Furthermore, teachers reported significant limitations in terms of ready-to-use assessment instruments, valid rubrics, and technical guidelines for developing instruments targeting the affective and psychomotor domains. Only 6 out of 20 schools possessed comprehensive assessment guidance documents, while the remaining schools relied on individual teacher interpretations or simply followed the provided report card formats (Mistiawati et al., 2025). Interestingly, several teachers (9 out of 30) have developed adaptive strategies to integrate holistic assessment within existing constraints. These strategies include utilizing classroom activities such as discussions, presentations, and group projects as simultaneous opportunities for assessing affective and psychomotor dimensions, thereby reducing additional workload. Other approaches involve facilitating peer assessment, allowing students to provide mutual feedback on contributions and learning, which in turn reduces the teacher's assessment burden. Additionally, some teachers employ simplified documentation methods, such as simple checklists or rating scales, to record students' affective development rather than relying on comprehensive yet time-consuming rubrics. Collaboration among teachers has also emerged as a valuable strategy, with some schools developing shared assessment guidelines and exchanging instruments among colleagues (Zhou et al., 2023).

When asked about how the transformation toward holistic assessment should be carried out, participants including teachers, principals, and policymakers provided fairly consistent insights. From the teachers' perspective, the transformation necessitates intensive and ongoing training on holistic assessment techniques rather than one-off workshops, alongside the development of ready-to-use instruments adaptable to local contexts. Teachers also emphasized the need to reduce non-assessment administrative burdens to create space for more meaningful assessment, as well as the provision of technological support for more efficient assessment data management. Principals, meanwhile, advocated for clearer and more realistic curriculum policies regarding the proportional time allocation for various types of assessment, the establishment of incentive or recognition systems for teachers who implement high-quality holistic assessment, and enhanced collaboration with higher education institutions for research-based professional development. Finally, policymakers identified

the need to revise PAI competency standards and basic competencies to more explicitly reflect the three learning domains, to develop national assessment instruments that embody holistic assessment rather than focusing solely on testing, and to reorient accountability systems so that they measure not only cognitive achievement but also character and spiritual development (Al-Hakim et al., 2025).

Integration of Quantitative and Qualitative Results

The integration of quantitative and qualitative findings reveals convergence, divergence, and expansion. Quantitative results (81% of teachers agreeing on the importance of holistic assessment) and qualitative findings (virtually all teachers acknowledging its importance) converge in demonstrating a consensus on the ideal of holistic assessment. However, a significant divergence exists between the quantitative implementation data (only 34% integrating holistic assessment systematically) and conceptualization (81% agreeing on its importance). Qualitative analysis explains this divergence through substantial practical barriers limited time, administrative burdens, and lack of resources that create a gap between ideal aspirations and actual implementation (Soicher et al., 2024). Furthermore, the qualitative data provide an expansion, adding nuance that the obstacles are not merely structural but also conceptual. Teachers understand holistic assessment in general terms but struggle with specific operationalization, such as differentiating affective assessment from psychomotor assessment or developing valid yet efficient rubrics (Zhou et al., 2023).

Table 8. Mixed Methods Integration

Quantitative Results	Qualitative Results	Meta-Inference
81% support holistic assessment	Teachers state the importance of affective and psychomotor aspects	High conceptual awareness
34% implement holistic assessment	Teachers struggle to develop instruments	Operational barriers
Training has a significant effect	Teachers consider training very helpful	Competence becomes a key factor

