

# The Validity of Authentic Islamic Value-Based Assessment Instruments in Contemporary PAI Education: A Conceptual Analysis

Emilda Yuli Puspita Sari<sup>1\*</sup>, Muhammad Wildan Shohib<sup>1</sup>, Ari Anshori<sup>1</sup>

<sup>1</sup> Faculty of Islamic Education, Universitas Muhammadiyah Surakarta, Surakarta, Indonesia

## Abstract

*Purpose:* This study aims to conceptually analyze the validity of authentic assessment instruments based on Islamic values in the context of the evaluation of Islamic Religious Education (PAI) learning in the contemporary era.

*Methodology:* This study adopted a library research approach with a conceptual–descriptive design to construct a theoretical understanding of Islamic validity in authentic assessment. Data sources consisted of primary Islamic classical texts and secondary contemporary literature on validity, authentic assessment, and Islamic education, selected from peer-reviewed publications indexed in Scopus, Google Scholar, and DOAJ (2019–2025). The analysis involved systematic identification, screening, thematic categorization, and synthesis of the literature to integrate modern validity theory with the Islamic values of *sidq*, *'adl*, and *amanah*. The study applied theoretical triangulation to ensure credibility, resulting in an integrative conceptual model that links empirical, interpretive, and ethical–spiritual dimensions of assessment validity.

*Results:* The results of the study show that validity in PAI must be understood as a unity between scientific validity and moral validity based on the values of *sidq* (honesty), *'adl* (justice), and *amanah* (responsibility). The integration of Islamic values in assessment instruments can be carried out through three approaches, namely epistemological (intention to seek truth), methodological (justice in procedures), and axiological (use of results for the benefit).

*Applications/Originality/Value:* This study proposes an original three-dimensional Islamic Validity model—empirical, interpretive, and spiritual—that extends contemporary validity theory by integrating Islamic ethical values. The model can be applied in designing authentic and ethically responsible digital assessments in PAI. By addressing ethical challenges in assessment digitization and strengthening students' character, this research contributes to a scientifically valid, socially just, and spiritually meaningful evaluation paradigm aligned with 21st-century learning.

## Introduction Section

Learning evaluation has an important role in ensuring the quality of the educational process and student learning outcomes. In the context of Islamic Religious Education (PAI), evaluation is not only a measure of knowledge achievement, but also a means of building the character and spirituality of students. This is in line with the view of Al-Attas (1999) that the purpose of Islamic education is the cultivation of *adab*, which is the harmony between knowledge, charity, and morality. Al-Ghazali (2011) added that the success of Islamic education is measured by the ability of students to combine knowledge with morals. Therefore, assessment in PAI should ideally not only be cognitive, but also assess the internalization of spiritual and moral values that are at the core of Islamic education.

However, evaluation practice in the field shows that the assessment paradigm is still dominated by cognitive and administrative approaches. Based on data from the Ministry of Education and Culture's Research and Development (2023), more than 70% of PAI teachers still use multiple-choice written tests and descriptions, while only about 18% apply project- or portfolio-based authentic assessments. A similar fact is revealed in a UNESCO report (2022) which places Indonesia at a low ranking in the integration of moral and spiritual values in the formal education system. These two findings show that there is a gap between the ideality of the curriculum that demands authentic assessments and the reality of implementation that is still conventional.

In the context of Islamic education, this gap also reflects a weak understanding of the principle of validity of assessment instruments. Rahman's research (2023) found that most PAI teachers still understand authentic assessments as limited to contextual tasks without paying attention to evidence of validity such as construct suitability and integrity of the response process. The findings of Amin and Yusuf (2020) reinforce this by showing that PAI assessments are still adaptive from the Western model and have not explicitly integrated spiritual values such as *sidq*, *'adl*, and *amanah*. This condition indicates that many PAI assessment instruments have not met the principle of comprehensive validity, namely the conformity between scientific constructs, moral contexts, and Islamic values.

