BOCAH: Borneo Early Childhood Education and Humanity Journal
Volume 5 Nomor 2, July 2026
E-ISSN: 2807-7857, P-ISSN: 2807-9078

Development of an AI-Assisted Authentic Assessment Model to Enhance The Monitoring of Early Childhood Development

Kulsum Nur Hayati[1], Pria Dita Anis Wari[2]*, Dwi Enggal Wahyuni[3]
[1,2,3] Universitas Islam Negeri Sunan Kalijaga Yogyakarta, Indonesia

Received: June 14th, 2025; Revised: June 27th, 2026; Accepted: July 4th, 2026; Published: July 15th, 2026

Abstract

This study aimed to develop AI-assisted authentic assessment templates to support the monitoring of early childhood development. Preliminary findings from five early childhood education institutions in Bantul Regency, Special Region of Yogyakarta, Indonesia, revealed that teachers faced difficulties in implementing authentic assessment due to fragmented documentation and challenges in interpreting and narrating children's developmental records. To address these issues, this study employed a Research and Development (R&D) approach using the ADDIE model, involving teachers and school principals throughout the Analysis, Design, Development, Implementation, and Evaluation phases. The developed product consisted of structured, authentic assessment templates, including work assessment templates, observation checklists, anecdotal notes, and assessment rubrics integrated with NotebookLM to support AI-assisted analysis of children's developmental data. Expert validation yielded an average score of 4.4 out of 5.0 (88%), indicating that the developed assessment templates were feasible for implementation. The implementation results showed that the templates enabled teachers to document children's development more systematically, organize assessment data, and generate individualized narrative reports with AI assistance. Teachers remained responsible for reviewing, validating, and interpreting AI-generated outputs before using them to make decisions about appropriate stimulation strategies for children at school and at home.

Keywords: artificial intelligence, authentic assessment, developmental assessment, early childhood education

Abstrak

Penelitian ini bertujuan untuk mengembangkan template asesmen autentik berbantuan AI guna mendukung pemantauan perkembangan anak usia dini. Temuan awal dari lima lembaga PAUD di Kabupaten Bantul, Daerah Istimewa Yogyakarta, Indonesia, menunjukkan bahwa guru mengalami kesulitan dalam menerapkan asesmen autentik karena dokumentasi masih terfragmentasi serta adanya kendala dalam menafsirkan dan menyusun narasi perkembangan anak. Untuk menjawab permasalahan tersebut, penelitian ini menggunakan pendekatan Research and Development (R&D) dengan model ADDIE yang melibatkan guru dan kepala sekolah pada tahap Analysis, Design, Development, Implementation, dan Evaluation. Produk yang dikembangkan berupa template asesmen autentik terstruktur yang mencakup template penilaian hasil karya, lembar observasi, catatan anekdot, dan rubrik asesmen yang diintegrasikan dengan NotebookLM untuk mendukung analisis data perkembangan anak berbantuan AI. Hasil validasi ahli memperoleh skor rata-rata 4,4 dari 5,0 atau setara dengan 88%, yang menunjukkan bahwa template asesmen yang dikembangkan layak untuk diimplementasikan. Hasil implementasi menunjukkan bahwa template tersebut membantu guru mendokumentasikan perkembangan anak secara lebih sistematis, mengorganisasi data asesmen, dan menghasilkan laporan naratif perkembangan anak secara individual dengan bantuan AI. Guru tetap berperan dalam meninjau, memvalidasi, dan menafsirkan keluaran yang dihasilkan AI sebelum digunakan sebagai dasar pengambilan keputusan mengenai strategi stimulasi yang sesuai bagi anak, baik di sekolah maupun di rumah.

Kata kunci: artificial intelligence, asesmen autentik, asesmen perkembangan, pendidikan anak usia dini.

*** Correspondence Address**:
Email Address: priaditaaniswarii@gmail.com

A. Introduction

Early childhood education plays a crucial role as the initial foundation for the development of children's cognitive abilities, moral and religious values, social-emotional skills, language, and physical-motor skills (Atikah & Oktaviani, 2025). Monitoring children's holistic growth and development requires assessment activities that employ appropriate instruments capable of providing accurate information about children's developmental progress. Assessment not only functions to identify children's developmental achievements and detect developmental delays at an early stage but also serves as feedback for teachers and parents in supporting children's optimal development (Novianti et al., 2025). Therefore, authentic assessment becomes an essential component of early childhood education because it emphasizes continuous observation and documentation of children's learning experiences within natural contexts. In line with the Merdeka Curriculum, assessment should support holistic child development and provide meaningful information for educational decision-making (Khairunnisa Ulfadhilah, 2024).

However, the implementation of authentic assessment in Early Childhood Education (PAUD) institutions still faces various challenges. Assessment practices often rely heavily on teachers' subjectivity, while teachers frequently experience difficulties in documenting children's development and communicating assessment results to parents(Syafitri et al., 2025). Teachers also encounter time constraints and administrative burdens that limit their ability to conduct comprehensive observations of all children (Indalaillah & Rukayah, 2025). In addition, many teachers still have limited understanding of authentic assessment concepts and practices (Faizah, 2017). Consequently, assessment activities often become administrative routines rather than meaningful processes that support children's development.

Preliminary observations conducted in five early childhood education institutions in Bantul Regency, Yogyakarta, revealed several challenges in child development assessment practices. Teachers experienced difficulties in implementing authentic assessments that adequately reflected children's individual developmental processes. Assessment practices tended to be general, relied heavily on manual documentation, and often produced similar developmental descriptions despite children's diverse characteristics. In addition, teachers had limited knowledge of integrating digital technology, including Artificial Intelligence (AI), into assessment practices. Heavy administrative workloads also reduced opportunities for reflective analysis and the preparation of detailed developmental reports.

These challenges indicate the need for more practical and efficient approaches to authentic assessment. The integration of digital technology has the potential to support documentation, developmental monitoring, and communication with parents when implemented in accordance with children's developmental characteristics (Salim, 2022). Nevertheless, teachers remain responsible for interpreting assessment results and making pedagogical decisions. Despite these opportunities, existing studies have primarily emphasized digital documentation and technology integration in learning activities, while limited attention has been given to AI-assisted authentic assessment in early childhood education. Existing digital assessment systems generally focus on documentation and reporting rather than supporting data-driven developmental analysis.

The novelty of this study lies in the development of an AI-assisted authentic assessment model specifically designed for early childhood education. The model integrates authentic assessment instruments, structured assessment databases, and AI-supported developmental analysis using NotebookLM within the ADDIE framework while maintaining teachers as the primary pedagogical decision-makers. Interview findings from teachers in Bantul Regency revealed considerable challenges related to assessment

documentation and report preparation. Therefore, this study aims to develop an AI-assisted authentic assessment model that responds to these challenges and supports more objective, efficient, and comprehensive assessment practices in the digital era.