The findings of this study reveal the complex phenomenon underlying PAI learning evaluation practices, which can be understood through the lenses of constructivist theory and ecological systems theory. The significant gap between ideal perceptions (81% of teachers agreeing that holistic assessment is important) and actual implementation (34% implementing it systematically) reflects what is known in the educational literature as the "implementation gap" or the "knowing-doing gap" (Fikri et al., 2025). Based on the integration of quantitative and qualitative data, this study identifies multiple interrelated layers of barriers: At the micro level (individual teacher), the gap in teachers' operational knowledge constitutes a fundamental barrier. Although teachers are aware of the importance of holistic assessment, they experience difficulties in practical operationalization namely, how to develop valid and reliable instruments for the affective and psychomotor domains, and how to integrate multiple assessments within a limited schedule. This aligns with previous research findings regarding the importance of targeted and intensive professional development in enhancing teacher competence (Mistiawati et al., 2025). At the meso level (school/institution) , school policies and infrastructure often fail to support the implementation of holistic assessment. Excessive administrative burdens, resource limitations, and the absence of efficient assessment data management systems create an uncondusive environment. Data analysis indicates that schools with visionary leadership regarding the importance of holistic assessment and adequate administrative support have higher implementation rates (53% vs. 28% in schools with minimal support), $p < 0.05$ (Samsuddin & Siregar, 2024).

At the macro level (national education system), national curriculum policies and accountability systems continue to prioritize measurable cognitive achievement through standardized examinations, while affective and psychomotor dimensions remain inadequately captured in the national evaluation

system. This creates a perverse incentive for schools to focus on cognitive assessment in order to "improve performance" on national examinations (Al-Hakim et al., 2025). One noteworthy finding is that some teachers have developed adaptive strategies to integrate holistic assessment within existing constraints. These strategies demonstrate pedagogical creativity and can serve as a foundation for scaling up best practices. The characteristics of effective strategies include: (1) integration affective and psychomotor assessments are embedded within regular learning activities rather than treated as add-ons, thereby reducing time burdens; (2) simplicity the use of simple yet valid instruments (checklists, simple rubrics) rather than comprehensive but time-consuming instruments; (3) collaboration sharing instruments and guidelines among teachers, and collaborative development of school-based assessment guidelines; and (4) technology utilization of simple technology (spreadsheets, Google Forms) for more efficient data management (Nurohmah & Ma'rifah, 2025; Zhou et al., 2023). From the perspective of learning theory, these findings reinforce the importance of a constructivist approach to learning evaluation. Constructivist theory acknowledges that learning is an active process in which students construct knowledge through experience, reflection, and social interaction (Arega & Hunde, 2025). The implication for assessment is that instruments must be designed to capture this learning process not merely the final outcomes and should involve multiple modalities, including observation, student reflection, peer feedback, and self-assessment, which together provide a holistic picture of student learning (Doan et al., 2025).

More specifically, within the PAI context, holistic assessment aligns with the concepts of *tarbiyah* (nurturing), *ta'lim* (instruction), and *ta'dib* (character formation) in the Islamic educational tradition (Fikri et al., 2025). These three concepts are inseparable and cannot be evaluated through cognitive tests alone; rather, they require assessment that captures holistic change in students. The finding that 81% of teachers support holistic assessment while only 34% implement it indicates the presence of an implementation gap. This phenomenon aligns with the knowing-doing gap theory, which explains that changes in practice do not always follow changes in understanding. In the PAI context, this gap is exacerbated by the deeply entrenched culture of number-based evaluation within the national education system. The findings of this study have direct implications for both practice and policy: For schools and teachers: (1) Development of school-based assessment frameworks that are realistic and context-responsive, rather than merely copying generic best practices. (2) Investment in sustained, focused, and practice-based teacher professional development. (3) Restructuring of teacher workload to create space for meaningful assessment practices. (4) Collaboration with universities for research-based capacity building. At the systemic level (district education offices, ministry): (1) Revision of PAI competency standards and basic competencies to more explicitly reflect the three learning domains. (2) Development of national assessment guidelines and rubrics that are practical and applicable across diverse contexts. (3) Reorientation of accountability systems to measure not only cognitive achievement but also character development and spiritual awareness. (4) Policy support for teacher professional development, technology infrastructure, and school autonomy in assessment practices (Al-Hakim et al., 2025).

