



Assessing The Quality of Learning Assessment in Islamic Theology and Ethics: A Psychometric and Cognitive-Level Analysis of Grade X Textbook Questions in Indonesian Islamic Senior High Schools

Dalvah Nur Azizah

UIN Sunan Gunung Djati Bandung, Indonesia

dalvaahnz@gmail.com

Article history: Received: 10, 2023; Accepted: 11, 2023; Published: 12, 2023

DOI: <https://doi.org/10.15575/jkgk.v1i2.343>

Abstract: This study aims to examine the quality of assessment items in the Grade X Islamic Theology and Ethics (Akidah Akhlak) textbook used in Indonesian Islamic senior high schools by analyzing their cognitive levels and classification into Lower Order Thinking Skills (LOTS), Middle Order Thinking Skills (MOTS), and Higher Order Thinking Skills (HOTS). The study employed a qualitative descriptive research design with document analysis as the primary data collection technique. The units of analysis consisted of 45 essay-type questions distributed across four chapters of the textbook. The findings reveal a strong dominance of LOTS-oriented items, with 39 questions (88%) categorized as LOTS, while only 3 questions (6%) reflected MOTS and 3 questions (6%) met the criteria for HOTS. In terms of cognitive level, most items were concentrated at the C2 (understanding) level, indicating limited opportunities for students to engage in analytical, evaluative, or creative thinking processes. These results suggest that the assessment items in the analyzed textbook have not yet adequately aligned with contemporary curriculum demands that emphasize higher-order thinking and 21st-century skills. The predominance of low-level cognitive questions may constrain students' critical thinking development and reduce the effectiveness of assessment in Islamic education. Therefore, teachers, textbook authors, and curriculum developers are encouraged to redesign assessment tasks by incorporating more HOTS-oriented items to enhance cognitive challenge, pedagogical quality, and alignment with modern educational standards in Islamic senior high schools.

Keyword: Cognitive Level Analysis; Higher Order Thinking Skills; Textbook Evaluation

Copyright: © 2023. The Authors.

Licensee: Jurnal Keprofesional Guru Keagamaan is licensed under the [Creative Commons Attribution License](https://creativecommons.org/licenses/by-sa/4.0/).

INTRODUCTION

Evaluation is an integral component of the learning process, serving as a systematic means to determine the extent to which instructional objectives have been achieved and to assess the overall quality of educational practices. In the context of formal education, evaluation not only functions as a tool for measuring learning outcomes but also provides critical information for improving curriculum design, instructional strategies, and assessment practices (Purwanto, 2008). One of the most commonly used instruments in educational evaluation is the test, which is designed to measure students' cognitive achievement across various levels of thinking.

High-quality test instruments are a key indicator of effective learning, as they reflect the alignment between learning objectives, instructional content, and expected student competencies. Cognitive assessment instruments typically aim to measure students' abilities across different cognitive domains, including knowledge, comprehension, application, analysis, synthesis, and evaluation (Arifin, 2016). Several studies have emphasized that students' intellectual development can be observed through their performance on cognitive assessments, particularly those that challenge higher levels of thinking (Mulyana, Rochman, & Salim, 2019).

In contemporary educational discourse, cognitive assessment is commonly classified into three levels: Low Order Thinking Skills (LOTS), Middle Order Thinking Skills (MOTS), and High Order Thinking Skills (HOTS). Among these, HOTS has gained significant attention due to its alignment with 21st-century learning goals, which emphasize critical thinking, creativity, problem-solving, and decision-making (Arifin & Retnawati, 2015). HOTS refers to the ability to connect prior knowledge with new situations, analyze complex information, and generate innovative solutions to unfamiliar problems (Rofiah, Aminah, & Ekawati, 2013). Conklin (2012) further explains that HOTS encompasses critical, creative, reflective, logical, and metacognitive thinking skills.

Despite the growing emphasis on HOTS-oriented assessment, several studies have reported that learning materials and evaluation instruments, particularly textbooks, still tend to emphasize lower cognitive levels. This condition is also observed in religious education, including Islamic Religious Education subjects such as Akidah Akhlak at the Madrasah Aliyah level. Akidah Akhlak plays a strategic role in shaping students' theological understanding and moral reasoning; therefore, its assessment instruments should not merely test memorization of concepts but also encourage analytical, evaluative, and reflective thinking aligned with Islamic values and contemporary educational demands.