B. Literature Review

1. Authentic Assessment in Early Childhood Education

Authentic assessment is an essential component of early childhood education because it emphasizes the continuous observation of children's developmental processes within natural learning contexts. Various assessment techniques, such as observations, anecdotal records, portfolios, checklists, and children's work samples, provide comprehensive information regarding children's cognitive, social-emotional, language, and physical development. Previous studies have shown that digital portfolios can support authentic assessment by facilitating continuous documentation and improving communication between teachers and parents (Warmansyah et al., 2022). Similarly, electronic portfolios provide more efficient and process-oriented assessments that emphasize children's developmental progress rather than final outcomes (Masluhah & Afifah, 2022). However, teachers often encounter difficulties in managing assessment data, integrating evidence from various instruments, and producing comprehensive developmental reports. The complexity of authentic assessment frequently increases teachers' administrative workload and limits opportunities for reflective analysis of children's development.

2. Artificial Intelligence in Educational Assessment

Recent advances in artificial intelligence (AI) have created opportunities to improve assessment practices through automated data processing, developmental analysis, and administrative efficiency. The integration of AI has been shown to simplify documentation processes and reduce teachers' workload while improving reporting efficiency (Siri & Azis, 2025). In addition, AI can support more personalized and data-driven educational decisions by identifying learning patterns and providing deeper analysis of assessment information(Oktavianus et al., 2023). Despite these advantages, previous studies have primarily focused on digital documentation, electronic portfolios, and general AI applications in education. Existing studies have not specifically addressed the development of AI-assisted authentic assessment models in early childhood education, particularly those that integrate various authentic assessment instruments into a structured developmental database.

3. Pedagogical and Ethical Considerations

The implementation of AI in educational assessment requires strong pedagogical and ethical foundations. Universal Design for Learning (UDL) emphasizes flexibility, inclusivity, and responsiveness to children's diverse developmental characteristics (Espada-Chavarria et al., 2023). Similarly, Assessment for Learning (AfL) positions assessment as an integral part of the learning process by promoting continuous feedback and supporting children's development through observation and reflection (Vries et al., 2022). The integration of technology in assessment also requires teachers' pedagogical competence. The TPACK framework highlights the importance of balancing technological, pedagogical, and content knowledge in educational practice (Ardiyanto, 2024; Mala et al., 2025).

Furthermore, data governance and algorithmic fairness have become increasingly important in AI-supported assessment because children's developmental data require protection, transparency, and responsible use (Harahap et al., 2025; Syafitri et al., 2025). Human oversight remains essential to prevent bias and ensure that AI-generated analyses

remain pedagogically appropriate (Hamdani et al., 2025; Sabri & Wais, 2025). Therefore, AI should function as a supportive tool rather than a replacement for teachers' professional judgment. Teachers remain responsible for validating assessment results, interpreting developmental information, and making educational decisions that consider children's individual characteristics (Carvalho et al., 2022; Salim, 2022).

Previous studies have primarily focused on digital portfolios, technology integration, and general applications of artificial intelligence in education. However, limited studies have developed AI-assisted authentic assessment models specifically designed for early childhood education, particularly those that support developmental analysis and pedagogical decision-making. The novelty of this study lies in the development of an AI-assisted authentic assessment model that integrates authentic assessment instruments with AI-assisted analysis using NotebookLM within the ADDIE framework. By maintaining teachers as the primary pedagogical decision-makers, the model aims to support more objective, efficient, and comprehensive developmental assessment. Furthermore, interview findings from teachers in Bantul Regency revealed substantial challenges related to assessment documentation and report preparation. Therefore, the development of this assessment innovation was intended to address these identified needs and support the implementation of the Merdeka Curriculum in the digital era.

C. Method

This study employed a Research and Development (R&D) approach using the ADDIE model, which consists of five stages: Analysis, Design, Development, Implementation, and Evaluation. The study was conducted in five pilot early childhood education (PAUD) institutions in Bantul Regency, Yogyakarta, Indonesia. Participants included 10 teachers and 5 school principals selected purposively based on three criteria: (1) implementation of the Merdeka Curriculum, (2) the need for innovation in authentic assessment through AI-assisted approaches, and (3) challenges in conducting developmental assessment, including constructing narrative reports, managing administrative workload, and analyzing assessment data.

Data were collected through interviews, focus group discussions (FGDs), and documentation. Interviews with teachers and school principals explored challenges in authentic assessment, teachers' needs, and perceptions of AI-assisted assessment. FGDs were conducted to discuss the design and refinement of the assessment model and to obtain feedback throughout the development process. Documentation included assessment forms, observation records, anecdotal notes, children's work samples, portfolios, and child development reports. The collected data were analyzed descriptively to identify assessment challenges and inform the iterative development of the AI-assisted authentic assessment model. Revisions to the assessment templates and AI-assisted procedures were made based on findings from interviews, FGDs, and implementation feedback.

NotebookLM was employed as the AI platform to support developmental assessment. Teachers uploaded structured assessment data, including observation records, anecdotal notes, checklists, and children's work samples. Based on these data, NotebookLM generated narrative descriptions of children's developmental progress and recommendations for follow-up learning and home-based stimulation. All AI-generated outputs were subsequently reviewed, verified, and, where necessary, revised by teachers before being incorporated into children's developmental reports. Thus, AI functioned solely as a supportive analytical tool, while teachers retained full responsibility for interpreting assessment results and making pedagogical decisions.

Ethical Considerations Permission to conduct the study was obtained from the participating early childhood education institutions. The use of children's developmental

assessment data for research purposes was conducted with institutional approval and parental consent. To protect participants' privacy, all children's data were anonymized by removing personally identifiable information before being processed in NotebookLM. Access to assessment data and AI-generated outputs was restricted to the research team and participating teachers. All AI-generated reports were reviewed and verified by teachers prior to their inclusion in children's developmental reports, ensuring that professional judgment remained central to the assessment process.

Table 1. Implementation of the ADDIE Model

ADDIE Stage	Activities	Data Source	Output	Revision
Analysis	Identifying assessment problems, teachers' needs, and challenges in authentic assessment	Interviews, documentation	Needs analysis	Identification of assessment challenges and user needs
Design	Designing authentic assessment instruments and AI-assisted assessment templates	Literature review and FGD	Initial product design	Revision of indicators and assessment templates
Development	Developing assessment templates and AI-assisted procedures	FGD and teacher feedback	Prototype model	Product improvement
Implementation	Trial implementation in five PAUD institutions	Documentation and teacher feedback	Initial implementation results	Product refinement
Evaluation	Evaluating the developed model and AI-assisted process	Interviews and FGD	Final product	Final revision and recommendations

D. Result and Discussion

1. Forms of Assessment Practice Applied in Early Childhood Education Institutions

The results of the study show that the practice of formative assessment in the PAUD institution where this research was conducted is not fully in accordance with the principles of authentic assessment. Based on the results of observations and interviews with the sources, the results were obtained in the table below:

Tabel 2. Needs Analysis Findings

No	Aspects Observed	Findings
1	Assessment Practice	Assessment is carried out in learning activities through an observation process, but recording is carried out when approaching the distribution of report cards
2	Assessment Instruments	The assessment was carried out using checklist instruments and anecdotal notes, and the assessment of children's works.