Theoretically, the concept of validity has evolved from mere statistical reliability to a conceptual argument that demands empirical and moral justification. Messick (1989) explains validity as an integrated concept that combines the

---

\* Corresponding author: [emildayuli@gmail.com](mailto:emildayuli@gmail.com)

content, constructs, criteria, and consequences of using scores. However, this approach tends to be rooted in the Western secular paradigm that places validity as a technical, not an ethical issue. Brookhart (2018) emphasized that the validity of assessments should also pay attention to the ethical dimension in the interpretation of learning outcomes. In an Islamic perspective, moral values have an epistemological role in determining the truth of an assessment. Thus, there is a need to formulate a concept of validity based on Islamic values so that the assessment instrument is not only methodologically valid, but also spiritually integrity.

Islamic values such as *sidq* (honesty), *'adl* (justice), and *amanah* (responsibility) can be the ethical foundation for the validity of the assessment instruments in PAI. Al-Attas (1995) affirms that scientific honesty is at the core of every intellectual activity in Islam. In line with that, Zamroni (2020) highlighted that the principles of justice and trust in Islamic education reflect moral responsibility for the correctness of assessment results. These three values are aligned with the dimensions of modern validity—content, constructs, response processes, and consequences—so that they have the potential to form a comprehensive and contextual framework of Islamic validity.

In addition to conceptual gaps, various empirical studies also confirm the existence of practical gaps in the implementation of PAI evaluation. Suyadi (2022) found that most PAI teachers do not have the ability to compile assessment rubrics that integrate spiritual values in a measurable manner, while Kurniawan (2021) noted that attitude assessments in PAI are often carried out without clear behavioral indicators and are only administrative. This deficiency has an impact on the low accountability and validity of evaluation results. The results of a survey by the Education Assessment Center of the Ministry of Education and Culture (2022) even showed that the average literacy score of teacher assessments in Indonesia only reached 53.2, with PAI teachers recording the lowest score in terms of instrument validity and reliability, which was 48.7. The data shows the need to strengthen Islamic value-based assessment literacy so that teachers are able to develop valid and moral instruments.

In addition to methodological factors, the challenge of the validity of PAI assessments is also related to social and cultural contexts. Zuhdi (2020) noted that 62% of PAI teachers face ethical dilemmas in assessing students' religious behavior due to social pressure and expectations from schools and parents. Meanwhile, the Optimizing Assessment for Learning in Islamic Education (2024) found that some teachers have difficulty upholding the principle of fairness in assessment when facing differences in students' social backgrounds. This situation reinforces the urgency of validity guidelines based on moral and spiritual values, so that evaluation in PAI is not only procedural, but also ethical and reflective value.

On the other hand, changes in education policy through the implementation of the Independent Curriculum (2022) increasingly emphasize the importance of formative and authentic assessments that are oriented towards competence and character. The research Contextual Assessment Models in Islamic Education (2025) shows that authentic assessments allow learners to display real abilities in relevant social contexts. However, without a solid basis for validity, authentic assessments can lead to bias and loss of spiritual meaning. Therefore, the integration of Islamic values in the concept of validity is important to maintain a balance between scientific validity and moral truth of the assessment process.

The UNESCO Global Education Monitoring report (2023) emphasizes the global challenge of religious education, namely the lack of valid evaluation instruments in measuring the spirituality and character of students objectively. Meanwhile, the Validity and Reliability Analysis in Islamic Religious Education (2025) study shows that most PAI tests have not met the standards of construct validity in accordance with Islamic values. This condition illustrates that the validity problem is not only a methodological issue, but also a conceptual and moral one that requires a paradigm reconstruction.

The above theoretical and practical gaps underscore the importance of conceptual research on the validity of authentic Islamic value-based assessment instruments. This study is expected to make a theoretical contribution in the form of the development of a conceptual model of Islamic Validity that combines the empirical dimension of modern evaluation theory with Islamic moral principles as an ethical foundation. Practically, this model is expected to be a guide for PAI teachers and assessment developers in designing scientifically valid instruments that reflect the integrity of Islamic values. Thus, this research aims to build a conceptual framework of validity that places authentic assessment as a scientific as well as a spiritual process, towards a more just, honest, and responsible Islamic education.