4. CONCLUSION

This study demonstrates that the transformation of Islamic Religious Education (PAI) learning evaluation from a cognitive approach toward holistic assessment is a necessity that has been recognized by most teachers; however, its implementation still faces various obstacles. The research findings reveal a significant gap between teachers' perceptions of the importance of holistic assessment and evaluation practices that remain dominated by cognitive assessment. The primary barriers affecting implementation include limited instructional time, low technical competence in developing affective and psychomotor instruments, administrative burdens, and limited institutional

support and school infrastructure. Theoretically, this research reinforces the understanding that the implementation of holistic assessment is not solely determined by teachers' conceptual awareness but is also influenced by factors operating simultaneously at the individual, institutional, and educational system levels. These findings yield a conceptual framework for the transformation of holistic PAI assessment, which positions teacher competence, school support, and educational policy as key factors in driving change toward more comprehensive evaluation practices.

Practically, this research provides a foundation for the development of policies and programs aimed at improving the quality of PAI evaluation through strengthening teacher competence, providing applicable assessment instruments, optimizing school leadership support, and aligning evaluation policies with the goals of character formation and student spirituality. Thus, holistic assessment can serve as a strategic instrument for realizing more authentic, meaningful, and student-centered PAI learning oriented toward the holistic development of learners. Although this study provides empirical evidence regarding the implementation gap in holistic PAI evaluation, its findings should be interpreted in light of several key limitations: (1) sample size and scope constraints, as it involved only 150 teachers across 20 secondary schools, which limits the generalizability of the findings to a broader national scale or other educational levels; (2) reliance on perception-based data and potential self-report bias, where participant responses in surveys and interviews may reflect normative expectations or social desirability rather than actual classroom practices; (3) geographic and institutional limitations, which may not fully capture the structural and infrastructural disparities present in underdeveloped or remote regions; and (4) a primary focus on formal school settings, without accounting for external factors such as family and community religious environments that significantly influence students' affective and character development.

5. REFERENCES

- Al-Hakim, D. N., Wijayanti, F., Wicaksana, A. S., & Annur, F. (2025). The Role Of Contemporary Islamic Education Teachers: Learning Evaluation As A Path Toward Public Trust In The 21st Century. *Al-Fikri Jurnal Studi Dan Penelitian Pendidikan Islam*. <https://doi.org/10.30659/Jspi.8.2.96-117>
- Alafnan, M. (2025). Enhancing Educational Outcomes Using Alafnan Taxonomy: Integrating Cognitive, Affective, And Psychomotor Domains. *International Journal Of Evaluation And Research In Education (IJERE)*. <https://doi.org/10.11591/Ijere.V14i3.33147>
- Arega, N. T., & Hunde, T. (2025). Constructivist Instructional Approaches: A Systematic Review Of Evaluation-Based Evidence For Effectiveness. *Revista De Educación*. <https://doi.org/10.1002/Rev3.70040>
- Arlinda, A. (2024). Challenges In Evaluating Islamic Education Learning In Schools: Implications For Educational Objectives. *Tadibia Islamika*. <https://doi.org/10.28918/Tadibia.V4i1.2277>
- Barorah, N. N., Ika, V. E., Nasrun, & Kadek, S. W. N. (2025). Holistic Approach In Education To Promote Critical Thinking And Creativity. *The New Educational Review*. <https://doi.org/10.15804/Tner.2025.79.1.11>
- Battista, A., Torre, D. M., & Konopasky, A. (2024). Essential Concepts For Effective Mixed Methods Research In The Health Professions: AMEE Guide No. 173. *Medical Teacher*. <https://doi.org/10.1080/0142159X.2024.2401464>
- Corr, C., Love, H. R., Snodgrass, M., Kern, J. L., & Chudzik, M. (2022). Methodological Training In Special Education Doctoral Programs: A Mixed-Methods Exploration. *Teacher Education And Special Education*. <https://doi.org/10.1177/08884064221103902>
- Doan, C. T., Vu, T. V., & Nguyen, A. N. (2025). University Students' Perceptions On Developing Constructivist Learning Approach In Classroom Settings. *International Journal Of Evaluation And Research In Education (IJERE)*. <https://doi.org/10.11591/Ijere.V14i6.35117>
- Fadilah, S. N. F., Anwar, S., & Nugraha, R. H. (2026). Implementation Of Psychomotor Assessment And Teacher Readiness In Islamic Religious Education Learning At Senior High Schools In Bandung City. *Ta'Dib: Jurnal Pendidikan Islam*. <https://doi.org/10.19109/Td.V30i2.31719>