		However, child development analysis is still carried out manually, so it takes a long time.
3	Assessment Constraints	Teachers have difficulty in summarizing a lot of assessment data to make a report on child development
4	Data Processing	Data processing and analysis are carried out manually so that there are often delays in time. Analytical assessment is still often done subjectively because there are no results that collapse from time to time
5	Technology Needs	Teachers need a digital system that can help process assessment data automatically

Based on the findings, teachers tend to assess children's achievements based on the final results of activities or momentary displays, rather than through continuous naturalistic observation. In fact, authentic assessments require a systematic process of collecting information about child development through consistent observation. This condition results in the development data obtained to be less accurate to be used in the preparation of daily, weekly, and report cards. This finding is in line with the research of Wahyu Hidayat and Andriani (2020) which shows that most PAUD teachers, namely 85.71%, have not implemented authentic assessments comprehensively and have not used the entire domain of child development in the assessment process (Hidayat & Andriani, 2020). Assessments that are carried out without paying attention to the child's learning process as a whole cause children's development to not be reflected in reality. In practice, teachers have not documented observations systematically so that assessments lose accuracy. Structured observation is a key component of authentic assessment because it provides a real picture of a child's abilities and needs (Yates & Oates, 2021). If observation is carried out in a limited way, teachers do not have valid data to understand the actual abilities of children.

In line with that, Seitz (2024) in Astana & Permatasari (2025) explained that the relevance of observation to the stage of child development is an important element in assessment. Because observation helps teachers capture the child's learning process in a natural context (Astana & Permatasari, 2025). However, the findings of this study indicate that observations have not been carried out in depth. The lack of teachers' understanding of authentic assessments as identified by Astana and Permatasari (2025) is also a factor inhibiting the implementation of effective authentic assessments. Authentic assessments are actually able to provide a real picture of children's development and character, but teachers' competence in implementing them is still limited (Hutapea et al., 2025).

In addition to the inaccurate formative assessment, the study also found that teachers often group children based on general impressions of apparent abilities, so that children who appear to have similar behaviors are assumed to achieve the same development. The practice of homogenization is contrary to the principles of early childhood development which emphasizes the uniqueness and rhythm of each individual's development. According to Yogi et al. (2022) in Berliana & Atikah (2024), assessments in the Independent Curriculum should describe the development of children individually and be used as the basis for future learning planning. When teachers generalize children's achievements, assessments no longer function as a valid developmental mapping tool (Berliana & Atikah, 2024). Maisura, Anggraeni, Rahardjo, and Anggriani (2022) in Berliana & Atikah (2024) emphasized that formative and summative assessments require precision in observing the developmental achievements of each child personall. Assessments carried out in groups cause teachers to miss the opportunity to understand the unique needs of children. Authentic assessments should be systematic, measurable, individualized, and comprehensive. Therefore, the practice of grouping assessment results shows that teachers have not applied the principle of individuality that is at the core of authentic assessment.

The inaccuracy of assessments carried out by teachers also has a direct impact on their ability to design daily learning activities. When observations are not carried out accurately and are not supported by adequate documentation, teachers do not have a solid basis for determining activities that are appropriate to children's developmental needs. Research by Mastikawati, Baharun, Wahid, Solehatin, and Adawiyah (2022) shows that authentic assessments through checklist techniques, anecdotal notes, and work analysis are very helpful for teachers to understand children's abilities and interests, which in turn affects the quality of learning planning (Mastikawati et al., 2022). According to Yogi et al. (2022) in Berliana & Atikah, 2024 in the context of the Independent Curriculum, assessments should provide accurate feedback so that teachers can improve learning strategies. Thus, the inauthenticity of the assessments found in this study not only has an impact on the accuracy of child development reports, but also affects the low quality of the design of learning activities that should be varied, meaningful, and in accordance with the individual needs of children (Berliana & Atikah, 2024).

2. Obstacles, Challenges, and Needs of Teachers in Conducting Authentic Assessments and Accurate Analysis of Assessment Results

The implementation of authentic assessments in PAUD still faces various obstacles that have an impact on the accuracy of assessment results and the effectiveness of learning follow-up. Field findings show that some teachers do not have a comprehensive understanding of the urgency of authentic assessment. Teachers tend to view that the assessment results are only used as the basis for reports on children's development to parents, not as a source of information to plan stimulation that suits children's learning needs. In fact, early childhood assessments aim to collect important information about child development which is then used to design meaningful learning experiences (Amanda & Hasni, 2024).

One of the most dominant challenges is time constraints. Teachers are often unable to observe the entire child in one day, mainly due to the unique and diverse characteristics of the child, so some developmental behaviors can be missed . This limitation was also found in other studies that confirmed that not all learners can be optimally observed in a single learning period. In addition, teachers also have difficulty narrating the results of the assessment specifically so that the description of child development becomes less in-depth. This is related to the lack of competence in analyzing and interpreting data from various assessment instruments. This condition is in line with the finding that some teachers still have a not comprehensive understanding of the concept of authentic assessment in the Independent Curriculum (Rahmawati & Riyadi, 2025).

In terms of instruments, the implementation of assessments in PAUD institutions is still limited to the use of checklists and work results, so it does not fully reflect the principles of holistic assessment. In fact, in the context of the Independent Curriculum, authentic assessments should be carried out through various techniques such as anecdotal notes, portfolios, process documentation, and analysis of children's works (Akbar et al., 2024). The limited use of these types of instruments makes teachers less likely to have rich learning evidence to analyze children's developmental achievements. Teachers also face confusion when it comes to data derived from anecdotal notes, works, portfolios, or checklists often not arranged in a format that facilitates comprehensive analysis. This obstacle is seen in various studies that show that even though teachers have implemented authentic assessments, they still need assistance in processing assessment data to produce complete information (Akbar et al., 2024).

