Challenges and Teacher Strategies in Implementing Authentic Assessment: A Case Study at a State Islamic Junior High School

Hanna Hanifa Hira[1], Ernawati[2], Rifky Praptama[3], Muhamad Restu Andhika[4], Elin Driana[5]

[1]*Universitas Muhammadiyah Prof. DR. HAMKA, Indonesia*
[2]*Universitas Muhammadiyah Prof. DR. HAMKA, Indonesia*
[3]*Universitas Muhammadiyah Prof. DR. HAMKA, Indonesia*
[4]*Universitas Muhammadiyah Prof. DR. HAMKA, Indonesia*
[5]*Universitas Muhammadiyah Prof. DR. HAMKA, Indonesia*

Corresponding Author: Ernawati,   **E-mail:** ernawati.pep@uhamka.ac.id

Article Information:	ABSTRACT
	This study aims to describe the challenges and strategies of teachers in implementing authentic assessment within the Merdeka Curriculum at State Islamic Junior High School 3 Jakarta. Using a qualitative approach with a case study design, the research involved teachers of grades 8 and 9 as participants. Data were collected through in-depth interviews, classroom observations, and document analysis of assessment rubrics and student performance results. The instruments used included interview guidelines, observation sheets, and rubric review forms. The findings revealed that teachers faced several challenges, such as limited time for rubric development, difficulties in determining appropriate indicators, and students' lack of understanding of authentic criteria. To overcome these challenges, teachers applied adaptive strategies including peer collaboration, rubric simplification, and reflective assessment practices to improve learning quality. **Keywords**: *Authentic Assessment, Merdeka Curriculum, Teacher Strategies.*

INTRODUCTION

The shift in assessment paradigms from traditional tests to authentic assessment has emerged as a response to the need to measure student competence more

comprehensively through meaningful and contextual tasks. Traditional assessment, which relies on multiple-choice written tests, often captures only the mastery of declarative knowledge, while higher-order thinking skills, problem-solving, collaboration, and reflective abilities remain underrepresented. Authentic assessment offers an alternative by placing learners in task situations that resemble real-world contexts, enabling performance that reflects their ability to apply knowledge rather than merely recall it. Wiggins (2018) asserts that an assessment is considered authentic when students' performance on real intellectual tasks is examined directly rather than through test items as proxies. Thus, assessment is no longer positioned merely as a selection tool but as an integral part of the learning process that provides rich information for teachers and learners.

Theoretically, the design of authentic assessment can be reviewed through the framework of five dimensions: task, physical context, social context, form or product, and assessment criteria. These five dimensions ensure alignment between the targeted competencies and the performance evidence collected. Tasks must represent challenges relevant to real-world practice, the physical and social contexts determine how the task is carried out, the form or product represents observable performance outcomes, and the assessment criteria describe the agreed-upon quality standards (Gulikers, 2019). This framework helps teachers develop rubrics, instruments, and indicators that are not only aligned with the syllabus but also accurately reflect the demands of 21st-century competencies. Global research over the past decade has also shown that authentic assessment significantly enhances students' motivation, learning engagement, and higher-order thinking skills (Mueller, 2018), particularly when assessment criteria are communicated explicitly to learners.

Within the context of the Merdeka Curriculum, assessment is positioned as part of an integrated cycle with learning, emphasizing formative assessment, differentiation, and learner-centered practice. Official guidelines state that assessment should provide constructive feedback, guide learners in understanding their achievement levels, and support the continuity of the learning process rather than merely produce a final score (Kemendikbudristek, 2022). The flexibility in planning learning activities and reporting learning outcomes within the Merdeka Curriculum provides broader opportunities for implementing more varied authentic assessments, such as cross-disciplinary projects, portfolio-based performance tasks, or contextual assignments related to students' social, cultural, and religious experiences. However, this flexibility simultaneously demands teachers' pedagogical capacity to translate curricular expectations into assessment designs that are clear, operational, and feasible to carry out in the classroom.