- Faelasup, F., & Astuti, A. (2025). Evaluasi Hasil Belajar Pendidikan Agama Islam (PAI) Melalui Library Research. *Jurnal Riset Rumpun Agama Dan Filsafat*. <https://doi.org/10.55606/Jurrafi.V4i1.5007>
- Fikri, M., Firdayeni, F., & Jamilus, J. (2025). Transformasi Evaluasi Pendidikan Islam Dari Penilaian Kognitif Ke Pembentukan Karakter Peserta Didik. *YASIN*. <https://doi.org/10.58578/Yasin.V5i6.8504>
- Fitrah, M. (2025). The Impact Of Integrated Project-Based Learning And Flipped Classroom On Students' Computational Thinking Skills: Embedded Mixed Methods. *Education Sciences*. <https://doi.org/10.3390/Educsci15040448>
- Ginanjari, A. Y., Anisah, A., Munggaran, S. W., & Lestari, P. (2025). Authentic Assessment For Holistic Learning: Evaluating The Implementation Of Formative Practices In Indonesian CBE Classrooms. *Eduscape : Journal Of Education Insight*, 3(3). <https://doi.org/10.61978/Eduscape.V3i3.1013>
- Hasniah, H., & Sugeng, S. (2025). Implementation Of PAI Evaluation In The Context Of Inclusive Learning. *ALSYS*. <https://doi.org/10.58578/Alsys.V5i6.7952>
- Hutson, B. L., & He, Y. (2024). Mixed Methods Research Centering On Minoritized Students In Higher Education: A Literature Review. *Innovative Higher Education*. <https://doi.org/10.1007/S10755-024-09748-Y>
- Jayanti, S., Ma'rifah, M., Nurrahma, N., & Syurgawi, A. (2026). Evaluation Of Planning, Implementation, And Assessment Of Islamic Religious Education Learning In State Elementary School. *EDUSOSHUM Journal Of Islamic Education And Social Humanities*. <https://doi.org/10.52366/Edusoshum.V5i3.261>
- Liu, C., Wei, W., & Liu, Y. (2024). The Construction Of Hybrid Teaching Mode Of College Sports Under The Background Of Internet Plus. *Applied Mathematics And Nonlinear Sciences*. <https://doi.org/10.2478/Amns-2024-2340>
- Love, H. R., Fetting, A., & Steed, E. (2023). Putting The 'Mix' In Mixed Methods: How To Integrate Quantitative And Qualitative Research In Early Childhood Special Education Research. *Topics In Early Childhood Special Education*. <https://doi.org/10.1177/02711214231199268>
- Lubis, W. A., & Pratami, F. (2026). Toward Holistic Assessment In Adult Islamic Religious Education: Integrating Irfani Epistemology And Andragogy. *AL-ISHLAH: Jurnal Pendidikan*. <https://doi.org/10.35445/Alishlah.V18i1.8950>
- Matondang, A. R., Saleh, R., Hasan, A., & Addary, A. (2026). Integrating Learning Theories, Teaching Methods, And Educational Assessment Across Language, Mathematics, And Legal Awareness In Elementary Education. *Multidisciplinary Indonesian Center Journal (MICJO)*. <https://doi.org/10.62567/Micjo.V3i1.1845>
- Mistiawati, E. N., Inayati, N., Valarthodi, M. I., & Irshad, I. (2025). Teacher Competency In Applying Assessment For Islamic Education Learning In The Digitalisation. *SUHUF*. <https://doi.org/10.23917/Suhuf.V37i1.8644>
- Mukmin, M., & Nuraini, N. (2024). Integrasi Penilaian Tes Dan Non-Tes Dalam Pendidikan Agama Islam: Menuju Evaluasi Holistik Untuk Pembelajaran Berkelanjutan. *Journal Of Instructional And Development Researches*. <https://doi.org/10.53621/Jider.V4i5.384>
- Musanna, S., Jamaluddin, & Duskri, M. (2025). Test And Non-Test Evaluation Methods In Islamic Religious Education Learning In Senior High School: A Critical Analysis Based On Literature Review To Strengthen Student-Centered Learning. *DARUSSALAM: Scientific Journal Of Islamic Education*. <https://doi.org/10.62945/Darussalam.V2i1.752>
- Ndayisenga, J. d'Amour. (2025). The Impact Of Competency-Based Curriculum In Enhancing Learning Outcomes: A Study Of Rwandan Secondary Schools. *Journal Of Education, Learning, And Management*, 2(2). <https://doi.org/10.69739/Jelm.V2i2.828>
- Nisa, R. (2025). Transformation Of Islamic Education Learning Management: Integration Of Computer-Based Tests And Ethical Values Of Evaluation In Islamic Education. *IJGIE (International Journal Of Graduate Of Islamic Education)*. <https://doi.org/10.37567/Ijgie.V7i1.3973>
- Nurohmah, E. Y., & Ma'rifah, S. (2025). From Paper-Based To Digital Assessment: Adoption And Challenges Of Learning Evaluation Applications In Islamic Education. *Journal Of Educational Research And Practice*. <https://doi.org/10.70376/Jerp.V3i1.206>

