Arfannur: Journal of Islamic Education
Volume 7, Nomor 1, (2026) hlm. 1 - 14
eISSN 2775-2429

Terakreditasi Peringkat – 4 SK : NO 72/E/KPT/2024
https://e-journal.iainptk.ac.id/index.php/arfannur/article/view/5652
DOI: https://doi.org/10.24260/arfannur.v7i1.5652

Bridging Planning and Practice: Cognitive Assessment Based on Bloom's Taxonomy in Islamic Religious Education at Senior High Schools

Ayu Wulandari[1], Saepul Anwar[1]*, Achmad Faqihuddin[1]
[1] *Department of Islamic Religious Education, Universitas Pendidikan Indonesia, Bandung, Indonesia*
email Koresponden Penulis: saefull@upi.edu

Article Info

Article History

Received:
January 04, 2026

Revised:
February 17, 2026

Accepted:
February 17, 2026

ABSTRACT

This study investigates the pedagogical dissonance between curriculum expectations and classroom reality in the implementation of cognitive assessment within Islamic Religious Education (IRE). While existing literature highlights the gap between Higher-Order Thinking Skills (HOTS) planning and practice, a concise theoretical explanation for this divergence remains underexplored. Using an instrumental case study design involving IRE teachers at a senior high school in Bandung, this research deconstructs the assessment culture through in-depth interviews, observations, and document analysis. The study contributes to the literature by identifying two distinct, conflicting paradigms that govern assessment practices: the systematic–diagnostic paradigm, rooted in strict adherence to Bloom's Taxonomy for formative improvement, and the contextual–summative paradigm, which prioritizes flexibility and social relevance for accountability. Theoretically, this research argues that the failure to institutionalize HOTS is not merely a technical competency issue but a result of the lack of integration between these two paradigms. The findings imply that fostering HOTS requires a "hybrid assessment model" that harmonizes systematic rigor with contextual adaptability, moving beyond the binary of diagnostic versus summative functions. This framework offers a novel lens for evaluating teacher competence in 21st-century religious education.

Keywords: Bloom's Taxonomy, Cognitive Assessment, Islamic Religious Education

ABSTRAK

Studi ini menginvestigasi disonansi pedagogis antara harapan kurikulum dan realitas kelas dalam implementasi asesmen kognitif pada Pendidikan Agama Islam (PAI). Sementara literatur yang ada menyoroti kesenjangan antara perencanaan dan praktik Higher-Order Thinking Skills (HOTS), penjelasan teoretis yang ringkas mengenai divergensi ini masih belum banyak dieksplorasi. Menggunakan desain studi kasus instrumental yang melibatkan guru PAI di sebuah SMA di Bandung, penelitian ini mendekonstruksi budaya asesmen melalui wawancara mendalam, observasi, dan analisis dokumen. Studi ini berkontribusi pada literatur dengan mengidentifikasi dua paradigma berbeda yang saling bertentangan dalam mengelola praktik penilaian: paradigma sistematis–diagnostik, yang berakar pada kepatuhan ketat terhadap Taksonomi Bloom untuk perbaikan formatif, dan paradigma kontekstual–sumatif, yang memprioritaskan fleksibilitas dan relevansi sosial untuk akuntabilitas. Secara teoretis, penelitian ini berargumen bahwa kegagalan pelembagaan HOTS bukan semata-mata masalah kompetensi teknis, melainkan akibat dari tidak terintegrasinya kedua paradigma tersebut. Temuan ini menyiratkan bahwa pengembangan HOTS menuntut sebuah "model asesmen hibrida" yang

menyelaraskan ketetapan sistematis dengan kemampuan adaptasi kontekstual, melampaui biner fungsi diagnostik versus sumatif. Kerangka kerja ini menawarkan lensa baru untuk mengevaluasi kompetensi guru dalam pendidikan agama abad ke-21.

Kata Kunci: *Asesmen Kognitif, Pendidikan Agama Islam, Taksonomi Bloom*

How to Cite:
Bridging Planning and Practice: Cognitive Assessment Based on Bloom's Taxonomy in Islamic Religious Education at Senior High Schools. (2026). *Arfannur: Journal of Islamic Education*, 7(1), 1-14.
https://doi.org/10.24260/arfannur.v7i1.5652

A. INTRODUCTION

Islamic Religious Education (IRE) is a conscious and planned effort to prepare students to understand, practice, and master the teachings of Islam. Beyond transferring knowledge, IRE plays a strategic role in shaping character and fostering critical thinking skills (Demirtaş et al., 2023). One of the essential aspects of IRE instruction is cognitive assessment, which functions not only as a tool for measuring learning outcomes but also as a pedagogical instrument to cultivate higher-order thinking skills. However, empirical realities in the field reveal a significant gap between curriculum design and assessment implementation. Teachers' instruments remain dominated by lower-order cognitive questions, while the application of items targeting higher-order thinking skills is still very limited (Zana et al., 2024).