Essay-type questions are one form of assessment instrument that has the potential to measure higher-order thinking, as they require students to construct arguments, integrate ideas, and articulate reasoning (Mardapi, 2012). According to the revised Bloom's taxonomy, cognitive levels C1 (remembering) and C2 (understanding) are categorized as lower-order thinking, C3 (applying) represents middle-order thinking, while C4 (analyzing), C5 (evaluating), and C6 (creating) constitute higher-order thinking skills (Nafiati, 2021). Ideally, assessment instruments in senior secondary education should proportionally

include HOTS-oriented items to support students' critical and reflective capacities.

However, preliminary observations and a review of related literature indicate that systematic analyses of cognitive-level distribution in Akidah Akhlak textbooks, particularly at the Madrasah Aliyah level, remain limited. Existing studies tend to focus on general curriculum implementation or instructional strategies, while empirical investigations into the cognitive quality of textbook-based assessment items are still scarce. This gap raises concerns regarding whether the evaluation instruments embedded in Akidah Akhlak textbooks adequately support the development of higher-order thinking skills required by current curriculum standards and 21st-century competencies.

Therefore, this study aims to analyze the cognitive levels of essay questions contained in the Akidah Akhlak textbook for Grade X Madrasah Aliyah by examining their distribution across LOTS, MOTS, and HOTS categories, as well as their alignment with revised Bloom's taxonomy. Through this analysis, the study seeks to contribute to the improvement of assessment quality in Islamic education by providing empirical evidence that can inform teachers, textbook authors, and curriculum developers in designing more cognitively balanced and meaningful evaluation instruments.

METHODOLOGY

This study employed a qualitative descriptive approach using content analysis to examine the cognitive quality of essay questions contained in the Akidah Akhlak textbook for Grade X Madrasah Aliyah. Content analysis was selected as it is an appropriate method for systematically evaluating textual materials, particularly assessment items, in order to identify patterns, classifications, and alignment with established theoretical frameworks.

The object of analysis consisted of essay-type questions embedded in the nationally standardized Akidah Akhlak textbook published by the Directorate of Curriculum, Facilities, Institutions, and Student Affairs (KSKK) Madrasah, Ministry of Religious Affairs of the Republic of Indonesia. This textbook was selected because it serves as the official reference used widely across Madrasah Aliyah institutions in Indonesia, making it pedagogically representative and relevant for evaluating assessment quality at the national level.

1. The bibliographic details of the textbook are as follows:
2. Title: Akidah Akhlak MA/SMA
3. Author: Dra. Nurul Hidayah
4. Editor: Dr. Siswanto, M.Pd.I
5. Publisher: Directorate of KSKK Madrasah, Directorate General of Islamic Education, Ministry of Religious Affairs of the Republic of Indonesia
6. Year of Publication: 2019
7. Curriculum: 2013 Curriculum
8. ISBN: 978-623-6729-12-0 (complete edition); 978-623-6729-15-1 (Volume 3)

A total of 45 essay questions were analyzed, distributed across four chapters: 8 questions in Chapter 1, 10 in Chapter 2, 15 in Chapter 3, and 12 in Chapter 4. These questions constituted the units of analysis in this study.

The analysis of cognitive levels was conducted using the Revised Bloom's Taxonomy framework developed by Anderson and Krathwohl, which categorizes cognitive processes into six levels: remembering (C1), understanding (C2), applying (C3), analyzing (C4), evaluating (C5), and creating (C6). Based on this framework, the questions were classified into three cognitive categories:

1. LOTS (Lower Order Thinking Skills): C1 and C2
2. MOTS (Middle Order Thinking Skills): C3
3. HOTS (Higher Order Thinking Skills): C4, C5, and C6

Each essay question was carefully examined based on its operational verbs, cognitive demands, and expected student responses to ensure accurate classification. The distinction between MOTS and HOTS was determined by whether the question required procedural application of knowledge (C3) or higher-level analytical, evaluative, or creative thinking (C4–C6).

To enhance analytical rigor, the qualitative coding process followed the stages of data condensation, data display, and conclusion drawing, as proposed by Miles, Huberman, and Saldaña. Initially, all questions were coded individually according to their cognitive level. The results were then tabulated to identify patterns and distributions across LOTS, MOTS, and HOTS categories.