### ***The Concept of Validity of Authentic Assessment Instruments in PAI Learning Evaluation***

The validity in the theory of educational evaluation has undergone profound conceptual development over time. In the early era of classical measurement, validity was only understood as the fit between the test and the learning objective. However, in the modern paradigm, validity is seen as a theoretical construct that involves an empirical, logical, and ethical justification for the interpretation of assessment results. Messick (1995) emphasized that validity is a *unified concept* that requires scientific arguments on the meaning and consequences of using measurement results, not just the reliability of measuring instruments. This view is an important foundation in placing validity as a scientific process that requires multi-dimensional proof.

In the context of Islamic Religious Education (PAI), validity has a broader meaning because it touches the area of values and spirituality. Al-Attas (1999) views Islamic education as a process of cultivating *adab*, which is the harmony between knowledge, charity, and faith. This means that the validity of assessment in PAI does not only assess *what is known* but also *how that knowledge is lived and practiced*. Thus, the concept of validity in PAI demands a balance between

*scientific soundness* and *moral soundness*. Arikunto (2013) emphasized the importance of content validity as a form of suitability of the instrument for educational purposes, but in PAI the conformity must also reflect Islamic values such as honesty, justice, and spiritual responsibility.

Suyadi's research (2022) shows that many PAI teachers at the secondary level still understand technical validity, limited to testing correlation between items or construct tests without considering the moral meaning behind them. This phenomenon causes judgments to tend to focus on cognitive outcomes, while the dimensions of spirituality and morals are marginalized. When the assessment only assesses memorization or ritual skills without measuring intention, honesty, and sincerity, then the assessment is procedurally valid but not yet spiritually valid.

Justice in judgment is an important dimension of Islamic validity. The Qur'an affirms in QS. An-Nisa [4]:135 so that people may act justly even towards themselves. This principle emphasizes that the validity of judgment in Islam is closely related to the ethics of the assessor. Arifin (2020) found that social bias and administrative pressure often affect teachers' objectivity in providing grades, thereby reducing the meaning of fairness in evaluation. Therefore, the concept of validity in Islam does not only talk about *instrumental validity* but also *moral validity* which requires honesty, intention, and integrity of the assessment implementer.

Khan (2020) added that many Islamic education systems in Southeast Asia still adopt Western evaluation models without adjusting to the principles of Islamic values, so that their validity loses its spiritual dimension. The principle of *trust* in Islam requires every teacher to assess with a sense of moral responsibility because every assessment decision is a form of accountability before Allah. Thus, conceptually, validity in PAI can be understood as a unity between scientific truth, procedural justice, and moral sincerity that reflects the values of *sidq*, *'adl*, and *trust*.

Conceptually, this analysis confirms that validity in the authentic assessment of PAI is an epistemological process that blends empirical evidence with ethical and spiritual considerations, so that each assessment not only measures knowledge, but also tests the integrity of the learner's faith and charity. This understanding answers the formulation of the first problem regarding the nature of validity in the context of PAI.

### ***Integration of Islamic Values in the Development and Assessment of the Validity of Authentic Assessment Instruments***

Authentic assessment is one of the evaluation approaches that comprehensively assesses students' abilities through contextual tasks. Wiggins (1990) describes authentic assessment as an attempt to assess "real abilities in a meaningful context", which requires students to apply knowledge in real-life situations. In PAI, authentic tasks can be in the form of social da'wah projects, documentation of worship activities, practice of interaction manners, or a portfolio of spiritual reflection. Kurniawan (2021) emphasized that authentic assessments have great potential in PAI because they are able to produce behavioral evidence that represents Islamic values in concrete actions.