Addressing this issue requires a critical review of teachers' practices in planning, implementing, and utilizing cognitive assessments in schools. Arbeni et al. (2025) emphasize that the systematic development of instruments, accompanied by tests of construct validity, internal consistency, and test–retest reliability, is a crucial step to ensure assessment quality. Furthermore, assessment strategies that combine written tests, oral examinations, and classroom observations are considered effective in enriching contextual learning aligned with students' characteristics. Other findings indicate that authentic assessments through various activities such as role-playing, storytelling, the use of stimuli, questioning, presentations, and language games can provide a more comprehensive picture of students' abilities (Ramlan et al., 2025). Ideally, assessment utilization should be directed toward a diagnostic–formative function to improve instructional strategies, which can be further strengthened through the application of digital formative assessments in e-learning systems to provide rapid, relevant, and continuous feedback (Ariza & Afifah, 2024).

A synthesis of existing literature reveals that research on cognitive assessment in IRE predominantly clusters around two themes: the quantitative evaluation of instrument quality and the effectiveness of specific interventions. Studies by Anwar et al. (2021) and Fakhruddin et al. (2025) have extensively documented the dominance of Lower-Order Thinking Skills (LOTS) in teacher-made tests, highlighting a persistent gap between curricular ideals and practical realities. Meanwhile, research by Insani et al. (2020) and Ariza and Afifah (2024) focuses on technical solutions, such as remedial programs and digital tools, to enhance learning outcomes. However, these studies largely adopt a normative–evaluative approach, focusing heavily on the product of assessment rather than the pedagogical process. In contrast to these prior quantitative or survey-based studies, this research provides an in-depth qualitative examination of how assessment paradigms are implemented and negotiated in daily instructional practice, thereby moving beyond mere identification of competency deficits to

understanding the underlying systemic dissonance.

To dissect the complexity of assessment practices, this study adopts a theoretical framework distinguishing two dominant assessment paradigms. First, the diagnostic–formative paradigm, rooted in the views of Black and Wiliam (1998), positions assessment as a continuous feedback mechanism to improve instructional strategies and systematically close students' cognitive gaps. In the context of Bloom's Taxonomy, this paradigm demands strict structural adherence to ensure construct validity. Second, the summative–applicative paradigm, which is more oriented toward learning outcome accountability and social relevance (Taras, 2005). This paradigm emphasizes contextual flexibility where assessment serves as a bridge between textual material and students' lived realities. The dialectical tension between these two paradigms serves as the primary theoretical lens of this study in analyzing why the implementation of HOTS is frequently impeded in IRE classrooms.

Addressing these identified gaps, this study specifically aims to describe the practices of IRE teachers in planning cognitive assessments at the senior high school level, analyze their classroom implementation strategies, and identify patterns in the utilization of assessment results. This study also evaluates the extent to which these three aspects support the development of students' higher-order thinking skills. Theoretically, this research is grounded in Bloom's Taxonomy, which categorizes six levels of thinking: remembering (C1), understanding (C2), applying (C3), analyzing (C4), evaluating (C5), and creating (C6).

The findings of this study are expected to provide both practical and theoretical contributions. Practically, they may serve as a reference for teachers to improve the quality of HOTS-based assessment instruments, enrich implementation strategies through contextual approaches, and optimize the use of assessment results for instructional improvement. Theoretically, this study contributes to strengthening the literature on cognitive assessment in IRE learning and serves as a foundation for developing more adaptive and transformative assessment models to support the goals of 21st-century education.

B. METHOD

This study employed a qualitative approach with an instrumental case study design (Afista et al., 2020) to explore cognitive assessment practices in Islamic Religious Education (IRE) at the senior high school level. The research was conducted at a state senior high school in Bandung, West Java, which serves as a representative model of urban public education in Indonesia. The school is characterized by a heterogeneous student body from diverse socio-economic backgrounds and operates under the tight regulatory framework of the National Curriculum. This setting was purposively selected because it encapsulates the central tension of the study: the institutional demand for administrative standardization versus the dynamic need for contextual religious pedagogy in a pluralistic urban environment. Regarding participant selection, this study employed a purposive sampling strategy based on the logic of "information-rich cases" (Patton, 2015). The recruitment of two IRE teachers (GPAI01 and GPAI02) is epistemologically justified not by the pursuit of statistical representativeness, but by the objective of analytical generalization (Yin, 2018). These two subjects were selected because they exhibit contrasting pedagogical approaches one representing the systematic-structural pole and the other the flexible-contextual pole. This 'dialectic contrast' allows the study to provide a thick description of how different assessment paradigms are negotiated within the same structural constraints, offering depth of insight that a larger, surface-level survey could not capture.