Although the benefits of authentic assessment have been widely reported, various studies in Indonesia indicate that its implementation presents a number of complex challenges. Teachers often face limited time and resources to develop detailed rubrics, difficulties in formulating indicators truly aligned with targeted competencies, and students' confusion in understanding the authentic assessment criteria used (Lubis, 2022). Additionally, administrative workload and large class sizes frequently lead

teachers to perceive authentic assessment as an additional burden. On the other hand, several studies have reported that when authentic assessment is implemented consistently, students demonstrate increased motivation, engagement, and learning responsibility—for example, through active participation in discussions, improved task product quality, and willingness to revise work based on feedback (Rahmawati & Putra, 2021). This contradiction between potential benefits and practical challenges suggests that the main issue does not lie in the concept of authentic assessment itself, but in how the concept is translated into classroom practice.

Based on this review, a clear gap emerges between authentic assessment as recommended in the Merdeka Curriculum and its implementation in schools, particularly at the lower secondary level. This gap is not only about whether teachers implement authentic assessment, but more importantly about how they understand it, design its tasks and rubrics, integrate it into learning flows, and manage various challenges in the field. Most previous research has tended to focus on the effectiveness of authentic assessment models or their impact on student learning outcomes, while process-related dimensions—such as how teachers negotiate criteria with students, balance curriculum demands with time constraints, and conduct ongoing assessment— remain insufficiently explored. In madrasah settings, including MTs, this complexity is further heightened by the need to integrate academic competencies with religious and character values.

Departing from this gap, the present study focuses on how teachers at MTs Negeri 3 Jakarta understand, design, and implement authentic assessment in Merdeka Curriculum learning, as well as the strategies they use when facing implementation challenges. MTs Negeri 3 Jakarta was chosen as the study context because it is a State Islamic Junior High School currently implementing the Merdeka Curriculum in Grades VIII–IX with diverse student characteristics. This situation provides an opportunity to examine how teachers interpret assessment policies in the Merdeka Curriculum and operationalize them in everyday classroom practice. A qualitative case study approach was used because it allows exploration of teachers' meanings, pedagogical considerations, and adaptive strategies throughout the entire authentic assessment cycle, from planning and implementation to the use of assessment results.

Accordingly, the purpose of this study is to describe in detail the challenges teachers face and the strategies they use in implementing authentic assessment within Merdeka Curriculum learning at MTs Negeri 3 Jakarta. More specifically, this study aims to: (1) reveal teachers' understanding of the concepts and principles of authentic assessment within the Merdeka Curriculum framework; (2) map how teachers design authentic assessment tasks and rubrics that are contextual for Grade VIII–IX students; and (3) identify teachers' strategies for overcoming pedagogical, technical, and structural implementation challenges. The results are expected to provide practical contributions for enhancing teacher capacity through assessment training and mentoring programs, as well as strengthen the empirical foundation for implementing authentic

assessment as an integral part of learning in secondary schools, particularly within madrasah settings.

The shift in assessment paradigms from traditional tests to authentic assessment has emerged as a response to the need to measure student competence more comprehensively through meaningful and contextual tasks. Traditional assessment, which relies on multiple-choice written tests, often captures only the mastery of declarative knowledge, while higher-order thinking skills, problem-solving, collaboration, and reflective abilities remain underrepresented (Vo & Simmie, 2024). Authentic assessment offers an alternative by placing learners in task situations that resemble real-world contexts so that their performance reflects their ability to apply knowledge rather than merely recall it (Churiyah et al., 2019). Wiggins (2018) asserts that an assessment is considered authentic when students' performance on real intellectual tasks is examined directly rather than through test items as proxies. Thus, assessment is no longer positioned merely as a selection tool but as an integral part of the learning process that provides rich information for teachers and learners (Pantiwati & Nyono, 2020).

Theoretically, the design of authentic assessment can be reviewed through the framework of five dimensions: task, physical context, social context, form or product, and assessment criteria. These five dimensions ensure alignment between targeted competencies and the performance evidence collected. Tasks must represent challenges relevant to real-world practice, and the physical and social contexts determine how the task is carried out (Gulikers, 2019). This framework helps teachers develop rubrics, instruments, and indicators that are not only aligned with the syllabus but also accurately reflect the demands of 21st-century competencies. Global research over the past decade has also shown that authentic assessment significantly enhances students' motivation, learning engagement, and higher-order thinking skills, particularly when assessment criteria are communicated explicitly to learners (Gedera, 2023).