To improve the credibility of the findings, the classification results were reviewed through peer examination by referencing established cognitive indicators within the taxonomy framework and comparing interpretations with existing literature on HOTS-based assessment. Any ambiguous items were re-examined to ensure consistency and theoretical alignment.

Finally, the analyzed data were presented in tabular form and interpreted descriptively to address the research objectives, focusing on the extent to which the textbook's assessment items support higher-order thinking skills in Islamic education at the Madrasah Aliyah level.

RESULTS AND DISCUSSION

Results

This study examined the cognitive characteristics of essay questions embedded in the *Sejarah Kebudayaan Islam (SKI)* textbook for Grade IX Madrasah Tsanawiyah, published by the Indonesian Ministry of Religious Affairs and officially designated as a national standard reference. A total of 35 essay questions across all chapters were analyzed using the revised Bloom's Taxonomy framework proposed by Anderson and Krathwohl (C1–C6). To enhance analytical clarity and methodological efficiency, the findings are presented in aggregated form rather than item-by-item descriptions.

Overall, the results reveal a clear dominance of lower-order cognitive demands within the textbook's assessment profile. Of the 35 analyzed questions, 19 items (54.3%) were classified as Lower Order Thinking Skills (LOTS), encompassing the cognitive levels of remembering (C1) and understanding (C2). A smaller proportion of items, 7 questions (20.0%), were categorized as Middle Order Thinking Skills (MOTS), primarily representing the application level (C3). Meanwhile, only 9 questions (25.7%) met the criteria of Higher Order Thinking

Skills (HOTS), corresponding to the cognitive levels of analyzing (C4), evaluating (C5), and creating (C6).

This distribution indicates that more than half of the assessment tasks are designed to measure students' ability to recall and comprehend historical facts, events, and figures in Islamic history. In contrast, relatively limited opportunities are provided for students to engage in deeper cognitive processes such as historical analysis, critical evaluation of events, or creative reconstruction of historical narratives. Although the presence of HOTS-oriented items is notable compared to some previous studies, their proportion remains insufficient to support a balanced development of higher-order thinking competencies.

When analyzed across chapters, the cognitive imbalance appears consistently throughout the textbook. While certain chapters include questions employing higher-level operational verbs such as analyze, decide, or create, these items are not systematically or evenly distributed. Most chapters continue to rely heavily on descriptive and factual prompts that require students to restate historical information rather than critically interpret Islamic historical phenomena or connect them to broader social, cultural, and moral contexts. As a result, the overall assessment structure does not reflect a progressive cognitive scaffold from lower- to higher-order thinking across chapters.

Further analysis of cognitive process distribution shows that the most dominant level is C1 (remembering), which accounts for 51.4% of all analyzed items. Questions at this level typically require students to recall names, dates, events, or definitions related to Islamic history. Higher cognitive levels, particularly C4 (analyzing) and C5 (evaluating), appear only sporadically. Although C6 (creating) is identified more frequently than C4 and C5, many of these items lack complex stimuli, authentic historical problems, or multi-perspective contexts. Consequently, despite being labeled as C6 based on operational verbs, their actual cognitive demand often remains limited.

These findings suggest that while the textbook demonstrates partial awareness of higher-order cognitive objectives, the pedagogical implementation of HOTS remains inconsistent. The occasional use of higher-level verbs is not always supported by instructional contexts that genuinely require complex reasoning, critical judgment, or creative historical thinking. Therefore, the assessment profile of the SKI textbook still predominantly reflects a traditional knowledge-transmission orientation rather than a fully developed higher-order thinking framework.

Discussion

The findings of this study reveal a clear and consistent pattern in the cognitive characteristics of assessment items contained in the Grade X Akidah Akhlak textbook used in Madrasah Aliyah. The dominance of Lower Order Thinking Skills (LOTS) items, accounting for 88% of the analyzed questions, indicates that the assessment design within the textbook remains largely oriented toward recall and basic comprehension. While such cognitive levels are essential as foundational learning stages, their overwhelming prevalence raises critical concerns regarding the textbook's alignment with contemporary educational demands, particularly those emphasizing Higher Order Thinking Skills (HOTS).

