GAJIE; Global Journal of Islamic Education, Vol 2, Number: 3, September 2026. P 241- 262
E-ISSN: 3110-651X

REVERSED PYRAMID: HISTORY OF ASSESSMENT, MUHASABAH, AND THE FUTURE OF ISLAMIC EDUCATION ASSESSMENT

Raikhan [1], Abdul Hakim [2]
[1,2] Universitas Tarbiyatut Tholabah lamongan, Indonesia

Email Corresponding Author: raikhan@iai-tabah.ac.id.
Received: 01 August 2026. Accepted: 25 September 2026. Published: 30 September 2026

ABSTRACT
Assessment in Islamic education remains dominated by the Assessment of Learning (AoL) paradigm, which prioritizes final outcomes and memorization while neglecting affective and spiritual dimensions. This article aims to map the historical evolution of learning assessment from conventional measurement to Assessment as Learning (AaL) analyze its impact on Islamic education practices, and reconfigure the assessment pyramid. This study employs a systematic literature review with a qualitative-descriptive approach based on content analysis of primary and secondary literature from 1967 to 2026. The analysis is conducted through textual criticism, comparative analysis of key figures' thoughts, and synthesis of changes in the evaluation pyramid structure. Findings reveal that the history of assessment has evolved through seven periods, from measurement toward learning improvement (AfL) and ultimately learning as assessment (AaL). The AaL paradigm aligns with the concept of muhasabah in Islam, yet assessment practices in madrasahs remain dominated by AoL. The traditional assessment pyramid, which places AoL at its foundation, needs to be deconstructed and reconfigured by placing AaL as the primary foundation, AfL as the middle layer, and AoL at the apex for certification purposes. This reconceptualization offers a theoretical and practical framework for teachers and Islamic educational institutions to integrate AaL through reflection, self-assessment, and holistic value-based assessment. Future research is recommended to empirically test the effectiveness of this reconceptualization in madrasahs and pesantren.
Keywords: Assessment as Learning, Assessment pyramid, History of assessment, Islamic education, Muhasabah,

ABSTRAK
Penilaian dalam pendidikan Islam masih didominasi oleh paradigma Assessment of Learning (AoL) yang berorientasi pada hasil akhir dan hafalan, sehingga dimensi afektif dan spiritual terabaikan. Artikel ini bertujuan memetakan sejarah evolusi penilaian hasil belajar dari pengukuran konvensional hingga Assessment as Learning (AaL) serta menganalisis dampaknya terhadap praktik pendidikan Islam dan merekonfigurasi piramida penilaian. Kajian ini merupakan studi literatur sistematis dengan pendekatan kualitatif-deskriptif berbasis analisis isi terhadap literatur primer dan sekunder dari tahun 1967 hingga 2026. Analisis dilakukan melalui kritik teks, komparasi pemikiran tokoh, dan sintesis perubahan struktur piramida evaluasi. Temuan menunjukkan bahwa sejarah penilaian telah berevolusi melalui tujuh periode, dari pengukuran menuju perbaikan pembelajaran (AfL) hingga pembelajaran sebagai asesmen (AaL). Paradigma AaL sejalan dengan konsep muhasabah dalam Islam, tetapi praktik penilaian di madrasah masih didominasi AoL. Piramida

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

penilaian tradisional yang menempatkan AoL sebagai fondasi perlu didekonstruksi dan direkonfigurasi dengan menempatkan AaL sebagai fondasi utama, AfL sebagai lapisan tengah, dan AoL sebagai puncak untuk sertifikasi. Rekonfigurasi ini menawarkan kerangka teoretis dan praktis bagi guru dan lembaga pendidikan Islam untuk mengintegrasikan AaL melalui refleksi, penilaian diri, dan asesmen holistik berbasis nilai Islam. Penelitian lanjutan disarankan untuk menguji efektivitas rekonfigurasi ini secara empiris di madrasah dan pesantren.
Kata kunci:, Assessment as Learning, Muhasabah, Piramida asesmen, Pendidikan Islam. Sejarah penilaian

INTRODUCTION

Assessment in the learning process plays a critical role in determining the effectiveness of education. However, assessment practices across various parts of the world, including Indonesia, remain dominated by Assessment of Learning (AoL), which emphasizes final outcomes rather than the learning process itself (Setyawarno et al., 2026). This paradigm has long served as the foundation of modern education systems, where examinations and standardized tests are regarded as the sole measure of student success. Consequently, assessment is often reduced to a social formality in the relationship between teachers and institutions, where formative assessment is carried out merely for accountability purposes, meaning the assessment process fails to meet proper assessment standards (Raikhan, 2020).

The history of learning assessment shows that assessment paradigms have undergone a long and continuous evolution. Globally, there have been seven periods of assessment change reflecting shifts in how the nature of learning and teaching is understood. Indonesia itself has experienced seven changes in assessment policy, and many of these changes have been influenced by paradigm shifts occurring at the global level (Raikhan, 2020). Curriculum changes are also not widely understood by the public as educational dynamics but rather as mere political projects; even changes in evaluation are still not understood as paradigm shifts but merely as attempts to create differences (Raikhan, 2020). This condition indicates that the history of assessment is not merely a chronological record of methods and instruments but a reflection of philosophical struggles over what is considered valuable in education.

In the context of Islamic education, the history of assessment has a more complex dimension because it is closely tied to spiritual values and character formation. Islamic education aims not only to transfer knowledge but also to shape the whole person (insan kamil) who is balanced in intellectual, spiritual, and social aspects (Fatmawati, 2025). Therefore, assessment in Islamic education should not stop at measuring cognitive aspects alone but should also encompass affective and psychomotor dimensions that constitute the core of Islamic educational goals (Idris

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

& Asyafah, 2020). However, reality on the ground shows that assessment practices in madrasahs remain dominated by cognitive approaches and memorization-oriented learning outcome measurement, while affective, spiritual, and social dimensions have not received adequate attention in the evaluation system (Fatmawati, 2025).

The gap between the ideal and the reality of assessment in Islamic education is a pressing issue to examine. On the one hand, there is theoretical awareness that assessment must be able to measure students' holistic competencies, going beyond the limits of traditional cognitive tests. Authentic assessment, for instance, has great potential to measure students' holistic competencies, and in Islamic Religious Education (PAI), which emphasizes character formation and value internalization, the application of authentic assessment is essential to ensure that evaluation aligns with the spiritual goals of education (Mutammimah et al., 2026). On the other hand, practice in the field still tends to focus on cognitive aspects through written tests, which does not reflect holistic assessment (Zahra et al., 2026).