Within the context of the Merdeka Curriculum, assessment is positioned as part of an integrated cycle with learning, emphasizing formative assessment, differentiation, and learner-centered practice (Wulandari et al., 2025). Official guidelines state that assessment should provide constructive feedback, guide learners in understanding their achievement levels, and support the continuity of the learning process rather than merely produce a final score (Kemendikbudristek, 2022). The flexibility in planning learning activities and reporting learning outcomes within the Merdeka Curriculum provides broader opportunities for implementing more varied authentic assessments, such as cross-disciplinary projects, portfolio-based performance tasks, or contextual assignments related to students' social, cultural, and religious experiences. However, this flexibility simultaneously demands teachers' pedagogical capacity to translate curricular expectations into assessment designs that are clear, operational, and feasible for classroom implementation (Suttrisno & Yulia, 2022).

Although the benefits of authentic assessment have been widely reported, various studies in Indonesia indicate that its implementation presents a number of complex

challenges. Teachers often face limited time and resources to develop detailed rubrics, difficulties in formulating indicators truly aligned with targeted competencies, and students' confusion in understanding the authentic assessment criteria used (Lubis, 2022). Additionally, administrative workload and large class sizes frequently lead teachers to perceive authentic assessment as an additional burden. On the other hand, several studies have reported that when authentic assessment is implemented consistently, students demonstrate increased motivation, engagement, and learning responsibility, such as through active participation in discussions, improved task quality, and willingness to revise work based on feedback (Rahmawati & Putra, 2021). This contradiction between potential benefits and practical challenges suggests that the main issue does not lie in the concept of authentic assessment itself, but in how the concept is implemented in classroom practice (Nirman et al., 2025). Therefore, this study aims to explore teachers' adaptive strategies in addressing the complexities associated with the implementation of authentic assessment in the Merdeka Curriculum (Rahmatika et al., 2025).

Based on this review, a clear gap emerges between authentic assessment as recommended in the Merdeka Curriculum and its implementation in schools, particularly at the lower secondary level. This gap is not only about whether teachers implement authentic assessment, but more importantly about how they understand it, design its tasks and rubrics, integrate it into learning flows, and manage various challenges in the field (Prawisanthi et al., 2023). Most previous research has tended to focus on the effectiveness of authentic assessment models or their impact on student learning outcomes, whereas process-related dimensions—such as how teachers negotiate criteria with students, balance curriculum demands with time constraints, and conduct ongoing assessment—remain insufficiently explored. In the madrasah context, including MTs, this complexity is heightened by the need to integrate academic competencies with religious and character values (Ashari, 2022).

Departing from this gap, this study focuses on how teachers at MTs Negeri 3 Jakarta understand, design, and implement authentic assessment within Merdeka Curriculum learning, as well as the strategies they use when facing implementation challenges. MTs Negeri 3 Jakarta was chosen as the study context because it is a State Islamic Junior High School currently implementing the Merdeka Curriculum in Grades VIII–IX with diverse student characteristics. This situation provides an opportunity to examine in depth how teachers interpret assessment policies in the Merdeka Curriculum and operationalize them in classroom practice. A qualitative case study approach was used because it allows exploration of teachers' meanings, pedagogical considerations, and adaptive strategies throughout the entire authentic assessment cycle, from planning and implementation to the use of assessment results.

Accordingly, the purpose of this study is to describe in detail the challenges teachers face and the strategies they use in implementing authentic assessment within Merdeka Curriculum learning at MTs Negeri 3 Jakarta. More specifically, this study aims to: (1) reveal teachers' understanding of the concepts and principles of authentic

assessment within the Merdeka Curriculum framework; (2) map how teachers design authentic assessment tasks and rubrics that are contextual for Grade VIII–IX students; and (3) identify teachers' strategies for overcoming pedagogical, technical, and structural implementation challenges. The results are expected to provide practical contributions for strengthening teacher capacity through assessment training and mentoring programs, as well as provide an empirical foundation for implementing authentic assessment as an integral part of learning in secondary schools, particularly within madrasah settings.

RESEARCH METHODOLOGY
Research Design

This study employed a qualitative approach with a case study design to explore in depth the practice of authentic assessment in Merdeka Curriculum learning at MTs Negeri 3 Jakarta. The case study design was chosen because it allows the researcher to understand the processes and contextual dynamics of authentic assessment implementation within a specific school setting, including the interactions between teachers and students, school policies, and real classroom conditions that influence teachers' pedagogical decisions (Abdurrohim & Khuriyah, 2022).