From the perspective of Bloom's revised taxonomy (Anderson & Krathwohl), effective assessment should progressively guide learners from remembering and understanding toward applying, analyzing, evaluating, and creating. However, the results of this study suggest that this cognitive progression is not adequately reflected in the textbook's essay questions. Most items rely on operational verbs such as explain, mention, and describe, which predominantly activate C1 and C2 cognitive processes. Although these verbs are pedagogically valid, their excessive use limits opportunities for students to engage in deeper cognitive processing, such as critical reasoning or reflective judgment. As a result, assessment practices risk reinforcing surface learning rather than fostering meaningful understanding of Islamic theological and ethical concepts.

The limited presence of Middle Order Thinking Skills (MOTS) and HOTS items further accentuates this issue. Only 6% of the questions were categorized as MOTS and another 6% as HOTS, indicating that students are rarely challenged to apply concepts, analyze moral dilemmas, or evaluate ethical issues within Islamic teachings. This imbalance suggests that the textbook assessment design prioritizes content transmission over cognitive engagement. In the context of Islamic education, such an approach is particularly problematic, as Akidah Akhlak is inherently value-oriented and reflective, requiring learners not only to understand doctrines but also to internalize and apply ethical principles in real-life situations.

The findings also highlight a discrepancy between curriculum aspirations and assessment implementation. Indonesia's educational policies, including the 2013 Curriculum and subsequent revisions, emphasize the development of critical thinking, creativity, and character education. These objectives align closely with the goals of Islamic education, which seeks to cultivate *insan kamil*—individuals who are intellectually competent, spiritually grounded, and morally responsible. However, the predominance of LOTS-oriented questions suggests that the textbook does not fully operationalize these curricular goals within its assessment components. Consequently, students may excel in memorizing religious concepts but struggle to contextualize and apply them in complex moral and social contexts.

Another critical issue revealed by this study is the limited use of contextual and stimulus-based questions. Effective HOTS assessment typically requires rich stimuli—such as case studies, narratives, or real-life scenarios—that prompt students to analyze, evaluate, and synthesize information. In contrast, many of the essay questions analyzed in this study were presented without sufficient context, reducing their cognitive demand even when higher-order verbs were employed. This finding supports previous research indicating that the presence of higher-level verbs alone does not guarantee HOTS quality; rather, cognitive complexity emerges from the interaction between verbs, content, and contextual stimuli.

The dominance of LOTS items may also reflect broader challenges in teacher assessment literacy and textbook development practices. Textbooks often serve as primary instructional and evaluative references in Madrasah Aliyah, particularly in contexts where teachers rely heavily on standardized materials. When textbooks predominantly feature low-level assessment tasks, they

inadvertently shape teachers' assessment practices and reinforce traditional evaluation models focused on rote learning. This cyclical pattern may hinder pedagogical innovation and limit teachers' capacity to design assessments that promote higher-order thinking.

Comparatively, the findings of this study are consistent with previous research on assessment practices in Islamic education textbooks. Studies by Ansori (2019) and Cahyaningsih (2020) similarly reported a disproportionate emphasis on LOTS items in religious education assessments, attributing this trend to conservative curriculum interpretations and limited training in HOTS-based assessment design. However, this study extends prior research by providing a systematic content analysis of essay questions and explicitly linking cognitive-level distribution to assessment quality and pedagogical implications. In doing so, it contributes to the growing body of literature advocating for assessment reform within Islamic education.

From a pedagogical standpoint, the implications of these findings are significant. Essay questions, by nature, offer considerable potential to assess higher-order thinking, as they allow students to articulate reasoning, construct arguments, and reflect on values. Yet, this potential remains underutilized when essay prompts merely require descriptive responses. By redesigning essay questions to incorporate analytical prompts, evaluative criteria, and authentic moral dilemmas, textbook authors and educators can transform assessments into tools for deeper learning and character formation. This is particularly relevant for Akidah Akhlak, where ethical reasoning and value-based judgment are central learning outcomes.

Furthermore, the limited representation of HOTS items raises concerns about students' preparedness for higher levels of education and contemporary societal challenges. In an era characterized by rapid social change, digitalization, and ethical complexity, students must be equipped with critical thinking and problem-solving skills rooted in Islamic values. Assessments that fail to cultivate these competencies risk producing learners who possess theoretical knowledge but lack the ability to navigate moral ambiguities and real-world challenges effectively.

