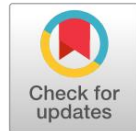


<https://research.adra.ac.id/index.php/ijen/>

P - ISSN: 2988-1579

E - ISSN: 2988-0092



Narratives of Developing Authentic Assessment Instruments for Islamic Religious Education in Secondary Schools Through a Participatory ADDIE Approach

Susanna¹, Jamaluddin Idris², Zainal Abidin³

¹Kantor Pendidikan Kabupaten Nagan Raya, Indonesia

²Universitas Islam Negeri Ar-Raniry Banda Aceh, Indonesia

³Universitas Islam Negeri Ar-Raniry Banda Aceh, Indonesia

ABSTRACT

Background. Authentic assessment has become a central component of the Merdeka Curriculum implementation in Indonesia, particularly in Islamic Religious Education (PAI), which emphasizes holistic student development across cognitive, affective, and psychomotor domains.

Purpose. This study aimed to develop a valid and practical participatory-based module for developing authentic assessment instruments in secondary school PAI learning through the integration of the ADDIE model and participatory professional collaboration within the MGMP forum.

Method. This study employed a Research and Development (R&D) design using the ADDIE model consisting of Analysis, Design, Development, Implementation, and Evaluation phases. The research involved 35 PAI teachers from 23 senior secondary schools in Nagan Raya Regency, Aceh Province.

Results. The findings revealed that the developed module achieved an average expert validity score of 85% categorized as *Very Valid* and an average practicality score of 80% categorized as *Practical*. The participatory mechanism implemented through structured FGDs effectively positioned teachers and supervisors as active co-constructors in developing authentic assessment instruments that were contextual, relevant, and applicable to classroom practice.

Conclusion. The integration of the ADDIE model with a participatory approach provides an effective framework for developing authentic assessment competencies among PAI teachers.

KEYWORDS

Authentic Assessment, Islamic Religious Education, Participatory Module

INTRODUCTION

Islamic Religious Education (Pendidikan Agama Islam/PAI) carries a mission that extends beyond the mere transmission of religious knowledge. As a subject responsible for shaping students' character, internalizing Islamic values, and fostering holistic Islamic personality development, PAI requires an assessment system capable of capturing students' cognitive, affective, and psychomotor development comprehensively and authentically (Zacometti, 2024).

Citation: Susanna, Susanna., Idris, J & Abidin, Z. (2026). Narratives of Developing Authentic Assessment Instruments for Islamic Religious Education in Secondary Schools Through a Participatory ADDIE Approach. *International Journal of Educational Narrative*, 4(3), 939–958. <https://doi.org/10.70177/ijen.v4i3.3896>

Correspondence:

Susanna,
susanna@gmail.com

Received: January 08, 2026

Accepted: March 11, 2026

Published: June 25, 2026



This demand has become increasingly significant within the framework of the Merdeka Curriculum, which has been widely implemented in Indonesia since 2022. Under this curriculum paradigm, assessment is no longer conceptualized merely as a terminal activity (assessment of learning), but rather as an integral component of the learning process itself through assessment for learning and assessment as learning (Hou, 2024).

Jiang, (2022) emphasized that authentic assessment requires learners to engage in tasks that reflect the standards and complexities of real-life professional and social contexts. In the context of PAI, this means that assessment instruments should not be limited to testing memorization abilities; instead, they must be able to evaluate the extent to which students internalize and actualize Islamic values in their daily lives. However, classroom realities indicate a substantial gap between these ideal expectations and teachers' actual competencies. More recent scholarship has further highlighted that authentic assessment should not only measure academic performance but also connect learning with real-world application, personal development, and social relevance (Broadbent, 2023).

A preliminary study conducted through the Islamic Religious Education Subject Teachers' Forum (Musyawarah Guru Mata Pelajaran/MGMP) in Nagan Raya Regency, Aceh Province, revealed concerning findings. Approximately 65% of PAI teachers from 23 senior secondary schools relied solely on cognitive-based assessment instruments, while assessments targeting attitudes and practical skills were rarely found in their instructional planning documents. Furthermore, most assessment items remained concentrated at the lower levels of the revised Bloom's Taxonomy, particularly remembering and understanding (C1–C2), whereas analytical, evaluative, and creative thinking skills (C4–C6) were largely neglected. These findings are consistent with previous national studies. Widiastuti and Basuki (2019) reported the dominance of cognitive assessment practices among religious education teachers in East Java, while Song, (2022) identified limited understanding of competency-based assessment as a major obstacle in implementing the Merdeka Curriculum within Islamic schools.

The gap between the ideal demands of authentic assessment and the actual capacities of PAI teachers is not merely rooted in individual limitations; rather, it is systemic in nature. Contributing factors include the absence of contextualized practical assessment modules, limited professional development forums specifically designed to train teachers in authentic assessment practices, and the lack of structured collaboration between teachers and supervisors. Shi, (2024) argued that effective teacher professional development must be collaborative, sustainable, and grounded in teachers' real working contexts characteristics that are often absent from incidental and top-down training programs. Although some teachers in Nagan Raya had participated in Higher Order Thinking Skills (HOTS) training programs, these initiatives primarily focused on cognitive dimensions and did not adequately address the broader scope of authentic assessment, which includes the evaluation of attitudes, practical skills, and students' abilities to actualize Islamic values in real-life contexts.

From a broader perspective of PAI teacher competence, the ability to develop authentic assessment instruments constitutes a multidimensional sub-competency. It is not limited to pedagogical competence, but is also closely connected to professional competence (mastery of subject content to determine what should be assessed), personal competence (integrity and consistency in applying assessment standards), and leadership competence, which is distinctive to PAI teachers as regulated by the Indonesian Ministry of Religious Affairs Regulation Number 16 of 2010. Li, (2022) referred to this capability as assessment literacy, namely the ability to understand why assessment is conducted, what should be assessed, how assessment should be implemented,

and how assessment results can be utilized to improve learning quality. The low level of assessment literacy among PAI teachers in Nagan Raya, as revealed in the preliminary study, therefore represents a fundamental and urgent professional development need. Recent evidence from a systematic review of formative assessment literacy also indicates that teachers frequently experience difficulties in applying assessment concepts in practice, particularly when professional development opportunities are limited or disconnected from classroom realities. This finding is consistent with (Sun, 2023), who reported that teachers perceive assessment literacy development as one of the most important areas of professional learning needed to support effective instructional and assessment practices.

Several previous studies have addressed this issue from various perspectives. Li, (2022) explored the challenges and opportunities of authentic assessment in PAI under the Merdeka Curriculum; Cai, (2022) developed project-based authentic assessment instruments for PAI; Southwell, (2022) investigated teacher professional development through MGMP forums in Islamic schools; while Perkins, (2022) conducted a systematic review of Participatory Action Research in the development of PAI teacher competencies. Nevertheless, none of these studies specifically produced a participatory-based authentic assessment instrument development module for secondary-level PAI learning using the ADDIE model. This gap provides a strong academic justification for the present study.

