International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

Evaluation Of The Islamic Religious Education Curriculum: A Critical Analysis Of The Assessment And Measurement Of Student Learning Outcomes In Relation To The Alignment Of Curriculum Objectives

Cici Mulia Sari[1], Alfauzan Amin[2] Zubaedi[3] Rohimin[4]

[1,2,3,4] Fatmawati Sukarno State Islamic University, Bengkulu

[1]cicimuliasari@gmail.com [2]alfauzan_amin@mail.uinfasbengkulu.ac.id, [3]zubaedi@mail.uinfasbengkulu.ac.id, [4]rohimin@mail.uinfasbengkulu.ac.id

Abstract

This study critically examines the alignment between the objectives of the Islamic Religious Education (Pendidikan Agama Islam/PAI) curriculum and the system of assessment and measurement of students' learning outcomes. Normatively, the PAI curriculum—as articulated in Law Number 20 of 2003 on the National Education System—aims to cultivate individuals who are faithful, pious, and morally upright. Nevertheless, in practice, assessment systems tend to privilege cognitively measurable outcomes. Against this backdrop, the research adopts a qualitative approach with an evaluative case study design conducted in two secondary schools implementing the 2013 Curriculum and the Merdeka Belajar policy. Data were collected through document analysis (core competencies, teaching modules, assessment instruments, and report cards), in-depth interviews with PAI teachers and curriculum administrators, participatory classroom observations, and artifact analysis of assessment tools. Data analysis followed the interactive model of Miles, Huberman, and Saldaña (2014), involving data reduction, data display, and conclusion drawing/verification.The findings reveal that approximately 68% of assessment instruments emphasize the cognitive domain, predominantly at the levels of remembering and understanding according to the revised Bloom's Taxonomy (Anderson & Krathwohl, 2001). Although spiritual and social attitudes are formally assessed through scaled descriptions in report cards, the indicators remain largely generic and insufficiently operationalized. Consequently, the alignment between curricular objectives and assessment practices appears formal-procedural rather than substantively transformative. The study underscores the need to strengthen authentic and reflective assessment paradigms so that evaluation in Islamic Religious Education does not

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

merely quantify learning outcomes but genuinely captures the internalization of religious values.

Keywords: curriculum evaluation, Islamic Religious Education, authentic assessment, learning outcomes, goal alignment

INTRODUCTION

Islamic Religious Education (PAI) is, in essence, not merely a process of transferring religious knowledge but a process of forming the whole person through the internalization of values, moral guidance, and the strengthening of spiritual character. PAI is therefore not oriented solely toward academic achievement; it is directed at shaping learners who are able to live ethically, religiously, and responsibly. In this context, the PAI curriculum represents the vision of Islamic education regarding humanity, morality, and social life.

This orientation calls for a learning assessment system that measures not only cognitive achievement but also learners' attitudinal development, spiritual reflection, and behavioral change. In practice, however, PAI assessment still often focuses on memorizing concepts and definitions and on the ability to answer test questions. Yet the core of religious education lies in the internalization of values and moral transformation, which unfolds gradually and cannot always be measured directly.

Normatively, the national education goals emphasize the importance of developing learners holistically. Law Number 20 of 2003 on the National Education System states that education aims to develop learners into human beings who are faithful, pious, morally upright, knowledgeable, capable, creative, independent, democratic, and responsible. In line with this, the regulations on National Education Standards position assessment as an important component in ensuring that educational goals are attained. Assessment should therefore produce not only numbers but also a comprehensive picture of learners' development.

However, there is a tension between the ideal goals of PAI and its assessment practices in schools. Assessment still tends to be reduced to an administrative activity oriented toward written tests and numerical achievement, while the affective and psychomotor aspects do not receive balanced attention. This condition is reinforced by various obstacles, such as teachers' administrative burdens, limited time, minimal training in authentic assessment, and a school culture that still prioritizes academic results. As a result, the success of PAI learning is more easily judged by learners' ability to answer questions than by their ability to apply religious values in daily life.

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

Theoretically, curriculum evaluation must examine the alignment among objectives, content, learning processes, and the assessment system. Tyler (1949) asserted that evaluation must be grounded in clearly formulated educational objectives, whereas Stake (1975), through the responsive evaluation approach, emphasized attention to context, learners' experiences, and the realities of learning. Both perspectives indicate that PAI evaluation needs to consider not only the attainment of formal objectives but also learners' actual experiences and their developmental process in internalizing religious values.

In this context, authentic assessment is a relevant approach because it enables teachers to assess learners through behavioral observation, portfolios, reflective journals, self-assessment, and socio-religious practice. This approach can provide a more complete picture of learners' development. Nevertheless, its implementation still faces obstacles, particularly in constructing affective assessment instruments that are objective and measurable. The development of the 2013 Curriculum and the Merdeka Belajar paradigm has in fact created room for more holistic assessment, but practice in schools is still often trapped in an orientation toward numbers and administration.

Given these conditions, it is important to analyze the assessment and measurement of learning outcomes in the PAI curriculum. This study aims not only to examine the technical effectiveness of assessment but also to investigate the relevance and coherence between curriculum objectives and assessment practices in schools. Assessment is, in essence, a mirror of educational orientation; when assessment rewards only the cognitive aspect, education tends to produce a narrow academic orientation. Conversely, assessment that attends to the moral, spiritual, social, and behavioral dimensions can support a PAI education that is more humane, reflective, and consistent with the nature of Islamic education, which places the formation of moral character (akhlak) at its core.