Data were collected using three primary techniques: (1) structured in-depth interviews to obtain information regarding strategies for planning, implementing and utilizing cognitive assessment (Achmad et al., 2022); (2) classroom observations to record assessment practices directly, including test formats, modes of delivery, and feedback mechanisms (Mustofa et al., 2023); and (3) document analysis of assessment instruments used by teachers, focusing on the distribution of cognitive levels based on Bloom's Taxonomy (C1–C6) (Zakkiyah et al., 2024). Data analysis followed the interactive model of Miles et al. (2014) which consists of three stages: data reduction, data display, and conclusion drawing. To ensure a more systematic analysis, each informant was assigned an identity code: GPAI01 for Teacher 1 and GPAI02 for Teacher 2. To enhance data validity, methodological triangulation was applied by comparing the results of interviews, observations, and document analysis, as well as conducting member checking to confirm interview findings with participants (Sofa et al., 2025) and aligning them with previous research findings.

C. RESULT AND DISCUSION

Planning of Cognitive Assessment in Islamic Religious Education at Senior High School

The planning of cognitive assessment by Teacher 1 and Teacher 2 demonstrates fundamental differences in approach, particularly in terms of systematic design, determination of cognitive levels, and focus of content. These differences reflect the variation in teachers' strategies for aligning learning objectives with assessment instruments. A summary of the comparison is presented in table 1.

Table 1. Comparative analysis of cognitive assessment planning

Aspect	GPAI01	GPAI02
Planning approach	Structured and sequential: objectives, test blueprint, indicators, rubric, distribution of questions across Bloom's Taxonomy (C1–C6).	Simple and contextual: questions based on the textbook, divided by chapter, adjusted to classroom conditions.
Content focus	Conceptual foundation and theoretical-philosophical aspects.	Relevance to students' real-life experiences.
Limitations	Dominance of LOTS items; difficulty in designing HOTS-based stimuli.	Uneven distribution of questions; limited stimulation of HOTS.

Source: interviews and document analysis

Table 1 highlights significant differences in approach. Based on interview results, Teacher 1 emphasized a systematic structure by designing assessments sequentially, beginning with the determination of learning objectives, followed by the development of test blueprints, indicators, rubrics, and the distribution of questions across Bloom's cognitive levels. This approach ensured coherence between learning outcomes and assessment instruments, so that assessments measured not only memorization but also critical thinking skills. As Teacher 1 stated: *"The planning process begins with determining learning objectives based on the basic competencies and indicators. Next, I prepare the test blueprint, question types, and distribution across cognitive levels C1–C6, complete with rubrics and trial testing before implementation"* (GPAI01, Personal Communication, May 2025). This indicates strong

methodological awareness of the alignment between objectives, indicators, and assessment instruments. These findings are consistent with Gunartha (2024) who concluded that Bloom's Taxonomy based test construction fosters mastery of critical thinking skills and significantly differentiates between high and low achieving students.

Teacher 2, conversely, prioritized simplicity and flexibility, adopting a strategy of designing questions directly from the textbook and dividing them proportionally by chapter, and adapting them to classroom conditions and student characteristics. This was reflected in the statement: *"I look at the number of chapters; if there are three chapters, then 30 questions are divided equally, ten per chapter. For the questions, I take them from the textbook, since that is what the students use"* (GPAI02, Personal Communication, April 2025). Such planning illustrates a more practical and contextual practice, aligned with the spirit of flexibility in learning. This finding resonates with Saa (2024) who emphasized that the Merdeka Curriculum provides teachers with greater autonomy to design adaptive assessments according to students' needs.

Document analysis of assessment instruments, contrasting with the relatively ideal planning suggested by interviews, revealed a divergence between design and reality. The instruments developed by both teachers demonstrated a significant quality gap between the test blueprint and the final product. To provide a deeper analytical overview of the instrument characteristics, table 2 presents a comparison of representative question items developed by both teachers.

Table 2.
Comparison of Assessment Items (Representative Samples)

Aspect	Real Question Item	Cognitive Analysis
GPAI01	"Ilmu Kalam is the science that discusses..." (Answer options: A. History of Islam... C. Islamic Aqidah with rational arguments...)	C1 (Remembering): This item strictly demands the recall of a textbook definition without requiring contextual processing or deep reasoning.
GPAI02	"After graduating, Farhan got a decent job... He wants to marry to avoid prohibited acts. The ruling on marriage for Farhan is..." (Answer options: A. Makruh... B. Wajib...)	C3 (Applying): This item presents a social case (contextual), but the cognitive demand is limited to applying a standard fiqh rule to a specific situation.

Source: document analysis

The data in table 2 visualizes the distinct paradigms of the two teachers. The item constructed by GPAI01 (Teacher 1) purely demands the recall of a theoretical definition (C1), reflecting a structural-academic approach. Conversely, GPAI02 (Teacher 2) utilizes a case narrative ('Farhan') that is closer to students' reality; however, the core question merely requires the application of a basic legal rule (C3). Although Teacher 2's style is more contextual, neither item reaches the level of deep analysis (C4) or evaluation (C5).

This phenomenon indicates 'cognitive stagnation' within the teachers' instruments. The author's interpretation suggests that teachers struggle to transfer theoretical understanding of Bloom's Taxonomy into technical item construction. Teacher 1 is trapped in definitional formalism, while Teacher 2, despite attempting contextuality, is trapped in applicative simplification. Both fail to trigger critical reasoning (HOTS). This failure is not due to a lack of familiarity with HOTS terminology, but rather an inability to deconstruct complex

material into challenging stimuli. These findings confirm studies by Zana et al. (2024) and Kosasih et al. (2022), which report that despite curricular demands for HOTS, teacher assessment products often degrade into mere memory tests due to weak technical skills in instrument construction.

Teacher 1 tended to emphasize theoretical–philosophical aspects with a small proportion of analytical questions (C4), while Teacher 2 emphasized contextual questions but remained limited to the same levels (C3–C4). These findings indicate that, despite efforts to integrate Bloom's Taxonomy into planning, the actual implementation of test items remained confined to LOTS (Lower Order Thinking Skills). This condition is consistent with Mardatillah et al. (2025) who asserted that Islamic education practices in schools are still largely dominated by traditional approaches based on memorization and authority, and thus have not fully encouraged creativity and critical thinking skills.

Beyond differences in cognitive levels, analysis of assessment instruments also revealed variation in content focus. Teacher 1 was oriented toward strengthening conceptual foundations, which remained within lower-order domains, while Teacher 2 emphasized the relevance of content to students' real-life experiences, though without fully accommodating higher-order cognitive levels. This variation not only reflects contextual adaptation but also underscores the challenge of designing genuinely HOTS-based items. These findings align with Kosasih et al. (2022) who noted that although the 2013 Curriculum normatively demands HOTS-based learning, IRE practice in schools remains dominated by LOTS items, with a significant gap between curricular expectations and classroom implementation.

These findings collectively underscore the critical importance of integrating systematization and flexibility in cognitive assessment planning. Teachers must be able to design instruments that are not only methodologically structured but also responsive to classroom contexts and student characteristics. Strengthening teachers' competence in developing HOTS-based items is crucial to bridging the gap between curricular demands and classroom practice. Enhancing teacher capacity thus becomes a key strategy in promoting more meaningful assessments, while supporting Islamic Religious Education that is adaptive, contextual, and transformative in line with the goals of 21st-century education.

Fundamentally, this gap between neat planning and weak item production stems from the issue of assessment literacy. In the context of this study, teachers' assessment literacy appears to have reached the stage of 'administrative literacy' (knowing how to create lesson plans and blueprints) but has not yet attained 'substantial literacy' (the ability to construct valid measurement tools to stimulate reasoning). The inability to operationalize HOTS operational verbs into meaningful questions demonstrates a fundamental competence deficit. This aligns with the argument of Susilawati et al. (2023), who emphasize that the main obstacle to HOTS implementation is not the availability of the curriculum, but the low assessment literacy of teachers in designing instruments that go beyond rote memorization.

Implementation Strategies of Cognitive Assessment in Islamic Religious Education at Senior High Schools

The strategies for implementing cognitive assessment in Islamic Religious Education (IRE) at the senior high school level reflect how teachers translate planned instruments into classroom practice. These strategies encompass test formats, modes of delivery, instructional methods, and the provision of feedback to students. The findings reveal contrasting approaches between Teacher 1 and Teacher 2, highlighting the tension between

systematization and flexibility. A comparison is presented in table 3.

Table 3.
Comparative analysis of cognitive assessment implementation

Aspect	GPAI01	GPAI02
Test format	Written tests (multiple choice, short essays), formal instruments with rubrics.	Combination of written tests, oral tests, direct observation and spontaneous questioning.
Cognitive focus	Conceptual understanding (C1–C3) with limited analysis (C4).	Analysis and evaluation (C4–C5) through contextual discussion.
Question characteristics	Theoretical and conceptual, emphasizing definitions and memorization.	Contextual, linking material to contemporary issues and real-life experiences.

Source: interviews, classroom observations and document analysis

Table 3 illustrates the contrast between systematic and flexible strategies. Based on interviews, Teacher 1 implemented a more structured and systematic approach, relying on written tests, formal instruments, and predetermined rubrics. This was complemented by oral tests to assess aspects not fully captured by written formats, particularly Qur'anic memorization. As Teacher 1 explained: *"Technically, I use several strategies in designing tests. There are written tests such as multiple choice or essays, but also oral tests, especially to assess students' memorization of Qur'anic verses. The aim is to ensure that the assessment truly measures students' cognitive abilities comprehensively"* (GPAI01, Personal Communication, May 2025). This reflects not only an emphasis on the accuracy of formal instruments but also an effort to complement them with alternative approaches for a more holistic evaluation.