-
- Onwuegbuzie, A., Forzani, E. E., & Corrigan, J. A. (2023). A Fully Integrated Systematic Review Of Mixed Methods Design-Based Research. *Journal Of Mixed Methods Studies*. <https://doi.org/10.59455/Jomes.34>
- Pham, T.-T., & Le, T.-T. (2024). Implementation Of Ungrading In Vietnamese EFL Education: A Qualitative Study Of Teachers' Perspectives. *Journal Of Contemporary Language Research*. <https://doi.org/10.58803/Jclr.V3i2.103>
- Qolbi, S. K. (2025). Pengembangan Evaluasi Pembelajaran PAI Di Era Modern Pada Kampus Politeknik Berbasis Pemikiran Imam Al Ghazali. *At-Tadzkir: Islamic Education Journal*. <https://doi.org/10.59373/Attadzkir.V4i1.210>
- Ridho, A. R., Sayidah, A., & Mustofa, A. A. (2025). Instrumen Penilaian Holistik: Fondasi Evaluasi Hasil Belajar Yang Valid Dan Objektif. *Tsaqofah*. <https://doi.org/10.58578/Tsaqofah.V6i1.8341>
- Ristiliana, R., Husniati, L., & Zaki, A. (2025). Beyond Exams: Alternative Assessment Methods For Holistic Student Evaluation. *Journal Of Educational Teaching And Trends*. <https://doi.org/10.64840/Jett.V1i1.23>
- Samsuddin, S., & Siregar, I. S. (2024). Optimizing Islamic Religious Education: A Case Study On Curriculum Management In Planning, Implementation, And Evaluation. *AL-ISHLAH Jurnal Pendidikan*. <https://doi.org/10.35445/Alishlah.V16i4.6389>
- Soicher, R. N., Baker, A. R., & Thomas, R. C. (2024). A Mixed-Methods Research Design To Advance Inclusive And Equitable Teaching. *Innovative Higher Education*. <https://doi.org/10.1007/S10755-024-09741-5>
- Suranto, Mafariha, N., & Fuadi, D. (2024). Harnessing Technology For Educational Assessment: An In-Depth Analysis Of Quizizz In Economics Classrooms. *Jurnal Varidika*. <https://doi.org/10.23917/Varidika.V36i1.5431>
- Uri, F., & Saputra, D. (2025). Pengembangan Evaluasi Pelaksanaan Pendidikan Islam. *Imtiyaz: Jurnal Ilmu Keislaman*, 9(2). <https://doi.org/10.46773/Imtiyaz.V9i2.2111>
- Zhou, Y., Zhou, Y., & Machtmes, K. (2023). Mixed Methods Integration Strategies Used In Education: A Systematic Review. *Methodological Innovations*. <https://doi.org/10.1177/20597991231217937>
-