In light of these considerations, this study underscores the urgent need for a paradigm shift in Islamic education assessment. Textbook developers, curriculum designers, and educators must collaboratively re-envision assessment practices to ensure a balanced distribution of cognitive levels. Integrating HOTS-oriented questions that are contextually rich, value-driven, and analytically demanding can enhance the role of assessment as a catalyst for learning rather than merely a measure of recall. Such reforms would not only align assessment practices with curriculum standards but also strengthen the transformative potential of Islamic education in shaping reflective, ethical, and critically minded learners.

CONCLUSION AND IMPLICATION

Conclusion

This study concludes that the Akidah Akhlak textbook for Grade X Madrasah Aliyah is critically limited in its capacity to promote higher-order thinking skills

among students. The analysis reveals a substantial dominance of Lower Order Thinking Skills (LOTS), accounting for 88% of the total essay questions, with most items concentrated at the levels of remembering and understanding (C1–C2). In contrast, Middle Order Thinking Skills (MOTS) and Higher Order Thinking Skills (HOTS) are minimally represented, each constituting only 6% of the assessed items. This imbalance indicates that the textbook assessment framework remains largely oriented toward factual recall and basic comprehension, rather than encouraging analytical, evaluative, and creative thinking as mandated by 21st-century learning standards and contemporary Islamic education curricula. Consequently, the limited presence of HOTS-oriented items may hinder students' development of critical reasoning and reflective moral judgment in learning Islamic theology and ethics. These findings underscore an urgent need for systematic revision of assessment components within Akidah Akhlak textbooks, particularly through the integration of context-based, problem-oriented, and reflective HOTS questions that align with modern curriculum demands and support the holistic goals of Islamic education.

Implication

Based on the findings of this study, which examined the cognitive characteristics of essay questions in the *Sejarah Kebudayaan Islam* textbook for Grade IX Madrasah Tsanawiyah, several important implications can be drawn for Islamic education assessment and textbook development. Although the textbook still demonstrates a dominant emphasis on lower-order thinking skills, the presence of several higher-order thinking items indicates its potential to be further developed as a learning resource that supports critical and analytical thinking. The HOTS-oriented questions identified in this study may serve as initial references for teachers in designing assessment instruments that emphasize analysis, evaluation, and historical reflection within Islamic education contexts.

Furthermore, this study highlights the urgent need for textbook authors, curriculum developers, and educators to reconstruct assessment designs to better align with 21st-century learning demands. Integrating more contextual, problem-based, and reflective questions would enable students not only to recall historical facts but also to interpret, evaluate, and draw meaningful lessons from Islamic historical events. In this regard, improving the quality and balance of cognitive levels in textbook assessments is essential to fostering students' critical thinking skills and strengthening the relevance of Islamic education in contemporary educational settings.

REFERENCES

- Arifin, Z. (2016). *Evaluasi Pembelajaran*. Bandung: Remaja Rosdakarya.
- Arifin, Z., & Retnawati, H. (2015). Analisis Instrumen Pengukur Higher Order Thinking Skills (HOTS) Matematika Siswa SMA.
- Asep Mulyana, A., Rochman, C., & Salim, A. (2019). Tantangan Kompetensi Kognitif pada Mata Pelajaran Pendidikan Agama Islam di Sekolah Menengah Atas Alfa Centauri Bandung. *At-Tajdid: Jurnal Ilmu Tarbiyah*, 344-368.
- Eka Putri Umacina, N., Mondolang, A. H., & Rondonuwu, I. T. (2020). Analisa Butir Soal Sumatif Semester Ganjil Mata Pelajaran Fisika. *Jurnal Pendidikan Fisika Unima*, 33-38.
- Nafiati, D. A. (2021). Revisi taksonomi Bloom: Kognitif, Afektif, dan Psikomotik. *Humanika, Kajian Ilmiah Mata Kuliah Umum*, 151-172.
- Purwanto, M. N. (2008). *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*. Bandung: Remaja Rosdakarya.
- Rofiah, E., Aminah, N., & Ekawati, E. (2013). Penyusunan Instrumen Tes Kemampuan Berpikir Tingkat Tinggi Fisika Pada Siswa SMP. *Jurnal Pendidikan Fisika Universitas Sebelas Maret*, 17-22.
- Wirandani, T., Kasih, A. C., & Latifah. (2019). Analisa Butir Soal HOTS (High Order Thinking Skill) Pada Soal Ujian Sekolah Kelas XII Mata Pelajaran Bahasa Indonesia Di SMK An-Nahl. *Parole: Pendidikan Bahasa dan Sastra Indonesia*, 488.