Participants

The participants in this study consisted of teachers and students directly involved in the implementation of authentic assessment in Grades VIII and IX. Participants were selected using purposive sampling with criteria that teachers had implemented at least one cycle of authentic assessment in the current semester and students had participated in tasks or projects assessed using authentic assessment in the relevant subjects. Based on these criteria, six subject teachers and twenty students from Grades VIII–IX were selected as participants. The emphasis in participant selection was not on the number but on the relevance of their experiences to the research focus and the richness of information they could provide regarding the challenges and strategies in implementing authentic assessment (Prawisanthi et al., 2023).

Data Collection

Primary data were collected through semi-structured interviews. Two sets of interview guides were developed separately for teachers and students, referring to the authentic assessment framework and the implementation principles of the Merdeka Curriculum (Prof. Dr. Toni Toharudin, 2025). Interviews with teachers focused on their understanding of authentic assessment concepts, their experiences in designing tasks and rubrics, the implementation of assessments in the classroom, and the strategies they used to address challenges.

Meanwhile, student interviews explored their perceptions of authentic tasks they had completed, their understanding of assessment criteria, and their experiences receiving feedback from teachers. Individual interviews were conducted at the school at times agreed upon with participants. All interviews were audio-recorded with participant consent and transcribed verbatim for analysis. In addition, field notes were

compiled throughout the data collection process to document interaction contexts and additional information not captured in audio recordings (Rutakumwa et al., 2019).

Data Analysis

Data were analyzed using Miles and Huberman's interactive model, which includes three main stages: data reduction, data display, and conclusion drawing and verification. At the data reduction stage, the researcher reread interview transcripts and field notes repeatedly and coded segments relevant to the research focus, such as types of challenges, teacher responses, and adaptive strategies used (Munadzifah & Fradana, 2025).

The data display stage involved developing matrices and narrative summaries to group similar codes into broader categories; at this point, categories and relationships between them were organized through axial coding to identify patterns and connections among aspects of authentic assessment. The conclusion drawing and verification stage was conducted using constant comparison across participants and data sources, followed by re-examining raw data to ensure that interpretations were consistent with empirical evidence (Patton, 1999).

Trustworthiness of the Study

The credibility of the findings was maintained through several strategies. Source triangulation was carried out by comparing information obtained from teachers and students regarding the same authentic assessment practices. The researcher also conducted limited member checking by asking several teachers to review preliminary findings to ensure that the interpretations did not deviate from their actual experiences. Additionally, an audit trail was maintained through systematic documentation of data collection and analysis procedures, ensuring that the research process could be traced and verified (Ahmed, 2024).

RESULT AND DISCUSSION

Results

The findings of this study are presented to address three major aims: to explore teachers' understanding of authentic assessment within the Merdeka Curriculum, to describe how authentic assessment is designed and implemented across subjects, and to identify teachers' adaptive strategies in addressing implementation challenges. Evidence was obtained from interviews with six teachers and twenty students, classroom observations, and analysis of rubrics, student work, and instructional videos.

Teachers' Understanding of Authentic Assessment

Teachers exhibited a shared foundational understanding that authentic assessment involves evaluating students through meaningful performance tasks situated in real-life contexts. Their interpretations, however, varied depending on subject orientation, prior assessment practices, and disciplinary expectations.

An Indonesian Language Teacher emphasized contextual relevance:

> *"Authentic assessment is a task close to students' lives and shows how they apply concepts, not memorization" (T-BI1).*

A Mathematics Teacher highlighted inquiry-based validity:

> *"It becomes authentic when students collect their own data and validate it before analyzing" (T-MT2).*

A Science Teacher focused on procedural accuracy:

> *"What matters is that the experiment results are correct and follow procedures" (T-IPA1).*

These differences reflect a spectrum of assessment literacy from performance-focused to outcome-driven yet all teachers upheld the essential qualities of authentic assessment: relevance, real-world alignment, and applied understanding. Such variation is characteristic of educational systems transitioning from traditional testing to competency-oriented evaluation.

Design and Implementation of Authentic Assessment Across Subjects

Authentic assessment was implemented through performance-based tasks tailored to each subject's epistemic structure. Teachers applied analytic rubrics, iterative feedback, revision cycles, and structured scaffolds to enhance learning processes.

Indonesian Language: Observational Report and Oral Presentation

The Indonesian Language Teacher designed an observational report supported by an oral presentation, assessed through a four-level analytic rubric addressing content, organization, linguistic accuracy, and delivery.

A student reported:

> *"Because the rubric was discussed at the beginning, we knew exactly what we had to achieve" (S-08).*

The integration of writing, speaking, and feedback conferences fostered deep engagement with argument development and text structure. The task assessed both process and product, illustrating robust authentic assessment design.

Science: Laboratory Experiment and Science Poster

The Science Teacher implemented laboratory experiments culminating in a science poster, supported by microdemo videos to simplify procedural instructions.

> *"The microdemo videos made the steps clearer and reduced confusion in the lab" (T-IPA1).*

Students demonstrated improved data accuracy, systematic documentation, and clearer understanding of variable relationships. The alignment between laboratory practice, visual modeling, and multimodal reporting exemplified high task authenticity.

Islamic Religious Education: Funeral Prayer Practice and Reflection

The Islamic Religious Education Teacher integrated procedural practice with reflective writing, using detailed checklists and microteaching to standardize student performance.

Teacher:

> *"The checklist and microteaching ensure the movements are uniform and correct" (T-PAI1).*

Student:

"I became aware of the meaning of worship, not just the movements" (S-12).

This approach revealed the multilayered nature of authentic assessment in religious education, capturing cognitive, psychomotor, and affective learning dimensions simultaneously.

Mathematics: Primary-Data Statistical Project

The Mathematics Teacher required students to collect and validate primary data, maintain work logs, and create visualizations.

Teacher:

"The work log ensures that students' data are representative" (T-MT2).

Student:

"Statistics felt connected to real issues around us" (S-19).

Students engaged deeply in empirical inquiry, demonstrating improved reasoning, visualization accuracy, and contextual interpretation. The project shifted mathematics learning from abstract computation to real-world data analysis.

Social Studies: Local Economic Profile Interview

The Social Studies Teacher assigned students to interview local business owners and create infographics synthesizing key findings.

"The note-taking template helped us include clearer evidence" (S-06).

The task strengthened inquiry skills, evidence-based reasoning, and community engagement, enabling students to situate economic concepts within real-life contexts.

English: Debate and Speech Performance

The English Teacher implemented debates and speeches assessed through rubrics evaluating content, fluency, coherence, and audience engagement. Students engaged in self-review using video recordings.

"The recording helped me see my own progress over time" (S-14).

Performance-based assessment allowed learners to develop confidence, oral fluency, and metacognitive awareness, reducing code-mixing tendencies and promoting effective communication.

Challenges in Implementing Authentic Assessment

Teachers experienced major challenges, including time constraints, rubric design complexity, inconsistent interpretation across teachers, and limited laboratory or digital resources.

Science Teacher:

"Grading complex tasks in large classes takes an extraordinary amount of time" (T-IPA1)

Mathematics Teacher:

"Designing indicators that are specific yet practical is extremely challenging" (T-MT2).

Islamic Religious Education Teacher:

"Standards differ across teachers even when using the same rubric" (T-PAI1).

Student:

>*"I find it difficult to distinguish between 'adequate' and 'good' levels" (S-07).*

These constraints highlight systemic obstacles rather than conceptual limitations. Authentic assessment demands time, shared understanding, and adequate resources—conditions that remain uneven across subject areas.

Adaptive Strategies Employed by Teachers

Teachers adopted several adaptive strategies that balanced instructional goals with contextual limitations. These included simplifying rubrics, co-constructing criteria with students, collaborating through MGMP forums, moderating student work, and using staged assessment with digital portfolios.

Student:

>*"Once the rubric was discussed together, we immediately understood the expectations" (S-10).*

Social Studies Teacher:

>*"The MGMP helped us build rubric banks and sample works" (T-IPS1).*

Mathematics Teacher:

>*"Reviewing student work together made the standards much clearer" (T-MT2).*

These strategies demonstrate teacher agency and pedagogical adaptability. Rubric co-construction enhanced transparency, MGMP collaboration increased scoring reliability, and checkpoints with digital portfolios strengthened formative monitoring—showing how teachers negotiated structural constraints while preserving assessment integrity.