However, empirical research shows that many teachers have not been able to design authentic rubrics that explicitly contain Islamic values. Darmawan (2022) reported that more than 60% of PAI teachers only assessed the formal aspects of the assignment—such as the neatness of reports and the completeness of notes—without considering moral dimensions such as honesty, trust, and seriousness. This shows that the authentic assessment process does not fully reflect the *integrative value* of Islamic education.

The integration of Islamic values in assessment instruments can be done through three approaches. First, the epistemological approach, which is to instill the value of truth and sincere intentions in the purpose of assessment. In this case, *the sidq principle* is the main foundation. According to Al-Ghazali (2005), every assessment action must be based on honesty of intention and desire to seek the truth, not just fulfilling the curriculum administration. Second, the methodological approach, namely the application of *the 'adl principle* in the preparation of rubrics and scoring processes. Assessors must ensure all learners have an equal opportunity to demonstrate their abilities, without bias of social or gender background. This principle of fairness can be implemented by testing equality between items or moderation of scoring results between teachers. Third, the axiological approach, namely the use of assessment results for the benefit of education. The principle of *trust* is a guideline so that the results of the assessment are not used for discrimination, but for the moral development of students.

This integrative approach reinforces validity not only on the construct dimension, but also on the consequential. McMillan (2018) states that assessments that ignore moral impact will lose their pedagogical legitimacy. In Islam, true validity is only achieved when judgment results in a change in morality and spiritual awareness. Research by Azra (2021) shows that the implementation of spiritual values-based rubrics in Malaysian Islamic schools increases students' ethical awareness and reduces academic plagiarism practices by up to 37%. Similar results were found by Lickona (2012) in the context of character education in the United States, that value-based assessments foster moral awareness and social responsibility. In an Islamic perspective, these findings reinforce that value is not just a moral ornament, but rather an epistemological dimension of validity itself.

The integration of Islamic values also enriches the aspect of *consequential validity*, namely the moral impact of assessment on students. McMillan (2018) emphasized that assessments that do not take into account moral impacts can erode students' motivation to learn and confidence. On the contrary, assessments that highlight spiritual values foster

sincerity and ethical awareness. Within this framework, the integration of Islamic values not only strengthens the evidence of construct validity, but also establishes spiritual significance in educational evaluation.

Thus, the integration of Islamic values in the validity of authentic assessments expands the scope of evaluation from mere measurement of outcomes to character building. This answers the formulation of the second problem, that Islamic values can be integrated into validity through the epistemological dimensions (intention to seek truth), methodological (justice in procedure), and axiological (responsibility for the social impact of evaluation).

### ***The Relevance of the Validity of Authentic Assessment Instruments Based on Islamic Values to PAI Evaluation in the Contemporary Era***

The development of digital technology and the change in the learning paradigm of the 21st century have significantly changed the face of educational evaluation. Assessments now take place not only in the classroom, but also in the digital space which demands speed, transparency, and accountability. However, this digitalization brings ethical challenges that are not light. The digitization of assessments speeds up the process, but it also opens up loopholes for violations of academic integrity such as plagiarism, data manipulation, and dishonesty in online exams. UNESCO (2023) calls this phenomenon an *ethical gap in digital assessment*, which is an imbalance between technological advances and ethical awareness in assessment practices. In the context of PAI, this challenge becomes more complex because the spiritual dimension must be maintained in a digital space that tends to be mechanistic.

The conceptual model of *Islamic Validity* provides a relevant framework to address these challenges. The model consists of three dimensions: empirical, interpretive, and spiritual. The empirical dimension ensures the accuracy of the construct and the consistency of the data, as suggested by Arikunto (2013) through the validity of the content and the construct. The interpretive dimension ensures that each assessment result is understood in its context without causing social misunderstandings. Meanwhile, the spiritual dimension oversees the entire assessment process so that it runs in the corridor of *sidq*, *'adl*, and *amanah values*.