Based on these challenges, PAUD teachers have a number of important needs to improve the quality of assessment implementation. First, teachers need consistent instrument templates, so they can apply a variety of assessment techniques without having

to build instruments from scratch. A standard template also helps ensure uniformity of the data collected, while also facilitating further analysis. Second, teachers need assessment analysis tools that are able to integrate data from various instruments and produce comprehensive, specific, and accurate narratives of child development. This tool is expected to support teachers in uniting daily observation notes, work results, portfolios, and progress checks so that the information produced is more informative and can be used as a basis for future learning planning. This need is in line with the finding that authentic assessment results will only be meaningful if analyzed through the process of data reduction, data presentation, and verification as described in the Miles and Huberman analysis model (Akbar et al., 2024). Thus, meeting the needs of consistent instruments and integrated analysis tools is a strategic step to improve the quality of authentic assessments in PAUD. This effort not only helps teachers overcome time constraints and analytical competence, but also supports the creation of more accurate and meaningful assessments for children's development.

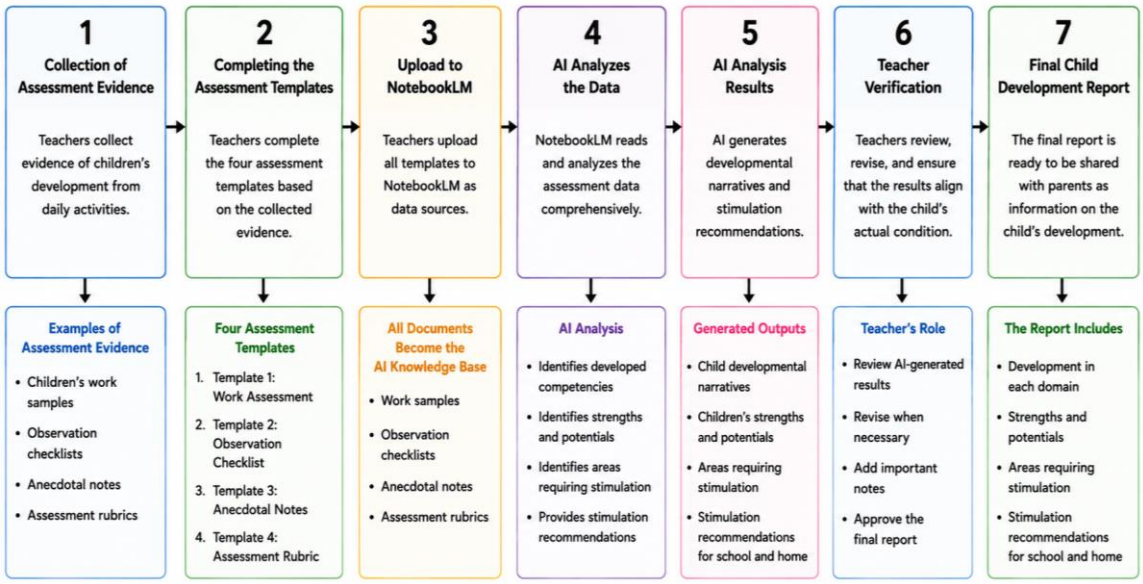

Figure 1. Workflow of Notebook LM implementation for output results

3. Development of AI Assisted Authentic Assessment Models

Table 3 presents the work assessment template developed in this study. The template is intentionally presented in a blank format as it serves as an assessment instrument for teachers during authentic assessment implementation. The result of this study is a template for assessing the development of the child with the following view:

Table 3. Developed Work Assessment Template

Tema/ Sub Tema/ Topik :
Kelas :
Semester :

Nama Anak	Foto	Keterangan
		Deskripsi:
		Nilai Agama dan Budi Pekerti:
		Analisis Jati Diri:
		Analisis Literasi dan STEM:
		Deskripsi:

	Nilai Agama dan Budi Pekerti:
	Analisis Jati Diri:
	Analisis Literasi dan STEM:

The work assessment template is an important instrument in early childhood assessment that serves to help educators assess the products or works produced by children in a systematic, objective, and sustainable manner. The use of this template allows teachers to conduct assessments based on clear and structured criteria, so that assessment results are more consistent and reflect children's development authentically. In terms of documentation and accountability, the assessment template serves as authentic evidence of child development that can be stored in a developmental portfolio. This document makes it easier for teachers to compile child development reports in a narrative and communicative manner, as well as being an informative communication material for parents about the progress of their children's learning.

Table 4 presents the checklist/observation assessment template developed in this study. As an assessment instrument, it is intentionally presented in a blank format to be completed by teachers during classroom observation and authentic assessment activities.

Table 4. Checklist/Observation Assessment Template

Nama:

No	TP (Tujuan Pembelajaran)	KKTP (Kriteria Ketercapaian Tujuan Pembelajaran)	Muncul/ Tidak Muncul	Hasil Pengamatan
1.				
2.				
3.				

The check/observation assessment template serves as an instrument to systematically and objectively document children's development based on the behaviors seen during learning and play activities. Through the checklist format, teachers can make observations in a directed manner according to the development indicators that have been set, so that the assessment results become more consistent and easy to analyze. The check/observation assessment template also acts as authentic evidence of assessment that can be used in the preparation of child development reports and portfolios. With clear and structured documentation, the results of observations can be communicated informatively to parents and support holistic development-oriented pedagogical decision-making.

Table 5 presents the anecdotal note assessment template developed in this study for teachers to document significant learning events and children's developmental progress

Table 5. Anecdotal Note Assessment Template

Tema/ Sub Tema/ Topik:
Kelas :
Semester :

Nama Anak	Peristiwa	Analisis Capaian
		Nilai Agama dan Budi Pekerti
		Analisis Jati Diri:
		Analisis Literasi dan STEM:
		Umpan Balik
		Nilai Agama dan Budi Pekerti
		Analisis Jati Diri:
		Analisis Literasi dan STEM:
		Umpan Balik

Anecdotal note assessment templates serve as educators' tools for documenting meaningful behaviors and events that arise naturally during the child's learning and play process. The use of this template helps teachers record events systematically and objectively, so that child development data is more accurate and easy to analyze. In addition, anecdotal notes are the basis for pedagogical reflection and follow-up learning planning according to the individual needs of children. This template also serves as authentic evidence of child development that supports the preparation of reports and communication of child development to parents in an informative and sustainable manner.

Table 6 presents the assessment rubric template developed in this study as a reference for authentic assessment and AI-assisted analysis of children's developmental achievements.

Table 6. Assessment Rubric Template

Elemen Capaian Pembelajaran	Deskripsi	Tujuan Pembelajaran	KKTP (Kriteria Ketercapaian Tujuan Pembelajaran)
Nilai Agama dan Budi Pekerti	Anak percaya kepada Tuhan Yang Maha Esa, mulai mengenal dan mempraktikkan ajaran pokok sesuai dengan agama dan kepercayaannya.		
Jati Diri	Anak mengenali identitas diri, mampu menggunakan fungsi gerak, memiliki kematangan emosi dan sosial untuk berkegiatan di lingkungan belajar.		
Dasar-dasar Literasi dan STEAM	Anak memiliki kemampuan literasi dasar, matematika dasar, dan sains, mampu memanfaatkan teknologi dan rekayasa sederhana, serta menciptakan dan mengapresiasi karya seni.		