The problem is compounded when changes in assessment policy are not accompanied by adequate understanding of the historical and philosophical roots of those changes. Teachers and education practitioners often understand curriculum and assessment changes as administrative obligations rather than as paradigm shifts demanding a transformation of thinking and action (Raikhan, 2020). As a result, assessment innovations such as formative assessment, authentic assessment, and Assessment as Learning (AaL) are often adopted superficially without deep understanding of their theoretical and historical foundations.

Furthermore, there is a gap in Islamic education literature that specifically addresses the history of assessment. Most studies on assessment in Islamic education focus on practical aspects such as instrument design, rubrics, or implementation strategies but pay less attention to the historical and philosophical dimensions underlying the evolution of assessment (Fatmawati, 2025; Rachman et al., 2026). Yet a deep understanding of the history of assessment can help educators and policymakers understand why paradigm shifts occur, what their impact is on Islamic education practice, and how to navigate such changes wisely.

Chronologically, the concept of assessment in education has evolved through several stages. In the early stage, assessment was understood as measurement, focusing on quantifying learning outcomes through standardized tests. This stage was dominated by the behaviorist paradigm, which viewed learning as observable and objectively measurable behavioral change. In the next stage, awareness emerged that assessment serves not only to measure but also to improve the learning process. This became known as Assessment for Learning (AfL) or formative assessment, which emphasizes the role of feedback in enhancing learning (Black & Wiliam, 1998).

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

Further development brought us to the concept of Assessment as Learning (AaL), where assessment is not merely a tool to measure or improve learning but becomes an integral part of the learning process itself. AaL positions students as active participants in the evaluation process, enabling them to reflect on their understanding and develop more effective learning strategies (Setyawarno et al., 2026). In the AaL paradigm, students are encouraged to become independent learners capable of monitoring, regulating, and evaluating their own learning (Dann, 2014; Yan & Boud, 2022).

Interestingly, the AaL paradigm has a parallel with the concept of muhasabah in Islam. Muhasabah—literally meaning "taking account" or "self-evaluation" is a practice of introspection and self-evaluation encouraged in Islamic teachings as a means of self-improvement and drawing closer to Allah. Raikhan (2020) explicitly concludes that the Assessment as Learning paradigm aligns with the concept of muhasabah in Islam. This parallel indicates that historically, the roots of self-evaluation in Islam align with the global evolution of assessment toward AaL, without the need to set the two in opposition.

Based on the background and problems outlined above, this article aims to:

1. Map the historical evolution of learning assessment chronologically, from the era of conventional measurement to the emergence of the Assessment as Learning (AaL) paradigm, emphasizing the milestones of key figures such as Scriven, Bloom, Black, Wiliam, and Earl.
2. Analyze the impact of the historical changes in assessment approaches on educational practice, particularly in the context of Islamic education, highlighting how the paradigm shift from AoL to AaL has implications for the view of students, learning objectives, and the role of teachers.
3. Deconstruct the assessment pyramid model and propose its reconfiguration toward a contemporary assessment pyramid that places AaL as the primary foundation, linking it to Islamic values such as muhasabah and character formation.
4. Present a historical-philosophical synthesis of how the evolution of assessment in Islamic education reflects a broader paradigm shift, without getting caught up in technical comparisons between AoL, AfL, and AaL.

The significance of this article lies in its effort to bridge two domains that are often discussed separately: the history of educational assessment and the philosophy of Islamic education. Most literature on the history of assessment focuses on the Western context, while literature on assessment in Islamic education tends to be ahistorical and oriented toward practical aspects. By synthesizing the two, this article offers a more holistic perspective on how the evolution of assessment from cognitive measurement to reflective assessment can be interpreted within the framework of Islamic values.

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

The novelty of this article also lies in its use of the assessment pyramid metaphor as a tool for historical analysis. The assessment pyramid, which traditionally depicts a hierarchy from basic knowledge to complex performance, is deconstructed and reconfigured in this article to reflect the paradigm shift from AoL (which dominates the base of the pyramid) toward AaL (which should become the primary foundation). This reconfiguration is not merely a visual change but reflects a fundamental shift in how we understand the nature of learning, the role of students, and the purpose of assessment itself.

This article is organized into several main sections. Following this introduction, the Methods section will explain the approach used namely, systematic literature review and historical-philosophical content analysis. The Results section will present the chronology of the historical evolution of assessment, the thoughts of key figures, and data on the characteristics and visualization of changes in the assessment pyramid. The Discussion section will interpret these findings, analyze their impact on Islamic education, relate them to the context of the Merdeka Curriculum, and discuss the implications and limitations of the study. The Conclusion section will summarize the main findings and offer directions for future research.

METHOD

This article is a conceptual-historical study employing a qualitative-descriptive approach based on systematic literature review and content analysis. This approach was chosen because the main purpose of the article is to map the historical evolution of learning assessment, identify the thoughts of key figures, and analyze the impact of paradigm shifts in assessment on educational practice, particularly in the context of Islamic education. Systematic literature review allows researchers to collect, evaluate, and synthesize relevant literature sources in a structured and transparent manner (Raikhan, 2020; Fatmawati, 2025).

The primary data sources in this study are primary and secondary literature relevant to the history of assessment, the development of assessment paradigms, and assessment in Islamic education. Primary literature includes monumental works that mark milestones in the history of assessment, such as: Scriven (1967) on evaluation methodology, who coined the term "formative evaluation"; Bloom (1968, 1969) on Mastery Learning and the extension of formative concepts to classroom assessment; Black and Wiliam (1998) through "Inside the Black Box," which placed formative assessment at the center of learning improvement; Dann (2014) on AaL as a blurring of boundaries between assessment and learning; Yan and Boud (2022) on contemporary AaL conceptualization; and Raikhan (2020) on the history of learning assessment and its relation to muhasabah in Islam.

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

Secondary literature includes journal articles, books, and policy documents discussing the history of assessment, the Merdeka Curriculum, and assessment in Islamic education. These sources were selected based on criteria: (1) relevance to the topic of assessment history and Islamic education, (2) academic credibility (indexed in Scopus, Sinta, or published by reputable publishers), (3) representativeness of historical periods (1960s to 2026), and (4) diversity of contexts (global and Indonesian).