LITERATURE REVIEW

1. Curriculum Evaluation from a Theoretical Perspective

Curriculum evaluation is an important part of educational development because it serves not only to measure learning success but also to assess the effectiveness, relevance, and coherence of the education system. Tyler (1949) viewed evaluation as the process of determining the extent to which educational objectives have been achieved. This approach emphasizes clarity of objectives, learning experiences, and measurability of outcomes as the basis of curriculum evaluation.

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

The development of evaluation theory subsequently brought more contextual and interpretive approaches. Stake (1975), through responsive evaluation, emphasized attention to learners' experiences, pedagogical interactions, and the context of curriculum implementation. Meanwhile, Guba and Lincoln (1989), through a constructivist paradigm, viewed educational reality as dynamic and plural, so that evaluation must understand the various perspectives of educational actors.

In the context of Islamic Religious Education (PAI), an evaluation approach cannot rely on quantitative measurement alone. PAI has goals related to the formation of morality, spirituality, character, and social responsibility that develop gradually. Curriculum evaluation therefore needs to integrate objective and interpretive approaches so that it can depict academic achievement as well as learners' moral and spiritual development more fully.

2. Assessment and Measurement of Learning Outcomes: Concepts and Development

Assessment and measurement are different concepts. Measurement is oriented more toward assigning numbers to particular achievements, whereas assessment encompasses interpretation and decision-making based on information about learning outcomes (Nitko & Brookhart, 2011). Bloom (1956) later developed a classification of educational objectives into cognitive, affective, and psychomotor domains, showing that the learning process is multidimensional.

Subsequent developments were marked by the revision of Bloom's taxonomy by Anderson and Krathwohl (2001), which emphasizes thinking skills from remembering to creating. In addition, Black and Wiliam (1998) introduced the importance of formative assessment as part of a continuous learning process. Wiggins (1990), through the concept of authentic assessment, likewise emphasized assessment based on learners' ability to apply knowledge in real situations.

In PAI learning, these concepts are highly relevant because the success of education cannot be judged from examination scores alone. Religious attitudes, honesty, responsibility, discipline, empathy, and integrity need to be assessed through real behavior, observation, portfolios, reflective journals, self-assessment, and other forms of authentic assessment. PAI assessment systems therefore need to integrate the cognitive, affective, and psychomotor aspects in a balanced way.

3. The Islamic Religious Education Curriculum and Its Goal Orientation

The Islamic Religious Education (PAI) curriculum aims not only to meet academic needs but also to shape learners who possess intellectual, spiritual, moral, and social balance. Law Number 20 of 2003 on the National Education System affirms that education is directed at forming human beings who are faithful, pious, morally upright, knowledgeable, creative,

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

independent, democratic, and responsible. PAI learning must therefore encourage learners to understand, internalize, and apply Islamic values in daily life.

The 2013 Curriculum seeks to realize this orientation through four core competencies: spiritual attitudes, social attitudes, knowledge, and skills. This approach shows an effort to integrate the cognitive, affective, and psychomotor aspects in learning. In practice, however, a gap remains between curriculum design and implementation in schools because learning and evaluation are often oriented more toward mastery of content and academic achievement.

Majid (2014) asserted that Islamic religious education aims not only to transmit teachings but also to build religious awareness reflected in behavior. Muhaimin (2009) also showed the importance of pedagogical approaches and the evaluation system in the success of religious education. PAI assessment therefore needs to be directed toward authentic evaluation that can depict learners' knowledge, attitudes, skills, and behavioral change more comprehensively.

4. Previous Research and the Research Gap

Previous research on PAI curriculum evaluation has extensively discussed learning implementation, the effectiveness of methods, media development, learning outcomes, and the application of authentic assessment. Several studies show that PAI teachers still face difficulties in constructing affective assessment instruments, conducting continuous behavioral observation, and managing evaluation administration. As a result, assessment still tends to focus on the cognitive aspect, which is easier to measure quantitatively.

Nevertheless, most studies still emphasize technical and implementation aspects, such as competency attainment, curriculum effectiveness, and improvements in academic scores. Studies that specifically analyze the coherence between the philosophical and normative objectives of the PAI curriculum and the system for measuring learning outcomes remain relatively limited. Yet the goals of PAI relate to forming human beings who are faithful, pious, and morally upright, which cannot be fully represented by numbers.

This research gap points to the need for a critical-analytical approach in evaluating PAI assessment systems. Studies need to look not only at how effectively assessment is carried out but also at how the evaluation paradigm shapes the direction of learners' character formation. This study is therefore directed at analyzing the coherence and relevance of the learning-outcome measurement system in relation to the objectives of the PAI curriculum, and at contributing to the development of evaluation that is more holistic, contextual, and consistent with the substantive goals of religious education.

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

International Journal of Technology and Education Research

https://e-journal.citakonsultindo.or.id/index.php/IJETER

E-ISSN: 2986-0733

Research Method

This study employs a qualitative-critical approach with an evaluative case study design. The approach is used to analyze the alignment between the objectives of the Islamic Religious Education (PAI) curriculum and the system for assessing learning outcomes through document review and field data.

1. Research Design

The study uses a qualitative case study with an interpretive paradigm to understand PAI assessment practices in a real-world context (Yin, 2014). The evaluation draws on Tyler's (1949) objectives-based approach and Stake's (1975) responsive evaluation to examine the alignment among curriculum objectives, the learning process, and assessment practices.