Adopting a distinct approach, Teacher 2 employed a more practical and flexible strategy that combined written and oral tests with direct observation and spontaneous questioning. As noted: *"The strategy I usually use is simple, asking questions directly in class. I often ask spontaneously about material being studied or already covered. There is not always a formal pre-test, but sometimes I conduct informal initial checks to gauge their understanding"* (GPAI02, Personal Communication, April 2025). This indicates a contextual and adaptive orientation, prioritizing quick and practical mapping of students' understanding over rigid formal procedures. Such variation aligns with Syafii et al. (2021) who emphasized that assessment quality depends not only on test variety but also on the instrument's ability to proportionally differentiate student abilities.

Despite differences in planning and implementation, classroom observations revealed that both teachers integrated formative feedback as an essential component of their assessment strategies. Teacher 1 provided feedback after group presentations, focusing on factual accuracy, while Teacher 2 embedded feedback within interactive discussions and added rewards to enhance motivation. This underscores the pedagogical role of formative assessment, consistent with Amalia et al. (2024) who highlighted the importance of discussion and collaboration based formative assessment in the Merdeka Curriculum, where constructive feedback plays a crucial role in fostering engagement, motivation, and learning outcomes.

Document analysis, nevertheless, revealed that both teachers' formal instruments remained dominated by lower-order cognitive levels. Most items were at C1–C2, with limited

attempts at C3–C4, and almost no items at C5–C6. This finding aligns with Susilawati et al. (2023) who reported that school assessments are still largely dominated by LOTS, with HOTS-based items rarely employed due to teachers' limited understanding and skills in instrument design. Furthermore, differences in question characteristics were evident: Teacher 1 tended to design theoretical–conceptual items emphasizing definitions and memorization, while Teacher 2 developed contextual items linking IRE content to contemporary issues. This contextual approach resonates with Amirudin et al. (2025) who found that Problem-Based Learning (PBL) in IRE enhances content relevance to real life, increases motivation, and fosters critical thinking skills.

Nevertheless, both teachers faced challenges in implementing instruments capable of measuring higher-order cognitive levels. Teacher 1 cited time constraints and limited training in HOTS item construction as major obstacles, while Teacher 2 highlighted student diversity and reliance on personal experience as barriers to optimal assessment strategies. These challenges suggest that the barriers to HOTS implementation are not merely individual but also structural. This is consistent with Karwadi et al. (2024) who emphasized the need for teacher competence development and innovative instructional strategies to strengthen HOTS-based assessment in Islamic education. Similarly, Jima'ain et al. (2024) found that the application of HOTS in *Sirah* learning remained at a moderate level, particularly in lesson closure, stressing the importance of teacher competence in planning, content mastery, and reflective methods to optimize HOTS-based assessment (Ilmi et al., 2020).

The effective implementation of cognitive assessment strategies, ultimately, necessitates an integration of systematic planning and contextual flexibility. Enhancing teacher competence particularly in HOTS item construction and strategy diversification therefore becomes essential to ensure that cognitive assessment functions not only as a measurement tool but also as a pedagogical instrument that fosters students' critical thinking skills.

Utilization of Cognitive Assessment Results in Islamic Religious Education at Senior High Schools

The utilization of cognitive assessment results plays a strategic role, not only as the basis for grading but also as a pedagogical instrument to strengthen the learning process, design remedial and enrichment programs, and improve instructional strategies. The findings indicate that Teacher 1 and Teacher 2 adopted different patterns of utilization, reflecting two distinct paradigms. A summary of the comparison is presented in table 4.

Table 4.
Comparative analysis of cognitive assessment utilization

Aspect	GPAI01	GPAI02
Utilization orientation	Differentiated instruction and adaptive instructional development.	Grading, methodological reflection and accountability evaluation.
Form of feedback	Direct interaction, focusing on conceptual reinforcement.	Oral feedback in discussions and summative evaluation.
Approach characteristics	Theoretical-philosophical, emphasizing conceptual reinforcement and basic understanding.	Contextual-applicative, linking assessment to social issues and real-life contexts.

Source: interviews and classroom observations

Table 4 illustrates two distinct paradigms in the utilization of cognitive assessment. Based on interviews, Teacher 1 optimized assessment results as the basis for differentiated instruction by designing varied questions according to students' cognitive levels and providing targeted feedback. As Teacher 1 stated: *"I use cognitive assessment results primarily to improve the effectiveness of future learning by identifying students' strengths and weaknesses. From there, we can design more precise instructional planning"* (GPAI01, Personal Communication, May 2025). This indicates that Teacher 1 employed assessment not merely as a grading tool but as a formative mechanism to design adaptive strategies aligned with students' needs.