Several previous studies have addressed this issue from various perspectives. Revell, (2024) explored the challenges and opportunities of authentic assessment in PAI under the Merdeka Curriculum; Gao, (2023) developed project-based authentic assessment instruments for PAI; Zhou, (2022) investigated teacher professional development through MGMP forums in Islamic schools; while Sultoni and Zainiyati (2022) conducted a systematic review of Participatory Action Research in the development of PAI teacher competencies. Recent research trends have shown increasing attention to authentic assessment, assessment literacy, and collaborative professional development as important elements of curriculum reform. However, existing studies have largely focused on assessment challenges, individual instrument development, or teacher professional development separately. Nevertheless, none of these studies specifically produced a participatory-based authentic assessment instrument development module for secondary-level PAI learning using the ADDIE model. This gap provides a strong academic justification for the present study. A recent scoping review by James, (2022) further demonstrated the growing international interest in authentic assessment design across educational settings, while also highlighting the need for contextualized implementation frameworks that support teachers in translating assessment principles into classroom practice.

Theoretically, this study contributes to the literature by integrating authentic assessment theory, participatory professional development, and the ADDIE instructional design model into a single development framework. This integration provides a conceptual basis for understanding how collaborative professional learning can support the development of assessment literacy and authentic assessment competencies among PAI teachers.

Accordingly, this study aims to develop a valid and practical Participatory Approach Module for Developing Authentic Assessment Instruments in Secondary School PAI Learning, to describe the ADDIE-participatory-based development process, and to provide empirical evidence regarding the effectiveness of the MGMP forum as a professional empowerment platform for PAI teachers in the field of assessment. The study addresses the following research question:

How can the development of authentic assessment instruments for secondary school PAI learning through a participatory approach integrated with the ADDIE model produce a valid and practical module?

THEORETICAL FRAMEWORK

This study is grounded in three major theoretical constructs: authentic assessment, participatory approach, and the ADDIE instructional design model. (Heidari, 2024) defined authentic assessment as an assessment process that engages learners in tasks reflecting the complexity and standards of real-life situations rather than merely testing isolated factual recall. This perspective was further strengthened by Jon Mueller (2008), who emphasized that the authenticity of assessment lies in its ability to measure whether learners can meaningfully apply their knowledge and skills within relevant contexts.

Within the context of Islamic Religious Education (PAI), authentic assessment possesses a distinctive dimension. Authentic assessment should not merely evaluate students' abilities to retell prophetic stories or memorize religious texts; rather, it must assess the extent to which Islamic values related to *aqidah* (faith), *akhlaq* (morality), and *fiqh* (Islamic jurisprudence) are internalized and reflected in students' actual attitudes and behaviors. Therefore, authentic assessment in PAI is inseparable from the broader mission of Islamic education, namely the holistic formation of students' character and spirituality.

The participatory approach adopted in this study originates from the tradition of action research pioneered by Kurt Lewin (1946) and later developed by Paulo Freire (2000), as well as Peter Reason and Hilary Bradbury (2008). This approach is based on the premise that the most meaningful and sustainable knowledge is knowledge collectively constructed by individuals who directly experience the problem themselves. In teacher professional development, this principle positions teachers not as passive recipients of training, but as active agents involved in identifying needs, designing solutions, testing products, and reflecting on outcomes collaboratively. Previous studies have demonstrated that participatory approaches generate higher levels of internalization and more sustainable changes in professional practice compared to conventional top-down training models (Mate, 2022).

Meanwhile, the ADDIE model, conceptualized by Robert Maribe Branch (2009), provides a systematic yet flexible framework that can be effectively integrated with participatory principles. The iterative nature of the ADDIE model where evaluation is embedded throughout each phase rather than conducted solely at the end makes it highly compatible with participatory processes that emphasize continuous cycles of reflection and action. The integration of these three frameworks authentic assessment as the substantive focus, participatory approach as the developmental strategy, and ADDIE as the methodological framework forms a holistic and contextual research model that constitutes the primary novelty of this study.

RESEARCH METHODOLOGY

This study employed a Research and Development (R&D) method using the ADDIE model (Analysis, Design, Development, Implementation, and Evaluation) integrated with a participatory approach. According to (Pan, 2022), R&D is a research method used to develop and validate educational products. The ADDIE model was selected because it is systematic, iterative, and incorporates evaluation mechanisms within each phase, making it highly suitable for integration with participatory approaches that position teachers and supervisors as co-designers throughout the development process. In addition, the flexibility of the ADDIE model enables continuous formative

evaluation at every stage of product development, thereby allowing systematic and contextual revisions aligned with the actual needs of PAI teachers in the field.

Compared with other educational product development models, such as Borg and Gall's Research and Development model and the 4D model, ADDIE provides a practical and flexible framework for developing instructional products within limited institutional settings while maintaining systematic development procedures and continuous evaluation mechanisms. While Borg and Gall emphasize extensive field testing and multiple revision cycles, and the 4D model focuses primarily on instructional material development, ADDIE provides a balanced structure that combines systematic development, continuous evaluation, and adaptability. These characteristics make ADDIE particularly suitable for the present study, which aims to develop and validate a participatory-based module through iterative collaboration among teachers and supervisors within the MGMP forum.

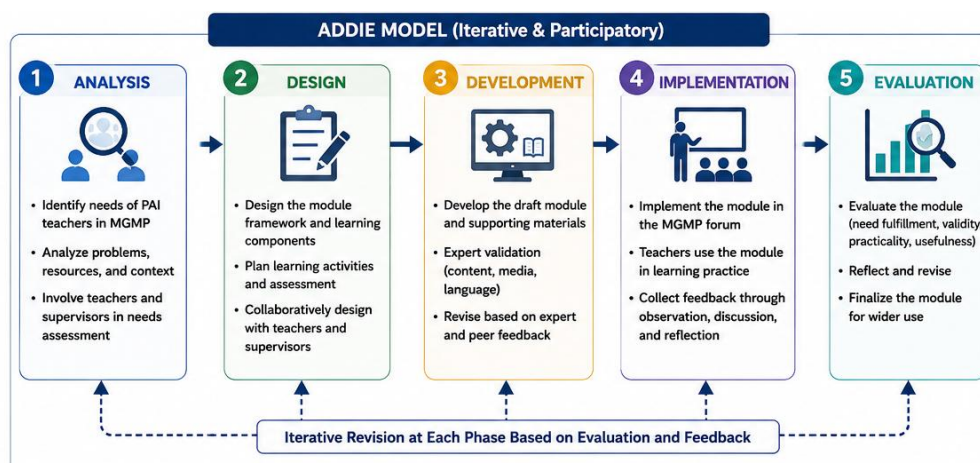


Figure 1. Research and development (r&d) framework using the addie model integrated with a participatory approach

Figure 1 illustrates the research framework employed in this study, which integrates the Research and Development (R&D) method with the ADDIE model and a participatory approach. The development process consists of five interconnected stages: Analysis, Design, Development, Implementation, and Evaluation. Throughout each stage, PAI teachers and supervisors actively participate as collaborators in identifying needs, designing learning components, validating the module, implementing the product, and evaluating its effectiveness. The iterative nature of the ADDIE model enables continuous formative evaluation and revision, ensuring that the developed module remains relevant to actual educational needs. Compared with other development models, ADDIE provides a practical, systematic, and flexible structure that supports collaborative product development while maintaining rigorous evaluation procedures. This integration aims to produce a valid, practical, and useful participatory-based module that can enhance the quality of PAI instruction within the MGMP forum (Lam, 2023).