Data were collected through observation, interviews, and documentation to understand how the assessment system is implemented, teachers' experiences, and its alignment with the goal of forming learners' faith, piety, and moral character (akhlak). Data were analyzed using the model of Miles, Huberman, and Saldaña (2014), through data reduction, data display, and conclusion drawing.

2. Research Site, Subjects, and Duration

The study was conducted at two senior high schools (SMA/MA) that implement the 2013 Curriculum and Merdeka Belajar together with authentic assessment. The research subjects comprised 4 PAI teachers, 2 vice principals for curriculum affairs, approximately 60 students in grades XI–XII, as well as curriculum documents, teaching modules, assessment instruments, and learning outcomes.

The study lasted four months, covering the pre-observation stage, data collection, and verification of findings. The researcher acted as a moderate participant observer in order to obtain field data without being involved in instructional decision-making.

3. Data Collection Techniques and Instrument Development

Data were collected through document analysis, in-depth interviews, participant observation, and the study of assessment artifacts. The documents analyzed included the curriculum, teaching modules, cognitive, affective, and psychomotor assessment instruments, and learning outcome reports. Interviews were conducted to explore teachers' understanding of curriculum objectives, their choice of instruments, and assessment constraints.

The research instruments consisted of an interview guide, observation sheets, and a document analysis matrix. The researcher served as the primary instrument (human instrument), with instruments developed on the basis of Tyler (1949), Wiggins (1990), and the

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

revised Bloom's Taxonomy by Anderson and Krathwohl (2001), and through expert judgment from two experts in educational evaluation.

4. Data Analysis Techniques

Data were analyzed using the interactive model of Miles, Huberman, and Saldaña (2014), which comprises data reduction, data display, and conclusion drawing/verification. The data were grouped by curriculum objective orientation, assessment instrument design, implementation practices, and the alignment between objectives and evaluation.

Thematic analysis was then conducted to identify patterns of alignment or misalignment between the objectives of PAI, particularly the formation of faith, piety, and moral character (akhlak), and the assessment indicators. The results were then interpreted critically on the basis of curriculum evaluation theory and authentic assessment.

5. Data Trustworthiness

Data trustworthiness was established through source triangulation, method triangulation, and member checking. Source triangulation compared data from teachers, students, and documents, whereas method triangulation compared the results of interviews, observations, and documentation.

In addition, the study used an audit trail to document the research process systematically and researcher reflexivity to minimize interpretive bias. These procedures were carried out to maintain the credibility and transparency of the research results.

RESULTS

This section presents empirical findings obtained through document analysis, in-depth interviews, participant observation, and the study of assessment artifacts conducted during the research. The presentation of results is directed not merely at describing the phenomenon but also at reading the deeper relationship between the goal orientation of the Islamic Religious Education (PAI) curriculum and the learning evaluation practices implemented in schools. The discussion in this study therefore does not stop at technical assessment issues but also examines the extent to which the evaluation system can represent the philosophical vision of religious education as formulated in national curriculum policy.

Fundamentally, the PAI curriculum is built on an orientation toward forming human beings who possess not only intellectual competence but also spiritual and moral integrity. The assessment system should therefore ideally measure not only cognitive mastery of religious content but also capture the dynamics of value internalization in learners' behavior. However, the results show that the relationship between the normative goals of the curriculum and

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

assessment practices in the field still exhibits a number of significant tensions.

1. Curriculum Goal Orientation and Teachers' Understanding

Normatively, all the PAI teachers who participated in the study demonstrated a relatively adequate understanding of the orientation of the PAI curriculum. The teachers generally referred to the structure of Core Competencies (KI) in the 2013 Curriculum, particularly KI-1 concerning spiritual attitudes and KI-2 concerning social attitudes, as the primary basis for designing both instruction and evaluation.

One teacher stated that *"the main goal of PAI is not only for students to know, but for them to become accustomed to behaving in accordance with Islamic teachings."* This statement shows that the teachers hold a normative awareness of the essence of religious education, which does not stop at the mere transfer of knowledge. This understanding is consistent with the mandate of Law Number 20 of 2003 on the National Education System, which affirms that education aims to shape human beings who are faithful, pious, and morally upright.

In addition, in-depth interviews showed that teachers view PAI learning as a process of forming religious character that unfolds gradually and cannot always be measured instantly. Some teachers even acknowledged that the success of religious education often becomes visible only in learners' daily habits, not merely in examination scores. This perspective reveals a reflective awareness that the dimension of religiosity is more complex than mastery of ordinary academic content.

However, when this philosophical understanding is connected to assessment practices in the classroom, a more problematic reality emerges. Teachers acknowledged that administrative pressure within the education system remains very strong, particularly regarding the obligation to report learning outcomes in quantitative form. This condition indirectly pushes teachers to prioritize evaluation instruments that are easy to measure, quick to grade, and simple to document.

Thus, there is a kind of epistemological distance between teachers' normative awareness of the goals of religious education and the technical evaluation practices they carry out daily. Teachers understand that religious education aims to shape character and spiritual awareness, yet the school administrative system provides more room for measuring quantitative academic achievement. In this context, PAI learning evaluation appears caught in a tug-of-war between pedagogical idealism and the demands of educational bureaucracy.

It cannot be denied that this condition reveals a more fundamental problem in the national education evaluation system. Assessment that overemphasizes administrative aspects ultimately risks simplifying the complexity of religious education into mere numerical

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

achievement. Yet the internalization of religious values is a gradual, contextual process that often cannot be fully represented by numbers.