In contrast, Teacher 2 emphasized the use of assessment for summative–operational purposes, focusing on evaluation and accountability. As expressed: *"After obtaining the assessment results, I record the grades in the report card. The scores are also communicated to students so they know their results"* (GPAI02, Personal Communication, April 2025). This shows that Teacher 2's utilization of assessment was oriented more toward administrative and accountability functions, prioritizing the reporting of student achievement rather than using assessment as a foundation for subsequent instructional planning. This distinction aligns with Sajja et al. (2023) who identified two paradigms of assessment utilization: formative–diagnostic for adaptive instructional development and summative–operational for educational accountability.

Classroom observations further revealed significant differences in the quality of feedback provided. Teacher 1 used assessment results to reinforce students' conceptual understanding by correcting misconceptions, allowing students to summarize material, and motivating them through creative approaches. Teacher 2, on the other hand, emphasized assessment results as benchmarks for students' analytical and evaluative skills in relation to contemporary issues, primarily through contextual discussions accompanied by oral feedback. These findings suggest that while both teachers integrated assessment into instruction, their feedback orientations differed: Teacher 1 focused on conceptual reinforcement and learning motivation, whereas Teacher 2 emphasized analytical skills within discussion contexts. This is consistent with Black and Wiliam (1998) who demonstrated that formative feedback specifically addressing students' errors strengthens conceptual understanding, enhances motivation, and supports independent learning skills.

Interviews and observations also identified several challenges hindering the optimal utilization of cognitive assessment results. Teacher 1 cited time constraints, administrative demands, and limited skills in assessment analysis as barriers to providing comprehensive feedback. Teacher 2 highlighted large class sizes, limited time, and the dominance of

summative functions as obstacles to maximizing assessment results for instructional development. These findings are consistent with Mustofa et al. (2023) who emphasized that time limitations, insufficient teacher training, and lack of resources are major barriers to implementing learning-oriented assessment. Nevertheless, these challenges also present opportunities for significant development. One potential avenue is to revisit curriculum design and instructional practices to be more student-centered, aligned with scientific developments, and integrated across cognitive, affective, and psychomotor domains. This aligns with Sidik et al. (2024) who stressed that integrating Islamic curriculum with modern scientific knowledge can serve as a model for strengthening contextual and sustainable IRE.

In conclusion, effective utilization of cognitive assessment results requires the integration of diagnostic–formative and contextual–applicative paradigms, supported by teachers' competence in analyzing assessment outcomes to design both remedial and enrichment programs based on higher-order thinking skills (HOTS). Optimizing this utilization necessitates continuous professional development programs that emphasize assessment analysis, differentiated strategies, and the use of technology to support more effective and transformative feedback.

D. CONCLUSION

This study demonstrates that cognitive assessment in Islamic Religious Education (IRE) at the senior high school level is implemented through different approaches by each teacher, reflecting two main paradigms: systematization based on Bloom's Taxonomy and contextual flexibility. Although assessment planning has been directed toward the integration of higher-order thinking skills, the implementation of instruments remains dominated by lower-order thinking skills. The utilization of assessment results also reveals divergent orientations, namely the diagnostic–formative function for instructional development and the summative–applicative function for evaluating learning outcomes. Structural constraints such as limited time, insufficient training, and the dominance of administrative functions constitute major challenges in optimizing assessment practices. Therefore, strengthening teacher competence through continuous professional development, leveraging educational technology, and implementing evaluation policies that emphasize formative functions are essential to ensure that cognitive assessment can play a more effective role in enhancing the quality of IRE learning in schools. This study advances the conceptual understanding of cognitive assessment in IRE by framing assessment practices as a dynamic negotiation between systematisation and contextual flexibility.

E. BIBLIOGRAPHY

Achmad, G. H., Ratnasari, D., Amin, A., Yuliani, E., & Liandara, N. (2022). Penilaian Autentik pada Kurikulum Merdeka Belajar dalam Pembelajaran Pendidikan Agama Islam di Sekolah Dasar. *Edukatif: Jurnal Ilmu Pendidikan*, *4*(4), 5685–5699. https://doi.org/10.31004/edukatif.v4i4.3280

Afista, Y., Priyono, A., & Huda, S. A. A. (2020). Analisis Kesiapan Guru PAI dalam Menyongsong Kebijakan Merdeka Belajar (Studi Kasus di MTSN 9 Madiun). *Journal of Education and Management Studies (JoESM)*, *3*, 53–60. https://ojs.unwaha.ac.id/index.php/joems/article/view/338

Amalia, E. R., Indri, D. B., Khoiriyati, S., Ummah, N., Oviani, M., Kusumawardhani, J., & Umayyah, U. (2024). Bridging Educational Reform and Faith: Evaluating Kurikulum