The Analysis phase was conducted through semi-structured interviews, observations of instructional documents, and preliminary discussions with PAI teachers to identify the needs related to authentic assessment development within the implementation of the Merdeka Curriculum. The Design phase involved constructing the module framework, determining learning outcomes, developing Focus Group Discussion (FGD) formats, and designing examples of authentic PAI assessment instruments. During the Development phase, the initial module prototype was created and subsequently validated by experts to obtain substantive and technical feedback. The Implementation phase was carried out through a small-group trial within the MGMP PAI forum

involving five structured FGD sessions. Finally, the Evaluation phase was conducted formatively throughout every stage of ADDIE through revisions based on validators' suggestions and users' responses.

The product developed in this study was a participatory-based module for developing authentic assessment instruments in PAI learning. The module consisted of five instructional units and was designed as a practical guide for PAI teachers in developing authentic assessment instruments collaboratively through the MGMP forum. The module was complemented with examples of authentic tasks, assessment rubrics, FGD worksheets, and examples of attitude, knowledge, and skills assessment instruments.

The research participants consisted of 35 PAI teachers from 23 senior secondary schools (SMA/SMK) who were members of the MGMP PAI forum in Nagan Raya Regency, Aceh Province. Participants were selected purposively based on the following criteria: actively teaching PAI at the secondary school level, implementing the Merdeka Curriculum, and willing to participate actively throughout the development process. During the analysis phase, 10 teachers served as initial informants. In the practicality testing phase (small-group trial), 20 teachers participated as module users in five FGD sessions. Most participating teachers had more than five years of teaching experience and had already implemented the Merdeka Curriculum in their respective schools. The implementation sessions were conducted gradually through the MGMP forum over five meetings, each lasting approximately 120–150 minutes.

Product validation involved three validators: (1) an Educational Assessment Expert specializing in educational evaluation and measurement, (2) an Islamic Religious Education Learning Expert with expertise in PAI content and pedagogy, and (3) a Senior Educational Supervisor experienced in supervising secondary schools. The validation instrument employed a four-point Likert scale covering aspects of content validity, construct validity, language clarity, and practicality. The module practicality was assessed through teachers' response questionnaires using the same four-point Likert scale, encompassing five aspects: ease of use, language clarity, suitability to teachers' needs, relevance of examples, and feasibility of participatory FGD activities. In addition to validation sheets and teacher response questionnaires, this study also utilized interview guidelines, observation sheets of FGD activities, and participants' reflective notes to strengthen the qualitative data.

The validation instruments were developed based on indicators related to the alignment of materials with PAI learning outcomes, module organization, language clarity, relevance of authentic assessment examples, and feasibility of participatory activities. Product validity was analyzed using percentage-based feasibility scoring techniques by comparing empirical scores with the maximum possible scores from all validators. Quantitative data were analyzed descriptively using means and percentages to determine product validity and practicality categories. Meanwhile, qualitative data were analyzed through thematic processes involving data reduction, categorization, data display, and conclusion drawing.

The feasibility criteria referred to the score conversion proposed by [Walland, \(2022\)](#): percentages ranging from 81–100% were categorized as Very Valid/Very Practical, 61–80% as Valid/Practical, and below 61% as Invalid/Impractical. The product was considered sufficiently valid and eligible for implementation if the average validation percentage reached at least 61%. Qualitative data obtained from validators' notes and teachers' reflections were analyzed using the interactive model of Matthew B. Miles and A. Michael Huberman, including data reduction, data presentation, and conclusion drawing, in order to complement and strengthen the quantitative findings.

All participants were informed about the objectives of the study and participated voluntarily throughout the development process. Participants' identities were kept confidential and used solely for academic research purposes.

RESULT AND DISCUSSION

The Analysis phase revealed a concerning picture regarding the condition of PAI assessment practices in Nagan Raya Regency. Through semi-structured interviews with 10 PAI teachers from 23 schools and observations of instructional planning documents, the study found that more than 80% of teachers still relied exclusively on written cognitive tests as their primary assessment instrument. Assessments targeting affective domains which should constitute the core objective of PAI learning as well as psychomotor assessments were largely absent from the reviewed instructional documents. Although teachers had been exposed to key concepts of the Merdeka Curriculum, such as Learning Outcomes (*Capaian Pembelajaran/CP*), Learning Objective Flow (*Alur Tujuan Pembelajaran/ATP*), and Teaching Modules (*Modul Ajar*), their assessment practices remained largely conventional, focusing primarily on end-of-semester written examinations as the sole evidence of student achievement.

Table 1. Summary of needs analysis findings among pai teachers

Analysis Aspect	Findings	Implications
Assessment Practices	More than 80% of teachers assessed only cognitive aspects through written tests	Diversification of assessment instruments is needed
Understanding of Authentic Assessment	Teachers lacked comprehensive understanding of authentic assessment concepts and procedures	Practical CP-based guidance is required
Rubric Development	Assessment rubrics had not been systematically developed	Training in rubric construction is needed
Analysis of Learning Outcomes	Teachers had difficulty translating CP into authentic assessment instruments	Strengthening CP understanding is necessary
Institutional Support	No specific module for authentic PAI assessment was available	A standardized assessment module is needed

These findings align with previous studies conducted by Widiastuti and Basuki (2019) in East Java, as well as research by Arifin (2022), which consistently reported that limited understanding of competency-based assessment constitutes a structural challenge rather than merely an individual weakness. Such limitations are strongly associated with the absence of practical assessment guidelines and the lack of relevant professional development forums. The local context of Nagan Raya adds another layer of complexity to this issue. No standardized module specifically designed for authentic PAI assessment had previously been available, while the existing MGMP forums had not yet been optimized as collaborative spaces for developing teachers' assessment competencies.

From an Islamic perspective, this gap also carries important ethical implications. The Qur'an explicitly emphasizes justice in judgment and evaluation, as reflected in Surah An-Nisa verse 58 (*wa idza hakamtum bain al-nas an takhumu bi al-'adl*). If assessment instruments measure only a small portion of students' actual competencies while neglecting affective and psychomotor dimensions the essential goals of PAI learning then the assessment process itself cannot fully

embody the principle of justice. Consequently, improving PAI teachers' competencies in authentic assessment is not merely a technical-pedagogical concern but also an ethical and religious responsibility.

One particularly important phenomenon identified in this analysis phase was what may be described as the “pseudo-implementation” of the Merdeka Curriculum. Administratively, all participating schools had formally adopted the curriculum; CP documents and teaching modules were already available. However, at the level of assessment practice, conventional approaches continued to dominate. This phenomenon is not unique to Nagan Raya. Vlachou et al. (2025), in their comparative study, found that resistance to changes in assessment practices is not necessarily caused by teachers' unwillingness, but rather by the absence of adequate scaffolding in the form of practical guidance and collaborative reflection forums capable of bridging curriculum policy and classroom implementation. This study seeks to address precisely this missing gap.