2. Assessment Instrument Design: Dominance of the Cognitive Domain

An analysis of 24 assessment instrument documents, consisting of multiple-choice items, essay questions, attitude observation sheets, and project assessment rubrics, showed a fairly strong dominance of the cognitive domain. Approximately 68% of the evaluation items focused on the ability to recall (*remembering*) and comprehend (*understanding*) according to the revised Bloom's Taxonomy developed by Lorin W. Anderson and David R. Krathwohl (2001).

Meanwhile, only about 12% of the instruments targeted higher-order thinking skills such as *analyzing* and *evaluating*. The *creating* aspect, which should be the highest indicator in the creative and reflective thinking process, was almost never found in the form of open-ended authentic tasks. This finding indicates that the PAI evaluation system still tends to be oriented toward the reproduction of knowledge rather than the development of learners' reflective capacity.

On the other hand, the assessment of spiritual and social attitudes has indeed been accommodated through observation sheets with a 1–4 scale, as recommended in the 2013 Curriculum guidelines. However, a deeper analysis showed that the indicators in the assessment rubrics remain general and abstract. For example, there are indicators such as *"demonstrates honest behavior"* or *"is active in religious activities"* without clear and measurable operational parameters.

As a result, attitude assessment becomes heavily dependent on teachers' subjective perceptions. In some cases, teachers admitted having difficulty determining clear boundaries between the categories of "good," "fair," or "very good" in affective assessment. This condition shows that the construct validity of the attitude assessment instruments is still not fully robust.

Participant observation in the classroom also showed that written tests were administered routinely at the end of each learning unit. In contrast, project assessments, portfolios, and reflective tasks were carried out only once per semester. Teachers generally explained that limited instructional time and heavy school administration were the main factors making authentic assessment difficult to implement intensively.

Although these reasons are understandable in practical terms, substantively this condition shows that the orientation of learning evaluation still emphasizes content mastery over value internalization. In other words, the PAI assessment system still displays a strong formal-academic tendency, while the reflective and transformative dimensions of religious

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

education have not been given proportionate room.

3. Affective Assessment Practices and Implementation Challenges

Affective assessment is one of the most complex aspects of Islamic Religious Education learning. Conceptually, the affective domain is in fact the core of religious education because it concerns the formation of learners' attitudes, morality, and spiritual awareness. However, the results show that the implementation of affective assessment still faces various epistemological and technical obstacles.

In in-depth interviews, a number of teachers acknowledged difficulty in measuring value internalization objectively. One informant stated that *"students' behavior in class does not necessarily reflect their actual attitudes outside school."* This statement reveals a critical awareness that morality and religiosity do not always appear in forms of behavior that can be observed directly.

Fundamentally, values and attitudes are psychological dimensions that develop gradually and contextually. Their measurement therefore cannot be carried out through simple instruments or momentary observation alone. Here arises an epistemological dilemma in religious education evaluation: how to assess something internal, personal, and dynamic without falling into administrative simplification.

Participant observation showed that some teachers in fact already use informal approaches such as anecdotal records to document learners' behavior. Teachers noted certain events considered to reflect learners' religious attitudes, responsibility, or honesty during the learning process. However, these records are often not integrated into the school's official reporting system.

Thus, there is a gap between teachers' individual pedagogical reflection and the formally applicable administrative mechanisms of evaluation. Teachers are aware of the importance of in-depth character assessment, but the school evaluation system does not yet fully provide room for such practice.

The findings also showed that students tend to understand success in the PAI subject through written examination scores. In *focus group discussions*, several students stated that their final PAI grade was influenced more by test results than by their participation in school religious activities. This shows that the evaluation culture that has developed still positions academic ability as the main indicator of success in religious learning.

It cannot be denied that this situation has the potential to shift students' learning orientation. When evaluation emphasizes memorization and written tests, learners tend to view

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

religious education as an ordinary academic subject rather than as a process of forming moral and spiritual awareness.

4. Coherence between Objectives and Measurement of Learning Outcomes

If the primary goal of the PAI curriculum is to shape human beings who are faithful, pious, and morally upright, then the assessment system should ideally be able to represent these dimensions proportionately and authentically. However, the results show an imbalance between the normative orientation of the curriculum and the practice of measuring learning outcomes in the field.

Administratively, learning outcome reports do include descriptions of learners' spiritual and social attitudes. However, an analysis of report card documents showed that these descriptions are often general and uniform across students. Sentences such as *"shows good religious attitudes"* or *"is active in learning activities"* appear repeatedly without clear differentiation.

This condition indicates that affective assessment is not yet fully based on systematic, in-depth observation. In other words, the presence of attitude descriptions in report cards functions more as administrative compliance than as an authentic representation of learners' character development.

Therefore, the alignment between curriculum objectives and the evaluation system can be said to remain at the formal-procedural level, not yet substantive. The ideal goals of religious education have indeed been accommodated in curriculum documents and assessment formats, but their implementation is still constrained by the logic of educational bureaucracy and technical limitations in the field.

On the other hand, this study also shows a reflective awareness among teachers regarding the need for a paradigm shift in evaluation. Some teachers realize that authentic assessment needs to be strengthened so that PAI learning does not become trapped in a mere orientation toward rote memorization. This awareness is an important asset for developing a more holistic evaluation system in the future.

5. Reflections on the Validity and Reliability of Assessment

In terms of *content validity*, cognitive assessment instruments are relatively aligned with the basic competencies stated in the curriculum. The test items generally represent the learning material taught by teachers in the classroom. However, *construct validity* in the affective domain still faces fairly fundamental problems because the assessment indicators lack clear operational

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

definitions.