Merdeka's Compatibility with Islamic Values in Madrasahs. *Munaddhomah: Jurnal Manajemen Pendidikan Islam*, *5*(4), 483–500. https://doi.org/10.31538/munaddhomah.v5i4.1413

Amirudin, A., Muzaki, I. A., & Nurhayati, S. (2025). Problem-Based Learning as a Pedagogical Innovation for Transforming Higher Education Students' Islamic Religious Comprehension. *Educational Process: International Journal*, *18*. https://doi.org/10.22521/edupij.2025.18.412

Anwar, S., Supriadi, U., & Hermawan, W. (2021). Analisis Instrumen Tes Kognitif Mata Pelajaran Pendidikan Agama Islam dan Budi Pekerti di SD: Aspek Substansi, Konstruksi, Bahasa, dan High Order Thinking Skill. *Taklim: Jurnal Pendidikan Agama Islam*, *19*(1), 27–64. https://ejournal.upi.edu/index.php/taklim/article/view/40789

Arbeni, W., Windiani, A., Sihotang, D. S. B., Anggraini, N., Wulandari, S., & Nugroho, A. (2025). Test Reliability Analysis in Educational Evaluation: A Quantitative Approach to Consistency and Validity. Jurnal Nasional Holistic Science, 5(1), 59–64. https://doi.org/10.56495/hs.v5i1.838

Ariza, N., & Afifah, Q. K. (2024). Penggunaan Teknologi dalam Pengembangan Asesmen Pembelajaran Pendidikan Agama Islam. Balajea: Jurnal Pendidikan Islam, 9, 25–44. https://journal.iaincurup.ac.id/index.php/belajea/article/view/8840

Basri, B., & Kaharuddin, K. (2022). Kualitas Pembelajaran Pendidikan Agama Islam di SMA Negeri 1 Jeneponto. Jurnal Al-Qiyam, 3, 151–159. https://ojs.staialfurqan.ac.id/alqiyam/article/view/249

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

Demirtaş, Z., Takunyaci, M., & Yeşil, R. (2023). Investigation of Religious Education Teacher Candidates' Cognitive Flexibility Levels and Critical Thinking Tendencies. Hitit Theology Journal, 22(2), 723–744. https://doi.org/10.14395/hid.1334350

Fakhruddin, A., Anwar, S., Romli, U., Firmansyah, M. I., & Faqihuddin, A. (2025). Pelatihan Implementasi Kurikulum Merdeka di Sekolah bagi Guru Pendidikan Agama Islam di Kabupaten Garut: Meningkatkan Kualitas Pembelajaran PAI di Era Kurikulum Merdeka. PengabdianMu: Jurnal Ilmiah Pengabdian Kepada Masyarakat, 10(2), 596–606. https://doi.org/10.33084/pengabdianmu.v10i2.8775

Gunartha, I. W. (2024). Pengembangan Penilaian Berorientasi HOTS: Upaya Peningkatan Kemampuan Berpikir Kritis Siswa di Era Global Abad Ke-21. Widyadari Jurnal Pendidikan, 25(1), 133–147. https://doi.org/10.59672/widyadari.v25i1.3660

Ilmi, A. M., Sukarmin, & Sunarno, W. (2020). Development of TPACK Based-Physics Learning Media to Improve HOTS and Scientific Attitude. Journal of Physics: Conference Series, 1440(1), 1–6. https://doi.org/10.1088/1742-6596/1440/1/012049

Insani, A. N., Anwar, S., & Supriadi, U. (2020). Efektivitas Program Remedial Terhadap Peningkatan Hasil Belajar Peserta Didik pada Mata Pelajaran Pendidikan Agama Islam dan Budi Pekerti di SMA (Studi Eksperimental Kelas X SMA Negeri 1 Banjaran Kabupaten Bandung). Taklim: Jurnal Pendidikan Agama Islam, 18(1), 65–74. https://ejournal.upi.edu/index.php/taklim/article/view/32820

Jima'ain, M. T. A., Mohamed, A. M., Hehsan, A., Saidalvi, A., Nasi, B. M., Faisal, M. S., & Awae, F. (2024). Embedding higher Order Thinking Skills in Islamic History (Sirah) Education in Malaysia. International Journal of Evaluation and Research in Education, 13(2), 952–959. https://doi.org/10.11591/ijere.v13i2.26431

Karwadi, K., Zakaria, A. R., & Syafii, A. (2024). A Review of The Effects of Active Learning on High Order Thinking Skills: A Meta-Analysis Within Islamic Education. Journal of

Education and Learning (EduLearn), 18(1), 97–106. https://doi.org/10.11591/edulearn.v18i1.20895

Kosasih, A., Supriyadi, T., Firmansyah, M. I., & Rahminawati, N. (2022). Higher-Order Thinking Skills in Primary School: Teachers' Perceptions of Islamic Education. Journal of Ethnic and Cultural Studies, 9(1), 56–76. https://doi.org/10.29333/ejecs/994