The Design phase resulted in the development of a module consisting of five hierarchically organized and interconnected units, each designed to be completed within a structured Focus Group Discussion (FGD) session lasting approximately 120–150 minutes. The five units comprised: (1) Analysis of PAI Learning Outcomes (*Capaian Pembelajaran/CP*); (2) Authentic Task Design; (3) Development of Assessment Instruments; (4) Construction of Assessment Rubrics; and (5) Classroom Implementation of Authentic Assessment. This design explicitly adopted Jerome Bruner's (1960) spiral curriculum framework. Unit I functioned as the conceptual foundation, ensuring that teachers clearly understood the direction and objectives of assessment before designing assessment strategies themselves. The subsequent units progressively developed teachers' competencies, moving from conceptual understanding (Unit II) toward technical application (Units III–IV), and ultimately to comprehensive classroom implementation (Unit V).

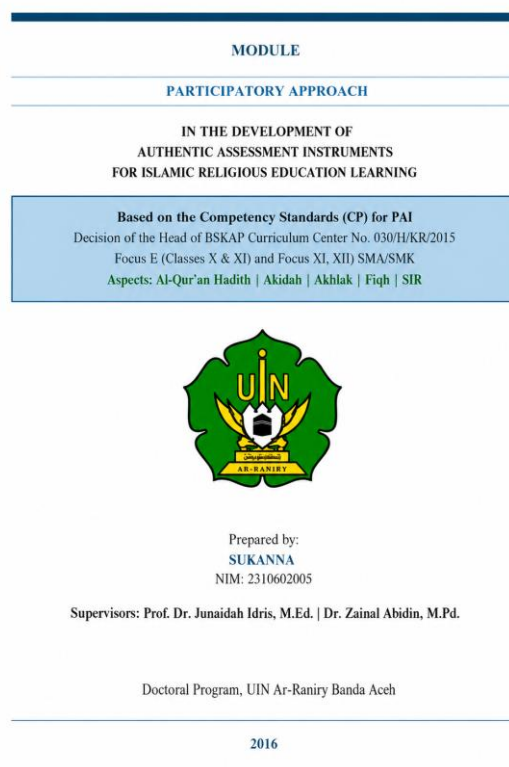


Figure 1. Cover of the authentic assessment instrument development module for pai

The product developed in this study was a participatory-based module for developing authentic assessment instruments in PAI learning for senior secondary school teachers

(SMA/SMK). The module was developed based on the 2025 PAI Learning Outcomes framework and designed as a practical guide for developing authentic assessment practices through the MGMP forum. The module was produced in both printed and digital formats and was complemented with worksheets, sample instruments, and FGD activity guidelines that could be implemented individually or collaboratively within MGMP activities.

Table 2. Structure of the developed module

Unit	Material Focus	Main Activities
Unit 1	Learning Outcomes Analysis	Identification of assessment elements
Unit 2	Authentic Tasks	Designing context-based authentic tasks
Unit 3	Assessment Instruments	Developing assessment formats
Unit 4	Rubrics	Constructing scoring indicators
Unit 5	Implementation	Simulating classroom assessment

The module structure was intentionally designed in a hierarchical and progressive manner, beginning with teachers' understanding of learning outcomes and culminating in the practical implementation of authentic assessment in classroom settings.

Several design variables appear to have played a particularly important role in shaping the outcomes of the module development process. These include the alignment of module content with PAI Learning Outcomes (CP), the integration of participatory FGD activities, and the use of contextual assessment examples relevant to teachers' classroom experiences. Among these variables, the participatory FGD mechanism was considered the most influential, as it enabled continuous interaction, collaborative problem-solving, and collective reflection among teachers and supervisors. These factors contributed substantially to the high practicality ratings observed during the implementation phase.

The participatory approach was operationalized through structured FGD mechanisms implemented in every unit following David A. Kolb's (1984) experiential learning cycle. Each session included conceptual introductions by facilitators (educational supervisors), open discussions based on real classroom cases, guided practice in developing assessment instruments, and collective reflection activities. Teachers and supervisors were positioned not as passive recipients of knowledge, but as active co-constructors collaboratively designing, discussing, and refining authentic assessment instruments together.

This positioning aligns closely with the principles of effective Professional Learning Communities (PLC) proposed by Richard DuFour et al. (2010), particularly the emphasis on collective orientation toward student learning, structured collaboration, and commitment to measurable outcomes. From an andragogical perspective, the module design also fulfilled the principles proposed by Malcolm Knowles (1980), which suggest that adult learners in this case teachers learn most effectively when instructional materials are perceived as directly relevant to their professional needs and immediately applicable in practice. Therefore, the selection of the MGMP forum as the implementation arena was not merely pragmatic, but theoretically grounded in the understanding that existing professional communities already possess essential social capital, including mutual trust, shared professional contexts, and collective commitment three critical prerequisites for successful participatory professional development (Reason & Bradbury, 2008).

An additional aspect that deserves attention is the ideological dimension of this design within the context of Islamic education. Participatory approaches rooted in Paulo Freire's (2000) tradition of critical pedagogy may initially appear incompatible with traditional Islamic educational

paradigms that tend to position teachers as the sole authority of knowledge. However, the Islamic principle of *shura* (consultation), which is explicitly emphasized in the Qur'an (Q.S. Asy-Syura: 38), provides a strong theological foundation for collaborative professional development practices. Thus, the MGMP forum designed as a professional *shura* arena for PAI teachers does not merely adopt Western participatory theories, but also embodies Islamic values concerning collective decision-making grounded in wisdom (*hikmah*) and deliberation (*musyawarah*). This integration constitutes one of the novel dimensions of the present study that distinguishes it from conventional assessment instrument development research.

Table 3. Specifications of the developed module

Aspect	Description
Target Users	PAI Teachers in Senior Secondary Schools
Basis	PAI Learning Outcomes (CP) 2025
Approach	MGMP-Based Participatory Approach
Assessment Types	Performance, Project, Portfolio
Assessment Domains	Cognitive, Affective, Psychomotor

The five units incorporated within the module also implicitly reflected the complex taxonomy of competencies required of PAI teachers. Unit I (CP Analysis) emphasized professional competence in understanding curriculum substance; Unit II (Authentic Task Design) activated pedagogical competence in designing meaningful learning experiences; Units III and IV (Instrument and Rubric Development) strengthened technical-pedagogical competencies that had previously represented teachers' most significant weaknesses; while Unit V (Assessment Implementation) required social and leadership competencies in communicating assessment standards fairly and transparently to students. Consequently, the module does not merely train a single technical competency, but instead seeks to establish an integrated ecosystem of PAI teacher competencies. This approach is highly consistent with E. Mulyasa's (2013) perspective that teacher competence should be understood as a holistic and multidimensional construct.

The module concept map further illustrates the interconnections among units in fostering PAI teacher competencies progressively and integratively. This structure was intentionally designed to ensure that teachers not only understand the conceptual foundations of authentic assessment, but are also capable of implementing such assessment practices effectively within real classroom contexts.