Table 1. Authentic Assessment Rubric

Assessment Aspect	Indicators	Score 4 (Excellent)	Score 3 (Good)	Score 2 (Fair)	Score 1 (Poor)
Conceptual Understanding	Ability to understand and relate concepts to real contexts	Demonstrates deep understanding with relevant real-life examples	Shows basic understanding with partially relevant examples	Limited understanding with shallow explanations	Does not demonstrate understanding and provides irrelevant explanations
Critical and Creative Thinking	Ability to analyze, evaluate, and generate solutions	Highly insightful analysis with innovative and realistic solutions	Reasonable analysis with some creative ideas	Limited analysis with weak ideas	No analytical ability or creativity demonstrated
Communicati	Ability to	Actively	Cooperates	Limited	No

on and Collaboration	work with peers and communicate clearly	contributes and communicates effectively and respectfully	well with occasional lapses in engagement	participation with weak communication	participation and poor collaboration
Quality of Product/Project	Neatness, completeness, and creativity	Product is highly neat, complete, creative, and engaging	Product is neat, sufficiently complete, and engaging	Product is incomplete and not sufficiently engaging	Product is unfinished or inconsistent with instructions
Self-Reflection and Responsibility	Ability to evaluate and manage learning processes	Provides detailed reflection on learning process, challenges, and improvements	Shows general awareness of strengths and weaknesses	Reflection is shallow and inconsistent	No reflection is provided
Affective Learning Attitudes	Discipline, enthusiasm, and positive learning behavior	Highly disciplined and consistently enthusiastic	Generally disciplined with minor inconsistencies	Low consistency and minimal enthusiasm	Shows poor discipline and negative learning behavior

This rubric provides a structured framework for evaluating student performance across cognitive, collaborative, creative, and affective dimensions. Its analytic structure guides teachers in making consistent scoring decisions while supporting students' understanding of quality performance.

Table 2. Final Score Criteria

Score Range	Category	Description
21–24	Excellent (A)	Demonstrates deep, reflective, and creative learning performance
17–20	Good (B)	Shows strong understanding and active participation
13–16	Fair (C)	Shows partial understanding and requires guidance
< 13	Poor (D)	Does not demonstrate adequate understanding and is minimally engaged

This scoring criterion categorizes students' holistic performance based on total rubric scores. It ensures that final judgments reflect not only academic mastery but also learning behaviors and reflective growth, aligning with the Merdeka Curriculum's emphasis on comprehensive competency development.

Discussion

The findings of this study demonstrate that the implementation of authentic assessment at State Islamic Junior High School 3 Jakarta aligns closely with the five-dimensional framework proposed by Gulikers et al. (2019), which emphasizes the coherence between task requirements, physical and social contexts, assessment forms, and scoring criteria. The use of observational reports, laboratory experiments, interview-based economic profiling, statistical projects, and performance-based English tasks reflects a strong orientation toward real-world relevance, consistent with the principles of authentic learning described by Churiyah et al. (2019) and Wiggins (2018). The varied physical and social contexts in which these tasks occurred such as science laboratories, Islamic ritual practice groups, community interviews, and collaborative mathematical analysis further strengthened the authenticity of the assessment environment in accordance with Simmie (2024).

The consistent use of analytic rubrics, structured feedback loops, and iterative revisions also demonstrates clear alignment with the core ideas of assessment for learning and assessment as learning embedded in the Merdeka Curriculum (Kemendikbudristek, 2022; Shodiq et al., 2023; Wulandari et al., 2025). Teachers' use of conferencing, annotated comments, and staged checkpoints reflects the educative spirit of assessment conceptualized by Wiggins (2018), in which assessment functions as a learning tool rather than merely a measurement instrument. These practices were found to encourage deeper cognitive engagement and accountability among students.

The learning dispositions strengthened through authentic assessment including increased motivation, confidence in communication, critical thinking, and collaboration mirror the outcomes reported by Gedera (2023) and Mueller (2018), who highlight the potential of performance-based tasks to cultivate higher-order thinking. In the context of Islamic Religious Education, the integration of procedural knowledge, practical performance, and reflective meaning-making aligns with Ashari et al. (2022), who emphasize the need for holistic assessment models in madrasah settings. Reflection activities in particular enabled students to engage with the affective dimensions of learning, reinforcing the multidimensional nature of competence expected in Islamic educational environments.

Despite these positive developments, teachers faced considerable structural and pedagogical barriers similar to those identified in previous studies by Lubis (2022), Nirman et al. (2025), and Prawisanthi et al. (2023). Heavy workload, limited time for rubric development, inconsistency in interpreting performance indicators, and insufficient laboratory or ICT facilities complicated the implementation of authentic assessment. The use of the Miles and Huberman qualitative analysis model, as

explained by Munadzifah and Fradana (2025), helped reveal patterns in these challenges, showing that most obstacles stemmed from systemic constraints rather than conceptual rejection of authentic assessment itself.

In response to these challenges, teachers developed a set of adaptive strategies that, when synthesized, form an emergent operational model referred to in this study as the 4S model. The first component, simplification, involves reducing rubrics to essential criteria to avoid cognitive overload for both teachers and students. The second, socialization or co-construction, refers to jointly discussing and negotiating rubric criteria with students to ensure transparency and shared understanding. The third, synchronization or moderation, relies on cross-teacher collaboration through MGMP forums to enhance scoring consistency. The final component, staging, includes the use of checkpoints and digital portfolios to strengthen formative monitoring. This model resonates with Wiggins' (2018) principles of educative assessment and parallels the adaptive strategies observed in Merdeka Curriculum settings discussed by Rahmatika et al. (2025).

The overall findings underscore the importance of institutional support in enabling meaningful and sustainable implementation of authentic assessment. Schools need to provide collaborative planning time, high-quality exemplars, shared rubric banks, and continuous professional development on task design, as emphasized by Toharudin (2025). Future research would benefit from multi-site comparative studies and longitudinal data collection to capture variations across diverse school contexts and to examine long-term impacts on students' academic and character development. A deeper integration of performance artifacts and digital portfolios could also enrich the assessment landscape in subsequent studies.

CONCLUSION

This study concludes that the implementation of authentic assessment at State Islamic Junior High School 3 Jakarta is shaped by teachers' ability to align meaningful tasks, clear criteria, and staged feedback within the framework of the Merdeka Curriculum. The findings highlight that authentic assessment not only strengthens students' cognitive performance but also fosters essential learning dispositions such as motivation, collaboration, responsibility, and self-regulation—demonstrating its pedagogical value beyond conventional testing. The study affirms the necessity of coherent rubric design, consistent teacher moderation, and collaborative professional inquiry as foundational conditions for reliable and educative assessment practices. It further proposes the 4S operational model—Simplification, Socialization, Synchronization, and Staging—as a practical framework that schools can adopt to enhance assessment quality. Future research should extend this inquiry through multi-site comparative studies and longitudinal designs to better understand contextual variations and long-term impacts, particularly in strengthening teacher assessment literacy and sustaining student-centered learning cultures.

ACKNOWLEDGEMENT

REFERENCES

Abdurrohim, M. I., & Khuriyah. (2022). Penilaian autentik memadukan pembelajaran dengan penemuan. Jurnal Inovasi Riset Akademik, 2(3), 258–265. https://mail.jurnalp4i.com/index.php/academia/article/download/1555/1501

Ahmed, J. (2024). Strategi keabsahan data kualitatif: Triangulasi, member check, dan audit trail. Jurnal Penelitian Teknik Audit dan Manajemen, 5(1), 1–15. https://jptam.org/index.php/jptam/article/download/21803/15330/38118

Ashari, M., dkk. (2022). Integrasi kompetensi akademik dan nilai keagamaan dalam penilaian autentik di madrasah. Jurnal Pendidikan Islam. https://ejournal.iainmadura.ac.id/index.php/re-jiem/article/download/6205/3030/

Churiyah, M., dkk. (2019). Penilaian autentik: Mengintegrasikan konteks dunia nyata dalam pembelajaran. Jurnal Pendidikan. https://files.eric.ed.gov/fulltext/EJ977576.pdf

Gedera, D. (2023). Dampak asesmen autentik terhadap motivasi dan berpikir tingkat tinggi. Journal of Authentic Learning. https://journal.unj.ac.id/unj/index.php/ishel/article/download/57721/20498/16522
5

Gulikers, J., Bastiaens, T., & Kirschner, P. (2019). A five-dimensional framework for authentic assessment. Educational Technology Research and Development. https://www.acsouth.edu/wp-content/uploads/2023/07/Guilikers-et-al.-2004-Five-Dimensional-Framework-for-Authentic-Assessment.pdf

