

To cite this article: Sunhaji and Muhamad Muzaki (2026). RECONSTRUCTING COMPETENCY-BASED LEARNING EVALUATION IN ISLAMIC EDUCATION IN THE SOCIETY 5.0 ERA: INTEGRATING AUTHENTIC ASSESSMENT, DIGITAL LITERACY, AND ARTIFICIAL INTELLIGENCE, International Journal of Education and Social Science Research (IJESSR) 9 (4): 225-254 Article No. 1268, Sub Id 1944

RECONSTRUCTING COMPETENCY-BASED LEARNING EVALUATION IN ISLAMIC EDUCATION IN THE SOCIETY 5.0 ERA: INTEGRATING AUTHENTIC ASSESSMENT, DIGITAL LITERACY, AND ARTIFICIAL INTELLIGENCE

Sunhaji¹ and Muhamad Muzaki²

¹Islamic Education Management, Faculty of Tarbiyah and Teacher Training, Prof. K.H. Saifuddin Zuhri State Islamic University,

²Islamic Education, Faculty of Tarbiyah and Teacher Training, Prof. K.H. Saifuddin Zuhri State Islamic University, Purwokerto, Indonesia

DOI : <https://doi.org/10.37500/IJESSR.2026.9409>

ABSTRACT

The evaluation of Islamic Education learning in many educational institutions remains dominated by written-test instruments that emphasize the cognitive domain, causing the affective and psychomotor domains which in fact constitute the core of Islamic Education success in shaping learners' character and religious practice to be largely neglected (Abdillah, 2016; Sholahudin et al., 2025). This condition becomes increasingly inadequate when confronted with the demands of the Merdeka Curriculum, which is oriented toward competency attainment, and with the Society 5.0 era, which requires a human-centered integration of cyber space and physical space (Cabinet Office, Government of Japan, 2016). This study aims to reconstruct a competency-based model for evaluating Islamic Education learning through the integration of three elements authentic assessment, digital literacy, and artificial intelligence (AI) while simultaneously offering a novel contribution in the form of a single framework that explicitly combines these three elements, rather than treating them separately as is common in previous studies. The study employs a qualitative library-research approach that adapts PRISMA-style literature search and selection procedures (Page et al., 2021) to enhance methodological transparency, drawing on primary data consisting of 40 selected national and international journal documents as well as curriculum policy documents published over the past ten years, which were analyzed using content-analysis techniques. The findings indicate that the reconstruction of Islamic Education evaluation must rest on four interrelated pillars: competency as the foundation of learning outcomes, authentic assessment as the measurement approach, digital literacy as the medium of implementation, and artificial intelligence as an accelerator of personalization and assessment efficiency synthesized into a conceptual model called the Integrated Islamic Education Competency Evaluation Model. This model enables Islamic Education evaluation to address the cognitive, affective, and psychomotor domains simultaneously, objectively, and sustainably. The implications are that Islamic Education teachers are required to possess digital pedagogical competence and AI literacy, schools need to provide supporting infrastructure and policies, and curriculum developers need to integrate digital-literacy indicators and

AI-use ethics into Islamic Education learning outcomes. This study is expected to make a conceptual contribution to the development of Islamic Education assessment policy that remains relevant to contemporary demands without sacrificing the Islamic values that constitute the spirit of religious education.

KEYWORDS: learning evaluation; Islamic Religious Education; authentic assessment; digital literacy; artificial intelligence; Society 5.0; systematic literature review

1. INTRODUCTION

Islamic Education occupies a strategic position within the national education system because it functions not only to transmit religious knowledge but also to shape learners' attitudes, beliefs, and behavior so that they align with Islamic values in daily life. It is precisely this dual function that makes the evaluation of Islamic Education learning fundamentally different from the evaluation of other subjects. Its success cannot be measured adequately by mastery of scriptural evidence (dalil) and concepts alone; it must also encompass the internalization of values in the affective domain and the practice of worship and moral conduct in the psychomotor domain as an integrated whole (Sholahudin et al., 2025; Siregar et al., 2025).

The problem is that the rapid transformation of the curriculum from the 2013 Curriculum to the Merdeka Curriculum, which places Learning Outcomes and the Pancasila Student Profile at the center of its orientation (Kemendikbudristek, 2022) has not yet been fully matched by an equivalent transformation in classroom evaluation practice. This shift has occurred alongside a broader restructuring of society toward what the Japanese government has termed Society 5.0, a society that integrates cyber space and physical space in a human-centered manner in order to resolve social problems while simultaneously driving economic growth (Cabinet Office, Government of Japan, 2016). In the context of education, this era is marked by the massive use of digital devices, big data, and artificial intelligence permeating nearly every facet of learning practice, including assessment (Halkiopoulous & Gkintoni, 2024).

Islamic education cannot afford to take a passive stance toward this change. Several studies show that learners of the Society 5.0 generation grow up with patterns of interaction, learning resources, and ways of thinking that differ from those of previous generations, thereby demanding corresponding adjustments in Islamic Education learning strategies and assessment (Sari & Fitriana, 2022; Bahri, 2022). If Islamic Education evaluation continues to rely on conventional written-test instruments, the considerable potential offered by authentic assessment, digital literacy, and artificial intelligence to capture learners' competencies in a more complete and meaningful way will simply be missed.

A number of empirical studies reveal similar patterns across various levels of education. Abdillah (2016) found that Islamic Education learning evaluation at the senior-high-school level remains heavily dependent on teachers' professional and pedagogical competence in developing instruments,

yet the instruments produced generally remain cognitively oriented. A study on the integration of Qur'anic verses into Islamic Education evaluation likewise reveals that the cognitive domain is often confined to the level of basic knowledge, the affective domain has yet to probe the internalization of Islamic values in any depth, and the psychomotor domain, although displaying fairly good practical ability, remains weak in higher-order thinking skills such as creativity and moral evaluation (Sholahudin et al., 2025).

The same pattern is found in the implementation of the Merdeka Curriculum. Achmad et al. (2022) report that although elementary-school Islamic Education teachers have applied a range of authentic-assessment techniques written and oral tests for the cognitive domain; observation, self-assessment, and journals for the affective domain; and projects, performance tasks, and portfolios for the psychomotor domain implementation still faces constraints of limited time, resources, and teacher training. Teachers themselves acknowledge that they still require further training to apply this kind of alternative assessment consistently. This phenomenon indicates a gap between the curriculum's ideal of holistic, competency-oriented evaluation and the reality of evaluation practice in the field, which has not yet fully moved away from the old pattern centered on rote memorization and written tests.

At the same time, the potential of digital technology and artificial intelligence to support Islamic Education evaluation has not yet been harnessed systematically to any great extent. M. R. Fauzi et al. (2025) identify that integrating AI into Islamic Education learning evaluation has the potential to improve the efficiency, objectivity, and personalization of assessment, but that its implementation remains overshadowed by teachers' limited digital literacy, uneven infrastructure, and ethical concerns regarding data privacy and the compatibility of AI output with Islamic values.

The literature addressing Islamic education evaluation, authentic assessment, digital literacy, and AI in Islamic education is in fact fairly abundant, yet it remains scattered and tends to treat each element separately. Studies on authentic assessment in Islamic Education generally stop at describing assessment techniques at a single educational level (Achmad et al., 2022; Hanafi et al., 2025). Studies on digital literacy in PAI focus more on its use as a learning medium than as an evaluation instrument (Pratiwi et al., 2024; Lisyawati et al., 2023). Studies on AI in Islamic Education have only begun to explore opportunities and challenges in general terms without formulating a concrete evaluation framework (M. R. Fauzi et al., 2025; Muchlis, 2025). As far as this search has been able to establish, no study has yet been found that explicitly reconstructs these three elements authentic assessment, digital literacy, and artificial intelligence—into a single, coherent, competency-based Islamic Education evaluation framework that is responsive to the demands of Society 5.0. It is this gap that the present study seeks to fill.

Compared with the studies cited above, the novelty of this research lies in three respects. First, in terms of the scope of integration, this study constructs the first framework as far as the systematic search described in the Method section can establish that explicitly reconstructs three elements at once,

namely authentic assessment, digital literacy, and artificial intelligence, into a single, coherent, competency-based Islamic Education evaluation model, rather than discussing the three separately as found in studies on authentic assessment (Achmad et al., 2022; Hanafi et al., 2025), digital literacy (Pratiwi et al., 2024; Lisyawati et al., 2023), or AI (M. R. Fauzi et al., 2025; Muchlis, 2025) in PAI. Second, in terms of the positioning of digital literacy, this study assigns digital literacy a dual role not found in comparable studies: it serves simultaneously as the medium through which evaluation is carried out and as a competency being evaluated in its own right, namely tabayyun-based Islamic media literacy in contrast to international studies on digital-competency evaluation, which generally treat digital literacy merely as an object to be measured using generic instruments such as DigCompEdu, without any religious-value content (López-Núñez et al., 2024). Third, in terms of the position assigned to AI, the proposed model explicitly positions AI as a complementary rather than a substitutive element, with the PAI teacher functioning as the principal validator of algorithmic output a firmer ethical boundary than that found in the general literature on AI integration in educational assessment, which tends to stop at identifying efficiency benefits without asserting any limit on technology's authority over the spiritual domain (Fu et al., 2025; Garzón et al., 2025; Halkiopoulou & Gkintoni, 2024). These three elements of novelty are synthesized into the Integrated Islamic Education Competency Evaluation Model (EKPT), which constitutes the principal contribution of this study and is positioned in more detail against comparable frameworks in the Findings and Discussion section. Based on the discussion above, this study aims to: (1) identify the problems inherent in conventional Islamic education learning evaluation; (2) reconstruct the concept of competency-based Islamic Education evaluation that integrates authentic assessment, digital literacy, and artificial intelligence; and (3) formulate a conceptual model together with its practical implications for teachers, schools, and PAI curriculum development in the Society 5.0 era.

To achieve these aims, the study is guided by three research questions. First, what are the fundamental problems in the practice of conventional Islamic Education learning evaluation? Second, how can the concept of competency-based Islamic education evaluation be reconstructed through the integration of authentic assessment, digital literacy, and artificial intelligence? Third, how can the resulting conceptual model for the reconstruction of Islamic education evaluation be depicted, and what are its implications for teachers, schools, and Islamic education curriculum development in the Society 5.0 era?

2. LITERATURE REVIEW

2.1 Islamic Education Learning Evaluation

Conceptually, learning evaluation needs to be distinguished from mere measurement or assessment. Measurement is quantitative in nature, consisting of numbers obtained from a measuring instrument, whereas evaluation and assessment are qualitative in nature because they involve making judgments about the meaning of those numbers (Arifin, 2016). Learning evaluation is defined as a systematic process for determining the effectiveness and efficiency of a learning system, encompassing objectives, materials, methods, media, and learning resources, as well as the assessment system itself

(Arifin, 2016). The principles of sound evaluation include objectivity, validity, reliability, comprehensiveness, and continuity (Sholahudin et al., 2025).

In the context of Islamic Education, evaluation has a distinctive character because it must address three domains simultaneously and in an integrated manner: cognitive (religious knowledge), affective (attitudes and the internalization of values, or *qolbiyah*), and psychomotor (the practice of worship and moral conduct in daily life) (Sholahudin et al., 2025; Siregar et al., 2025). Ideally, all three domains are evaluated on an ongoing basis rather than solely at the moment of a final examination, since the purpose of religious education is fundamentally the formation of a complete personality rather than the mere mastery of information (Siregar et al., 2025). Unfortunately, a number of studies show that Islamic Education evaluation practice in the field still leans toward the cognitive domain through written tests, while the affective and psychomotor domains which are in fact the core of religious education's success often escape systematic attention (Abdillah, 2016; Achmad et al., 2022).

Attention to this distinctive character of Islamic Education evaluation has also drawn notice at the international level, so that the problem outlined above cannot be regarded as merely a local issue. A self-assessment instrument for Islamic-education teachers' competence developed and validated in Malaysia, for example, confirms that evaluation in Islamic education demands indicators different from those used in general subjects, encompassing the dimensions of skill, knowledge, personality, and environment all at once (Masuwai et al., 2024). Another study in Indonesia found that Islamic Education learning models that emphasize only cognitive-theoretical aspects risk leaving a spiritual void in learners, thereby requiring a model specifically designed to address the internalization of values in a measurable way (Jenuri et al., 2025). These findings reinforce the argument that the dominance of the cognitive domain in PAI evaluation is a challenge also faced by Islamic religious education across various national contexts, rather than merely a structural problem of education in Indonesia.

2.2 Competency-Based Evaluation

Competency-based evaluation shifts the orientation of assessment away from merely measuring mastery of material and toward measuring learners' ability to apply knowledge, attitudes, and skills in an integrated way to solve real problems. In Indonesia, this orientation has been formalized through the Merdeka Curriculum policy, which places phase-based Learning Outcomes as the reference for the competencies learners must achieve, replacing the more rigid, grade-by-grade Core Competencies and Basic Competencies of the previous curriculum (Kemendikbudristek, 2022). Through this approach, mastery of learning is no longer measured by a single quantitative score such as the Minimum Mastery Criterion, but rather through diagnostic, formative, and summative assessment oriented toward the gradual achievement of learning objectives (Kemendikbudristek, 2022).

The Merdeka Belajar (Freedom to Learn) policy has also driven a redesign of Islamic Education learning evaluation. Arifin et al. (2021) affirm that this policy requires Islamic Education teachers to design evaluation instruments that no longer stop at rote memorization of *dalil*, but instead assess

learners' ability to apply Islamic values in real-life contexts. Even so, this paradigm shift still leaves practical challenges. Research shows that project- and observation-based assessment within the Merdeka Curriculum for the Islamic education subject has only recently begun to be adopted, while most teachers report that they still need further training to design valid and applicable competency instruments (Achmad et al., 2022). Competency-based evaluation fundamentally places the question of what learners can do with what they know above the question of how much they have memorized a shift that is highly relevant for Islamic Education, since the practice of Islamic values matters far more than the mere ability to recall dalil.

2.3 Authentic Assessment

Authentic assessment emerged from criticism of standardized tests' limitations in capturing learners' deep understanding. Wiggins and McTighe (2005), through their Understanding by Design framework, emphasize that genuine understanding is demonstrated when learners are able to explain, apply, and transfer knowledge to new situations, rather than merely recall facts. The backward-design principle they developed requires that evidence of understanding be designed first, based on learning objectives, before learning activities themselves are devised, so that assessment reaches authentic tasks reflecting real-world performance. Subsequent developments affirm that authentic assessment is not merely a matter of technical method but is also connected to learners' work readiness and well-being after their education, since authentic tasks cultivate competencies relevant to real-world demands (McArthur, 2023).

In Indonesia, authentic assessment has been formalized as an approach to evaluating learning outcomes that measures readiness, process, and results as a whole through various techniques such as performance tasks, projects, portfolios, and self-assessment (Kunandar, 2018). In the context of Islamic Education, authentic assessment carries special relevance because it is able to address all three evaluation domains cognitive, affective, and psychomotor in an integrated manner, while simultaneously strengthening the formation of religious character and encouraging the practice of Islamic values in daily life (Firda Maghfiratus S. et al., 2026). Several techniques already applied by Islamic Education teachers include observation of worship practice, peer assessment, reflective journals, religious social projects, and portfolios of religious practice (Achmad et al., 2022; Hanafi et al., 2025). A holistic assessment model based on multiple intelligences has even been proposed so that authentic assessment in Islamic Education accommodates the diversity of learners' intelligences, rather than only the logical-verbal intelligence typically measured by written tests (Suryatina et al., 2026).

2.4 Digital Literacy

Digital literacy was first popularized by Gilster (1997) as the ability to critically understand and use information from a variety of digital sources. As the concept developed, it expanded to encompass not only the technical skill of operating devices but also the ability to evaluate the credibility of information, communicate ethically in digital space, and create content responsibly (Hamidi &

Chavoshi, 2022). Contemporary frameworks even position digital literacy as a competency that can be measured systematically through its own dedicated assessment instruments, rather than merely as a supporting skill (Hamidi & Chavoshi, 2022).

In Islamic education, digital literacy acquires an additional, theological layer of meaning. A. Fauzi (2022) shows that the principle of *tabayyun* the exhortation to verify the truth of information before believing or disseminating it can serve as a theological foundation for digital-literacy competence, so that digital literacy in Islamic Education is not merely a matter of technical skill but also part of media-related *akhlak*. A number of studies have found that strengthening digital literacy in Islamic Education learning is able to improve learners' access to information, creativity, and critical-thinking ability (Pratiwi et al., 2024), although implementation still faces constraints of limited facilities and teachers' low digital competence (Lisyawati et al., 2023; Naufal, 2021). So far, the use of digital literacy in Islamic Education has functioned more as a learning medium than as an evaluation instrument a gap that needs to be bridged, given that digital platforms offer considerable opportunity to document digital portfolios, record worship practice through video, and manage formative-assessment databases more efficiently (Kurnia et al., 2019).

At the global level, a systematic review of instruments for evaluating educators' digital competence in higher education found that the DigCompEdu framework is the most dominant point of reference, yet its validity remains geographically uneven, with research concentrated in Europe and few instruments specifically validated for the Asian region (López-Nuñez et al., 2024). This kind of cross-context validation gap is one reason why digital literacy in Islamic education cannot simply adopt generic instruments from secular-Western educational contexts, but instead needs to be reformulated with Islamic-value content such as *tabayyun* (A. Fauzi, 2022).

2.5 Artificial Intelligence (AI) in Education

Artificial intelligence (AI) in education refers to the use of machine-learning algorithms to support the teaching and learning process, including the personalization of materials, learning analytics, and the automation of assessment. Systematic studies show that AI is able to provide adaptive feedback in real time, adjust the difficulty level of material to learners' ability, and improve the efficiency and consistency of assessment compared with manual grading, which is prone to subjective bias (Halkiopoulous & Gkintoni, 2024). The use of AI in assessment has moved beyond the mere automation of multiple-choice scoring toward essay-answer analysis, detection of conceptual-error patterns, and even personalized learning-intervention recommendations.

In Islamic education, research on AI integration remains at the stage of exploring opportunities and challenges. M. R. Fauzi et al. (2025) find that integrating AI into Islamic education learning evaluation has the potential to improve the efficiency, objectivity, and personalization of assessment through automated grading, learning-data analysis, and adaptive feedback, but that implementation is constrained by teachers' still-limited digital literacy, uneven infrastructure, and issues of ethics and

data privacy. Another study affirms that the use of AI in Islamic education must take its compatibility with *maqashid syariah* (the objectives of Islamic law) into account, so that technological efficiency does not come at the expense of the law's objectives, such as the preservation of intellect (*hifz al-'aql*), in the educational process (Akbar et al., 2025). This means that AI integration in Islamic education evaluation cannot be carried out in a value-neutral manner; it must instead be accompanied by an ethical framework that ensures AI output remains aligned with Islamic values while preserving the teacher's essential role as an educator of values that cannot be replaced by machines.

At the global scale, a meta-review of 126 systematic reviews on AI in education (AIED) shows that AI's impact is most extensively documented in the aspects of teaching, learning, and administration, while its application to affective-domain assessment and to religious-value-based education contexts remains a comparatively under-researched area (Fu et al., 2025). Another systematic review of 155 empirical studies likewise confirms the benefits of AI in terms of personalization and improved learning outcomes, while also cautioning about the risk of digital dependency and teacher resistance that need to be anticipated from the earliest stages of design (Garzón et al., 2025). Both of these large-scale reviews display a pattern consistent with the findings of M. R. Fauzi et al. (2025) in the Islamic education context, while also confirming that the issues of ethics and teachers' digital-literacy gaps are not merely challenges specific to Islamic education, but rather global challenges of AI integration in educational assessment that remain fully unresolved in any body of literature.

2.6 The Society 5.0 Era and Islamic Education

Society 5.0 is a concept first put forward by Japan through the Fifth Science and Technology Basic Plan approved by the Japanese Cabinet in January 2016, described as a human-centered society that aligns economic progress with the resolution of social problems through a system that deeply integrates cyber space and physical space (Cabinet Office, Government of Japan, 2016). The concept is positioned as the successor to hunter-gatherer society (1.0), agrarian society (2.0), industrial society (3.0), and information society (4.0), with the emphasis that technological progress—including AI, the Internet of Things, and big data—must remain oriented toward human well-being, rather than the reverse (Cabinet Office, Government of Japan, 2016).

For Islamic education, this era brings distinctive opportunities as well as challenges. On one hand, technology expands access to Islamic learning resources and facilitates the personalization of learning (Sari & Fitriana, 2022; Yusuf, 2024). On the other hand, the swift current of digitalization risks shallowing spiritual appreciation if it is not balanced by strengthened character education (Bahiyah, 2022; Maharani & Nursikin, 2025). Several studies affirm that Islamic education in the Society 5.0 era must not merely adapt technically to digital devices, but must instead place Islamic values including the concept of humankind as *khalifah* (steward), responsible for the use of technology as the ethical framework guiding that technology's use (Bahri, 2022; Amalia & Munif, 2023). The human-centered principle within Society 5.0 is thus, in fact, aligned with Islamic education own orientation, which consistently places the formation of the complete human being, or *insan kamil*, above merely technical

or material achievement. This point of convergence is one of the arguments for why the reconstruction of competency-based, authentic, digital, and AI-integrated Islamic education evaluation is relevant precisely within the framework of Society 5.0, rather than standing in opposition to it.

A systematic review of the Education 5.0 literature within the Society 5.0 landscape likewise confirms that the transition toward future competencies demands a realignment of learning objectives, processes, and, above all, evaluation an agenda that, at the level of macro policy, is consistent with the urgency of the Islamic education evaluation reconstruction proposed by this study, even though that review does not address the technical details of evaluation for a specific subject such as religious education (Shahidi Hamedani et al., 2024).

2.7 Synthesis and Mapping of Previous Studies

To clarify the relationship between prior research, its principal findings, and the gap on which this study is founded, Table 1 presents a synthesis of a number of key studies already discussed narratively in the preceding subsections. This mapping also confirms that the study's four core themes Islamic education evaluation, authentic assessment, digital literacy, and artificial intelligence have consistently been examined separately at both the national and international levels, thereby reinforcing the urgency of the integrated reconstruction proposed by this study.

Table 2.7.1. Synthesis and Mapping of Previous Studies

Researcher(s) (Year)	Focus & Method	Key Findings	Relevance/Gap to This Study
Abdillah (2016)	Islamic education evaluation at senior-high-school level; quantitative correlational	Islamic education evaluation is highly dependent on teacher competence; instruments remain cognitively centered	Provides the empirical basis for the conventional-evaluation problem; does not yet address digital/AI solutions
Sholahudin et al. (2025)	Islamic education evaluation based on Qur'anic	Basic cognitive domain, shallow affective domain, psychomotor domain	Reinforces the three-domain problem; does not offer an integrative model
Researcher(s) (Year)	Focus & Method	Key Findings	Relevance/Gap to This Study
	verses; content analysis	weak in higher-order thinking skills	

Achmad et al. (2022); Hanafi et al. (2025)	Authentic assessment in the Merdeka Curriculum & Islamic education; qualitative descriptive	Authentic techniques beginning to be applied but constrained by time, resources, and teacher training	Confirms authentic assessment's relevance for Islamic education, but not yet integrated with digital literacy/AI
Pratiwi et al. (2024); Lisyawati et al. (2023)	Digital literacy in Islamic education learning; qualitative descriptive	Digital literacy improves access, creativity, and critical thinking, but remains confined to a learning medium	Digital literacy positioned as a medium, not an evaluation instrument—the main gap filled by the EKPT Model
M. R. Fauzi et al. (2025); Muchlis (2025)	Opportunities and challenges of AI in Islamic education; conceptual study	AI has potential to improve efficiency & objectivity, but is constrained by teachers' digital literacy & ethics	Only at the stage of identifying opportunities; has not yet formulated an operational evaluation framework
Suryatina et al. (2026)	Holistic Islamic education assessment model based on multiple intelligences; conceptual	Accommodates the diversity of learners' intelligences in Islamic education assessment	Enriches the authentic-assessment dimension, but does not yet integrate digital literacy & AI
Halkiopoulos & Gkintoni (2024)	AI for personalization & adaptive assessment; international systematic review (general education)	AI is effective for real-time feedback & assessment efficiency	General education context without religious-spiritual value content
Fu et al. (2025)	Meta-review of 126 PRISMA-based AIED systematic reviews & bibliometric analysis (international)	AI's impact is substantial on teaching & administration; affective-domain assessment and religious-education contexts remain under-researched	Confirms the AI research gap in the affective/spiritual domain that the EKPT Model seeks to address

Researcher(s) (Year)	Focus & Method	Key Findings	Relevance/Gap to This Study
Garzón et al. (2025)	Systematic review of 155 empirical studies on AI in education (international)	Benefits of personalization & motivation; challenges of ethics, teacher resistance, digital dependency	Provides the basis for the argument that the teacher serves as principal validator in the EKPT Model
López-Nuñez et al. (2024)	Systematic review of digital-competence evaluation (DigCompEdu) in higher education (international)	Instruments dominated by European models; little validation in the Asian region	Shows the absence of a digital-literacy instrument carrying Islamic-value content
Masuwai et al. (2024)	Validation of a Islamic education teacher self-assessment instrument, Malaysia; psychometric	Instrument content- and face-valid for Islamic education teacher competence, covering skill, knowledge, and personality	Focuses on teacher competency assessment, not learners; does not yet integrate digital/AI
Jenuri et al. (2025)	Islamic education learning model integrating Al-Ghazali's concepts, Indonesia; mixed-method experimental	Model effectively improves learners' spiritual understanding (N-Gain 0.85)	Focuses on a learning model, not an evaluation model; does not yet address digital/AI

3. METHOD

3.1 Research Design

This study employs a qualitative approach in the form of library research, that is, research whose data and discussion rest on the searching, critical reading, and synthesis of written sources, without direct field data collection. This approach was chosen because the study's aim is conceptual in nature namely, to reconstruct a competency-based Islamic education evaluation framework through the synthesis of ideas from various fields of study, including educational evaluation, authentic assessment, digital literacy, artificial intelligence, and curriculum policy, fields that have thus far developed separately.

To enhance the transparency and accountability of the search process, this study adapts part of the reporting procedure used in systematic reviews, without claiming to be a full systematic review or meta-analysis. This methodological position is known, within the typology of literature reviews, as a systematized review a study that adopts one or more elements of the systematic-review process, such as an explicit search strategy and documented selection criteria, while still retaining a narrative-interpretive character in the synthesis of its findings (Grant & Booth, 2009). The systematic elements adapted in this study include the determination of keywords and search databases, explicit inclusion-exclusion criteria, and a PRISMA-style document-selection flow (Page et al., 2021), as elaborated in the following subsections.

3.2 Literature Search Strategy and Procedure

The literature search was conducted across six databases and indexing engines—Google Scholar, Scopus, ERIC (Education Resources Information Center), DOAJ (Directory of Open Access Journals), Garuda/SINTA of the Ministry of Education, Culture, Research, and Technology, and Crossref—in order to achieve balanced coverage of both national and international publications. The search used a combination of Indonesian- and English-language keywords linked with Boolean operators, comprising: (1) Islamic education learning evaluation OR Islamic religious education assessment; (2) authentic assessment AND competency; (3) digital literacy AND Islamic education; (4) artificial intelligence AND learning evaluation; (5) “authentic assessment” AND “religious education”; and (6) Society 5.0 AND education.

The search was restricted to documents published between 2016 and 2026 to ensure currency, with the exception of a number of classic conceptual references that serve as the study's principal theoretical foundations namely, the original definition of digital literacy (Gilster, 1997) and the backward-design framework for authentic assessment (Wiggins & McTighe, 2005) which were retained on account of their status as seminal references that continue to be widely cited in contemporary literature.

3.3 Inclusion and Exclusion Criteria

Every document obtained from the initial search was screened using explicit inclusion and exclusion criteria, as presented in Table 2, in order to safeguard the credibility and relevance of the sources synthesized.

Table 3.3.1. Literature Inclusion and Exclusion Criteria

Criterion	Requirements
Inclusion	1. Journal articles accredited by Sinta (national) or reputable international journals indexed in Scopus, Web of Science, or DOAJ. 2. Published between 2016–2026, or constituting a classic conceptual reference serving as a principal theoretical foundation. 3. Addressing at least one of four core themes: Islamic education evaluation/assessment, authentic assessment/competency-based evaluation, digital literacy in education, or artificial intelligence in education/assessment. 4. Full text accessible openly or through institutional access. 5. Written in Indonesian or English.
Exclusion	1. Non-peer-reviewed publications, including opinion pages, popular blogs, and un-reviewed preprints. 2. Articles addressing learning evaluation in subjects other than PAI with no conceptual connection to the four core themes. 3. Duplicate reports of the same study appearing in more than one publication outlet. 4. Published in journals identified as predatory or not indexed in a reputable database. 5. Full text inaccessible by the data-collection deadline.

3.4 Literature Selection Procedure

The literature-selection procedure adapts the PRISMA flow (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) developed by Page et al. (2021), which is commonly used to document the process of identification, screening, eligibility assessment, and the final determination of documents included in a synthesis (Fu et al., 2025; López-Nuñez et al., 2024). This adaptation does not make the present study a systematic review in the full sense; rather, it merely borrows PRISMA's reporting framework for the sake of transparency, consistent with the systematized-review character already explained above (Grant & Booth, 2009). Figure 1 presents the literature-selection flow followed in this study.

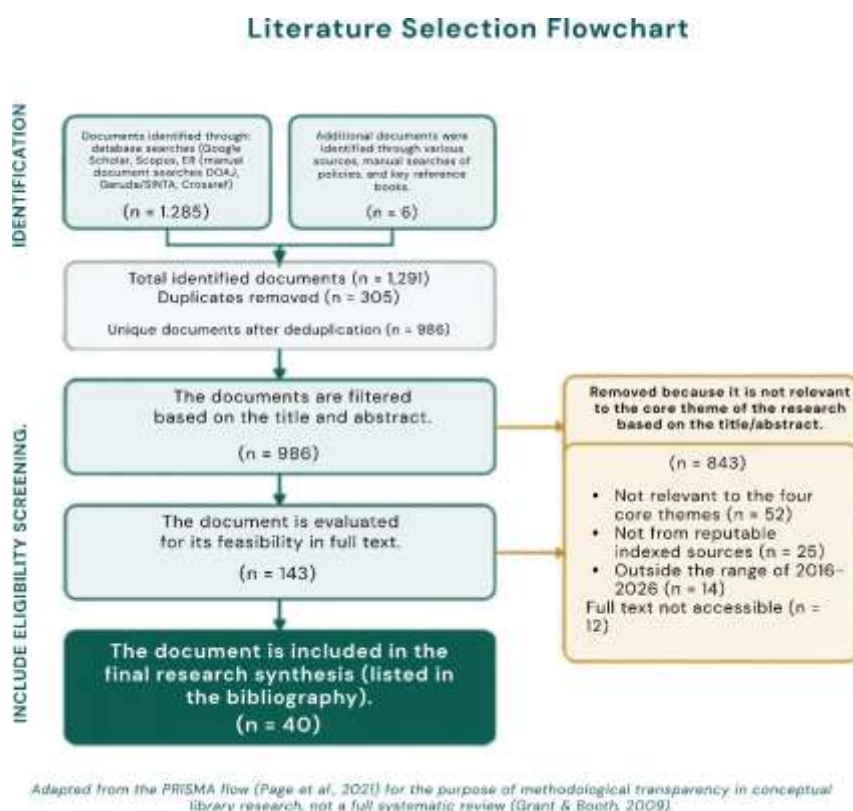


Figure 3.4.1. Literature Selection Flow Diagram

As shown in Figure 1, the database search yielded 1,285 documents, plus 6 documents from other sources namely, a manual search of official policy documents and key reference books for a total of 1,291 documents identified. After the removal of duplicates (305 documents), 986 documents were screened by title and abstract, of which 843 were excluded as irrelevant to the core themes. A total of 143 documents were then assessed for eligibility in full text, and 103 documents were excluded, broken down as follows: 52 documents did not explicitly address competency-based Islamic education evaluation or any one of its three integrative elements; 25 documents originated from publishers not indexed in a reputable database; 14 documents were published outside the specified year range without qualifying as classic references; and 12 documents had full texts that could not be accessed. This process yielded 40 documents included in the final synthesis and listed in this study's reference list.

4. DATA SOURCES

This study's data sources are divided into primary and secondary data. Primary data consist of national and international scholarly journal articles addressing Islamic education learning evaluation, authentic assessment, digital literacy, artificial intelligence in education, and Islamic education in the Society 5.0 era, with emphasis placed on publications from the past ten years (2016 to 2026) so that the study reflects the most recent developments. Secondary data consist of learning-evaluation textbooks,

Merdeka Curriculum policy documents, and official documents on the concept of Society 5.0, used as a conceptual reference framework.

4.1 Data Collection Technique

Data collection was carried out through documentary-study techniques, that is, the systematic searching and recording of the 40 documents that passed selection, as described in the preceding subsection. Each document that passed selection was read in depth for extraction into a synthesis worksheet containing the researcher's identity, focus and method, key findings, and their relevance to the study's themes, as summarized in Table 1.

4.2 Data Analysis Technique

The data were analyzed using content-analysis techniques, a procedure used to identify concepts, principles, and patterns of argument within the sources collected, and then to synthesize them into a new, coherent framework. Operationally, the data analysis followed three stages. First, data reduction, that is, sorting relevant ideas from the sources collected and setting aside information not relevant to the study's focus. Second, data display, that is, organizing these ideas into thematic categories comprising the problems of conventional evaluation, competency, authentic assessment, digital literacy, and AI. Third, conclusion drawing, that is, synthesizing these categories into a complete reconstruction of the concept of Islamic education evaluation and into the conceptual model proposed in this study, including the formulation of operational indicators and the visualization of the model, as presented in the Findings and Discussion section.

5. FINDINGS AND DISCUSSION

5.1 Problems of Conventional PAI Evaluation

The synthesis of the various studies reviewed reveals at least four fundamental problems in conventional Islamic education evaluation practice. First, the dominance of the cognitive domain. PAI teachers tend to rely on multiple-choice or essay written tests as their principal instrument, so that the affective and psychomotor domains which are substantially far more important in religious education are often neglected or assessed carelessly (Abdillah, 2016; Sholahudin et al., 2025). Second, an orientation toward rote memorization rather than applied understanding. Evaluation that stops at the ability to recall dalil and definitions fails to capture the extent to which learners genuinely internalize and practice Islamic values in daily life a gap between constructivist pedagogical ideals and a classroom reality still dominated by lecturing and memorization.

Third, teachers' limited competence and time for designing alternative instruments. Although teachers recognize the importance of authentic assessment such as portfolios and observation, limited time, administrative burdens, and insufficient training make its implementation inconsistent (Achmad et al., 2022). Fourth, a gap in technology utilization. The potential of digital literacy and artificial intelligence to support the objectivity and efficiency of Islamic education evaluation has not yet been harnessed systematically to any great extent, partly because of limited infrastructure and teachers' limited digital

literacy, and partly because of concerns over the ethical compatibility of such technology use with Islamic values (M. R. Fauzi et al., 2025). These four problems are mutually entangled: cognitively centered evaluation makes it difficult for teachers to measure holistic competency, while limited digital competence hinders teachers from making use of assessment alternatives that are, in fact, already available. It is this condition that underlies the need for a thoroughgoing reconstruction, rather than merely a partial improvement to a single aspect of evaluation.

This pattern of problems is not merely a local phenomenon. Recent studies in Malaysia as well as Indonesia likewise report similar symptoms in a somewhat different guise: a gap between theoretical-cognitive Islamic education learning models and learners' need for spiritual internalization (Jenuri et al., 2025), and a need for more comprehensive and validated instruments for assessing Islamic education teachers' competency (Masuwai et al., 2024). Yet, unlike these two studies, each of which focuses respectively on improving the learning model and the teacher-competency instrument, the present study specifically targets learner evaluation as the point of leverage, on the argument that improvements on the evaluation side will drive backward alignment on the learning and teacher-competency side—consistent with the logic of backward design introduced by Wiggins and McTighe (2005).

5.2 Reconstruction of Competency-Based PAI Evaluation

The reconstruction proposed by this study starts from a shift in perspective: Islamic education evaluation needs to be redirected toward answering the question of what learners can do and practice, rather than how much they have memorized. This shift is consistent with the spirit of the Learning Outcomes within the Merdeka Curriculum, which places cross-phase competency, rather than merely per-grade material, as the benchmark of success (Kemendikbudristek, 2022). Within this reconstruction framework, competency-based Islamic education evaluation is designed through three steps. First, deriving Islamic education Learning Outcomes into operational, observable competency indicators that address the cognitive, affective, and psychomotor domains at once, rather than indicators emphasizing only a single domain. Second, developing instruments that explicitly test the application of Islamic values in contextual situations, rather than the mere reproduction of dalil (Arifin et al., 2021). Third, establishing competency-attainment criteria that are qualitative-descriptive in nature rather than a single numerical score, in line with the Merdeka Curriculum's principle of having abandoned the Minimum Mastery Criterion as the sole benchmark (Kemendikbudristek, 2022).

This competency-based reconstruction requires Islamic education teachers to change their view of the function of evaluation itself—from a mere tool for generating report-card scores into an instrument of continuous feedback that helps learners recognize the extent to which Islamic values have become part of the way they think, feel, and act.

5.3 Integration of Authentic Assessment

Authentic assessment is the operational approach best suited to realizing competency-based Islamic education evaluation, because it is inherently designed to measure real performance rather than mere knowledge reproduced on a test (Wiggins & McTighe, 2005). The integration of authentic assessment into Islamic education evaluation can be realized through several concrete forms: observation of worship practice for example, the quality of recitation and devotion during salat, the practice of wudu, or Qur'anic recitation for the psychomotor domain; reflective journals and self-assessment for tracing the process of value internalization in the affective domain; and religious social projects, such as charitable activities, tolerance campaigns, or community-service work grounded in Islamic values, which require learners to integrate knowledge, attitudes, and skills at once (Achmad et al., 2022; Hanafi et al., 2025).

A multiple-intelligences approach further enriches this integration by ensuring that learners who excel in kinesthetic, musical, or interpersonal intelligence have an appropriate assessment channel available to them, rather than being invariably forced into a written-test format that favors logical-verbal intelligence (Suryatina et al., 2026). Authentic assessment of this kind also answers the criticism that educational assessment should not stop at merely administrative concerns, but should also help prepare learners' life readiness and well-being beyond their formal education (McArthur, 2023) an orientation consistent with PAI's ultimate goal of forming individuals capable of practicing Islamic values throughout their lives, not only while they are at school.

5.4 Integration of Digital Literacy

Digital literacy functions as the medium that enables authentic assessment to be carried out in a more practical and well-documented manner. Digital portfolios, for instance, allow teachers to document the development of learners' worship practice and reflective work over time without being constrained by physical storage space. Video recordings of Qur'anic-recitation practice or worship simulations can be re-assessed more objectively than a one-off judgment relying solely on the teacher's memory. Digital learning platforms likewise facilitate the ongoing management of formative assessment, consistent with the principle of continuity demanded by PAI evaluation (Kurnia et al., 2019).

Even so, the integration of digital literacy into Islamic education evaluation must not be reduced to the mere technical mastery of applications. Drawing on the principle of *tabayyun*, digital literacy in Islamic education should also instill the habit of verifying the truth of information and adopting a critical attitude toward digital content before accepting it (A. Fauzi, 2022) a competency that is simultaneously part of what is being evaluated, not merely a tool for evaluation. In other words, digital literacy within this reconstruction framework carries a dual role: it serves as the medium through which evaluation is implemented, via the digitalization of instruments and documentation, while also constituting a competency evaluated in its own right, namely the ability to engage with digital media in an Islamic manner. This dual role requires Islamic education teachers to possess adequate digital competence a

challenge that still needs to be overcome, given that teachers' low digital competence has repeatedly emerged as a principal constraint in previous studies (Lisyawati et al., 2023; Naufal, 2021).

This issue also serves as a response to the validation gap in digital-literacy instruments, which are concentrated in Western contexts, as found in the international systematic review discussed earlier (López-Nuñez et al., 2024). This means that the EKPT version of digital literacy cannot simply adopt generic instruments such as DigCompEdu wholesale, but instead needs to be developed and specifically validated within the context of Islamic values in future research.

5.6 Integration of AI in PAI Evaluation

Artificial intelligence functions as an accelerator that strengthens the objectivity, efficiency, and personalization of implementing the competency-based evaluation and authentic assessment described above. At least three AI functions can be integrated into Islamic education evaluation. First, the automation of cognitive-domain grading for example, the analysis of short essay answers concerning understanding of fiqh or aqidah concepts so that teachers can redirect time that has thus far been consumed by manual correction toward observation of the affective and psychomotor domains, which in fact require the teacher's human presence far more (Halkiopoulous & Gkintoni, 2024). Second, learning analytics that map patterns in learners' competency attainment over time, enabling teachers to identify learners who require more intensive affective mentoring before attitudinal or moral (akhlak) problems develop further. Third, personalized adaptive feedback for example, recommendations for remedial material on a particular point of adab for learners displaying a specific gap in understanding (M. R. Fauzi et al., 2025).

Even so, the integration of AI into Islamic education evaluation must be kept from exceeding the bounds of its proper authority. Assessment of the affective and psychomotor domains, which are laden with spiritual dimensions such as devotion (khusyuk), sincerity of intention, or depth of value internalization, still requires qualitative judgment by the teacher as a human being who understands the learner's personal context something that has not yet been, and may never be, fully replaceable by an algorithm. This principle of caution is consistent with the requirement that the use of AI in Islamic education always be aligned with maqashid syariah, so that technological efficiency does not come at the expense of the very essence of value education (Akbar et al., 2025). AI's position within this reconstruction of evaluation is therefore complementary in nature strengthening efficiency and objectivity in quantitatively measurable domains rather than substitutive of the teacher's role in assessing learners' spiritual dimension.

This position is consistent with the general recommendation found in the AI-in-education literature, which likewise emphasizes the need for human oversight of algorithmic output (Garzón et al., 2025); however, this study places a more specific emphasis: the Islamic education teacher's oversight functions not only in a technical-pedagogical capacity, as commonly discussed in the general literature, but also in a theological capacity—namely, ensuring that AI output does not contradict Islamic values

in accordance with the principle of maqashid syariah (Akbar et al., 2025). It is this dual emphasis that distinguishes AI's position within the EKPT Model from the human-in-the-loop framework found in the general AI-in-education literature, which tends to stop at purely technical-ethical considerations.

5.5 Conceptual Model of the Evaluation Reconstruction

Based on the synthesis of the problems, the competency-based reconstruction, and the integration of the three elements above, this study proposes a conceptual model called the Integrated Islamic education Competency Evaluation Model, hereafter abbreviated EKPT. This model is composed of four mutually integrated components and one final orientation, as elaborated in Table 3.

Table 5.5.1. Components of the Integrated Islamic education Competency Evaluation Model

Component	Role in the Model	Principal Function
Competency	Foundation	Establishes the cognitive, affective, psychomotor, and spiritual outcomes to be evaluated, derived from Islamic education Learning Outcomes
Authentic Assessment	Approach	Translates competency outcomes into real performance tasks: observation of worship, portfolios, religious social projects, self-assessment
Digital Literacy	Medium	Digitalizes instruments, documentation, and assessment-data management, while also constituting a competency being evaluated (Islamic media literacy)
Artificial Intelligence (AI)	Accelerator	Automates cognitive-domain grading, maps attainment patterns through learning analytics, and provides adaptive feedback
Final Orientation	Goal	The formation of learners' holistic competency and Islamic character, adaptive to the demands of Society 5.0

These four components operate cyclically, rather than in a single, one-way linear fashion, as visualized in Figure 2. Once competencies have been established, authentic assessment translates them into performance tasks. These performance tasks are carried out and documented through the medium of digital literacy. The data collected are then processed by artificial intelligence to generate an attainment map and feedback. This feedback, in turn, becomes the basis for refining the formulation of competencies and authentic-assessment instruments in the next cycle.

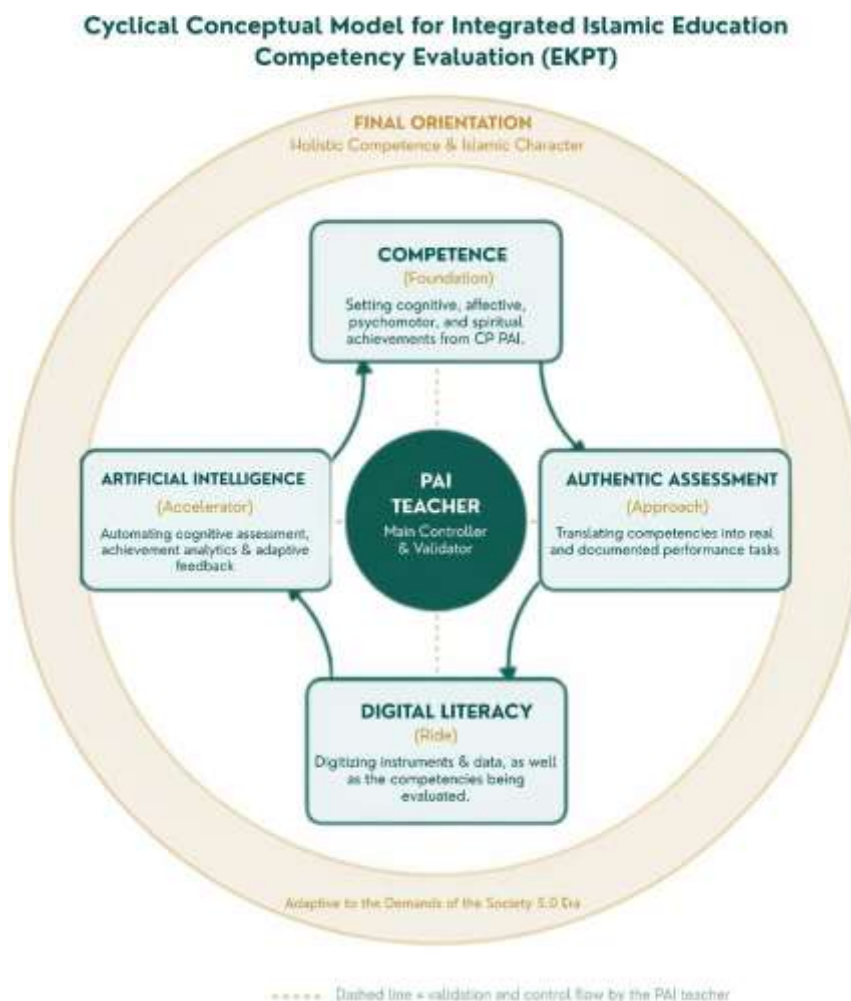


Figure 5.5.1. Cyclical Conceptual Model of Integrated Islamic education Competency Evaluation (EKPT)

As depicted in Figure 2, the PAI teacher acts as the principal controller at every stage of this cycle, particularly in validating AI output and providing qualitative judgment on the spiritual dimension that is not fully measurable algorithmically represented by the dashed lines connecting the Islamic education teacher to all four components. The EKPT Model thus positions technology as an instrument that strengthens, rather than replaces, the essence of Islamic education evaluation as a means of forming the insan kamil (the complete human being).

5.7 Operational Indicators of the EKPT Model

So that the EKPT Model does not remain at a purely conceptual and abstract level, each component needs to be translated into observable and measurable operational indicators. Table 4 formulates these operational indicators for the five measurement dimensions cognitive, affective, psychomotor, digital

literacy, and artificial intelligence—complete with examples of instruments or techniques relevant to each indicator.

Table 5.7.1. Operational Indicators of the EKPT Model by Component

Component	Example Operational Indicators	Example Instruments/Techniques
Cognitive	• Explains fiqh/aqidah/akhlak concepts in one's own words • Relates dalil to contemporary contextual issues • Analyzes simple ethical-religious case studies	Structured written tests, analytical essays, adaptive online quizzes (AI-assisted grading)
Affective	• Demonstrates consistent honesty and discipline in daily interactions • Expresses reflections on Islamic values in a personal journal • Demonstrates empathy and tolerance in social interaction	Digital reflective journals, self- and peer-assessment, periodic attitude observation
Psychomotor	• Quality of recitation and devotion during salat/tilawah • Consistency of wudu practice according to prescribed guidance • Participation in religious social projects	Structured direct observation/video recording with rubric, portfolio of worship practice, project assessment
Digital Literacy	• Verifies the truth of religious information before sharing it (tabayyun) • Compiles and manages a digital portfolio of Islamic education learning • Demonstrates Islamic social-media ethics (digital adab)	Religious-news fact-checking tasks, digital portfolio, digital-ethics rubric
Artificial Intelligence (AI)	• Degree of agreement between AI-generated automatic scores and teacher grading in the cognitive domain • Use of AI adaptive feedback for independent learning improvement • Teacher's ability to critically validate/correct AI output	Learning-analytics dashboard, adaptive-feedback usage logs, teacher validation sheet for AI recommendations

The formulation of indicators in Table 4 is illustrative in nature and needs to be further adapted to the Islamic education Learning Outcomes for each phase as established by Kemendikbudristek (2022), given that the scope and depth of Islamic education material differ across educational levels. Even so,

this indicator framework provides a more concrete starting point than previous Islamic education evaluation models, which generally stop at naming the cognitive, affective, and psychomotor domains without detailing their measurable indicators (Nafiati, 2021; Sholahudin et al., 2025).

5.8 Position and Novelty of the EKPT Model Compared with Similar Frameworks

To reaffirm the novelty already touched upon in the introduction, Table 5 positions the EKPT Model directly against a number of the most relevant comparable frameworks or models from prior studies. This comparison shows that each comparable framework, whether drawn from the Islamic education literature or from the broader international education literature, tends to address only one or two integrative elements at a time, and that almost none explicitly place Islamic education spiritual-affective domain, Islamic-value-laden digital literacy, and artificial intelligence within one and the same cyclical framework.

Table 5.8.1. Comparison of the EKPT Model with Similar Evaluation Frameworks

Framework/Model (Source)	Principal Focus	Elements Integrated	Limitations Compared with EKPT
Holistic Multiple-Intelligence Assessment Model for Islamic education (Suryatina et al., 2026)	Diversity of intelligences in authentic Islamic education assessment	Competency + Authentic Assessment	Does not yet integrate digital literacy or AI as medium/accelerator
Al-Ghazali-Concept-Based Islamic education Learning Model (Jenuri et al., 2025)	Spiritual internalization through a learning model	Spiritual (affective) competency	Focuses on learning design, not an evaluation model; does not address digital/AI
AI Framework for Personalization & Adaptive Assessment (Halkiopoulous & Gkintoni, 2024)	Efficiency & personalization of assessment based on cognitive neuropsychology	Artificial Intelligence	General education context without religious-spiritual value content; teacher not explicitly positioned as ethical validator

DigCompEdu Digital-Competence Evaluation Instrument (López-Nuñez et al., 2024)	Measurement of educators' digital competence	Digital Literacy	Generic, cross-subject instrument; not yet validated for the context of Islamic values
Islamic education Teacher Competency Self-Assessment Instrument (Masuwai et al., 2024)	Content and face validity of the Islamic education teacher competency instrument	Competency (teacher)	Target of assessment is the teacher, not the learner; does not yet integrate digital/AI
EKPT Model (this study)	Evaluation of Islamic education learners' competency addressing the cognitive, affective, and psychomotor domains simultaneously	Competency + Authentic Assessment + Digital Literacy + Artificial Intelligence (four elements integrated, cyclical)	—

This comparison also reveals a pattern worth critically examining: internationally scaled frameworks (Halkiopoulou & Gkintoni, 2024; López-Nuñez et al., 2024) generally excel in methodological maturity and the scale of empirical testing, yet are weak in sensitivity to religious-value contexts; conversely, frameworks grounded in Islamic values (Jenuri et al., 2025; Masuwai et al., 2024; Suryatina et al., 2026) excel in value sensitivity, yet have not optimally harnessed the potential of digital technology and AI. The EKPT Model is positioned to bridge these two poles. Even so, this study acknowledges that this bridging position is, as yet, proposed only at the conceptual level and has not been empirically tested in the way that several of the comparator frameworks above have been—a limitation elaborated further in the Limitations of the Study section.

6. IMPLICATION

6.1 Implications for Teachers

The reconstruction of competency-based Islamic education evaluation requires Islamic education teachers to develop three new competencies simultaneously. First, authentic-assessment competency, namely the ability to design valid and applicable performance instruments such as worship-observation rubrics, portfolio formats, and religious-social-project criteria rather than merely recycling multiple-choice questions year after year. Second, pedagogical digital literacy, namely the ability to use digital platforms to systematically document and manage evidence of learners' learning. Third, critical AI literacy, namely the ability to understand the basic workings and limitations of AI in assessment, so

that teachers can make critical use of AI recommendations rather than accepting them uncritically, particularly for matters concerning the assessment of learners' affective and spiritual domains. Without strengthening these three competencies, no matter how well the evaluation model is formulated, it will remain confined to the conceptual level.

6.2 Implications for Schools

At the level of the educational institution, this reconstruction has implications for the need to provide adequate digital infrastructure including devices, networks, and learning platforms as well as internal policies that support the consistent implementation of authentic assessment, for example through the specific allocation of time for observation and performance assessment, rather than merely adding to teachers' administrative burden outside teaching hours. Schools also need to facilitate ongoing training for Islamic education teachers on designing authentic-assessment instruments and on the ethical use of digital platforms and AI, given that limited training has repeatedly emerged as a principal constraint in previous studies (Achmad et al., 2022). In addition, schools need to formulate clear learner-data governance policies when adopting AI-based assessment systems, in order to safeguard data privacy and security while also building parents' trust in the application of such technology.

6.3 Implications for Curriculum Development

For curriculum developers, this reconstruction presupposes the need for explicit integration between Islamic education Learning Outcomes and indicators of digital literacy and AI-use ethics, rather than treating these as cross-subject competencies separate from Islamic education. Future Islamic education curriculum documents should ideally contain technical guidance on forms of authentic assessment suited to the characteristics of Islamic education material at each phase, examples of digital instruments that teachers can adapt, and ethical guidelines for the use of AI in assessment consistent with maqashid syariah (Akbar et al., 2025). In this way, the curriculum would not only direct what is taught, but also explicitly guide how that competency is evaluated in a manner relevant to learners of the Society 5.0 generation, without sacrificing the Islamic spirit of religious education.

7. LIMITATIONS OF THE STUDY

7.1 Limitations of the Literature Review

As a library research study, this research is not free from the potential for selection bias and language bias. Although the search procedure has been adapted following the PRISMA framework (Page et al., 2021) and the inclusion-exclusion criteria have been explicitly established (Table 2), the process of reducing 1,291 identified documents down to the 40 documents included still involved the interpretive judgment of a single researcher in determining thematic relevance, so that the possibility cannot be ruled out that relevant documents were overlooked or, conversely, that undue emphasis was placed on documents that happened to be more easily accessible (accessibility bias). The restriction to literature written in Indonesian and English also risks overlooking relevant studies published in other languages, such as Arabic, Malay, Urdu, or Turkish, which are, notably, important languages in global Islamic-education discourse. Publication bias is also a potential concern, since this study reaches only

documents that have been published openly or through institutional access, meaning that studies in the form of theses, dissertations, or internal policy reports not widely published (grey literature) are not fully captured, even though such documents may potentially contain more contextually grounded field data.

7.2 Limitations of the Data Sources

The second limitation rests on the nature of the data sources themselves. This study's primary data are drawn entirely from documents that have already been published, so that implementation challenges that are local, contextual, and not yet documented in the scholarly literature for example, digital-infrastructure constraints in peripheral schools or cultural resistance to the use of AI in particular religious contexts may not yet be fully captured. In addition, the composition of the data sources reveals an imbalance between national literature, which discusses Islamic education in depth but generally has not yet addressed digital-AI integration in technical terms, and international literature, which excels in the methodological maturity of its treatment of AI and digital literacy but shows little sensitivity to the context of Islamic religious education, as shown in Table 5. This imbalance makes it difficult to draw conclusions that are truly balanced between these two literary traditions, and the resulting synthesis remains potentially inclined toward one side or the other, depending on the researcher's interpretive emphasis.

7.3 Limitations of Model Implementation

The third, and most fundamental, limitation is that the EKPT Model is entirely conceptual in nature and has not been empirically tested at any educational institution, so that its validity, reliability, and practicality in the field cannot yet be confirmed. Several implementation challenges that can be anticipated, although not yet empirically verified in this study, include: disparities in digital-infrastructure readiness across Indonesia's highly diverse regions and educational levels; variation in Islamic education teachers' digital competence and AI literacy, which may not be as uniform as the model assumes; potential cultural or theological resistance to involving AI in assessing aspects considered sacred, such as devotion (*khushyuk*) in worship; and the absence of technical guidelines or a policy umbrella that explicitly governs learner-data management when AI-based assessment is applied within a religious-education context. The claimed strengths of the EKPT Model in this study must therefore be understood as a theoretical proposition requiring further verification through research and development, case studies, or limited trials at particular educational institutions, as recommended in the Conclusion section.

8. CONCLUSION

This study concludes that conventional Islamic education learning evaluation, still dominated by cognitively oriented written tests, is no longer adequate to meet the demands of the holistic competency championed by the Merdeka Curriculum or the challenges of Society 5.0. The reconstruction of competency-based Islamic education evaluation that integrates authentic assessment, digital literacy, and artificial intelligence synthesized into the Integrated Islamic education Competency Evaluation

Model, or EKPT offers an alternative framework that enables evaluation to address the cognitive, affective, and psychomotor domains simultaneously, objectively, and sustainably, while still positioning the teacher as the principal controller over the spiritual dimension that is not fully measurable algorithmically.

This study makes a conceptual contribution in the form of an integrative framework that links four fields of study that have thus far developed separately competency evaluation, authentic assessment, digital literacy, and artificial intelligence into a single, coherent Islamic education evaluation model that can serve as an initial reference for developing Islamic education assessment policy at educational institutions as well as for curriculum development on a wider scale. As elaborated in the Limitations of the Study section, this contribution is conceptual in nature and needs to be followed up empirically. Future research is therefore recommended to empirically test the implementation of the EKPT Model through research and development or case studies at particular educational institutions, while further exploring the ethical framework for using AI in Islamic education assessment so that its application remains consistent with Islamic values. For policymakers, this study recommends the development of technical guidelines for competency-based Islamic education assessment that explicitly accommodate authentic assessment and the use of digital technology and AI, accompanied by an ongoing teacher capacity-building program as a prerequisite for successful implementation.

REFERENCES

- Abdillah, H. T. (2015). Pengaruh kompetensi profesional dan pedagogik guru PAI terhadap evaluasi pembelajaran PAI di SMA [The influence of professional and pedagogical competence of PAI teachers on PAI learning evaluation in senior high school]. *Jurnal Pendidikan Ilmu Sosial*, 24(2). <https://doi.org/10.17509/jpis.v24i2.1450>
- Achmad, G. H., Ratnasari, D., Amin, A., Yuliani, E., & Liandara, N. (2022). Penilaian autentik pada kurikulum merdeka belajar dalam pembelajaran pendidikan agama Islam di Sekolah Dasar [Authentic assessment in the Merdeka Belajar curriculum in Islamic religious education learning in elementary school]. *Edukatif: Jurnal Ilmu Pendidikan*, 4(4), 5685–5699. <https://doi.org/10.31004/edukatif.v4i4.3280>
- Akbar, A., Malarangan, H., & Nurkhaerah, S. (2025). Etika Islam terhadap kecerdasan buatan: Kajian maqashid syariah dalam implementasi AI di lembaga pendidikan Islam [Islamic ethics toward artificial intelligence : A maqashid syariah study on AI implementation in Islamic educational institutions]. *Prosiding Kajian Islam dan Integrasi Ilmu di Era Society 5.0 (KIIES 5.0)*, 4, 87–92. <https://doi.org/10.46870/jstain.v7i2.1927>
- Amalia, N. F., & Munif, M. V. M. (2023). Tantangan dan upaya pendidikan dalam menghadapi era Society 5.0 [Challenges and educational efforts in facing the Society 5.0 era]. *MAANA: Jurnal Pendidikan Islam Anak Usia Dini*, 1(2), 1–13. <https://doi.org/10.52166/mjpiaud.v2i1.4741>
- Arifin, S., Abidin, N., & Al Anshori, F. (2021). Kebijakan Merdeka Belajar dan implikasinya

- terhadap pengembangan desain evaluasi pembelajaran Pendidikan Agama Islam [The Merdeka Belajar policy and its implications for the development of Islamic Religious Education learning-evaluation design]. *Dirasat: Jurnal Manajemen dan Pendidikan Islam*, 7(1), 65–78. <https://doi.org/10.26594/dirasat.v7i1.2394>
- Arifin, Z. (2016). *Evaluasi pembelajaran: Prinsip, teknik, dan prosedur* [Learning evaluation: Principles, techniques, and procedures] (8th ed.). Remaja Rosdakarya.
- Bahiyah, U. (2022). Urgensi pendidikan karakter dalam pendidikan agama Islam di era Society 5.0: Studi pendekatan filosofis [The urgency of character education in Islamic religious education in the Society 5.0 era: A philosophical-approach study]. *Edukatif: Jurnal Ilmu Pendidikan*, 4(6). <https://doi.org/10.31004/edukatif.v4i6.4179>
- Bahri, S. (2022). Konsep pembelajaran Pendidikan Agama Islam di era Society 5.0 [The concept of Islamic Religious Education learning in the Society 5.0 era]. *Edupedia: Jurnal Studi Pendidikan dan Pedagogi Islam*, 6(2), 133–145. [10.35316/edupedia.v6i2.1592](https://doi.org/10.35316/edupedia.v6i2.1592)
- Cabinet Office, Government of Japan. (2016). Society 5.0. https://www8.cao.go.jp/cstp/english/society5_0/index.html
- Fauzi, A. (2022). Tabayyun sebagai landasan teologis literasi digital: Menelusuri relevansi ajaran Islam dengan kompetensi digital [Tabayyun as the theological foundation of digital literacy: Exploring the relevance of Islamic teachings to digital competence]. *Journal of Islamic Education Studies*, 10(1), 78–96. [10.56832/edu.v5i3.2477](https://doi.org/10.56832/edu.v5i3.2477)
- Fauzi, M. R., Zahro, S. F., Bashith, A., & Hakim, M. J. (2025). Peluang dan tantangan integrasi artificial intelligence dalam evaluasi pembelajaran Pendidikan Agama Islam [Opportunities and challenges of integrating artificial intelligence into Islamic Religious Education learning evaluation]. *Educasia: Jurnal Pendidikan, Pengajaran, dan Pembelajaran*, 10(2).
- Firda Maghfiratus S., Sahlan, M., & Arifin, M. (2026). Mengukur kompetensi secara holistik: Optimalisasi asesmen autentik dalam evaluasi pembelajaran PAI [Measuring competency holistically: Optimizing authentic assessment in PAI learning evaluation]. *Sinergi: Jurnal Ilmiah Multidisiplin*, 2(6), 3536–3545. <https://doi.org/10.66914/pgkzdm70>
- Fu, Y., Weng, Z., & Wang, J. (2025). Examining AI use in educational contexts: A scoping meta-review and bibliometric analysis. *International Journal of Artificial Intelligence in Education*, 35(3), 1388–1444. <https://doi.org/10.1007/s40593-024-00442-w>
- Garzón, J., Patiño, E., & Marulanda, C. (2025). Systematic review of artificial intelligence in education: Trends, benefits, and challenges. *Multimodal Technologies and Interaction*, 9(8), Article 84. <https://doi.org/10.3390/mti9080084>
- Gilster, P. (1997). *Digital literacy*. John Wiley & Sons.
- Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, 26(2), 91–108. <https://doi.org/10.1111/j.1471-1842.2009.00848.x>
- Halkiopoulous, C., & Gkintoni, E. (2024). Leveraging AI in e-learning: Personalized learning and adaptive assessment through cognitive neuropsychology, a systematic analysis. *Electronics*,

13(18), Article 3762. <https://doi.org/10.3390/electronics13183762>

Hamidi, H., & Chavoshi, A. (2022). Digital literacy in education: A framework and assessment tool. *Computers & Education*, 183, Article 104595. <https://doi.org/10.1016/j.compedu.2022.104595>

Hanafi, S. S. A. N., Fadhila, A. S., & Bashith, A. (2025). Pendekatan penilaian autentik dalam evaluasi pembelajaran Pendidikan Agama Islam (PAI) [The authentic-assessment approach in Islamic Religious Education (PAI) learning evaluation]. *Educan: Jurnal Pendidikan Islam*, 9(2), 51–67.

<https://doi.org/10.21111/educan.v9i2.14679>

Jenuri, Faqihuddin, A., Suresman, E., & Abdullah, M. (2025). Overcoming the spiritual emptiness of students in the modern era through the integration of Al-Ghazali's human concepts in the Islamic religious education learning model. *Cogent Education*, 12(1), Article 2497147.

<https://doi.org/10.1080/2331186X.2025.2497147>

Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi [Ministry of Education, Culture, Research, and Technology]. (2022). Panduan pembelajaran dan asesmen: Pendidikan anak usia dini, jenjang pendidikan dasar, dan jenjang pendidikan menengah [Guidelines for learning and assessment: Early childhood education, primary education, and secondary education]. Badan Standar, Kurikulum, dan Asesmen Pendidikan.

https://kurikulum.kemdikbud.go.id/file/1755668120_manage_file.pdf

Kunandar. (2018). Penilaian autentik (Penilaian hasil belajar peserta didik berdasarkan Kurikulum 2013): Suatu pendekatan praktis disertai dengan contoh [Authentic assessment (assessment of learners' learning outcomes based on the 2013 Curriculum): A practical approach with examples]. *Rajawali Pers*. <https://opac.ut.ac.id/detail-opac?id=38499>

Kurnia, N., Astuti, I., & Winarko, E. (2019). Integrasi literasi digital dalam pembelajaran di era revolusi industri 4.0 [Integration of digital literacy into learning in the fourth industrial revolution era]. *Jurnal Teknologi Pendidikan*, 21(2), 101–117.

<https://doi.org/10.25299/jpim.2024.19628>

Lisyawati, E., Mohsen, M., Hidayati, U., & Taufik, O. A. (2023). Literasi digital pembelajaran Pendidikan Agama Islam pada MA Nurul Qur'an Bogor [Digital literacy in Islamic Religious Education learning at MA Nurul Qur'an Bogor]. *Edukasi: Jurnal Penelitian Pendidikan Agama dan Keagamaan*, 21(2), 224–242.

<https://doi.org/10.32729/edukasi.v21i2.1618>

López-Núñez, J.-A., Alonso-García, S., Berral-Ortiz, B., & Victoria-Maldonado, J.-J. (2024). A systematic review of digital competence evaluation in higher education. *Education Sciences*, 14(11), Article 1181. <https://doi.org/10.3390/educsci14111181>

Maharani, A., & Nursikin, M. (2025). Pendidikan karakter di era Society 5.0 pada lembaga pendidikan [Character education in the Society 5.0 era at educational institutions]. *Journal of Innovative and Creativity*, 5(2), 9089–9095. <https://doi.org/10.31004/joecy.v5i2.1576>

Masuwai, A., Zulkifli, H., & Hamzah, M. I. (2024). Evaluation of content validity and face validity of secondary school Islamic education teacher self-assessment instrument. *Cogent*

- Education, 11(1), Article 2308410. <https://doi.org/10.1080/2331186X.2024.2308410>
- McArthur, J. (2023). Rethinking authentic assessment: Work, well-being, and society. *Higher Education*, 85(1), 85–101. <https://doi.org/10.1007/s10734-022-00822-y>
- Muchlis. (2025). Penggunaan artificial intelligence (AI) dalam pembelajaran Pendidikan Agama Islam: Manfaat dan tantangan [The use of artificial intelligence (AI) in Islamic Religious Education learning: Benefits and challenges]. *Kreatif: Jurnal Pemikiran Pendidikan Agama Islam*, 23(1).
- Nafiati, D. A. (2021). Revisi taksonomi Bloom: Kognitif, afektif, dan psikomotorik [Revision of Bloom's taxonomy: Cognitive, affective, and psychomotor]. *Humanika*, 21(2), 151–172. <https://doi.org/10.21831/hum.v21i2.29252>
- Naufal, H. A. (2021). Literasi digital [Digital literacy]. *Perspektif*, 1(2), 195–202. <https://doi.org/10.53947/perspekt.v1i2.32>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, Article n71. <https://doi.org/10.1136/bmj.n71>
- Pratiwi, H., Ariyani, M., Elisa, M., & Harahap, M. (2024). Literasi digital sebagai inovasi pembelajaran dalam pendidikan agama Islam [Digital literacy as a learning innovation in Islamic religious education]. *Jurnal Pendidikan Islam Muta'allimin*, 1(2), 79–92. <https://doi.org/10.25299/jpim.2024.19628>
- Sari, M. N., & Fitriana, D. (2022). Ledakan baru pendidikan Islam: Bagaimana kecerdasan siswa disiapkan memasuki era Society 5.0 [The new explosion of Islamic education: How students' intelligence is being prepared to enter the Society 5.0 era]. *Jurnal Riset Rumpun Agama dan Filsafat*, 1(1).
- Shahidi Hamedani, S., Aslam, S., Mundher Oraibi, B. A., Wah, Y. B., & Shahidi Hamedani, S. (2024). Transitioning towards tomorrow's workforce: Education 5.0 in the landscape of Society 5.0: A systematic literature review. *Education Sciences*, 14(10), Article 1041. <https://doi.org/10.3390/educsci14101041>
- Sholahudin, T., Abid, I., Ikhwanudin, M., Arrizky, M. N., & Al-Ghozali, U. M. (2025). Evaluasi hasil pembelajaran Pendidikan Agama Islam (PAI): Tinjauan terhadap ayat Al-Qur'an dalam aspek kognitif, afektif, dan psikomotorik [Evaluation of Islamic Religious Education (PAI) learning outcomes: A review of Qur'anic verses in the cognitive, affective, and psychomotor aspects]. *Ainara Journal (Jurnal Penelitian dan PKM Bidang Ilmu Pendidikan)*, 6(1), 165–171. <https://doi.org/10.54371/ainj.v6i1.808>
- Siregar, I., Izzat, H., Al Hafizh, M., & Wulandari, S. (2025). Konsep dasar evaluasi dan implikasinya dalam evaluasi pembelajaran Pendidikan Agama Islam di sekolah [Basic concepts of evaluation and their implications for Islamic Religious Education learning

evaluation in schools]. Jurnal

Manajemen dan Pendidikan Agama Islam, 3(4), 275–282.

<https://doi.org/10.61132/jmpai.v3i4.1300>

Suryatina, Z. V. A., Muslih, M., & Rokhmah, A. (2026). Model asesmen holistik dalam pembelajaran PAI berbasis multiple intelligence [A holistic assessment model in PAI learning based on multiple intelligences]. Jurnal Simki Pedagogia, 9(1), 200–208.

<https://doi.org/10.29407/jsp.v9i1.1471>

Wiggins, G., & McTighe, J. (2005). Understanding by design (2nd ed.). Association for Supervision and Curriculum Development.

Yusuf, B. (2024). Teknologi dan personalisasi pembelajaran pendidikan Islam untuk generasi Z [Technology and the personalization of Islamic-education learning for Generation Z].

Journal of Instructional and Development Researches, 4(4).

<https://doi.org/10.53621/jider.v4i4.344>