In total, this study analyzed more than 30 literature sources consisting of journal articles, books, book chapters, and policy documents. These sources were grouped by theme: (a) history and evolution of assessment, (b) thoughts of key figures, (c) AaL and AfL paradigms, (d) assessment in Islamic education, and (e) assessment policy in Indonesia.

Data collection was carried out through literature searches using academic databases such as Scopus, Google Scholar, ERIC, and Portal Garuda (for Indonesian literature). Keywords used in searches included: "History of educational assessment," "Formative assessment evolution," "Assessment as Learning" or "AaL," "Assessment pyramid" or "Piramida penilaian," "Assessment in Islamic education," "Muhasabah," and "Kurikulum Merdeka" and "asesmen."

The search procedure was conducted in stages. In the first stage, researchers identified primary literature that marked milestones in the history of assessment through citation tracking and cross-referencing. In the second stage, researchers searched secondary literature discussing interpretations, critiques, and developments of primary literature. In the third stage, researchers specifically sought literature connecting the history of assessment with the context of Islamic education, including articles from Islamic education journals such as JIEM: Journal of Islamic Education and Management, Al-Alim: Jurnal Pendidikan Agama Islam, Jurnal PAI Raden Fatah, and Journal of Quality Assurance in Islamic Education.

Inclusion criteria in data collection were: (1) literature discussing the history of assessment, the evolution of assessment paradigms, or assessment in Islamic education; (2) published between 1967 and 2026; (3) available in Indonesian or English; and (4) directly relevant to the study's objectives. Exclusion criteria included: (1) literature discussing only technical aspects of instruments without connection to history or paradigm; (2) literature that could not be accessed in full; and (3) literature that did not undergo peer review (except official policy documents).

Data analysis was conducted through three main stages: textual criticism, comparative analysis of key figures' thoughts, and synthesis of changes in the evaluation pyramid structure. The first stage was textual criticism, namely reading primary and secondary literature critically to understand the original meaning of concepts developed by key figures. Textual criticism included: (a) identifying main

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

arguments and the historical context of their emergence; (b) tracing changes in the meaning of concepts over time, Raikhan (2020), for instance, identified parallels between AaL and muhasabah, indicating that the Islamic tradition has its own historical roots in self-evaluation practices.

The second stage was comparative analysis of key figures' thoughts, namely comparing the contributions of Scriven, Bloom, Black, Wiliam, Earl, Dann, and Yan & Boud to identify continuities and shifts in the history of assessment. Comparison was conducted using a comparison matrix covering dimensions: (a) main focus (measurement, improvement, or learning); (b) role of students (object, partner, or agent); (c) role of teachers (assessor, facilitator, or co-regulator); (d) instruments used (standardized tests, feedback, or self-reflection); and (e) philosophical foundation (behaviorism, constructivism, or socio-cultural).

The third stage was the synthesis of changes in the evaluation pyramid structure. This synthesis was conducted using the assessment pyramid metaphor as an analytical tool. The traditional assessment pyramid, which typically depicts a hierarchy from basic knowledge (at the base of the pyramid) to complex performance (at the apex) is deconstructed and reconfigured in this study to reflect the paradigm shift from AoL to AaL. The synthesis process was carried out through the following steps: identifying the structure of the traditional pyramid; analyzing shifts in proportion; reconfiguring the pyramid; and contextualizing within Islamic education. In this context, AaL is viewed not merely as a pedagogical innovation but as an actualization of Islamic values emphasizing self-evaluation and continuous improvement.

To ensure the validity and reliability of the study, researchers applied several strategies: source triangulation using various types of literature sources (primary, secondary, and policy documents) to verify findings and reduce bias; audit trail systematically documenting the literature search process, inclusion-exclusion criteria, and analysis steps so that they can be replicated by other researchers; and reflexivity—acknowledging the researcher's position as part of the Islamic education tradition and recognizing that the resulting interpretation is not entirely value-free. This reflexivity is important to avoid confirmation bias and ensure that the analysis remains open to alternative perspectives.

This study has several methodological limitations that must be acknowledged. First, as a literature study, this study does not generate new empirical data, so findings depend on the quality and availability of existing literature. Second, literature on the history of assessment in Islamic education remains limited, so the analysis of the impact of assessment history on Islamic education may be less profound than the analysis of assessment history in general. Third, this study does not conduct a comprehensive bibliometric analysis, so the mapping of research trends may not be fully representative. Fourth, interpretation

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

of primary and secondary texts always contains an element of subjectivity, although mitigation has been attempted through triangulation and reflexivity.

Nevertheless, these limitations do not reduce the value of the study's contribution, because the main purpose of the article is to present a historical-philosophical synthesis that can serve as a foundation for further research and practice, rather than to produce empirical generalizations that can be widely generalized.

RESULTS AND DISCUSSION

Chronology of the Historical Evolution of Assessment and the Thoughts of Key Figures

Before 1967, educational assessment practices were dominated by intelligence tests, aptitude tests, and achievement tests that developed in various countries with different emphases. In England, assessment emphasized statistical approaches and early achievement tests as research tools. In France, the main focus was intelligence testing and large-scale class investigations, reflecting a scientific approach in educational psychology. Germany developed aptitude testing and the leaderless group discussion technique, initially used for military selection and later for educational entrance examinations (Idris & Asyafah, 2020).

Although the term "formative assessment" had not yet been formalized, the seeds of this concept had already appeared in continuous testing practices in the 1930s. Recurring and comparable tests were advocated to guide long-term educational development, and early feedback mechanisms began to be used to inform students and teachers. Thus, the pre-1967 period laid the foundation for formative assessment, although there was not yet a unified terminology or framework (Black, 1996).

This period was marked by the formalization of the concept of formative assessment and its differentiation from summative assessment. Michael Scriven (1967) coined the term "formative evaluation" to distinguish it from summative evaluation, emphasizing the role of formative evaluation in improving educational programs during development. Benjamin Bloom (1968, 1969) extended the formative concept to classroom assessment through the Mastery Learning model, advocating feedback to bridge the gap between actual and desired performance (Smith et al., 2018).

In the 1990s, Paul Black and Dylan Wiliam, through the meta-analysis "Inside the Black Box" (1998), demonstrated that formative assessment significantly improves classroom learning. This work transformed the landscape of assessment research and encouraged the emergence of the Assessment Reform Group (1999), which developed the concept of Assessment for Learning (AfL). In England, the Task Group on Assessment and Testing (TGAT), led by Paul Black (1987–1988),

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

advocated formative assessment as a means of improving teaching and learning, influencing national curriculum reform (Black & Wiliam, 1998).