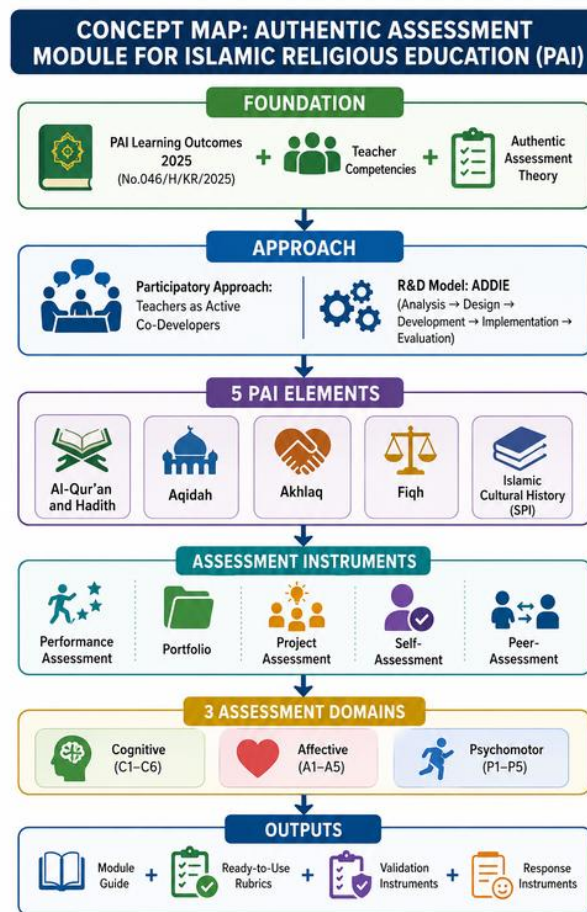


Figure 2. Concept map of the authentic pai assessment module

The Development phase resulted in a module product that subsequently underwent an expert judgment process involving three validators with complementary areas of expertise. Overall, the validation results demonstrated highly satisfactory outcomes, with an average validity percentage of 85%, categorized as *Very Valid*.

Table 4. Expert validation results of the authentic pai assessment module

No	Validator	Area of Expertise	Score (%)	Category
1	Educational Assessment Expert	Educational Evaluation & Assessment	80%	Valid
2	PAI Learning Expert	Islamic Religious Education	85%	Very Valid
3	Senior Educational Supervisor	Educational Supervision	90%	Very Valid
Average			85%	Very Valid

Source: *Expert Validation Data, 2025*

The distribution pattern of validation scores provides insights beyond the average numerical value alone. Interestingly, the highest score was assigned by the Senior Educational Supervisor (90%) rather than the academic experts. This finding indirectly confirms the module's contextual relevance: a supervisor who routinely engages with the realities of teachers' capacities in schools perceived the module as genuinely applicable rather than merely academically sophisticated. This finding supports the argument proposed by Thomas J. Sergiovanni and Robert J. Starratt (2007), who emphasized that effective instructional supervision requires practical instruments that can be integrated into teacher professional development programs.

Although the revision notes provided by the three validators did not alter the feasibility status of the product, they offered substantial directions for refinement. The recommendation from the PAI Learning Expert regarding the need for more specific examples of spiritual and social attitude assessment rubrics reflects a complexity widely acknowledged in the literature. Heidi Andrade and Gregory J. Cizek (2010) argued that assessing affective domains, particularly those related to beliefs and religious values, requires highly contextual and culturally sensitive instruments. This complexity also represents one of the unresolved research gaps in previous studies concerning the development of PAI assessment instruments.

Table 5. Product revisions based on validators' feedback

Validator	Feedback	Product Revision
Assessment Expert	Affective rubrics need further clarification	Added attitude observation indicators
PAI Expert	Instrument examples need to be more contextual	Added contextual examples for senior secondary PAI learning
Supervisor	Language should be more operational	Simplified technical terminology

It is important to critically acknowledge that expert validation, as conducted in this study, possesses inherent limitations. Rick Stiggins (1991) warned that validity judgments from academic experts do not always correlate directly with practical relevance in classroom contexts. For this reason, expert validation in this study was not treated as a standalone measure, but was complemented by practicality data obtained from teachers as the direct users of the module. The combination of these two perspectives academic and practical resulted in a more comprehensive and accountable evaluation of product quality. Furthermore, the involvement of a Senior Educational Supervisor as one of the validators introduced a perspective that is often absent from conventional development research: the perspective of a professional mentor who possesses direct access to the realities of teachers' capacities in schools.

Product revisions based on validators' feedback were conducted before the module entered the implementation phase, in accordance with the formative evaluation principles proposed by Lam, (2023). According to Scriven, the primary purpose of formative evaluation is to improve a product during the development process rather than merely evaluate it after completion. This principle was consistently implemented throughout every phase of the ADDIE model in the present study, ensuring that each developmental decision was evidence-based rather than assumption-driven.

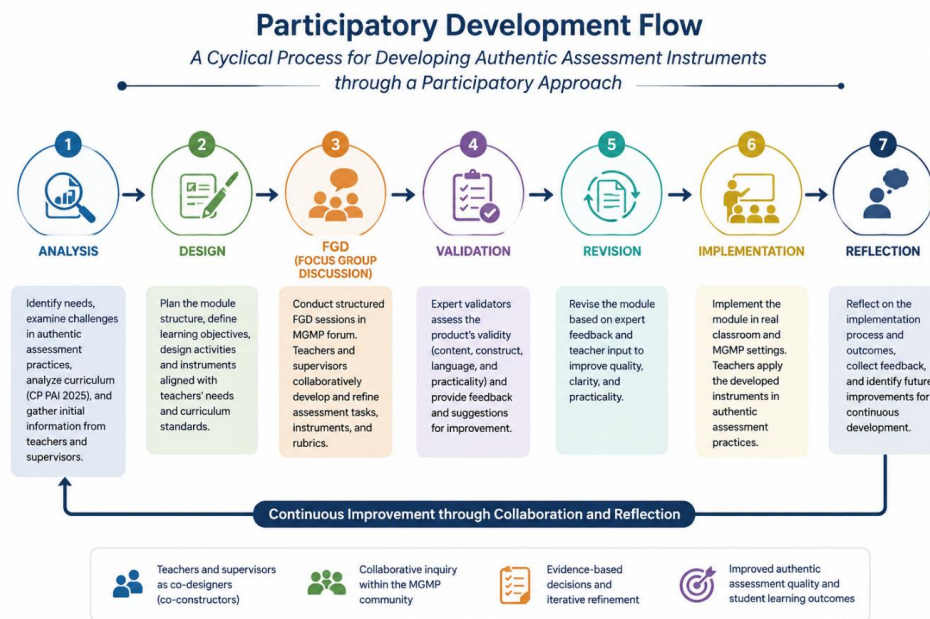


Figure 3. Participatory development flow of the authentic assessment module

The figure illustrates the participatory and iterative development process implemented in this study, integrating the ADDIE model with collaborative FGD activities involving teachers and supervisors within the MGMP forum. As shown in Figure 3, the implementation phase was conducted not in a linear manner, but iteratively through cycles of discussion, validation, revision, and reflection.