This assessment rubric template is compiled as a systematic reference in the implementation of assessment and pedagogical decision-making on early childhood development. The rubric structure is designed to ensure alignment between national policies, learning planning, and assessment practices in educational units. Therefore, the elements of learning outcomes and their descriptions refer directly to the applicable national government regulations, so they are standard and cannot be changed. This is intended to maintain the consistency of child development standards and ensure the conformity of assessments with national curriculum policies. Meanwhile, learning objectives and Learning Goal Achievement Criteria (KKTP) are prepared flexibly by educators according to the needs, characteristics, and context of student development. This flexibility provides space for teachers to tailor assessments to children's real conditions, while supporting the implementation of child-centered and individual development-oriented learning. By integrating these standard and adaptive components, this assessment rubric template functions as an objective, directed, and academically accountable child development analysis instrument. This rubric not only serves as a tool for recording assessment results, but also as a basis for pedagogical reflection for teachers in planning learning follow-up in accordance with children's needs and potentials holistically.

This assessment template and rubric was developed as a structured database used in the use of artificial intelligence (AI) to support data-driven early childhood development analysis. In the context of education, AI plays an important role in systematically processing assessment data, identifying developmental patterns, and producing more objective and consistent analysis than manual analysis (Holmes et al., 2019). Therefore, the availability of structured and standardized data is a key prerequisite for AI to work optimally. The learning outcomes elements and their descriptions in this template are prepared based on applicable national government regulations, so that they function as a normative and standard framework in the processing of assessment data. This standardization is important to ensure the conformity of assessments with national curriculum policies while increasing the validity of data analyzed by AI systems.

In the Independent Curriculum, assessment is positioned as an integral part of learning that is oriented towards the growth and development of children, not just the measurement of final outcomes. Meanwhile, learning objectives and Learning Goal Achievement Criteria (KKTP) are designed flexibly by educators according to the needs, characteristics, and context of children's development. This approach is in line with the principles of authentic and holistic assessment in early childhood education, which emphasizes the importance of understanding the child's learning process individually and contextually (Wortham & Hardin, 2021). The flexibility of this component allows assessment data to represent the child's developmental conditions more completely, so that AI-assisted analysis becomes more meaningful and pedagogically relevant. The use of AI in the processing of assessment data allows teachers to obtain an automatic, efficient, and evidence-based analysis of child development. AI can help reduce the administrative burden on teachers, improve time efficiency, and support more accurate and reflective pedagogical decision-making. In the context of early childhood education, this is important considering that teachers are required to conduct continuous assessments of many aspects of child development simultaneously.

Thus, this assessment rubric template serves not only as a conventional assessment instrument, but also as a foundation for the integration of AI in early childhood development assessment systems. This integration supports the transformation of assessment towards an objective, adaptive, and development-oriented data-driven approach, and is in line with the direction of national and global education policies in the digital era.

4. Implementation of the Developed AI-Assisted Assessment Model

To demonstrate the implementation of the developed assessment model, the work assessment, checklist/observation, anecdotal note, and assessment rubric templates were integrated into NotebookLM as a structured assessment database. Based on the assessment data entered by teachers, the AI automatically generated narrative reports describing children's developmental achievements in the three learning outcome elements: Religious and Moral Values, Self-Identity, and Literacy and STEAM. The generated reports also provided individualized recommendations for follow-up learning and home-based stimulation. Figure 2 presents an example of an AI-generated narrative report produced using the developed assessment templates.

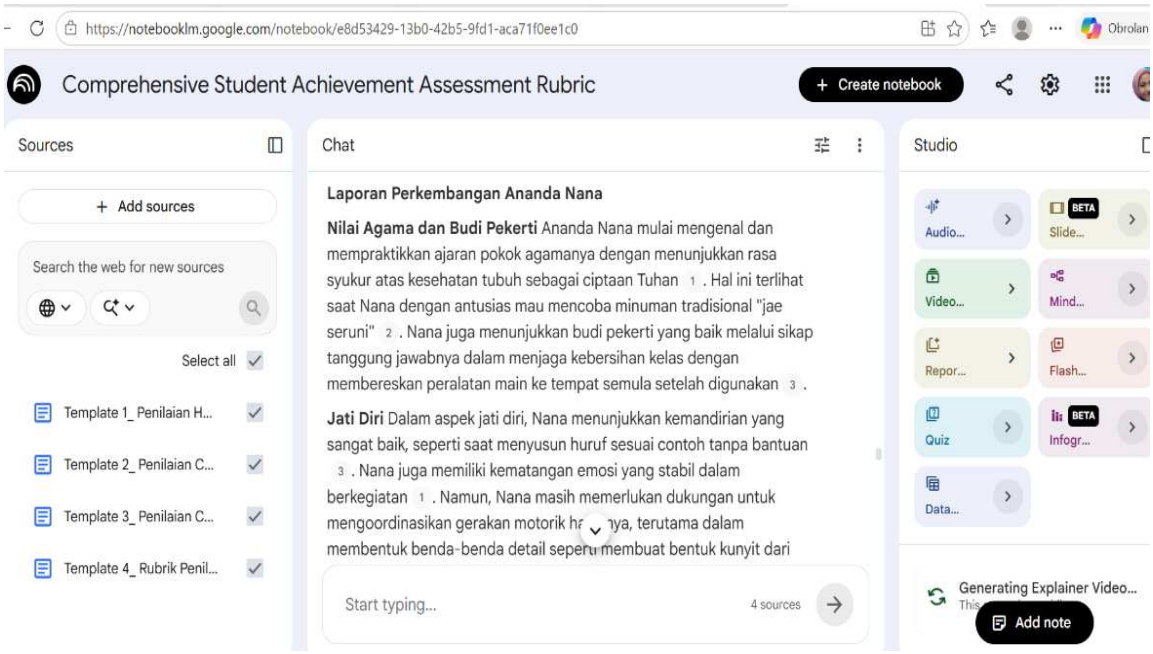

Figure 2. Example of an AI generated narrative child development report produced using the developed assessment templates in NotebookLM

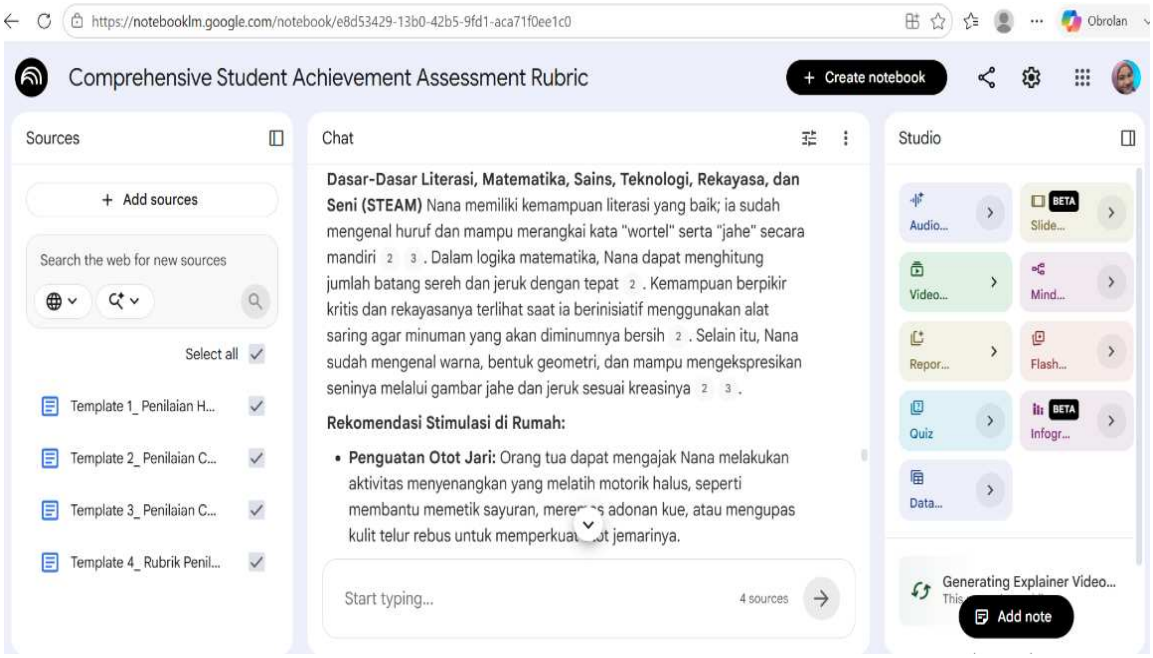

Figure 3. Example of an AI generated narrative child development report produced using the developed assessment templates in NotebookLM

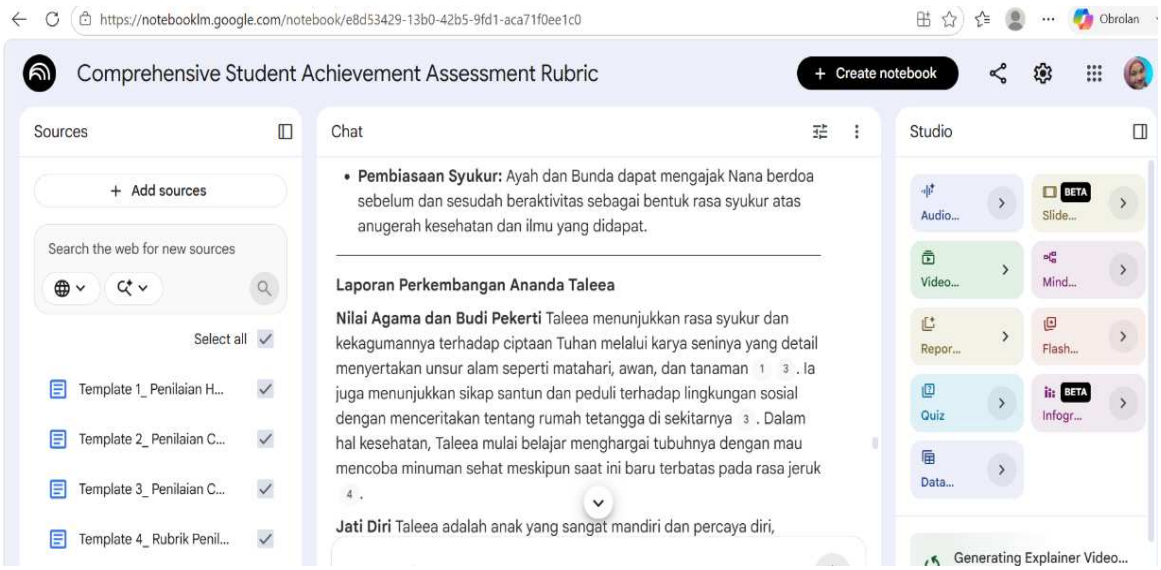

Figure 4. Example of an AI generated narrative child development report produced using the developed assessment templates in NotebookLM

Table 7. Comparison Between Manual Assessment and AI-Assisted Assessment

Aspect	Manual Assessment	AI Assisted Assessment
Time Required for Data Analysis	Data analysis is conducted manually and generally requires more time	AI-assisted analysis enables a more timely processing of assessment data
Data Processing	Assessment data are processed, organized, and summarized manually	Assessment data are organized and analyzed digitally with AI assistance.
Quality of Analysis Results	Analysis results rely primarily on teachers' interpretation and professional judgment.	Analysis results are generated based on structured assessment data and AI-assisted analysis.
Teacher's Role	Teachers perform all stages of the assessment process independently	Teachers serve as reviewers, validators, and final pedagogical decision-makers
Administration and Reporting	Assessment documentation and report preparation are conducted manually.	AI supports the preparation of assessment reports and documentation processes.

5. Development Process of the AI-Assisted Authentic Assessment Model Using the ADDIE Framework

ADDIE's instructional framework (Analysis, Design, Development, Implementation, Evaluation) details how AI-based assessment templates and rubrics are built from the planning stage to the evaluation of their effectiveness in the context of Early Childhood Education (PAUD).

a. Analysis Phase

The analysis phase focused on identifying the problems, user needs, and system requirements necessary for the integration of artificial intelligence (AI) into authentic assessment practices in early childhood education. The preliminary analysis revealed that early childhood teachers face a substantial administrative burden when conducting narrative and holistic assessments manually. Conventional assessment practices are often time-consuming, subjective, and susceptible to inconsistencies in interpretation. Furthermore, AI-based systems require structured and standardized data to generate

meaningful analyses , whereas existing assessment data in schools are frequently fragmented and unorganized, making them difficult to process computationally.

Policy analysis was also conducted to ensure that the developed system aligned with national educational regulations, particularly the Merdeka Curriculum. In this regard, Learning Outcomes (Capaian Pembelajaran/CP) were positioned as the normative standards underlying the assessment framewor. In addition, user analysis indicated that teachers require flexibility in determining Learning Objectives (Tujuan Pembelajaran/TP) and Learning Objective Achievement Criteria (Kriteria Ketercapaian Tujuan Pembelajaran/KKTP). Therefore, the assessment system was designed to remain adaptive to children's developmental contexts and classroom needs rather than imposing rigid assessment criteria.

b. Design Phase

The design phase focused on developing a database architecture capable of bridging pedagogical needs and technological functions. A structured assessment template was designed in the form of a digital spreadsheet containing separate fields for static and dynamic data. Static variables included Learning Outcomes (CP) and their descriptors, serving as standardized anchors for the assessment process. In contrast, dynamic variables consisted of teachers' formulated TP, KKTP, and observational records that reflected contextual and individualized learning experiences.

The AI logic was also designed at this stage. The system was structured to receive observational data and achievement criteria as inputs, process the information through pattern recognition mechanisms aligned with the CP framework, and produce outputs in the form of objective developmental narratives and recommendations for subsequent stimulation activities. Pedagogically, the assessment format emphasized authentic assessment principles by enabling teachers to provide descriptive behavioral observations rather than relying solely on numerical checklists.

c. Development Phase

The development phase involved constructing the assessment instruments and supporting system components. A master template and assessment rubric were developed to ensure compatibility with AI processing. Each assessment component was assigned a unique code to facilitate data recognition and analysis by the system. In addition, prompt engineering guidelines were developed to enable AI to transform structured assessment data into narrative reports that reflect educational language and developmental perspectives. To improve the validity of the system, the database incorporated Learning Outcomes from the Merdeka Curriculum as reference standards for automatic validation. Furthermore, a user guide was developed to assist teachers in completing the assessment template accurately, including guidance on formulating measurable learning objectives and operational achievement criteria**.**

Expert validation indicated that the developed assessment instruments were appropriate for implementation, supporting the feasibility of the proposed assessment model. *Expert validation was carried out by two media experts specializing in early childhood education, namely Atik Solehati, S.Pd.AUD. and Sri Pamuji, S.Pd.AUD. The validation process assessed five dimensions: content appropriateness, construct validity, presentation, language, and design. A five-point Likert scale was used as the evaluation instrument, ranging from 1 (highly inappropriate) to 5 (highly appropriate).*

Table 8 Expert Validation

Aspect	Validator 1	Validator 2	Mean Score (%)
Content Validity	5	5	5.0 (100%)
Construct Validity	3	3	3.0 (60%)
Presentation Validity	5	5	5.0 (100%)
Language Validity	4	4	4.0 (40%)
Design Validity	4	4	4.0 (40%)
Mean Expert Validation Score	4,4	4.4	4.4 (88%)

Based on the expert validation results presented in Table 7, the developed product obtained a mean validation score of 4.4 out of 5.0, equivalent to 88%. According to the predetermined feasibility criteria, this percentage falls within the "Highly Feasible" category. The results indicate that the product meets the required standards in terms of content validity, construct validity, presentation, language, and design. Although the construct validity aspect received a relatively lower score (3.0), the remaining aspects were rated as appropriate to highly appropriate by both experts. Overall, the findings demonstrate that the developed product is suitable for implementation in early childhood education and can proceed to the next stage of field testing with only minor revisions based on the experts' recommendations.

d. Implementation Phase

The implementation phase involved applying the developed assessment system in authentic educational settings. The process began with data entry, in which teachers completed the assessment template based on daily or weekly observations of children's activities and developmental progress. Teachers also formulated TP and KKTP according to the developmental characteristics and learning needs of individual children. This structured documentation process enabled teachers to systematically record developmental evidence and provided a strong basis for subsequent analysis.

After data entry, the completed template was processed by the AI system to identify developmental patterns and generate analytical descriptions of children's strengths and areas requiring further support. Nevertheless, the system maintained a human-in-the-loop approach by positioning teachers as the final validators of AI-generated analyses. Teachers reviewed and adjusted the outputs to ensure that the interpretations accurately reflected children's contextual and emotional conditions that may not be fully captured through data alone.

The outputs generated by the system can be utilized to prepare developmental reports for parents and to support teachers' reflective planning of future learning activities. Consequently, pedagogical decisions become increasingly evidence-based because they are derived from systematically documented and analyzed assessment data. In the context of early childhood education, the implementation of this system has the potential to improve the quality of assessment practices while facilitating a more comprehensive understanding of children's developmental trajectories.

d. Evaluation Phase

The evaluation phase examined the implementation of the AI-assisted authentic assessment model in supporting teachers' assessment practices. The results showed that the structured assessment templates enabled AI to systematically process assessment data and generate authentic narrative reports accompanied by individualized recommendations for follow-up learning and home-based stimulation. Teachers remained responsible for interpreting assessment results and making pedagogical decisions, while AI functioned solely as a supportive analytical tool. The findings suggest that the model can facilitate

assessment data analysis and support the preparation of authentic narrative reports. However, teacher verification remains essential to ensure that assessment results are interpreted accurately and used appropriately in pedagogical decision-making.

E. Conclusion

This study developed an AI-assisted authentic assessment model for early childhood education using the ADDIE framework. The developed product consists of structured authentic assessment templates, including work assessment, observation checklists, anecdotal notes, and assessment rubrics integrated with NotebookLM to support AI-assisted analysis of children's developmental data. Expert validation indicated that the developed assessment instruments were appropriate for implementation, supporting the feasibility of the proposed assessment model. The implementation results showed that the developed model enabled teachers to document children's development more systematically, facilitate the organization of assessment data, and generate individualized narrative reports with AI assistance. Teachers also reported that the model supported the assessment process while maintaining their role in reviewing, validating, and interpreting AI-generated outputs before they were used for pedagogical decision-making.This study has several limitations. The implementation involved only five early childhood education institutions with a limited number of participants, and the evaluation focused solely on the initial implementation of the developed model.Future studies are recommended to evaluate the effectiveness of the model using larger and more diverse samples and to examine its long-term impact on assessment quality, teacher workload, and children's developmental monitoring in early childhood education.

F. Acknowledgment

The authors would like to express their sincere gratitude to all parties who contributed to this research. Special thanks are extended to the participants and the institution that provided support and facilitated the implementation of this study. The authors also appreciate the valuable guidance and constructive feedback from colleagues and reviewers, which helped improve the quality of this manuscript.

References

Akbar, R. F., Ulumiyah, M., Wirdah, C., & Sulistiyaningrom, R. (2024). Implementasi Instrumen Penilaian Pembelajaran pada Anak Usia Dini di RA Sholahiyah Bae Kudus. *Jurnal Raudhah*, *12*(1), 67. https://doi.org/10.30829/raudhah.v12i1.3260

Amanda, R. S., & Hasni, U. (2024). *Analisis Penggunaan Authentic Assesment sebagai Alat Pengukuran Perkembangan Anak Usia Dini*. *7*(01).

Ardiyanto, F. R. (2024). Designing a Game-Based English Learning Application for Early Childhood Learners Using TPACK Framework. *Jurnal Caksana : Pendidikan Anak Usia Dini*, *7*(2).

Astana, A. C., & Permatasari, T. (2025). Asesmen Autentik di PAUD: Mengungkap Pemahaman Guru untuk Mendukung Perkembangan Anak yang Optimal. *Al-Athfal Stai Muhammadiyah Probolinggo*, *6*.

Atikah, C., & Oktaviani, R. (2025). Manajemen Pendidikan Sebagai Fondasi Penting dalam Perkembangan Anak. *Jurnal PAUD Agapedia*, *9*.

Berliana, D., & Atikah, C. (2024). Implementasi Asesmen dalam Kurikulum Merdeka di Pendidikan Anak Usia Dini. *Jurnal Citra Pendidikan (JCP)*, *4*.

Carvalho, L., Martinez-Maldonado, R., Tsai, Y.-S., Markauskaite, L., & De Laat, M. (2022). How can we design for learning in an AI world? *Computers and Education: Artificial Intelligence*, *3*, 100053. https://doi.org/10.1016/j.caeai.2022.100053

Espada-Chavarria, R., González-Montesino, R. H., López-Bastías, J. L., & Díaz-Vega, M. (2023). Universal Design for Learning and Instruction: Effective Strategies for Inclusive Higher Education. *Education Sciences*, *13*(6), 620. https://doi.org/10.3390/educsci13060620

Faizah, U. (2017). Pemanfaatan Asesmen Otentik untuk Menilai Karakter Peserta Didik melalui Pembelajaran Alquran pada Lembaga Pendidikan Islam Anak Usia Dini di TK/RA. *Proceedings of The 2nd Annual Conference on Islamic Early Childhood Education*.

Hamdani, A., Suherman, A., Darmawan, B., & Permana, E. (2025). Assessment and evaluation of an artificial intelligence-based deep learning curriculum. *Curricula: Journal of Curriculum Development*, *4*(2).

Harahap, I. N., Segara, K. G., Hutasuhut, S. A. H., & Nurbaiti. (2025). Konsep Dasar Sistem Data Governance dalam Era Digitalisasi. *Jurnal Sains Student Research*, *3*.

Hidayat, W., & Andriani, A. (2020). Pelaksanaan Penilaian Autentik Guru Pendidikan Anak Usia Dini. *Cakrawala Dini: Jurnal Pendidikan Anak Usia Dini*, *11*(2), 88–95. https://doi.org/10.17509/cd.v11i2.24922

Hutapea, R. H., Wahyuni, H. T., & Feronika, F. (2025). Penilaian Otentik Pendidikan Agama Kristen Berbasis Kearifan Lokal "Nyalah Karewau Napait Hang Urung". *DUNAMIS: Jurnal Teologi dan Pendidikan Kristiani*, *9*(2), 1014–1034. https://doi.org/10.30648/dun.v9i2.1725

Indalaillah, N., & Rukayah, I. (2025). Persepsi Guru Paud terhadap Penerapan Asesmen Autentik dalam Pembelajaran. *Jurnal Edueco Universitas Balikpapan*, *8*.

Khairunnisa Ulfadhilah. (2024). Penanaman Karakter Disiplin di Lingkungan Ramah Anak. *Kiddo: Jurnal Pendidikan Islam Anak Usia Dini*, 153–165. https://doi.org/10.19105/kiddo.v1i1.12713

Mala, L. S., Hibana, & Divanti, A. I. (2025). Technological Pedagogical and Content Knowledge (TPACK) Abilities of Early Childhood Education Teachers in Digital-Based Learning. *Lectura: Jurnal Pendidikan*, *16*.

Masluhah, M., & Afifah, K. R. (2022). Electronic Portofolio Sebagai Instrumen Penilaian Pembelajaran Siswa di Era Digital. *Jurnal Basicedu*, *6*(2), 1883–1896. https://doi.org/10.31004/basicedu.v6i2.2236

Mastikawati, M., Baharun, H., Wahid, A. H., Solehatin, S., & Adawiyah, R. (2022). Strengthening Early Childhood Learning Outcomes through Authentic Assessment of Students. *Jurnal Obsesi : Jurnal Pendidikan Anak Usia Dini*, *6*(5), 4547–4556. https://doi.org/10.31004/obsesi.v6i5.2525

Novianti, R., Firdayani, T., Rizki Intan, F., Sari, M., Padilah, P., Puspita Sari Jaya, M., Andriani, D., & Idayana Sinaga, S. (2025). Pelatihan Asesmen Awal di TK. *Wahana Dedikasi: Jurnal PkM Ilmu Kependidikan*, *8*(1), 47–52. https://doi.org/10.31851/wdk.v8i1.18393

Oktavianus, A. J. E., Naibaho, L., & Rantung, D. A. (2023). Pemanfaatan Artificial Intelligence pada Pembelajaran dan Asesmen di Era Digitalisasi. *Jurnal Kridatama Sains dan Teknologi*, *5*(02), 473–486. https://doi.org/10.53863/kst.v5i02.975

Rahmawati, I., & Riyadi, R. (2025). Implementasi Penilaian Autentik dalam Kurikulum Merdeka di RA Darul Abror Kabupaten Brebes. *Jurnal Asghar*, *5*.

Sabri, M., & Wais, A. (2025). Artificial Intelligence dalam Pengukuran dan Penilaian Pendidikan: Kajian Literatur dari Perspektif Inovasi Evaluatif. *Jurnal Evaluasi Pendidikan*, *2*(16).

Salim, N. A. (2022). Integrasi teknologi dalam pendidikan anak usia dini: Menilai dampaknya pada perkembangan kognitif. *Jurnal Warna : Pendidikan dan Pembelajaran Anak Usia Dini*, *7*(2), 96–107. https://doi.org/10.24903/jw.v7i2.1533

Siri, A., & Azis, A. (2025). Manajemen Asesmen Formatif PAUD Berbasis AI-Powered Assessment: Studi Komparatif Efisiensi Guru dan Akurasi Pelaporan Perkembangan di RA TA Kottah Daleman. *WALADI: Wawasan Belajar Anak Usia Dini*, *3*(1).

Syafitri, I. C., Anawaty, M. F., & Nisa', T. F. (2025). Membangun Dasar yang Kuat: Pelaksanaan dan Tantangan Asesmen pada Fase Fondasi. *Indonesian Journal of Early Childhood: Jurnal Dunia Anak Usia Dini*, *7*.

Vries, J. A. de, Dimosthenous, A., Schildkamp, K., & Visscher, A. J. (2022). The impact of an assessment for learning teacher professional development program on students' metacognition. *School Effectiveness and School Improvement An International Journal of Research, Policy and Practice*, *34*.

Warmansyah, J., Sari, M., Halif, A., Ismandela, A., Fatma, D., & Niarman, A. (2022). E-Portfolio as an Early Childhood Assessment Tool in Kindergarten. *Jurnal Pendidikan Anak Usia Dini Undiksha*, *10*(3).

Yates, E., & Oates, R. (2021). "It's nice to know you might make a difference": Engaging students through primary research as an authentic assessment. *Student Engagement in Higher Education Journa*, *3*.