In the 2010s, Lorna Earl developed the concept of Assessment as Learning (AaL) as a mirror of reflection, distinguishing AoL, AfL, and AaL. AaL was positioned as an aspect of formative assessment that blurs the boundary between assessment and learning, with emphasis on self-regulation, metacognition, and active participation of students in the evaluation process (Dann, 2014). Ruth Dann (2014) further conceptualized AaL from a socio-cultural perspective, highlighting three new views: understanding feedback, understanding the learning gap, and exploring the vocabulary of assessment.

Raikhan (2020) noted that globally there have been seven periods of assessment change, and Indonesia has experienced seven changes in assessment policy. These changes reflect not only shifts in methods but also paradigm shifts in how the nature of learning, the role of students, and the purpose of assessment are understood.

Recent developments show the integration of digital technology and artificial intelligence (AI) in assessment. Digital platforms such as Quizizz, Quiz Whizzer, and Wayground are used as interactive assessment media that enhance student motivation, engagement, and learning outcomes (Andika et al., 2024; Nairoh et al., 2026; Dzulfikar & Fediyanto, 2026). AI enables personalized feedback, automated scoring, and adaptive learning, although challenges such as algorithmic bias, data privacy, and teacher readiness remain obstacles (Setyawarno et al., 2026).

Table 1. Chronology of the Historical of Assessment and Key Figures

Period	Key Figures	Main Contribution	Dominant Paradigm
Pre-1967	—	Intelligence, aptitude, and achievement tests; continuous testing	Measurement
1967–2000	Scriven (1967), Bloom (1968), Black & Wiliam (1998)	Formalization of formative evaluation; Mastery Learning; Inside the Black Box	Formative assessment (AfL)
2000–2010s	Earl (2003), Dann (2014)	Conceptualization of AaL; blurring of assessment-learning boundaries	AaL
2010s–2026	Yan & Boud (2022), Gaskaree &	Contemporary AaL; integration of technology and AI	AaL with technology

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

The Impact of the Historical Changes in Assessment Approaches on Islamic Education Practice

The study's findings show that assessment practices in Islamic education remain dominated by cognitive approaches and memorization-oriented learning outcome measurement. Fatmawati (2025) asserts that in madrasahs, affective, spiritual, and social dimensions have not received adequate attention in the evaluation system. Assessment practices tend to focus on written tests measuring cognitive aspects, while the goals of Islamic education emphasizing character formation and value internalization are not yet reflected in the assessment instruments used.

Idris and Asyafah (2020) add that the assessment used in schools so far is traditional assessment, which tends to assess cognitive aspects and neglect affective and psychomotor aspects. Yet assessment should be able to assess all aspects, namely attitudes, skills, and knowledge. This gap indicates that the global history of paradigm shifts in assessment has not fully permeated assessment practices in Islamic education.

Raikhan (2020) explicitly concludes that the Assessment as Learning paradigm aligns with the concept of muhasabah in Islam. Muhasabah the practice of introspection and self-evaluation has historical roots in the Islamic tradition as a means of self-improvement and drawing closer to Allah. This parallel indicates that historically, the roots of self-evaluation in Islam align with the global evolution of assessment toward AaL.

This finding is reinforced by Zahra et al. (2026), who assert that assessment and evaluation in PAI learning play a strategic role in measuring learning achievement while simultaneously shaping students' religious character. However, practice in the field still tends to focus on cognitive aspects through written tests, which does not reflect holistic assessment. Therefore, an assessment design is needed that covers cognitive, affective, and psychomotor domains in a balanced manner, grounded in philosophical, psychological, and religious aspects emphasizing harmony between knowledge and action.

Fatmawati (2025) emphasizes the need for recontextualization of the concept of assessment in Islamic education to align with 21st-century demands. Assessment is not only a tool to measure academic achievement but also a means to build character, morals, and spiritual awareness. This study asserts that recontextualization of madrasah assessment is a strategic step toward improving the quality of Islamic education, with theoretical implications for strengthening the concept of value-based evaluation and practical implications for developing

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

assessment models that are humanistic and oriented toward the formation of the whole person.

Rachman et al. (2026) found that the Islamic Cultural History (SKI) curriculum at MTs and MA is generally aligned with students' developmental stages but has limitations in contextualizing content, depth of analysis, and the use of assessment strategies that encourage critical understanding. At the MTs level, SKI materials emphasize basic knowledge and chronological understanding, while at MA the curriculum is broader and deeper, with emphasis on analysis, reflection, and higher-order thinking skills. However, the assessment strategies used do not yet fully encourage HOTS.

Hafifah and Alam (2026) add that the Grade XII SKI textbook has limitations in assessment activities that do not yet fully encourage HOTS, in addition to problems of uneven content depth, linguistic inaccuracies, and limited integration of visual elements. This finding indicates that although the textbook meets basic curriculum requirements, improvements in assessment are needed to support more meaningful, contextual, and analytical SKI learning.

Mutammimah et al. (2026) examined the principles of authentic assessment in the context of PAI and found that the principles of contextuality and performance focus are highly relevant, but objectivity poses methodological challenges in assessing sincerity and moral character. The proposed solution involves the application of spiritual data triangulation and continuous assessment. The study concludes that authentic assessment in PAI requires fundamental adaptation to maintain ecological validity without compromising assessment integrity.

Table 2. The Impact of the Historical Changes in Assessment on Islamic Education

Aspect	Finding	Reference
Dominance of AoL	Assessment practices in madrasahs remain dominated by cognitive and memorization approaches	Fatmawati (2025); Idris & Asyafah (2020)
AaL-Muhasabah Parallel	AaL aligns with the concept of muhasabah in Islam	Raikhan (2020)
Recontextualization	Need for an integrative, holistic, Islamic value-based assessment paradigm	Fatmawati (2025); Zahra et al. (2026)
SKI Curriculum	Aligned with developmental stages but needs strengthening of analysis and HOTS	Rachman et al. (2026); Hafifah & Alam (2026)

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

Authentic Assessment	Requires fundamental adaptation to assess spiritual and affective domains	Mutammimah et al. (2026)

Deconstruction and Reconfiguration of the Learning Assessment Pyramid

The traditional assessment pyramid most prominently represented by Miller's Pyramid depicts a hierarchy of competence from basic knowledge (knows) to actual performance (does). In this pyramid, AoL dominates the base because knowledge measurement is regarded as the foundation, while AaL sits at the apex, rarely achieved because it demands high self-regulation and metacognition. This structure reflects the assumption that learning moves linearly from content mastery to application, and that assessment primarily serves to measure achievement at each level.