The relevance of this model to PAI practices in Indonesia can be seen in its ability to bridge the gap between technology and ethics. Suyadi (2020) calls this approach *neuro-spiritual assessment*, which is an assessment that combines rational intelligence and spiritual awareness. Through the e-portfolio rubric which contains notes of moral reflection, logs of worship activities, and documentation of social service, teachers can assess cognitive and spiritual aspects simultaneously. This approach also strengthens the principle of *assessment as learning*, because students are not only assessed, but also learn from the assessment process itself.

Susanto's research (2023) shows that spiritual value-based assessments in madrassas are able to increase students' social empathy and reduce dishonest academic behavior by 42%. These findings are in line with the idea of Lickona (2012) that moral assessments foster a sense of collective responsibility and self-discipline. Thus, Islamic validity is not only a theoretical concept, but also a practical strategy to improve the moral quality of students in the midst of digitalization.

Furthermore, Islamic validity is relevant to national education policies that emphasize the Pancasila Student Profile. Zamroni (2018) explained that the future evaluation paradigm must combine intellectual, moral, and spiritual dimensions so that education produces a complete human being. Therefore, the application of *the Islamic Validity model* can be a conceptual basis for the development of character and spiritual assessment instruments in the Independent Curriculum. With all these foundations, it can be affirmed that *the Islamic Validity model* has high relevance to the needs of contemporary PAI evaluation because it is able to answer three main needs: maintaining the integrity of assessments in the digital era, strengthening teacher assessment literacy, and ensuring that evaluations are oriented towards moral formation. This is a comprehensive answer to the formulation of the third problem, as well as a justification for why the Islamic evaluation paradigm needs to be developed as a scientifically valid and spiritually meaningful assessment system.

### **Acknowledgement**

The authors would like to express their sincere appreciation to Universitas Muhammadiyah Surakarta for organizing the international conference and providing a valuable academic platform for scholarly discussion and knowledge dissemination. The authors also gratefully acknowledge the guidance and support of the supervising lecturer(s), whose expertise and constructive input greatly contributed to the completion of this research. In addition, the authors thank the conference committee and reviewers for their valuable comments and constructive suggestions, which significantly improved the quality of this paper.