The Implementation phase was carried out through a small-group trial involving 20 PAI teachers from various secondary schools in Nagan Raya Regency within the MGMP forum. The implementation process consisted of five FGD sessions, each focusing on one module unit and facilitated by the Senior Educational Supervisor acting as a co-facilitator. Overall, the module achieved an average practicality percentage of 80%, categorized as *Practical*.

Table 6. Practicality analysis results based on pai teachers' responses

No	Practicality Aspect	Mean Score	Percentage (%)	Category
1	Ease of module use	3.2	80%	Practical
2	Clarity of language and material presentation	3.2	80%	Practical
3	Suitability to PAI teachers' needs	3.3	82%	Very Practical
4	Relevance of examples and PAI contextual illustrations	3.2	80%	Practical
5	Feasibility of participatory activities (FGD)	3.3	82%	Very Practical
Average		3.24	80%	Practical

Source: Teachers' Response Data, 2025

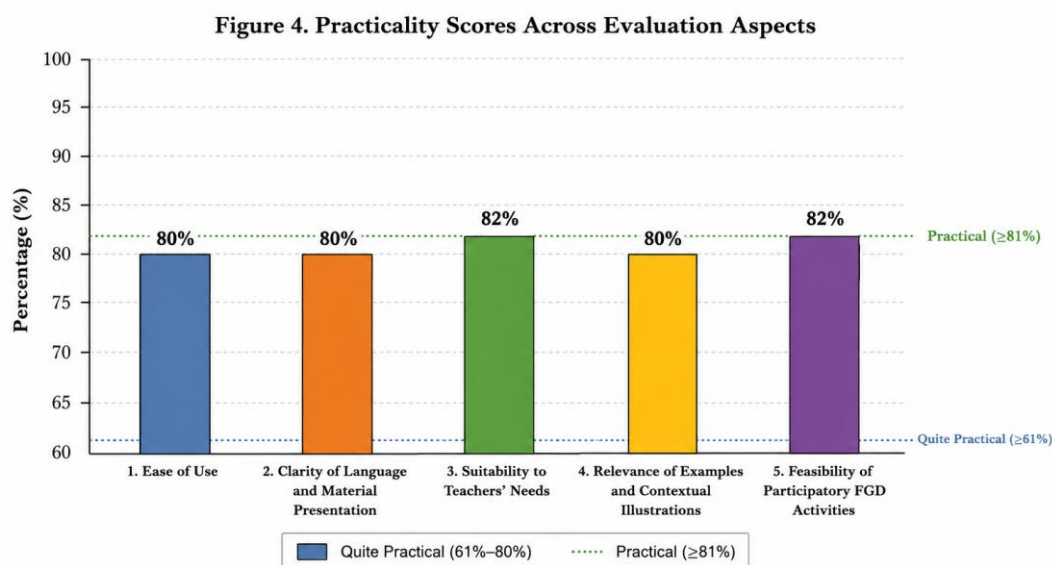


Figure 4. Practicality scores across evaluation aspects

Figure 4 presents the practicality scores across the five evaluation aspects of the developed module. Overall, the module demonstrated a high level of practicality, with scores ranging from 80% to 82%. The highest scores were obtained for suitability to teachers' needs (82%) and feasibility of participatory FGD activities (82%), indicating that the module successfully addressed the professional needs of PAI teachers while supporting collaborative learning processes within the MGMP forum. Meanwhile, ease of use (80%), clarity of language and material presentation (80%), and relevance of examples and contextual illustrations (80%) were also rated positively, reflecting the module's usability and contextual relevance for classroom practice. These findings suggest that the participatory design adopted in the module not only facilitated teachers' understanding of authentic assessment concepts but also promoted active engagement in the development of assessment instruments.

The practicality results are consistent with previous studies highlighting the effectiveness of collaborative professional development in enhancing teachers' assessment competencies. However, unlike earlier studies that primarily focused on assessment implementation or instrument development, the present study integrated participatory FGD activities and the ADDIE model into a structured module development process. This integration contributed to the high practicality ratings and demonstrates the potential of MGMP-based collaboration as a sustainable professional learning mechanism for PAI teachers. This finding also supports Winstone, (2022), who found that structured teacher development programs can significantly strengthen assessment literacy by helping teachers overcome practical and systemic barriers in assessment implementation.

The two aspects receiving the highest percentages suitability to PAI teachers' needs (82%) and feasibility of participatory FGD activities (82%) carry important theoretical implications. The high-rating for-content suitability confirms that the participatory needs analysis conducted during the Analysis phase successfully generated accurate information regarding teachers' actual professional needs. Consequently, the resulting product addressed genuinely experienced problems rather than researcher assumptions. Meanwhile, the high score for FGD feasibility validates the central argument of this study: the MGMP forum is not only logistically capable of accommodating participatory approaches, but also constitutes a socially and culturally supportive ecosystem for collaborative professional learning.

The high enthusiasm demonstrated by participants throughout the five implementation sessions can be explained through the self-determination theory proposed by (Lavoie, 2022). The participatory FGD design inherently facilitated the fulfillment of three fundamental human needs: competence (feeling capable of completing instrument development tasks), autonomy (having a voice in determining the characteristics of the instruments being developed), and relatedness (feeling meaningfully connected to a professional community). When these three needs are fulfilled simultaneously, they create optimal motivational conditions for learning and professional transformation.

Qualitative data obtained from observations and end-of-session reflections further strengthened the quantitative findings. Teachers reported that the module helped them better understand assessment instrument development procedures that had previously seemed confusing, while the collaborative FGD format made the learning process feel lighter and more engaging. One critical issue that emerged, however, was the need for more specific examples of spiritual attitude assessment instruments for various PAI topics. This concern was consistent with feedback previously provided by the PAI expert validator during the Development phase and further reinforces the urgency of developing a comprehensive bank of affective assessment instruments as an agenda for future research.

An unexpected finding emerging from the implementation phase was that, despite the overall positive practicality ratings, teachers consistently expressed difficulties in developing spiritual attitude assessment instruments. This challenge persisted even after participating in the module activities and receiving collaborative support through the FGD sessions. The finding suggests that affective assessment in Islamic education may involve complexities that extend beyond procedural knowledge and require deeper conceptual, contextual, and value-based considerations. This result indicates that future professional development programs should place greater emphasis on the design and implementation of affective assessment instruments, particularly those intended to assess students' internalization of Islamic values.

Another interesting phenomenon observed during the implementation process was the changing dynamic between teachers and supervisors during FGD sessions. Within the Indonesian educational context, teacher supervisor relationships are often characterized by rigid hierarchies, where supervisors act primarily as evaluators rather than facilitators. The participatory design implemented in this module explicitly deconstructed this hierarchy: supervisors functioned as co-facilitators rather than instructors, while teachers were positioned as co-constructors rather than passive participants. This positional shift created a psychologically safer environment for sharing difficulties, asking questions, and experimenting with new assessment instruments. Although seemingly simple, this transformation actually represents a profound cultural shift within the professional development ecosystem of PAI teachers one that could not easily be achieved through conventional top-down training programs, regardless of their quality.