However, Raikhan (2020) shows that this pyramid structure needs to be deconstructed because it contains a paradigm bias that places AoL as the foundation and AaL as a difficult-to-reach goal. In practice, the traditional pyramid often results in outcome-oriented assessment that neglects the learning process. Empirical data from Setyawarno et al. (2026) show that assessment practices in Indonesia remain dominated by AoL, which emphasizes final outcomes rather than the learning process.

Analysis of the literature reveals a shift in the proportion of AoL, AfL, and AaL throughout the history of assessment. In the pre-1967 period, AoL dominated almost the entire spectrum of assessment. In the 1967–2000 period, AfL began to gain ground, especially after the work of Black and Wiliam (1998). In the 2000–2010s period, AaL began to be conceptualized as a distinct paradigm, although its implementation remained limited. In the 2010s to 2026 period, the integration of technology and AI opened new opportunities for AaL, but empirical data show that practice in the field remains dominated by AoL (Setyawarno et al., 2026; Fatmawati, 2025).

Based on the study's findings, a reconfiguration of the assessment pyramid is proposed that places AaL as the primary foundation, AfL as the middle layer, and AoL at the apex, used only for certification and accountability purposes. This reconfiguration is based on the argument that:

1. AaL cultivates self-regulation and metacognition that serve as the foundation for lifelong learning (Dann, 2014; Yan & Boud, 2022).
2. AaL aligns with the concept of muhasabah in Islam, which emphasizes self-evaluation as a means of continuous improvement (Raikhan, 2020).
3. AaL enables the integration of assessment with learning, so that assessment is no longer a separate activity but an integral part of the learning process (Setyawarno et al., 2026).

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

Table 3. Comparison of Traditional and Contemporary Assessment Pyramids

Aspect	Traditional Pyramid	Contemporary Pyramid (Reconfiguration)
Foundation	AoL (knowledge measurement)	AaL (self-regulation and metacognition)
Middle Layer	AfL (formative)	AfL (formative)
Apex	AaL (rarely achieved)	AoL (certification and accountability)
Orientation	Final outcomes	Process and development
Role of Students	Object of assessment	Agent of learning
Role of Teachers	Assessor	Facilitator and co-regulator

Historical-Philosophical Synthesis of the Evolution of Assessment in Islamic Education

Raikhan (2020) identifies seven periods of global assessment change that parallel the seven changes in assessment policy in Indonesia. These changes reflect a paradigm shift from measurement toward learning, and ultimately toward learning as assessment (AaL). In the context of Islamic education, this shift finds a meeting point with the tradition of muhasabah, which has long been part of Islamic educational practice.

Fatmawati (2025) asserts that assessment in Islamic education should not only measure academic achievement but also build character, morals, and spiritual awareness. However, history shows that assessment practices in madrasahs remain dominated by cognitive approaches, so affective and spiritual dimensions have not received adequate attention. Recontextualization of madrasah assessment is a strategic step to address this gap.

Setyawarno et al. (2026) show that although digital technology offers great opportunities to optimize AaL, its implementation in Indonesia remains limited due to a lack of educator understanding, limited resources, and an assessment paradigm still oriented toward final grades. Dzulfikar and Fediyanto (2026) add that unstable internet connections are a major challenge in implementing digital platforms such as Wayground. These findings indicate that the history of assessment change does not proceed linearly but faces contextual resistance and obstacles.

Interpreting the History of Assessment Evolution: From Measurement to Learning

The study's findings show that the history of assessment has undergone a fundamental paradigm shift: from assessment as measurement to assessment as learning. This shift is not merely a change in methods or instruments but reflects a change in how the nature of learning, the role of students, and the purpose of

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

education are understood. In the pre-1967 period, assessment was understood as an activity of measuring students' abilities through intelligence, aptitude, and achievement tests. This paradigm was rooted in behaviorism, which viewed learning as observable and objectively measurable behavioral change. Students were positioned as objects of assessment, while teachers and institutions acted as assessors with full authority.

The formalization of formative evaluation by Scriven (1967) and Bloom (1968) marked the first shift in the history of assessment: from measuring final outcomes to improving the learning process. Bloom, through Mastery Learning, asserted that feedback can bridge the gap between actual and desired performance, so that assessment no longer serves as a final judgment but as a tool to improve learning. The work of Black and Wiliam (1998) through "Inside the Black Box" marked the second shift: from improving learning to learning as assessment (AaL). Black and Wiliam showed that feedback is not merely information given by teachers to students but a process that is received, interpreted, and used by students to regulate their own learning. This shift positioned students as agents of learning, not merely objects of assessment. Dann (2014) and Yan and Boud (2022) further conceptualized AaL as a blurring of the boundary between assessment and learning. In the AaL paradigm, assessment is no longer a separate activity conducted at the end of learning but becomes an integral part of the learning process itself. Students are encouraged to monitor, regulate, and evaluate their own learning, so that assessment becomes a mirror of reflection that helps them understand their strengths and weaknesses.

This paradigm shift has profound implications for Islamic education. The Islamic tradition has long recognized the concept of muhasabah self-evaluation conducted continuously to improve oneself and draw closer to Allah. Raikhan (2020) explicitly concludes that AaL aligns with the concept of muhasabah in Islam. This parallel indicates that historically, the roots of self-evaluation in Islam align with the global evolution of assessment toward AaL.

However, the study's findings also show that assessment practices in Islamic education remain dominated by AoL, which is oriented toward final outcomes and memorization (Fatmawati, 2025; Idris & Asyafah, 2020). This gap between the ideal and the reality indicates that the global history of paradigm shifts in assessment has not fully permeated Islamic education practice. Therefore, recontextualization of madrasah assessment is needed that integrates Islamic values with the contemporary AaL paradigm.

The Impact of the History of Assessment Change on Islamic Education Practice

The study's findings show that assessment practices in Islamic education remain dominated by cognitive approaches oriented toward memorization and

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

written tests. Fatmawati (2025) asserts that affective, spiritual, and social dimensions have not received adequate attention in the madrasah evaluation system. Yet the goal of Islamic education is to shape the whole person (insan kamil) who is balanced in intellectual, spiritual, and social aspects.

This gap has significant historical impact. When assessment only measures cognitive aspects, the other dimensions of Islamic education such as character formation, value internalization, and spiritual awareness are neglected. As a result, assessment no longer reflects the holistic goals of Islamic education but instead reduces Islamic education to mere knowledge transfer.

Idris and Asyafah (2020) add that traditional assessment used in schools tends to assess cognitive aspects and neglect affective and psychomotor aspects. Yet assessment should be able to assess all aspects, namely attitudes, skills, and knowledge. This finding indicates that the global history of paradigm shifts in assessment from AoL to AaL has not been fully adopted in Islamic education assessment practices. Raikhan (2020) offers a historical-philosophical bridge between AaL and the Islamic tradition through the concept of muhasabah. Muhasabah is not merely self-evaluation but a spiritual practice encouraged in Islam to improve oneself and draw closer to Allah. In the Islamic tradition, muhasabah is conducted continuously, not only at the end of a certain period, and covers all aspects of life, including intellectual, spiritual, and social aspects. The parallel between AaL and muhasabah lies in the emphasis on self-regulation and reflection. In AaL, students are encouraged to monitor, regulate, and evaluate their own learning. In muhasabah, individuals are encouraged to evaluate themselves, acknowledge mistakes, and strive to improve. Both place agency and reflection at the core of the evaluation process.

This finding has a significant theoretical contribution: it shows that AaL is not a concept entirely foreign to the Islamic educational tradition but has historical roots in the practice of muhasabah. Therefore, the adoption of AaL in Islamic education need not be viewed as Westernization or the importation of a foreign concept but as an actualization of Islamic values that have long existed. Fatmawati (2025) emphasizes the need for recontextualization of madrasah assessment to align with 21st-century demands. Recontextualization is not merely a change in methods or instruments but a paradigm shift from assessment as measurement to assessment as learning. Assessment must be viewed as a means of building character, morals, and spiritual awareness, not merely a tool to measure academic achievement.

However, the study's findings also show that this recontextualization faces contextual challenges. Setyawarno et al. (2026) found that the implementation of AaL in Indonesia remains limited due to a lack of educator understanding, limited resources, and an assessment paradigm still oriented toward final grades. Dzulfikar and Fediyanto (2026) add that unstable internet connections are an obstacle in

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

implementing digital platforms. These challenges indicate that the history of assessment change does not proceed linearly but faces resistance and obstacles that need to be overcome through professional development and institutional support.

Rachman et al. (2026) and Hafifah and Alam (2026) found that the Islamic Cultural History (SKI) curriculum and its textbooks have limitations in assessment strategies that encourage HOTS. Although SKI content at MA is broader and deeper than at MTs, the assessment strategies used do not yet fully encourage analysis, reflection, and critical thinking. This finding indicates that the history of assessment change which emphasizes the importance of HOTS and authentic assessment has not been fully reflected in SKI assessment practices. Mutammimah et al. (2026) add that authentic assessment in PAI requires fundamental adaptation to assess spiritual and affective domains. The principle of objectivity poses methodological challenges in assessing sincerity and moral character, so spiritual data triangulation and continuous assessment are needed. This finding indicates that the global history of assessment change needs to be contextualized with the values and characteristics of Islamic education.

Reconfiguration of the Assessment Pyramid: AaL as the Foundation

The traditional assessment pyramid—which places AoL as the foundation and AaL at the apex rarely achieved contains a paradigm bias that needs to be criticized. This structure reflects the assumption that learning moves linearly from content mastery to application, and that assessment primarily serves to measure achievement at each level. However, this assumption is not fully aligned with the study's findings showing that AaL cultivates self-regulation and metacognition that serve as the foundation for lifelong learning (Dann, 2014; Yan & Boud, 2022).

Raikhan (2020) shows that the traditional pyramid structure often results in outcome-oriented assessment that neglects the learning process. Empirical data from Setyawarno et al. (2026) show that assessment practices in Indonesia remain dominated by AoL, which emphasizes final outcomes rather than the learning process. Therefore, the traditional pyramid needs to be deconstructed and reconfigured to reflect the paradigm shift toward AaL. Based on the study's findings, a reconfiguration of the assessment pyramid is proposed that places AaL as the primary foundation, AfL as the middle layer, and AoL at the apex, used only for certification and accountability purposes. This reconfiguration is based on three main arguments:

First, AaL cultivates self-regulation and metacognition that serve as the foundation for lifelong learning. Dann (2014) asserts that AaL enables students to monitor, regulate, and evaluate their own learning, so that they become independent and adaptive learners. In the context of Islamic education, this self-

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

regulation capability aligns with the concept of muhasabah, which emphasizes continuous self-evaluation (Raikhan, 2020).

Second, AaL aligns with Islamic values that emphasize self-evaluation and continuous improvement. Fatmawati (2025) asserts that assessment in Islamic education should not only measure academic achievement but also build character, morals, and spiritual awareness. AaL, with its emphasis on reflection and self-regulation, offers a framework aligned with these goals.

Third, AaL enables the integration of assessment with learning, so that assessment is no longer a separate activity but an integral part of the learning process (Setyawarno et al., 2026). In the contemporary pyramid, AaL is no longer at the apex rarely achieved but becomes the foundation supporting the entire assessment process.

The reconfiguration of the assessment pyramid has practical implications for Islamic education. PAI teachers and madrasahs need to shift focus from outcome-oriented assessment toward process- and development-oriented assessment. This means:

1. Integrating AaL into daily learning, for example through reflective journals, self-assessment, and peer assessment (Kim, 2024; Panadero et al., 2016).
2. Developing assessment instruments covering cognitive, affective, and psychomotor domains, and linking them to Islamic values (Zahra et al., 2026).
3. Training teachers to implement AaL, including using digital technology to support reflection and feedback (Setyawarno et al., 2026).
4. Developing a school culture that supports muhasabah, where students are encouraged to evaluate themselves honestly and constructively.

Historical-Philosophical Synthesis: Contribution and Future Directions

This study offers a theoretical contribution in the form of a historical-philosophical synthesis that bridges two domains often discussed separately: the history of educational assessment and the philosophy of Islamic education. Most literature on the history of assessment focuses on the Western context, while literature on assessment in Islamic education tends to be ahistorical and oriented toward practical aspects. By synthesizing the two, this study offers a more holistic perspective on how the evolution of assessment from cognitive measurement to reflective assessment can be interpreted within the framework of Islamic values.