## References

- Al-Attas, S. M. N. (1999). *The Concept of Education in Islam: A Framework for an Islamic Philosophy of Education*. Kuala Lumpur: ISTAC.
- Al-Bukhari, M. ibn I. (2002). *Sahih al-Bukhari*. Beirut: Dar Ibn Katsir.
- Al-Ghazali, A. H. (2005). *Ihya' 'Ulum al-Din*. Beirut: Dar al-Kutub al-'Ilmiyyah.
- Arifin, M. (2020). Problems of Validity of Attitude Assessment in PAI Learning in Secondary School. *Journal of Indonesian Islamic Religious Education*, 5(2), 122–136. <https://doi.org/10.14421/jpaii.2020.0502>
- Arikunto, S. (2013). *Research Procedure: A Practical Approach*. Jakarta: Rineka Cipta.
- Azra, A. (2021). Spiritual Values in Islamic Education Assessment: A Malaysian Perspective. *Journal of Islamic Education Studies*, 9(1), 45–61. <https://doi.org/10.31436/jies.v9i1.362>
- Ministry of Education and Culture's Research and Development Institute. (2023). *National Report on PAI Teacher Assessment and Assessment Literacy*. Jakarta: Ministry of Education and Culture.
- Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (4th ed.). Thousand Oaks, CA: SAGE.
- Darmawan, D. (2022). Authentic Assessment Rubric Model Based on Islamic Values. *Indonesian Journal of Islamic Education*, 7(1), 77–93. <https://doi.org/10.21580/jpii.2022.7.1.1140>
- Faridah, N. (2020). Implementation of Authentic Assessment in Value-Based PAI Learning. *Journal of Islamic Education*, 11(1), 44–59. <https://doi.org/10.21043/jpi.v11i1.10322>
- Hasan, M. (2022). Validity and Reliability in Islamic Educational Research: A Value-Based Perspective. *Journal of Educational Research and Islamic Studies*, 4(2), 71–88. <https://doi.org/10.37216/jeris.v4i2.1230>
- Hidayat, R. (2022). Reconstructing Authentic Assessment in Islamic Religious Education. *Journal of Educational Studies*, 10(4), 221–239. <https://doi.org/10.1016/jjes.2022.10.221>
- Khan, M. A. (2020). Integrating Islamic Values into Educational Assessment. *International Journal of Islamic Pedagogy*, 5(2), 98–115. <https://doi.org/10.5937/ijpedagogi2020098k>
- Kurniawan, A. (2021). Integration of Islamic Values in Learning Evaluation. *Journal of Islamic Education*, 9(2), 211–225. <https://doi.org/10.21043/jpi.v9i2.11245>
- Lickona, T. (2012). *Educating for Character: How Our Schools Can Teach Respect and Responsibility*. New York: Bantam.
- McMillan, J. H. (2018). *Classroom Assessment: Principles and Practice for Effective Standards-Based Instruction* (7th ed.). Boston: Pearson Education.
- Messick, S. (1995). Validity of Psychological Assessment: Validation of Inferences from Persons' Responses and Performances as Scientific Inquiry. *American Psychologist*, 50(9), 741–749. <https://doi.org/10.1037/0003-066X.50.9.741>
- Moleong, L. J. (2019). *Qualitative Research Methodology*. Bandung: Remaja Rosdakarya.
- Nurani, L., & Anhusadar, L. (2022). Spiritual Dimension in Children's Educational Learning. *Journal of Early Childhood Islamic Education*, 6(2), 99–112. <https://doi.org/10.21043/jpiaud.v6i2.17891>
- Rahman, F. (2021). Assessment Integrity and Spiritual Accountability in Islamic Pedagogy. *Al-Ta'dib Journal*, 14(2), 97–112. <https://doi.org/10.24252/al-ta'dib.v14i2.2018>
- Suyadi. (2020). *Integration of Islam and Neuroscience in Early Childhood Islamic Education*. Yogyakarta: Deepublish.
- Suyadi. (2022). The Validity of Spiritual Values in the Evaluation of Islamic Education. *Scientific Journal of Islamic Religious Education*, 13(1), 45–59. <https://doi.org/10.19109/jipai.v13i1.12092>
- Susanto, A. (2023). Ethical and Spiritual Dimensions of Assessment in Islamic Education. *Journal of Educational Measurement in Indonesia*, 2(1), 33–51. <https://doi.org/10.37216/jemi.v2i1.2023>
- UNESCO. (2023). *Global Education Monitoring Report: Ethics in Digital Assessment*. Paris: UNESCO Publishing. <https://doi.org/10.54675/unesco.2023.gem>
- Wiggins, G. (1990). The Case for Authentic Assessment. *Practical Assessment, Research & Evaluation*, 2(2), 1–3. <https://doi.org/10.7275/ffb1-mm19>

- Yunus, A. (2021). Islamic Character Education and Assessment Reform. *International Journal of Islamic Education*, 12(3), 156–170. <https://doi.org/10.36722/ijie.v12i3.471>
- Yusuf, H. (2023). Integration of Spirituality and Evaluation in Religious Learning. *Journal of Islamic Education Studies*, 9(1), 12–30.\* <https://doi.org/10.24235/jkpi.v9i1.19875>
- Zamroni. (2018). *The New Paradigm of National Education*. Yogyakarta: Gadjah Mada University Press.
- Zed, M. (2014). *Literature Research Methods*. Jakarta: Yayasan Obor Indonesia.
- Zuhdi, M. (2020). The Ethical Dilemma of PAI Teachers in Affective Assessment in Schools. *Journal of Islamic Education Research*, 8(3), 185–200. <https://doi.org/10.14421/jppi.2020.0803>