Kemendikbudristek. (2022). Panduan pembelajaran dan asesmen Kurikulum Merdeka. Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi. https://kurikulum.kemdikbud.go.id/wp-content/uploads/2022/06/Panduan-Pembelajarn-dan-Asesmen.pdf

Lubis, M. (2022). Tantangan guru dalam implementasi penilaian autentik. Jurnal Pendidikan Dasar. https://journal.unpas.ac.id/index.php/pendas/article/download/34535/18117/12897
8

Mueller, J. (2018). The authentic assessment toolbox: Enhancing student learning. Journal of Online Learning and Teaching, 1(1). https://jolt.merlot.org/documents/VOL1No1mueller.pdf

Munadzifah, M., & Fradana, A. N. (2025). Analisis data kualitatif dengan model Miles dan Huberman dalam penelitian pendidikan. Jurnal Pendidikan dan Teknologi, 10(1),.https://jurnalfaktarbiyah.iainkediri.ac.id/index.php/proceedings/article/download/6723/1382/18174

Nirman, R., dkk. (2025). Praktik penilaian autentik: Hambatan dan solusi. Jurnal Pendidikan. https://jbasic.org/index.php/basicedu/article/view/3319

Pantiwati, D., & Nyono, S. (2020). Definisi dan prinsip penilaian autentik. Jurnal Pendidikan Teknologi dan Kejuruan. https://jurnal.uny.ac.id/index.php/jptk/article/download/20072/12099

Patton, M. Q. (1999). Meningkatkan kualitas penelitian kualitatif. Sage Publications. https://qdatraining.com/defining-the-constant-comparison-method/

Prawisanthi, N., dkk. (2023). Pemilihan partisipan purposive dalam studi penilaian autentik Kurikulum Merdeka. Jurnal Pendidikan Generasi Nusantara, 4(2), 120–135. https://samudrapublisher.com/index.php/jgenus/article/download/245/200/2774

Prawisanthi, W., dkk. (2023). Kesenjangan penerapan penilaian autentik di pendidikan menengah. Jurnal Pendidikan. https://e-journal.unair.ac.id/ETNO/article/download/49722/30846/358735

Rahmawati, C., & Putra, N. (2021). Implementasi penilaian autentik di sekolah Indonesia. Jurnal Pendidikan Bahasa Indonesia. https://journal.unj.ac.id/unj/index.php/jisae/article/download/23056/11278/60953

Rahmatika, S., dkk. (2025). Strategi adaptif guru asesmen autentik Kurikulum Merdeka. Jurnal Sakola. https://rayyanjurnal.com/index.php/sakola/article/view/5911

Rutakumwa, R., dkk. (2019). Wawancara semi-terstruktur: Teknik pengumpulan data kualitatif dengan transkripsi verbatim. Jurnal Penelitian Kualitatif, 19(3), 200–215. https://solusijurnal.com/wawancara-semi-terstruktur-memahami-konsep-implementasinya/

Shodiq, A., dkk. (2023). Penilaian autentik sebagai penilaian untuk/sebagai pembelajaran: Keselarasan perencanaan dan umpan balik. Jurnal Evaluasi Pendidikan, 10(2), 150–168. https://journal.unpas.ac.id/index.php/pendas/article/view/6930

Simmie, G. (2024). The role of authentic assessment in holistic student evaluation. Journal of Educational Research. https://ojs.unm.ac.id/IJEDI/article/viewFile/21899/12015

Dikpora. https://jurnaldikpora.jogjaprov.go.id/index.php/jurnalideguru/article/download/1542/811/

Toharudin, T. (2025). Kerangka penilaian autentik dan implementasi Kurikulum Merdeka. Badan Standar, Kurikulum, dan Asesmen Pendidikan, Kemdikbud. https://kurikulum.kemdikbud.go.id/file/1755668120_manage_file.pdf

Wiggins, G. (2018). Educative assessment: Designing assessments to inform and improve student performance (2nd ed.). Jossey-Bass.

Wulandari, D., dkk. (2025). Asesmen formatif dan diferensiasi dalam Kurikulum Merdeka. Jurnal Elementary. https://journal.ummat.ac.id/index.php/elementary/article/view/24370/0