Another theoretical contribution is the reconfiguration of the assessment pyramid that places AaL as the foundation. This reconfiguration is not merely a visual change but reflects a fundamental shift in how the nature of learning, the role of students, and the purpose of assessment are understood. In the context of Islamic education, this reconfiguration aligns with the concept of muhasabah and the goal of forming insan kamil.

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

Practically, this study offers a framework for teachers and Islamic educational institutions to implement AaL in learning. This framework includes:

1. Integrating AaL into the curriculum, placing reflection and self-regulation as integral parts of the learning process.
2. Developing holistic assessment instruments covering cognitive, affective, and psychomotor domains, and linking them to Islamic values.
3. Utilizing digital technology to support AaL, such as interactive quiz platforms, online reflective journals, and personalized feedback (Andika et al., 2024; Nairoh et al., 2026).
4. Continuous teacher professional development, focusing on understanding the AaL concept, designing reflection-based assessment instruments, and classroom implementation strategies (Setyawarno et al., 2026).

This study has several limitations that must be acknowledged. First, as a literature study, this study does not generate new empirical data, so findings depend on the quality and availability of existing literature. Second, literature on the history of assessment in Islamic education remains limited, so the analysis of the impact of assessment history on Islamic education may be less profound than the analysis of assessment history in general. Third, this study does not conduct a comprehensive bibliometric analysis, so the mapping of research trends may not be fully representative. Fourth, interpretation of primary and secondary texts always contains an element of subjectivity, although mitigation has been attempted through triangulation and reflexivity.

CONCLUSION

The history of learning assessment has evolved through seven periods from conventional measurement (pre-1967), the formalization of formative assessment by Scriven and Bloom (1967–2000), and the emergence of AaL by Earl and Dann (2000–2010s), to the integration of technology and AI (2010s–2026). Each period reflects a paradigm shift: from assessment as measurement, to learning improvement (AfL), and ultimately learning as assessment (AaL). The impact of this historical change on Islamic education is ambivalent. On the one hand, the AaL paradigm aligns with the concept of muhasabah in Islam, but on the other hand, assessment practices in madrasahs remain dominated by AoL, which is oriented toward cognitive aspects and memorization, so that affective and spiritual dimensions are neglected. This gap demands a recontextualization of madrasah assessment that is integrative and based on Islamic values. Furthermore, the traditional assessment pyramid, which places AoL as the foundation, needs to be deconstructed and reconfigured by placing AaL as the primary foundation, AfL as the middle layer, and AoL at the apex for certification and accountability. This reconfiguration aligns with the goals of Islamic education in forming insan kamil.

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

The historical-philosophical synthesis shows that the global evolution of assessment and the Islamic tradition have a meeting point in the emphasis on self-regulation, reflection, and continuous improvement. Thus, the adoption of AaL in Islamic education is not the importation of a foreign concept but the actualization of Islamic values that have long existed.

The main contributions of this article are threefold. Theoretically, the article synthesizes the history of global assessment with the philosophy of Islamic education, offering a conceptual bridge between AaL and muhasabah that has so far been discussed separately. Conceptually, the article deconstructs and reconfigures the assessment pyramid from a structure placing AoL as the foundation to a structure placing AaL as the foundation as a tool for historical analysis and a practical framework. Practically, the article offers a framework for teachers and Islamic educational institutions to integrate AaL into learning through reflective journals, self- and peer assessment, holistic assessment instruments, the use of digital technology, and continuous professional development.

Future research should prioritize several directions. First, empirical research on the implementation of AaL in madrasahs and pesantren is needed to test the effectiveness of the assessment pyramid reconfiguration in real contexts. Second, the development and validation of assessment instruments that specifically measure spiritual and affective domains in PAI should be pursued, taking into account the challenges of objectivity and ecological validity. Third, longitudinal studies on the impact of AaL on character formation, spiritual awareness, and long-term learning outcomes are essential. Fourth, the development of effective teacher professional development models to build capacity for AaL implementation, including the use of digital technology, deserves attention. Fifth, cross-cultural studies on the adaptation of AaL in Islamic education in various countries should be conducted, taking into account cultural, policy, and infrastructural differences.

BIBLIOGRAPHY

Andika, R., Erita, Y., & Ningsih, Y. (2024). Pelatihan aplikasi Quizizz sebagai media pembelajaran interaktif dan asesmen bagi pendidik sekolah dasar. *Dinamisia: Jurnal Pengabdian Kepada Masyarakat, 8*(3). https://doi.org/10.31849/dinamisia.v8i3.19041

Asrifan, A., Sudirman, R., Muhayyang, M., Tasni, N., & Isumarni, I. (2025). Foundations of assessment theories: A framework for higher education. In *AI, analytics, and assessment in higher education* (pp. 1–28). IGI Global Scientific Publishing. https://doi.org/10.4018/979-8-3373-7057-6.ch001

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

Black, P. (1996). Formative assessment and the improvement of learning. *British Journal of Special Education, 23*(2), 51–56. https://doi.org/10.1111/j.1467-8578.1996.tb00947.x

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Bloom, B. S. (1968). Learning for mastery. *Evaluation Comment, 1*(2), 1–12.

Brownlie, N., Burke, K., & van der Laan, L. (2024). Quality indicators of effective teacher-created summative assessment. *Quality Assurance in Education, 32*(1), 30–45. https://doi.org/10.1108/QAE-04-2023-0062

Chen, X., Shu, D., DeLuca, C., & Holden, M. (2025). Advancing e-assessment for learning in the primary EFL writing classroom: The role of collaborative teacher professional learning. *Asia Pacific Journal of Education*. Advance online publication. https://doi.org/10.1080/02188791.2025.2465296

Dann, R. (2014). Assessment as learning: Blurring the boundaries of assessment and learning for theory, policy and practice. *Assessment in Education: Principles, Policy and Practice, 21*(2), 149–166. https://doi.org/10.1080/0969594X.2014.898128

Dzulfikar, A. Z., & Fediyanto, N. (2026). Persepsi guru dan siswa terhadap penggunaan Wayground sebagai asesmen pembelajaran dalam pengajaran bahasa Inggris. *Indonesian Journal of Education Methods Development, 21*(3).

Fatmawati, E. (2025). Reconceptualizing assessment in Islamic education: A critical review of madrasah evaluation practices in the 21st century. *Journal of Quality Assurance in Islamic Education (JQAIE), 5*(2). https://doi.org/10.47945/jqaie.v5i2.2376

Gaskaree, B. L., & Chahkandi, F. (2026). Fostering three-way mutuality: A framework for actualizing assessment as learning. *Journal of Research in Applied Linguistics, 17*(1), 173–187. https://doi.org/10.22055/rals.2025.47318.3326

Hafifah, A. N., & Alam, N. A. R. (2026). Analysis of the Islamic Cultural History textbook using a multi-component evaluation framework. *JIE: Journal of Islamic Education, 11*(2). https://doi.org/10.35723/jie.v11i2.725

Hampton, S. E., & Morrow, C. (2003). Reflective journaling and assessment. *Journal of Professional Issues in Engineering Education and Practice, 129*(4), 186–189. https://doi.org/10.1061/(ASCE)1052-3928(2003)129:4(186)

He, M., & Wang, L. (2025). Implementing assessment as learning in online EFL writing classes. *RELC Journal, 56*(2), 468–482. https://doi.org/10.1177/00336882231224966

Idris, M. M., & Asyafah, A. (2020). Authentic assessment in Islamic education. *Jurnal Kajian Peradaban Islam, 3*(1). https://doi.org/10.47076/jkpis.v3i1.36

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

Kim, K. H.-A. (2024). Assessment FOR students, OF students, BY students: Student-centred assessment with assessment as learning pedagogy. *English Australia Journal, 40*(3), 38–43. https://doi.org/10.61504/CFCS3902

Lam, R. (2016). Assessment as learning: Examining a cycle of teaching, learning, and assessment of writing in the portfolio-based classroom. *Studies in Higher Education, 41*(11), 1900–1917. https://doi.org/10.1080/03075079.2014.999317

Lee, I., Mak, P., & Yuan, R. E. (2019). Assessment as learning in primary writing classrooms: An exploratory study. *Studies in Educational Evaluation, 62*, 72–81. https://doi.org/10.1016/j.stueduc.2019.04.012

Mutammimah, D., Kamila, I. F., Kamilia, D., Juhairiyah, J., & Fitriyah, E. (2026). A critical review of authentic assessment principles in the context of Islamic Religious Education learning. *Al-Alim: Jurnal Pendidikan Agama Islam, 2*(1), 147–162. https://doi.org/10.65118/alim.v2i1.39

Nairoh, M., Pramudiyati, & Rinaldi, D. (2026). Pengaruh media pembelajaran Quiz Whizzer sebagai assessment as learning terhadap peningkatan hasil belajar siswa pada materi pemrograman bahasa C. *Journal of Classroom Action Research, 8*(1). https://doi.org/10.29303/jcar.v8i1.13461

Panadero, E., Jonsson, A., & Strijbos, J.-W. (2016). Scaffolding self-regulated learning through self-assessment and peer assessment: Guidelines for classroom implementation. *Enabling Power of Assessment, 4*, 311–326. https://doi.org/10.1007/978-3-319-39211-0_18

Rachman, R. S., Sabda, S., & Rosalia, S. (2026). Critical review of the Islamic Cultural History curriculum in MTs and MA: Analysis of material depth, level continuity, and learning assessment. *Al-Wadhih: Journal of Islamic History and Civilization, 2*(1), 68–82. https://doi.org/10.61166/alwadhih.v2i1.40

Raikhan, R. (2020). Sejarah penilaian hasil belajar; Urgensi dan perkembangan paradigma. *JIEM: Journal of Islamic Education and Management, 1*(1).

Schellekens, L. H., Bok, H. G. J., de Jong, L. H., van der Schaaf, M. F., Kremer, W. D. J., & van der Vleuten, C. P. M. (2021). A scoping review on the notions of Assessment as Learning (AaL), Assessment for Learning (AfL), and Assessment of Learning (AoL). *Studies in Educational Evaluation, 71*, Article 101094. https://doi.org/10.1016/j.stueduc.2021.101094

Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagné, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39–83). Rand McNally.

Setyawarno, D., Maryati, M., & Rahayu, D. P. (2026). Workshop optimalisasi assessment as learning dalam proses pembelajaran sebagai upaya rekonfigurasi piramida asesmen bagi pendidik di Sekolah Indonesia Kuala Lumpur/Malaysia. *Jurnal Pengabdian Masyarakat MIPA dan Pendidikan MIPA, 10*(1). https://doi.org/10.21831/jpmmp.v10i1.90476

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)

Smith, J. K., Berg, D. A. G., & Lipnevich, A. A. (2018). Formative assessment and assessment for learning: Distinctions, mileposts, and a view to the future. In *The Oxford handbook of educational psychology* (pp. 587–610). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199841332.013.27

Tseng, S.-S., & Er, E. (2024). Maximizing the impact of dialogic peer feedback on student engagement: The role of regulation support. *Educational Technology and Society*, *27*(2), 133–148. https://doi.org/10.30191/ETS.202404_27(2).RP04

Van Hoe, A., Wiebe, J., Slotta, J., Rotsaert, T., & Schellens, T. (2024). Designing dialogic peer feedback in collaborative learning: The role of Thinq Tank. *Education Sciences, 14*(11), Article 1231. https://doi.org/10.3390/educsci14111231

Wang, L., & Lee, I. (2021). L2 learners' agentic engagement in an assessment as learning-focused writing classroom. *Assessing Writing, 50*, Article 100571. https://doi.org/10.1016/j.asw.2021.100571

Xiang, X., Yuan, R., & Yu, B. (2022). Implementing assessment as learning in the L2 writing classroom: A Chinese case. *Assessment and Evaluation in Higher Education, 47*(5), 727–741. https://doi.org/10.1080/02602938.2021.1965539

Yan, Z. (2021). Assessment-as-learning in classrooms: The challenges and professional development. *Journal of Education for Teaching, 47*(2), 293–295. https://doi.org/10.1080/02607476.2021.1885972

Yan, Z., & Boud, D. (2022). Conceptualising assessment as learning. In *Assessment as learning: Maximising opportunities for student learning and achievement* (pp. 1–16). Routledge.

Yang, Y. (2026). Artificial intelligence-empowering formative assessment: Research progress and future directions. *Proceedings of 2025 4th International Conference on Artificial Intelligence and Education, ICAIE 2025*, 217–223. https://doi.org/10.1145/3797552.3797589

Zahra, L., Lathifah, L., Mansur, A., & Anggara, B. (2026). Design of assessment and evaluation of Islamic Religious Education (PAI) learning. *Jurnal PAI Raden Fatah, 8*(2). https://doi.org/10.19109/pair.v8i2.35176

Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)