When contrasted with the findings of (Hsiao, 2023), who conducted a systematic review of Participatory Action Research in PAI teacher competency development, the present study both confirms and extends their findings. While they found that participatory approaches effectively increased teacher engagement, they also noted limitations regarding the transferability of products due to their strong attachment to local contexts. The present study addressed this limitation by constructing a sufficiently systematic module structure that could be transferred to other contexts while still providing contextualization guidelines adaptable to local characteristics. This constitutes one of the significant methodological contributions of the study.

The module practicality percentage of 80% exceeded the minimum threshold of 61% established according to Widoyoko (2009) and consistently remained above the average practicality levels reported in similar Islamic education assessment instrument development studies. Nienke Nieveen (1999) argued that the practicality of a learning product is determined by the extent to which users can implement it easily in accordance with the developer's intentions. The findings of this study convincingly demonstrate that the developed module fulfilled these practicality criteria.

These findings are generally consistent with previous studies on participatory professional development and authentic assessment in Islamic education, which reported that collaborative learning activities improve teachers' assessment competencies and confidence in implementing curriculum reforms. However, unlike previous studies that primarily focused on assessment implementation or instrument development, the present study produced a structured participatory-based module integrating the ADDIE model and MGMP collaboration. This distinction highlights the study's contribution to bridging the gap between assessment theory and classroom practice. Qualitative reflections from participating teachers further supported these findings, indicating that the module was practical, relevant, and directly applicable to authentic assessment development in PAI learning.

Evaluation in this study was formative and iterative in nature. Rather than being conducted only once at the end of the process, evaluation functioned as a continuous improvement mechanism embedded throughout every phase of the ADDIE model. This approach aligns with Michael Scriven's (1967) formative evaluation concept and was consistently operationalized through design revisions during the Design phase, substantive revisions based on feedback from the three validators during the Development phase, and technical revisions based on feedback from 20 teachers during the Implementation phase. This comprehensive formative evaluation cycle ensured that the final product was not merely "validated," but rather a product that had undergone multilayered testing from academic, Islamic-substantive, and practical-field perspectives.

Cumulatively, the five ADDIE phases produced a product that was both valid (85%) and practical (80%), while the participatory mechanism proved effective in fostering active involvement among teachers and supervisors. Kırkıç et al. (2025), in their study on teacher professional development competency scales, found that the effectiveness of professional development programs is highly dependent on design quality. Programs that address actual needs, provide constructive feedback, and encourage collaboration tend to generate more significant and sustainable competency improvement. The present study provides empirical evidence supporting this claim within the context of Indonesian Islamic education.

From a broader perspective, this study offers important methodological insights regarding how research and development in Islamic education can be conducted in more contextual and participatory ways. Too many studies on PAI assessment instrument development have produced technically valid products that were never genuinely implemented in classrooms because they were developed without meaningful involvement from teachers as primary users. Michael Fullan (2007) referred to this phenomenon as the *implementation gap* the discrepancy between well-designed policies or products on paper and their weak implementation in practice. The participatory approach adopted in this study directly addressed the root causes of this implementation gap. By involving teachers as active participants in the creation process, their sense of ownership toward the product increased substantially, thereby enhancing the likelihood of sustainable implementation.

The practical implications of these findings extend beyond the immediate context of this study. The developed module can be implemented through MGMP forums as a structured professional development resource to support teachers in designing authentic assessment

instruments aligned with the Merdeka Curriculum. Educational supervisors may utilize the module as a mentoring tool during academic supervision activities, while school administrators can integrate it into in-service teacher development programs. Furthermore, the participatory design of the module makes it adaptable to different educational contexts, enabling other regions and subject areas to adopt similar collaborative approaches for strengthening assessment literacy and instructional quality.

The findings also carry broader implications for the field of Islamic education research. First, they provide empirical support for the integration of participatory professional development approaches into authentic assessment training programs. Second, the study demonstrates that assessment literacy among PAI teachers can be strengthened more effectively when professional learning is embedded within existing collaborative structures such as MGMP forums. Finally, the study contributes to the growing body of literature emphasizing that the successful implementation of curriculum reform depends not only on policy adoption but also on the availability of contextualized professional support systems that bridge the gap between educational theory and classroom practice (Salinas-Navarro, 2024).

Nevertheless, one methodological limitation must be acknowledged transparently. This study has not yet measured the third dimension of educational product quality proposed by Nieveen (1999), namely effectiveness. Measuring effectiveness would require pre-test/post-test designs or longitudinal case studies, which remain important agendas for future research.

Walter Dick and Lou Carey (2015) clearly distinguished between *small-group evaluation* and *field evaluation*, emphasizing that effectiveness claims become scientifically valid only after field evaluations are conducted under representative real-world conditions. Similarly, Bruce Joyce and Beverly Showers (2002), in their seminal study, found that training without classroom mentoring produced skill transfer rates of only approximately 5–10%, whereas combining training with coaching increased transfer rates to 80–95%. These findings provide a strong argument for the necessity of post-implementation mentoring programs as an inseparable component of comprehensive professional development design a highly concrete recommendation for future research as well as for policymaking in the professional development of PAI teachers.

CONCLUSION

This research and development study successfully produced a Participatory Approach Module for Developing Authentic Assessment Instruments in Islamic Religious Education (PAI) for secondary schools, consisting of five hierarchical and contextual units. Through a systematic development process based on the ADDIE model integrated with a participatory approach, the module achieved an average expert validity score of 85% (*Very Valid*) and an average practicality score of 80% (*Practical*), fulfilling two of the three quality dimensions commonly required in educational development research.

The participatory mechanism operationalized through structured Focus Group Discussions (FGDs) within the MGMP forum proved effective in positioning teachers and supervisors as authentic co-constructors. As a result, the developed assessment instruments were contextual, relevant, and practically applicable, rather than products created through top-down approaches that are often vulnerable to weak implementation in classroom practice.

The contributions of this study are twofold. Theoretically, this study introduces an ADDIE-participatory integration model as a novel approach in the development of PAI assessment instruments, reinforcing the argument that teacher empowerment through community-based professional collaboration (MGMP) constitutes a more effective strategy than incidental and top-

down training programs. Practically, the developed module provides a ready-to-use resource for supporting authentic assessment development among PAI teachers and has the potential to be adapted and implemented in broader educational contexts through MGMP-based professional development initiatives.

These findings are consistent with previous studies emphasizing the importance of collaborative professional development and participatory approaches in strengthening teachers' competencies and supporting curriculum reform. However, the present study extends existing research by demonstrating that the integration of the ADDIE model with MGMP-based participatory collaboration can produce not only valid and practical assessment resources but also a sustainable framework for professional learning within Islamic education contexts. In the long term, the widespread adoption of this approach has the potential to enhance assessment literacy among PAI teachers, improve the quality of authentic assessment practices, and contribute to the successful implementation of the Merdeka Curriculum through stronger alignment between curriculum objectives, assessment processes, and classroom learning experiences.

For school supervisors, this module offers a concrete instrument for integrating assessment competency development into regular academic supervision programs. For educational authorities and the Indonesian Ministry of Religious Affairs (*Kementerian Agama/Kemenag*), the findings of this study reinforce the urgency of policies that optimize MGMP forums as sustainable platforms for teacher professional development within the era of the Merdeka Curriculum.