Mardatillah, F., Gumilang, R. M., Wahyudi, M. A., Rawanita, M., & Muhammad, M. (2025). Epistemological Reconstruction of Islamic Education: Developing a Transformative Pedagogical Model to Foster Creativity. Jurnal Ilmiah Peuradeun: The Indonesian of the Social Sciences, 13(2), 1071–1094. https://doi.org/10.26811/peuradeun.v13i2.2200

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). Qualitative Data Analysis A Methods Sourcebook (H. Salmon, K. Perry, K. Koscielak, & L. Barrett, Eds.; 3rd ed.). SAGE Publications Asia-Pacific. https://doi.org/10.1177/239700221402800402

Mustofa, A., Oktavia, V., & Himami, A. S. (2023). Problematika Penerapan Kurikulum Merdeka dalam Mata Pelajaran Pendidikan Agama Islam. Irsyaduna: Jurnal Studi Kemahasiswaan, 3(3), 350–360. https://doi.org/10.54437/irsyaduna.v3i3.1350

Patton, M. Q. (2015). Qualitative research & evaluation methods: Integrating theory and practice (4th ed.). SAGE Publications.

Ramlan, N., Mohammad, W. M. R. W., & Mahamod, Z. (2025). Authentic Assessment Activities for Oral Malay Language Skills in Primary Schools. International Journal of Evaluation and Research in Education (IJERE), 14(1), 682–694. https://doi.org/10.11591/ijere.v14i1.32079

Saa, S. (2024). Merdeka Curriculum: Adaptation of Indonesian Education Policy in the Digital Era and Global Challenges. RGSA - Revista de Gestao Social e Ambiental, 18(3), 1–24. https://doi.org/10.24857/rgsa.v18n3-168

Sajja, R., Sermet, Y., Cwiertny, D., & Demir, I. (2023). Integrating AI and Learning Analytics for Data-Driven Pedagogical Decisions and Personalized Interventions in Education. Technology, Knowledge and Learning, 2, 1–31. https://doi.org/10.1007/s10758-025-09897-9

Sidik, M. F., Vachruddin, V. P., Rusydiyah, E. F., Pertiwi, A. S., & Darmawan, M. A. (2024). Conceptualization of the Integrated Islamic Religious Education Curriculum: A Literature Study at Imam Hatip Schools Turkey and MAN Insan Cendekia Indonesia. Jurnal Pendidikan Agama Islam, 21(1), 111–130. https://doi.org/10.14421/jpai.v21i1.7617

Sofa, A. R., Firdausiyah, J., Putri, I. D. I. S., Romli, M., Bukhori, M. I., & Syamsuddin. (2025). Pengembangan Penilaian Pembelajaran PAI Berbasis Learning Analyties, IBM Watson Education, Adaptive Learning AI: Motivasi dan Konsekuensi di MTs Mambaul Hikam. Indonesian Research Journal on Education Web Jurnal Indonesian Research Journal on Education, 5, 720–727. https://doi.org/10.31004/irje.v5i2.2279

Susilawati, S., Yaqin, M. Z. N., Wahidmurni, Ulfa, F. A., & Krismoneta. (2023). Training on Preparation of Islamic Integrated HOTS-Based Questions and Their Application in Online Learning in Elementary Schools. Munaddhomah: Jurnal Manajemen Pendidikan Islam, 4(4), 867–877. https://doi.org/10.31538/munaddhomah.v4i4.459

Syafii, A., Haryanto, Ahmad, I. F., & Fauziah, A. (2021). Analysis of Items with Item Response Theory (IRT) Approach on Final Assessment for Al-Quran Hadith Subjects. Jurnal Pendidikan Agama Islam, 18(1), 167–194. https://doi.org/10.14421/jpai.2021.181-09

Taras, M. (2005). ASSESSMENT – SUMMATIVE AND FORMATIVE – SOME THEORETICAL REFLECTIONS. British Journal of Educational Studies, 53(4), 466–478. https://doi.org/10.1111/j.1467-8527.2005.00307.x

Yin, R. K. (2018). Case study research and applications: Design and methods (6th ed.). SAGE Publications.

Zakkiyah, F., Armedi, R., & Zuhriyah, I. A. (2024). Analisis Instrumen Tes Kognitif Mata Pelajaran Pendidikan Agama Islam di Sekolah Dasar. Jurnal Evaluasi Pendidikan, 15(2), 63–75. https://doi.org/10.21009/jep.v15i2.50143

Zana, F. M., Sa'dijah, C., Susiswo, Anwar, L., & Zulnaidi, H. (2024). Curriculum and Teacher Assessment Practices in Mathematics Learning: Alignment with Higher Order Thinking Skills in Indonesian Secondary Schools. Journal on Mathematics Education, 15(4), 1311–1334. https://doi.org/10.22342/jme.v15i4.pp1311-1334