The reliability of attitude assessment also depends heavily on teachers' consistency in conducting observation. In practice, this consistency varies between teachers, and even between classes taught by the same teacher. This shows that affective assessment is still strongly influenced by evaluator subjectivity.

Nevertheless, triangulation among interviews, observations, and document analysis shows consistent findings that PAI learning evaluation is still dominated by a quantitative approach. At the same time, teachers also display a reflective awareness of the need to strengthen authentic assessment and character-based evaluation.

Thus, this study shows that the main problem in PAI evaluation lies not in the absence of normative awareness but in the lack of an assessment system capable of translating the transformative goals of religious education into consistent, reflective, and substantive evaluation practice.

DISCUSSION

The findings of this study reveal a latent tension between the normative orientation of the Islamic Religious Education (PAI) curriculum and the evaluation practices carried out in schools. This tension was initially not explicit in curriculum documents or learning administration instruments. However, when the data from interviews, participant observation, document analysis, and the study of assessment artifacts were read more deeply, a fairly real distance became apparent between the philosophical ideals of religious education and the learning-outcome measurement mechanisms applied in the classroom.

Conceptually, teachers understand that the primary goal of Islamic Religious Education is not merely to transmit religious knowledge but to shape learners who possess faith, piety, and moral character, as affirmed in Law Number 20 of 2003 on the National Education System. This understanding is evident in various teacher statements that position religious learning as a process of character formation and the habituation of values. Normatively, then, there is in fact a pedagogical awareness that religious education has a transformative, not merely academic, orientation.

However, the reality of implementation points in a different direction. The analysis of evaluation instruments showed that approximately 68% of test items remain at the *remembering* and *understanding* levels according to the revised Bloom's Taxonomy developed by Lorin W. Anderson and David R. Krathwohl (2001). Meanwhile, the proportion of items that encourage critical thinking skills such as *analyzing*, *evaluating*, and *creating* is relatively limited.

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

This finding shows that PAI learning evaluation still tends to position information mastery as the main indicator of learning success.

Fundamentally, this condition reflects the influence of the *goal-oriented evaluation* paradigm introduced by Ralph W. Tyler (1949). Tyler positioned evaluation as a mechanism for determining the extent to which educational objectives have been achieved systematically and measurably. In the history of modern curriculum development, Tyler's approach has made a very significant contribution because it succeeded in presenting a rational and organized evaluation structure. Through this approach, the educational process can be designed in a more directed way because each learning objective is linked to specific achievement indicators.

However, problems begin to emerge when this paradigm is applied linearly in the context of religious education, which is laden with dimensions of value, meaning, and spirituality. Religious education relates not only to the ability to recall normative concepts but also to a process of moral internalization that unfolds gradually and personally. Therefore, if evaluation overemphasizes academic measurability, religious education risks being reduced to mere mastery of religious content without a transformation of consciousness.

Thus, the main problem in PAI evaluation practice in fact lies not in the legitimacy of Tyler's approach itself but in the way the approach is operationalized in a learning context that possesses affective and spiritual complexity. When the measurement of learning outcomes is directed more toward aspects that are easy to count and document, the moral and reflective dimensions of religious education tend to be marginalized.

On the other hand, the dominance of cognitive assessment also cannot be separated from the structural pressures teachers face in the school environment. Modern education systems demand strong administrative accountability, including the reporting of learning achievement in standardized and verifiable numbers. In this context, written tests are regarded as the most practical instrument because they are easy to construct, quick to grade, and simple to document.

It cannot be denied that such administrative demands also shape the orientation of learning evaluation. Teachers are often in a situation that forces them to choose assessment instruments that are technically efficient over instruments that are more reflective but require more time and energy. The dominance of cognitive tests in PAI learning is therefore caused not merely by teachers' low pedagogical awareness but also by an educational bureaucratic culture that places greater value on quantitative measurability.

However, if this condition continues without critical reflection, the PAI curriculum risks losing its transformative power. Religious education may turn into a mere rote-learning subject

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

that emphasizes the reproduction of information rather than a process of forming learners' moral and spiritual awareness. In the long term, this situation certainly presents a paradox: the curriculum goals speak of forming religious character, yet evaluation practice in fact measures formal academic capacity more.

The findings also show that the assessment of spiritual and social attitudes using a 1–4 scale has administratively met the 2013 Curriculum guidelines. Teachers have completed observation sheets, attitude journals, and descriptions of students' character development in accordance with applicable provisions. However, when these instruments were analyzed more deeply, it became apparent that the assessment indicators remain very general and not yet fully operational.

Several indicators such as *"demonstrates honest behavior"* or *"is active in religious activities"* are not accompanied by clear parameters regarding forms of behavior that can be observed consistently. As a result, the assessment process depends heavily on teacher subjectivity. From the perspective of construct validity as explained by Anthony J. Nitko and Susan M. Brookhart (2011), instruments with indicators that are too general risk producing data that do not fully represent the construct they are intended to measure.

In other words, there is an epistemological distance between what is assessed and what religious education actually seeks to achieve. Attitude assessment often captures only the visible surface of behavior, whereas internal dimensions such as moral awareness, sincerity, and spiritual commitment are difficult to represent through standard evaluation procedures.

Furthermore, the finding regarding the relatively uniform attitude descriptions in students' report cards shows that affective assessment is not yet fully differentiated. Many descriptions in report cards use almost identical sentences across learners. This situation indicates that character assessment remains at the administrative-formal level and is not yet fully based on in-depth observation of individual students' development.

This condition raises a fairly fundamental epistemological question: to what extent can religious values be measured through standard evaluation procedures without losing the depth of their meaning? This question is important because religious education, in essence, concerns the formation of consciousness that is personal and contextual. An evaluation approach that is too mechanistic therefore risks overlooking the complexity of learners' religious experience.

In this context, the responsive evaluation approach developed by Robert E. Stake (1975) becomes relevant to consider. Stake affirmed that evaluation should focus not only on the attainment of formal objectives but also be sensitive to the real experiences of educational

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

actors. Evaluation needs to read the social and pedagogical context surrounding the learning process, rather than merely producing numbers and administrative categories.

If Stake's perspective is applied in PAI learning, assessment should be more dialogical and reflective. Teachers act not only as score-givers but also as companions in learners' character formation. Evaluation thus becomes a space for shared reflection on students' moral and spiritual development, rather than merely a mechanism of academic classification.

Nevertheless, this study also found a reflective awareness among teachers regarding the limitations of the existing evaluation system. Teachers realize that students' behavior in the classroom does not always reflect their actual attitudes in daily life. This awareness shows that teachers have a fairly mature understanding of the complexity of values education.

However, at the same time, teachers also acknowledged that no evaluation mechanism yet exists that can truly bridge the gap between the goal of character formation and the need for administrative measurement. Thus, the problem of PAI evaluation is not merely an individual teacher issue but also concerns the structure of the education system and the assessment culture that has developed in schools.

Referring to the concept of *authentic assessment* developed by Grant Wiggins (1990), assessment should provide room for learners to demonstrate competence in contextual and meaningful situations. Authentic assessment measures not only what students know but also how they use that knowledge in real life.

However, the results show that project and portfolio assessment in PAI learning is carried out only once per semester. This very limited frequency shows that the authentic approach has not yet become mainstream in evaluation practice. Teachers still rely more heavily on written tests, which are considered more practical and easier to administer.

Thus, although the 2013 Curriculum normatively accommodates a holistic assessment approach, its implementation remains partial. Authentic assessment appears as an administrative supplement and has not yet become the main paradigm in the learning evaluation process.

From the perspective of constructivist evaluation developed by Egon G. Guba and Yvonna S. Lincoln (1989), educational reality is understood as the result of complex and continually evolving social interaction. Evaluation therefore cannot be understood merely as a process of measuring final outcomes but must be seen as part of a process of meaning-making.

The findings of this study show that PAI evaluation practice still operates within the logic of *assessment of learning*, namely assessment oriented toward final outcomes. Yet, as affirmed by Paul Black and Dylan Wiliam (1998), evaluation should develop toward *assessment for*

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

learning, and even *assessment as learning*, in which the assessment process becomes an integral part of learning and learners' self-reflection.

Critically speaking, the alignment between curriculum objectives and the PAI assessment system currently remains at the administrative-formal level, not yet substantive-transformational. The goal of forming religious character is indeed explicitly stated in curriculum documents, syllabi, and student report card formats. However, the evaluation instruments used are not yet fully able to capture the dynamics of value internalization in depth.

Nevertheless, this gap cannot be immediately understood as a failure of the curriculum. On the contrary, the findings show considerable room for reflection to improve the evaluation system more substantively. Strengthening teachers' evaluative capacity, providing more operational authentic instruments, and changing the assessment culture in schools are important steps to consider.

Ultimately, PAI curriculum evaluation needs to be understood as a continually evolving reflective process. Evaluation is not merely a tool of quality control or an administrative mechanism but a space of dialogue between the normative ideals of religious education and the pedagogical realities unfolding in the classroom. Therefore, improving the assessment system cannot be achieved through revisions of technical formats alone but requires a more fundamental paradigm shift—that assessment is an integral part of the process of forming learners' character and moral awareness.

Thus, the coherence between curriculum objectives and the system for measuring learning outcomes is not only a methodological demand in educational evaluation but also a moral commitment in the praxis of Islamic Religious Education itself.

Conclusion and Recommendations

Conclusion

This study began from a fundamental question regarding the extent to which the assessment and measurement of learning outcomes in the Islamic Religious Education (PAI) curriculum align with its normative objectives. Based on document analysis, interviews, observation, and the study of assessment artifacts, it can be concluded that, formally and administratively, efforts have been made to integrate the dimensions of spiritual attitudes, social attitudes, knowledge, and skills as formulated in the structure of Core Competencies of the 2013 Curriculum. However, at the level of implementation, the dominance of cognitive assessment—particularly at the remembering and understanding levels according to Anderson and Krathwohl (2001)—remains very prominent, with approximately 68% of all instruments analyzed.

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

These findings show that the coherence between curriculum objectives and the measurement system is at the procedural level but not yet fully substantive. The goal of forming faith, piety, and moral character, as affirmed in Law Number 20 of 2003, has been articulated in planning documents and report card formats. However, the evaluation instruments used are not yet fully able to capture the complexity of value internalization. Attitude assessment that uses numerical scales and general descriptions, although compliant with regulations, still leaves problems of construct validity and observational reliability.

Thus, there is an epistemological tension between the transformative aspirations of the curriculum and evaluation practices that tend to be reductionistic. This situation is caused not merely by a lack of teacher commitment but also by structural pressures in the form of administrative demands, a number-based accountability culture, and limited development of authentic instruments. On the other hand, teachers' reflective awareness of the importance of character assessment is an encouraging finding, as it shows room for deeper improvement.

Overall, this study affirms that the quality of the PAI curriculum cannot be separated from the quality of its assessment system. If assessment remains dominated by the logic of cognitive measurement, the goal of forming religious character risks being reduced to symbolic achievement. PAI curriculum evaluation therefore needs to be understood as a continuous reflective process, not merely compliance with administrative standards.

Recommendations

Based on these conclusions, several recommendations can be proposed in a proportionate and constructive manner.

First, at the policy level, the paradigm of authentic assessment needs to be strengthened in the training and mentoring of PAI teachers. The *assessment for learning* approach, as proposed by Black and Wiliam (1998), should be integrated more systematically into school policy, so that assessment functions not only as a tool of classification but also as a means of character development.

Second, at the school level, developing affective assessment instruments requires collaboration among teachers and continuous academic supervision. Attitude assessment rubrics need to be formulated more operationally to ensure their construct validity. In addition, the frequency of portfolios, contextual projects, and student self-reflection needs to be increased so that the authentic dimension becomes a truly integral part of PAI evaluation.

Third, at the teacher level, the capacity for reflection in reading student behavior comprehensively needs to be strengthened. Anecdotal records and informal observations, which

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

have so far been carried out personally, should be integrated into the official reporting system. In this way, attitude descriptions in report cards would no longer be generic but would represent individual students' development more accurately.

Fourth, for future researchers, studies of PAI curriculum evaluation can be extended with a longitudinal approach to examine the long-term impact of assessment systems on the formation of students' religious character. In addition, comparative exploration across schools or educational levels can provide a more comprehensive picture of the dynamics of assessment implementation.

Ultimately, improving the PAI assessment system cannot be achieved through technical changes alone. It requires epistemological courage to re-examine fundamental assumptions about what is worth measuring and how to measure it. Religious education, with all the complexity of the values it contains, demands an evaluation system that is not only administratively precise but also morally honest and pedagogically profound. It is there that the coherence between objectives and measurement finds its true meaning.

REFERENCES

A Taxonomy for Learning, Teaching, and Assessing. Anderson, L. W., & Krathwohl, D. R. (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. New York: Longman.*

Assessment for Learning. Black, P., & Wiliam, D. (2018). Assessment for Learning. New York: Routledge.

Assessment in Education: Principles, Policy and Practice. Stobart, G. (2021). Assessment in Education: Principles, Policy and Practice. London: Routledge.

Authentic Assessment. Wiggins, G. (1990). The Case for Authentic Assessment. Washington DC: ERIC Clearinghouse.

Basic Principles of Curriculum and Instruction. Tyler, R. W. (1949). Basic Principles of Curriculum and Instruction. Chicago: University of Chicago Press.

Belajar dan Pembelajaran Pendidikan Agama Islam. Majid, A. (2014). Belajar dan Pembelajaran Pendidikan Agama Islam. Bandung: Remaja Rosdakarya. [Learning and teaching of Islamic religious education]

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education: Principles, Policy & Practice, 5(1), 7–74.

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

Bloom, B. S. (1956). Taxonomy of educational objectives: The classification of educational goals. New York: David McKay Company.

Brookhart, S. M. (2017). How to use grading to improve learning. Alexandria, VA: ASCD.

Case Study Research and Applications. Yin, R. K. (2014). Case Study Research: Design and Methods. Thousand Oaks: Sage Publications.

Curriculum Development in the Postmodern Era. Slattery, P. (2013). Curriculum Development in the Postmodern Era. New York: Routledge.

Curriculum: Foundations, Principles, and Issues. Ornstein, A. C., & Hunkins, F. P. (2018). Curriculum: Foundations, Principles, and Issues. Boston: Pearson.

Educational Assessment of Students. Nitko, A. J., & Brookhart, S. M. (2019). Educational Assessment of Students. Boston: Pearson.

Educational Evaluation and Policy Analysis. Stufflebeam, D. L., & Zhang, G. (2017). The CIPP Evaluation Model. New York: Guilford Press.

Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research. Creswell, J. W., & Creswell, J. D. (2018). Educational Research. Boston: Pearson.

Evaluasi Pembelajaran PAI. Majid, A. (2019). Penilaian Autentik Proses dan Hasil Belajar. Bandung: Remaja Rosdakarya. [Authentic assessment of learning processes and outcomes]

Evaluating the Arts in Education: A Responsive Approach. Stake, R. E. (1975). Evaluating the Arts in Education: A Responsive Approach. Columbus: Merrill.

Fourth Generation Evaluation. Guba, E. G., & Lincoln, Y. S. (1989). Fourth Generation Evaluation. Newbury Park: Sage Publications.

Guba, E. G., & Lincoln, Y. S. (1989). Fourth generation evaluation. Newbury Park, CA: Sage Publications.

Hidayat, T., & Syahidin, S. (2021). Pengembangan instrumen penilaian sikap spiritual dan sosial dalam pembelajaran pendidikan agama Islam. Jurnal Pendidikan Agama Islam, 18(1), 67–84. [Development of spiritual and social attitude assessment instruments in Islamic religious education learning]

Inside the Black Box. Black, P., & Wiliam, D. (1998). Inside the Black Box: Raising Standards Through Classroom Assessment. London: King's College London School of Education.

Islamic Education: Theory and Practice. Al-Attas, S. M. N. (1999). The Concept of Education in Islam. Kuala Lumpur: ISTAC.

Kemendikbud Republik Indonesia. (2017). Panduan penilaian oleh pendidik dan satuan pendidikan untuk sekolah menengah atas. Jakarta: Kementerian Pendidikan dan

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

Kebudayaan. [Guide to assessment by educators and education units for senior secondary schools]

Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. Beverly Hills, CA: Sage Publications.

Majid, A. (2014). Belajar dan pembelajaran pendidikan agama Islam. Bandung: Remaja Rosdakarya. [Learning and teaching of Islamic religious education]

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). Qualitative data analysis: A methods sourcebook (3rd ed.). Thousand Oaks, CA: Sage Publications.

Muhaimin. (2009). Rekonstruksi pendidikan Islam: Dari paradigma pengembangan, manajemen kelembagaan, kurikulum hingga strategi pembelajaran. Jakarta: RajaGrafindo Persada. [Reconstruction of Islamic education: From the paradigm of development, institutional management, and curriculum to learning strategies]

Naturalistic Inquiry. Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic Inquiry. Beverly Hills: Sage Publications.

Nitko, A. J., & Brookhart, S. M. (2011). Educational assessment of students (6th ed.). Boston, MA: Pearson Education.

Pendidikan Agama Islam. Muhaimin. (2015). Pengembangan Kurikulum Pendidikan Agama Islam. Jakarta: Raja Grafindo Persada. [Development of the Islamic Religious Education curriculum]

Pengembangan Kurikulum Pendidikan Agama Islam. Muhaimin. (2009). Pengembangan Kurikulum Pendidikan Agama Islam. Jakarta: Raja Grafindo Persada. [Development of the Islamic Religious Education curriculum]

Peraturan Pemerintah Republik Indonesia Nomor 13 Tahun 2015 tentang Perubahan Kedua atas Peraturan Pemerintah Nomor 19 Tahun 2005 tentang Standar Nasional Pendidikan. [Government Regulation of the Republic of Indonesia Number 13 of 2015 on the Second Amendment to Government Regulation Number 19 of 2005 on National Education Standards]

Peraturan Pemerintah Republik Indonesia Nomor 19 Tahun 2005 tentang Standar Nasional Pendidikan. [Government Regulation of the Republic of Indonesia Number 19 of 2005 on National Education Standards]

Peraturan Pemerintah Republik Indonesia Nomor 32 Tahun 2013 tentang Perubahan atas Peraturan Pemerintah Nomor 19 Tahun 2005 tentang Standar Nasional Pendidikan. [Government Regulation of the Republic of Indonesia Number 32 of 2013 on the Amendment to Government Regulation Number 19 of 2005 on National Education Standards]

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

Qualitative Data Analysis. Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). Qualitative Data Analysis: A Methods Sourcebook. Thousand Oaks: Sage Publications.

Qualitative Inquiry and Research Design. Creswell, J. W. (2017). Qualitative Inquiry and Research Design. Thousand Oaks: Sage Publications.

Rahman, F., & Anwar, S. (2022). Problematika evaluasi pembelajaran afektif pada mata pelajaran pendidikan agama Islam di sekolah menengah. Jurnal Tarbawi, 18(2), 201–219. [Problems of affective learning evaluation in Islamic religious education at secondary schools]

Research and Evaluation in Education and Psychology. Mertens, D. M. (2020). Research and Evaluation in Education and Psychology. Thousand Oaks: Sage Publications.

Research Design. Creswell, J. W., & Creswell, J. D. (2018). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Thousand Oaks: Sage Publications.

Responsive Evaluation. Stake, R. E. (1975). Evaluating the Arts in Education: A Responsive Approach. Columbus: Merrill.

Stake, R. E. (1975). Evaluating the arts in education: A responsive approach. Columbus, OH: Merrill.

Suyadi, S., & Widodo, H. (2020). Paradigma penilaian autentik dalam pendidikan agama Islam di era Merdeka Belajar. Jurnal Edukasi Islami, 9(1), 33–48. [The paradigm of authentic assessment in Islamic religious education in the Merdeka Belajar era]

Taxonomy of Educational Objectives. Bloom, B. S. (1956). Taxonomy of Educational Objectives: The Classification of Educational Goals. New York: Longmans, Green.

The Case for Authentic Assessment. Wiggins, G. (1990). The Case for Authentic Assessment. Washington DC: ERIC Clearinghouse.

The CIPP Evaluation Model. Stufflebeam, D. L., & Zhang, G. (2017). The CIPP Evaluation Model. New York: Guilford Press.

Tyler, R. W. (1949). Basic principles of curriculum and instruction. Chicago: The University of Chicago Press.

Undang-Undang Republik Indonesia Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional. [Law of the Republic of Indonesia Number 20 of 2003 on the National Education System]

Wiggins, G. (1990). The case for authentic assessment. ERIC Digest. Washington, DC: ERIC Clearinghouse on Tests, Measurement, and Evaluation.

Yin, R. K. (2014). Case study research: Design and methods (5th ed.). Thousand Oaks, CA: Sage Publications.

Zainuddin, M. (2021). Evaluasi pembelajaran pendidikan agama Islam berbasis authentic

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.

assessment dalam Kurikulum 2013. Jurnal Pendidikan Islam, 10(2), 145–160. [Evaluation of Islamic religious education learning based on authentic assessment in the 2013 Curriculum]

Copyright to the author of some rights reserved This work is licensed under aCreative Commons Attribution-ShareAlike 4.0 International License.