DECLARATION OF AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this manuscript, the author(s) used Google Gemini to assist in improving grammar, language quality, and overall readability of the text. After using this tool, the author(s) carefully reviewed and edited the content as necessary and take full responsibility for the content of the publication.

AUTHORS' CONTRIBUTION

Author 1: Conceptualization; Data curation; Investigation; Project administration.

Author 2: Conceptualization; Methodology; Supervision.

Author 3: Formal analysis; Methodology; Supervision; Validation.

DECLARATION OF COMPETING INTEREST

The author(s) declare that there are no financial, professional, or personal relationships that could be perceived as influencing the results or interpretation of this study.

REFERENCES

- Broadbent, J. (2023). Beyond emergency remote teaching: Did the pandemic lead to lasting change in university courses? *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00428-z>
- Cai, M. (2022). Boosted photocatalytic antibiotic degradation performance of Cd_{0.5}Zn_{0.5}S/carbon dots/Bi₂WO₆ S-scheme heterojunction with carbon dots as the electron bridge. *Separation and Purification Technology*, 300(Query date: 2026-06-05 15:31:23). <https://doi.org/10.1016/j.seppur.2022.121892>
- Gao, G. (2023). CNN-Bi-LSTM: A Complex Environment-Oriented Cattle Behavior Classification Network Based on the Fusion of CNN and Bi-LSTM. *Sensors*, 23(18). <https://doi.org/10.3390/s23187714>

- Heidari, A. (2024). Deepfake detection using deep learning methods: A systematic and comprehensive review. *Wiley Interdisciplinary Reviews Data Mining and Knowledge Discovery*, 14(2). <https://doi.org/10.1002/widm.1520>
- Hou, R. (2024). Automated Assessment of Encouragement and Warmth in Classrooms Leveraging Multimodal Emotional Features and ChatGPT. *Lecture Notes in Computer Science Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics*, 14829(Query date: 2026-06-05 15:31:23), 60–74. https://doi.org/10.1007/978-3-031-64302-6_5
- Hsiao, Y. P. (2023). Developing a framework to re-design writing assignment assessment for the era of Large Language Models. *Learning Research and Practice*, 9(2), 148–158. <https://doi.org/10.1080/23735082.2023.2257234>
- James, M. (2022). Collaborative case-based learning with programmatic team-based assessment: A novel methodology for developing advanced skills in early-years medical students. *BMC Medical Education*, 22(1). <https://doi.org/10.1186/s12909-022-03111-5>
- Jiang, T. (2022). Awe Motivates Authentic-Self Pursuit via Self-Transcendence: Implications for Prosociality. *Journal of Personality and Social Psychology*, 123(3), 576–596. <https://doi.org/10.1037/pspi0000381>
- Lam, R. (2023). E-Portfolios: What We Know, What We Don't, and What We Need to Know. *Relc Journal*, 54(1), 208–215. <https://doi.org/10.1177/0033688220974102>
- Lavoie, J. A. A. (2022). Developing Community Co-designed Scenario-Based Training for Police Mental Health Crisis Response: A Relational Policing Approach to De-escalation. *Journal of Police and Criminal Psychology*, 37(3), 587–601. <https://doi.org/10.1007/s11896-022-09500-2>
- Li, B. (2022). Blindly Assess Quality of In-The-Wild Videos via Quality-Aware Pre-Training and Motion Perception. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(9), 5944–5958. <https://doi.org/10.1109/TCSVT.2022.3164467>
- Li, L. (2022). Blind Image Quality Index for Authentic Distortions with Local and Global Deep Feature Aggregation. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(12), 8512–8523. <https://doi.org/10.1109/TCSVT.2021.3112197>
- Mate, K. (2022). Considerations and strategies for effective online assessment with a focus on the biomedical sciences. *FASEB Bioadvances*, 4(1), 9–21. <https://doi.org/10.1096/fba.2021-00075>
- Pan, Z. (2022). DACNN: Blind Image Quality Assessment via a Distortion-Aware Convolutional Neural Network. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(11), 7518–7531. <https://doi.org/10.1109/TCSVT.2022.3188991>
- Perkins, D. (2022). Changes in mental health, wellbeing and personality following ayahuasca consumption: Results of a naturalistic longitudinal study. *Frontiers in Pharmacology*, 13(Query date: 2026-06-05 15:31:23). <https://doi.org/10.3389/fphar.2022.884703>
- Revell, T. (2024). ChatGPT versus human essayists: An exploration of the impact of artificial intelligence for authorship and academic integrity in the humanities. *International Journal for Educational Integrity*, 20(1). <https://doi.org/10.1007/s40979-024-00161-8>
- Salinas-Navarro, D. E. (2024). Designing experiential learning activities with generative artificial intelligence tools for authentic assessment. *Interactive Technology and Smart Education*, 21(4), 708–734. <https://doi.org/10.1108/ITSE-12-2023-0236>
- Shi, J. (2024). Blind Image Quality Assessment via Transformer Predicted Error Map and Perceptual Quality Token. *IEEE Transactions on Multimedia*, 26(Query date: 2026-06-05 15:31:23), 4641–4651. <https://doi.org/10.1109/TMM.2023.3325719>
- Song, T. (2022). Blind Image Quality Assessment for Authentic Distortions by Intermediary Enhancement and Iterative Training. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(11), 7592–7604. <https://doi.org/10.1109/TCSVT.2022.3179744>
- Southwell, R. (2022). Challenges and Feasibility of Automatic Speech Recognition for Modeling Student Collaborative Discourse in Classrooms. *Proceedings of the 15th International*

- Conference on Educational Data Mining Edm 2022*, (Query date: 2026-06-05 15:31:23). <https://doi.org/10.5281/zenodo.6853109>
- Sun, W. (2023). Blind Quality Assessment for in-the-Wild Images via Hierarchical Feature Fusion and Iterative Mixed Database Training. *IEEE Journal on Selected Topics in Signal Processing*, 17(6), 1178–1192. <https://doi.org/10.1109/JSTSP.2023.3270621>
- Waland, E. (2022). E-portfolios in teaching, learning and assessment: Tensions in theory and praxis. *Technology Pedagogy and Education*, 31(3), 363–379. <https://doi.org/10.1080/1475939X.2022.2074087>
- Winstone, N. E. (2022). Discipline-specific feedback literacies: A framework for curriculum design. *Higher Education*, 83(1), 57–77. <https://doi.org/10.1007/s10734-020-00632-0>
- Zacometti, C. (2024). Authenticity assessment of ground black pepper by combining headspace gas-chromatography ion mobility spectrometry and machine learning. *Food Research International*, 179(Query date: 2026-06-05 15:31:23). <https://doi.org/10.1016/j.foodres.2024.114023>
- Zhou, J. (2022). Cohesive Multi-Modality Feature Learning and Fusion for COVID-19 Patient Severity Prediction. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(5), 2535–2549. <https://doi.org/10.1109/TCSVT.2021.3063952>

Copyright Holder :

© Susanna et al. (2026).

First Publication Right :

© International Journal of Educational Narratives

This article is under:

