

Reading Skill Test Analysis in the Grade X Indonesian Language Textbook under the Merdeka Curriculum Based on Barrett's Taxonomy

Salma Sunaiyah^{1*}, Nurul Dwi Lestari², & Ida Nanik Zuliatin³

¹ UIN Syekh Wasil Kediri, Kediri, email: salmasun@uinkediri.ac.id

² UIN Syekh Wasil Kediri, Kediri, email: nuruldwilestari@uinkediri.ac.id

³ UIN Syekh Wasil Kediri, Kediri, email: ninaida2022@gmail.com

ARTICLE INFO

Article History:

Received June 25, 2026

Accepted July 20, 2026

Available online July 30, 2026

Kata Kunci :

Taksonomi Barrett, Keterampilan Membaca, Buku Teks, Kurikulum Merdeka, Asesmen

Keywords:

Barrett's Taxonomy, Reading Skills, Textbook, Merdeka Curriculum, Assessment

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

Copyright ©2026 by Author. Published by Institute of Education and Social Research

ABSTRAK

Penelitian ini bertujuan untuk menganalisis butir soal keterampilan membaca pada Uji Kompetensi buku teks Bahasa Indonesia Kelas X Kurikulum Merdeka (Edisi Revisi, 2023) berdasarkan Taksonomi Barrett. Taksonomi Barrett membagi keterampilan membaca ke dalam lima kategori, yaitu pemahaman literal, reorganisasi, pemahaman inferensial, evaluasi, dan apresiasi. Penelitian ini menggunakan pendekatan deskriptif kualitatif dengan teknik analisis isi. Sumber data adalah buku teks Bahasa Indonesia untuk SMA/MA/SMK/MAK Kelas X terbitan Kemendikbudristek. Data berupa 31 butir soal keterampilan membaca yang tersebar di lima bab (Bab I, II, III, V, dan VI). Hasil penelitian menunjukkan bahwa soal didominasi oleh level pemahaman literal (Level 1), diikuti pemahaman inferensial (Level 3), reorganisasi (Level 2), evaluasi (Level 4), dan apresiasi (Level 5) sebagai level yang paling sedikit ditemukan. Analisis yang membedakan teks sastra dan nonsastra menunjukkan bahwa soal teks sastra memiliki distribusi level yang lebih beragam. Temuan ini mengindikasikan bahwa instrumen penilaian keterampilan membaca pada buku teks tersebut masih didominasi kemampuan berpikir tingkat rendah dan perlu diperkaya dengan soal-soal level evaluasi dan apresiasi. Temuan ini memiliki implikasi praktis bagi pengembang kurikulum dan penulis buku teks dalam merancang instrumen asesmen membaca yang lebih proporsional dan berorientasi HOTS untuk pembelajaran Bahasa Indonesia ke depan.

ABSTRACT

This study aims to analyze reading comprehension items in the Competency Test section of the Grade X Indonesian Language textbook (Merdeka Curriculum, Revised Edition, 2023) using Barrett's Taxonomy. Barrett's Taxonomy categorizes reading skills into five levels: literal comprehension, reorganization, inferential comprehension, evaluation, and appreciation. This study employed a qualitative descriptive approach with content analysis technique. The data source was the Grade X Indonesian Language textbook published by the Ministry of Education, Culture, Research, and Technology of Indonesia. The data consisted of 31 reading skill items spread across five chapters (Chapters I, II, III, V, and VI). The results show that the items were dominated by literal comprehension (Level 1), followed by inferential comprehension (Level 3), reorganization (Level 2), evaluation (Level 4), and appreciation (Level 5) as the least frequent. Literary text items showed more varied level distribution than non-literary texts. These findings indicate that the reading assessment instrument is still predominantly low-order thinking and needs to be enriched with evaluation and appreciation level items. These findings carry practical implications for curriculum developers and textbook authors in designing more proportional, HOTS-oriented reading assessment instruments for future Indonesian language instruction.

INTRODUCTION

The current practice of Indonesian language instruction in classrooms continues to face disparities in students' mastery of receptive language skills, particularly reading comprehension. Various national educational evaluation reports consistently indicate that the majority of secondary-level students are only capable of decoding written symbols or reading texts mechanically without being able to extract meaning comprehensively. This condition is reinforced by the 2022 Programme for International Student Assessment (PISA), which recorded Indonesia's reading literacy score at 359 points. This score was 12 points lower than the 2018 result and even lower than Indonesia's achievement in 2000, making it the lowest score since Indonesia first participated in PISA. More concerningly, this score places Indonesia at Level 1a, indicating that Indonesian students are generally only able to understand the literal meaning of short sentences or paragraphs, while still struggling to comprehend lengthy texts containing implicit or abstract information or to compare perspectives across texts.

This phenomenon becomes evident when students are presented with complex texts, as they tend to read fluently aloud but fail to identify implied information, reorganize ideas, or evaluate the author's message. The low level of reading literacy among secondary students is further emphasized by Dewi (2023), who found that most students remain confined to low literacy proficiency due to limited exposure to analytical texts. At a more concrete level, Himawan (2023), who analyzed 50 reading literacy items used for regional assessment preparation in junior secondary schools in Yogyakarta using Barrett's Taxonomy, found that inferential comprehension was the most dominant category, whereas evaluation and appreciation categories were scarcely represented. These findings indicate that the problem of low-level higher-order reading skills does not solely stem from students themselves but also from the types of questions that continue to dominate reading assessment instruments.

Within the framework of the Merdeka Curriculum, Indonesian language learning at the senior secondary level (SMA/MA/SMK/MAK) is designed to enable students to understand, analyze, evaluate, and appreciate various types of texts. Such competencies require assessment instruments capable of measuring multiple levels of reading ability, ranging from literal comprehension to aesthetic appreciation of texts. This need is consistent with the study conducted by Martutik et al. (2016), which emphasized the importance of developing critical reading assessment instruments based on complex texts to objectively map senior high school students' higher-order thinking skills. If the questions contained in nationally used textbooks do not reflect the range of reading competencies expected in the Merdeka Curriculum, students will not receive adequate opportunities to practice and have their reading abilities comprehensively assessed.

One academically recognized framework for classifying the cognitive levels of reading comprehension questions is Barrett's Taxonomy. Developed by Thomas C. Barrett in 1968, this taxonomy was specifically designed to classify questions in the context of reading instruction into five hierarchical categories, ranging from the simplest to the most complex abilities: literal comprehension, reorganization, inferential comprehension, evaluation, and appreciation (Burns et al., 1996). According to Amiruddin et al. (2025), the application of Barrett's Taxonomy demonstrates that this framework remains relevant not only for analyzing reading questions but also for broader text-based formal assessment instruments. Compared with more general taxonomies, the primary advantage of Barrett's Taxonomy lies in its specific orientation toward

reading skills, ensuring that each category accurately reflects the cognitive processes involved in text comprehension (Riry & Binnendyk, 2025).

Several recent international studies further demonstrate the continued relevance of Barrett's Taxonomy beyond the Indonesian context. Krismadayanti and Zainil (2022) applied the taxonomy to English reading comprehension test items and found a similar dominance of literal-level questions, suggesting that the imbalance toward lower-order comprehension is not unique to Indonesian-language instructional materials. Kalhoro et al. (2025), in a broader review of reading comprehension assessment methods, positioned Barrett's Taxonomy alongside other cognitive frameworks and highlighted its particular strength in capturing the affective and appreciative dimensions of reading that are often neglected in more generic taxonomies. Similarly, Uysal (2025) compared Barrett's Taxonomy with Bloom's Taxonomy in the construction of text-based reading questions and concluded that Barrett's categories offer more precise operational definitions for classifying reading-specific cognitive demands. These international findings reinforce the theoretical grounding of the present study and indicate that the underrepresentation of evaluation and appreciation items is a recurring pattern across different educational and linguistic contexts, rather than merely a localized shortcoming of the textbook analyzed here.

Textbooks constitute one of the most widely used instructional instruments by teachers and students in classroom learning. Kusumawardani et al. (2024), who analyzed non-literary reading assessments in Grade VII Indonesian language textbooks under the Merdeka Curriculum, found that the distribution of cognitive levels in the questions was still uneven, with a predominance of lower-order thinking skills. Similarly, Wicaksana and Basuki (2019) found that the evaluative domain of the questions remained largely centered on literal comprehension. In line with these findings, Sare et al. (2025), who analyzed reading comprehension questions in Grade VIII Indonesian language textbooks using Barrett's Taxonomy, found that although the question distribution reflected an orientation toward Higher-Order Thinking Skills (HOTS), certain categories were still not proportionally represented. However, similar research focusing on the Grade X Indonesian language textbook, particularly the 2023 revised edition of the Merdeka Curriculum, which serves as students' initial foundation for developing upper secondary literacy competencies, remains unavailable.

It is important to note that although Sare et al. (2025) also applied Barrett's Taxonomy to a Merdeka Curriculum textbook, their study was conducted on the Grade VIII textbook and reported reading comprehension items as an aggregate distribution without distinguishing between literary and nonfiction text types. The present study departs from that work in two respects. First, it targets the Grade X textbook, which occupies a different curricular position as the entry point to senior secondary literacy development rather than junior secondary reading instruction. Second, it explicitly separates the analysis of literary and nonfiction texts, allowing the distribution of Barrett's Taxonomy levels to be interpreted in relation to the distinct reading competencies each text type demands, a distinction not addressed in Sare et al. (2025). The present study thus extends the application of Barrett's Taxonomy to a different educational level while offering a more fine-grained, text-type-sensitive account of how cognitive levels are distributed within a single textbook.

Based on this background, this study aims to describe the distribution of Barrett's Taxonomy levels in reading comprehension test items contained in the competency test sections of the Grade X Indonesian Language textbook under the Merdeka Curriculum (2023 Revised

Edition). This study is expected to provide an accurate description of the range of reading abilities assessed by the textbook's summative assessment instruments while also serving as a basis for recommendations in developing more proportionate and HOTS-oriented reading comprehension questions.

METHOD

This study employed a qualitative descriptive approach using content analysis techniques. The qualitative descriptive approach was selected because the study aimed to describe phenomena systematically based on data derived from documents without intervention or manipulation of the data (Supriyono, 2024). Content analysis was utilized to identify, classify, and interpret test items systematically according to the categories of Barrett's Taxonomy (Tabatabaee-Yazdi, 2020). The research data consisted of test items contained in the competency test sections of each chapter that specifically measured reading comprehension skills. The data source was the Indonesian Language textbook for Grade X SMA/MA/SMK/MAK under the Merdeka Curriculum (Revised Edition, 2023), published by the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia.

The inclusion criteria for the analyzed questions were as follows: the questions had to contain a reading text presented as the primary stimulus, the answers had to be identifiable or inferable from the text, and the questions had to measure the ability to comprehend, interpret, evaluate, or appreciate the text (Sare et al., 2025). Questions excluded from the analysis included language-focused questions, text production questions, conceptual questions about text types without a stimulus text, and questions whose answers could not be identified or inferred from the provided text.

Based on these criteria, a total of 31 reading comprehension test items were identified across five chapters: Chapter I (6 items), Chapter II (5 items), Chapter III (8 items), Chapter V (7 items), and Chapter VI (5 items). Chapter IV (Negotiation) was excluded because all competency test items in that chapter measured knowledge about negotiation texts rather than reading comprehension skills based on textual stimuli. This exclusion was not merely a matter of data availability but followed directly from the genre-specific communicative purpose of Chapter IV. Unlike the other five chapters, which use a written text (an article, an anecdote, a hikayat, a biography, or a poem) as the primary stimulus that readers must comprehend, interpret, or evaluate, the negotiation chapter is oriented toward productive and procedural competencies, namely constructing and delivering negotiation utterances according to a structural pattern. Its competency test items accordingly asked students to identify structural components or pragmatic strategies of negotiation in the abstract, without requiring them to read and respond to a bounded text as the source of the answer. Because the first inclusion criterion in this study required a reading text functioning as the primary stimulus, none of the Chapter IV items could be classified under Barrett's Taxonomy regardless of their cognitive demand, which is why the chapter contributed zero items to the 31 analyzed rather than simply a low number.

Barrett's Taxonomy, which served as the analytical framework in this study, consists of five levels. Level 1 (Literal Comprehension) includes questions whose answers are explicitly stated in the text. Level 2 (Reorganization) includes questions requiring students to reorganize or summarize information from the text. Level 3 (Inferential Comprehension) includes questions whose answers are implied and require inference. Level 4 (Evaluation) includes questions requiring students to assess the content of the text. Level 5 (Appreciation) includes questions requiring students to provide aesthetic responses to the text (Himawan et al., 2025).

RESULT AND DISCUSSION

Result

This section presents the results and discussion of the analysis of reading comprehension test items contained in the competency test sections of the Grade X Indonesian Language textbook under the Merdeka Curriculum (Revised Edition). The discussion is organized into two interrelated subsections. The first subsection presents the distribution of reading comprehension questions across chapters as an initial overview of the distribution and scope of the analyzed data. The second subsection presents the classification of each test item based on Barrett's Taxonomy, which includes analyses of nonfiction texts, literary texts, as well as a comparison of the distribution of cognitive levels between the two text types. This sequence of discussion was selected to enable readers to understand the quantitative overview of the data before proceeding to a more in-depth analysis of the cognitive quality of the test items.

Discussion

Distribution of Reading Comprehension Test Items Across Chapters

Based on the analysis of all competency test sections in the Grade X Indonesian Language textbook under the Merdeka Curriculum (Revised Edition), a total of 31 test items met the criteria for reading comprehension questions. The distribution of the questions across chapters is presented in Table 1 below.

Table 1. Distribution of Reading Comprehension Test Items Across Chapters

Chapter	Chapter Theme	Reading Texts	Number of Items
I	Expressing Natural Facts Objectively	Tabebuya Text + Lorentz National Park Text	6
II	Expressing Criticism Through Humor	Anecdote Indah pada Waktunya + Si Juki Comic + Stand-up Comedy Text	5
III	Exploring Stories Across Time	Hikayat Maharaja Bikrama Sakti + Hikayat Panca Logam	8
V	Learning Exemplary Values from Inspirational Biographies	Biography of W.R. Soepratman	7
VI	Creating and Expressing Through Poetry	Poem 1 (Untitled) + Poem 2 (about nature)	5
Total			31

A total of 31 test items were designed to assess students' comprehension of texts and to support the development of their literacy skills (Porhiel, 2004). These findings indicate that the textbook is effective in supporting the achievement of reading comprehension competencies in accordance with the learning objectives of the Merdeka Curriculum (Akbariski et al., 2024). Continuous textbook development remains essential to enhance instructional effectiveness and to ensure that the materials remain aligned with the expected competencies (Fauni et al., 2025).

Table 1 also shows that Chapter IV (Negotiation) contributed no items at all to the 31 analyzed, a gap that warrants explanation rather than simple omission. This absence is not incidental but stems from the distinct communicative purpose of the negotiation material relative to the other five chapters. Whereas Chapters I, II, III, V, and VI each build their competency tests around a bounded reading text that students must comprehend, interpret, or evaluate, Chapter IV emphasizes productive and pragmatic competencies, namely constructing negotiation utterances and applying their structural conventions, so its test items assess procedural knowledge about negotiation rather than comprehension of a specific stimulus text. Consequently, none of the items in that chapter satisfied the inclusion criterion requiring a text-based stimulus, regardless of their underlying cognitive complexity. This finding suggests a broader structural gap in the textbook: because reading comprehension assessment is entirely absent from the negotiation unit, students receive no opportunity within that chapter to practice or be evaluated on any level of Barrett's Taxonomy, including the higher-order evaluation and appreciation levels that are already underrepresented elsewhere in the book. Curriculum developers may therefore wish to consider embedding a short reading text, such as a sample negotiation transcript or dialogue, as a stimulus for at least one comprehension-based item in future editions of this chapter.

Classification of Test Items Based on Barrett's Taxonomy

After all reading comprehension test items had been identified, the next stage involved classifying each item into one of the five levels of Barrett's Taxonomy. The classification process was conducted by examining the cognitive demands explicitly embedded in the wording of each question (Riry & Binnendyk, 2025). Each question was assigned to only one level, namely the level that most dominantly reflected the cognitive process required by the item (Kuleshova, 2025).

The classification process was carried out in several stages. First, the researcher reread the stimulus text underlying each question. Second, the researcher examined the formulation of the question and determined whether the expected answer could be identified explicitly from the text (literal comprehension), had to be reorganized from several parts of the text (reorganization), required inference from the textual context (inferential comprehension), demanded judgment regarding the quality or validity of the text (evaluation), or invited personal and aesthetic responses to the text (appreciation). Third, the researcher matched the classification results with the operational definitions of each level in Barrett's Taxonomy (Riry & Binnendyk, 2025). To maintain consistency and reliability in the classification process, the researcher adhered to two important principles. First, if a question appeared to contain characteristics situated between two levels, the higher level was selected as the final classification when additional cognitive processing was implicitly required by the question. Second, true-false and matching-format questions could still be classified as inferential or evaluative if the required comprehension extended beyond merely identifying explicitly stated information (Uysal, 2025). It should be acknowledged that the classification of all 31 test items was carried out by a single researcher, and no second rater was involved to independently verify the classification results through an inter-rater reliability procedure. Given that the assignment of an item to a Barrett category ultimately rests on the researcher's judgment of its dominant cognitive demand, this single-rater procedure carries an inherent risk of subjectivity, particularly for items situated at the boundary between two levels. To mitigate this risk, the researcher repeatedly cross-checked each classification against the operational definitions of Barrett's Taxonomy and the two consistency principles described above; nevertheless, the absence of an inter-rater check remains a limitation of this study, and

future research is encouraged to involve at least one additional rater together with a reliability coefficient such as Cohen's kappa to strengthen the objectivity of the classification process.

It is important to understand that the classification of Barrett's Taxonomy levels does not depend solely on the question format, but rather on the actual cognitive process demanded by the item. Even multiplechoice questions may belong to the evaluation level if the available options require students to judge the validity or quality of information presented in the text rather than merely identify explicitly stated facts (Riry & Binnendyk, 2025). Conversely, essay questions may still fall into the literal comprehension level if they only require students to restate information explicitly provided in the text. This principle served as an important foundation to ensure that the classification results accurately reflected the cognitive depth of the questions rather than merely their technical format. In this study, the principle was consistently applied to all 31 analyzed reading comprehension items.

The classification results of the 31 reading comprehension test items based on Barrett's Taxonomy are presented separately for nonfiction and literary texts. This distinction was made because the two text types fundamentally require different reading competencies (Kalhoru et al., 2025). Nonfiction texts require readers to comprehend factual, conceptual, and procedural information objectively, whereas literary texts require readers to interpret meaning, values, and aesthetic dimensions of the text in a more interpretative and appreciative manner (Damaiyanti & Mulyati, 2025). This distinction is important to ensure that the distribution of Barrett's Taxonomy levels can be interpreted more contextually and meaningfully, while also allowing the resulting recommendations to be directed more specifically according to the characteristics of each text type.

Reading Comprehension Questions on Nonfiction Texts

The nonfiction texts in this textbook include popular scientific articles in Chapter I (Tabebuaya Text and Lorentz National Park Text) and a biographical text in Chapter V (Biography of W.R. Soepratman). From these two chapters, a total of 13 reading comprehension questions were identified. The distribution of Barrett's Taxonomy levels for the nonfiction reading questions is presented in Table 2 below.

Tabel 2. Distribution of Nonfiction Reading Comprehension Questions Based on Barrett's Taxonomy

Level	Barrett Category	Chapter I (Article)	Chapter V (Biography)	Total
1	Literal Comprehension	3	3	6
2	Reorganization	1	1	2
3	Inferential Comprehension	1	2	3
4	Evaluation	1	0	1
5	Appreciation	0	1	1
	Total	6	7	13

In the nonfiction texts, literal comprehension questions dominated both chapters. The literal questions in Chapter I required students to identify factual information explicitly stated in

the text, such as determining the meaning of naphthoquinone based on the explanation provided in the Tabebuaya text. The dominance of literal-level questions in nonfiction texts is understandable because factual texts are inherently rich in explicit information that can easily serve as the basis for assessment items (Setyoningrum et al., 2025). However, Kusumawardani et al. (2024) argue that reading assessments for nonfiction texts should not be limited to surface-level comprehension but should also encourage students to process, evaluate, and reflect upon the information they read. The density of factual information in nonfiction texts should instead provide opportunities to develop questions requiring students to compare, synthesize, and critically evaluate information, competencies that are highly relevant in the contemporary digital information era.

Reorganization questions appeared once in Chapter I and once in Chapter V. The reorganization question in Chapter I required students to identify the health benefits of tabebuaya, information that was not concentrated in a single paragraph but distributed across several sections of the text, thus requiring students to gather and reorganize the information cohesively. In Chapter V, the reorganization question appeared in the form of matching biography paragraphs with their respective text structures (orientation, significant events, and reorientation), an activity requiring students to understand the function of each section within the overall text structure (Kusumawardani et al., 2024). The very limited presence of reorganization questions in nonfiction texts deserves attention, considering that the abilities to summarize, classify, and reorganize information constitute essential competencies for students when reading academic and scientific texts at higher educational levels.

A total of three inferential questions were identified in the nonfiction texts. One representative example was a question in Chapter I asking about the origin of the name Lorentz National Park. The answer was not explicitly stated in the text, requiring students to make inferences based on contextual clues and available information. In Chapter V, inferential questions appeared in the form of selecting logical predictions based on cause-and-effect relationships presented in the biography. Such questions could not be answered merely by reading a single sentence but instead required students to connect information throughout the text (Damaiyanti & Mulyati, 2025). One evaluation question was identified in Chapter I, namely a question asking students to determine what information should be added to make the article more accurate. This question required students to assess the completeness and quality of the information presented in the text, an evaluative activity highly consistent with the characteristics of factual texts. One appreciation question was found in Chapter V, requiring students to express their opinion regarding W.R. Soepratman's influence on the Dutch. Although derived from a nonfiction text, the question still demanded a personal response extending beyond factual comprehension (Gasong, 2016).

The absence of evaluation questions in Chapter V and the minimal presence of evaluative questions overall in the nonfiction texts constitute an important finding, considering that factual texts inherently possess strong potential for critical examination regarding the accuracy and completeness of their information (Sultan et al., 2023). Harsiati and Priyatni (2017) emphasize that high-quality reading comprehension questions should encourage students not only to extract information from texts but also to reflect upon and evaluate the relevance and reliability of that information, a standard that has not yet been fully achieved in the nonfiction reading questions contained in this textbook.

Reading Comprehension Questions on Literary Texts

The literary texts in this textbook include anecdotal texts and stand-up comedy texts in Chapter II, hikayat texts in Chapter III, and poetry texts in Chapter VI. From these three chapters, a total of 18 reading comprehension questions were identified. The distribution of Barrett's Taxonomy levels for the literary reading questions is presented in Table 3 below.

Table 3. Distribution of Literary Reading Comprehension Questions Based on Barrett's Taxonomy

Level	Barrett Category	Chapter II (Anecdote)	Chapter III (Hikayat)	Chapter VI (Poetry)	Total
1	Literal Comprehension	1	2	3	6
2	Reorganization	1	2	0	3
3	Inferential Comprehension	2	2	2	6
4	Evaluation	1	2	0	2
5	Appreciation	0	0	1	1
	Total	5	8	6	18

In the literary texts, the distribution of Barrett's Taxonomy levels appears more balanced than in the nonfiction texts. Literal and inferential questions each accounted for six items. This finding indicates that the reading comprehension questions for literary texts in this textbook have more actively encouraged students to think beyond explicitly stated information (Sare et al., 2025). Literal questions in literary texts appeared in various forms. In Chapter III, for example, students were required to match characters with their respective roles based on the content of Hikayat Panca Logam, information explicitly stated in the text and therefore not requiring inference. In Chapter VI, literal questions appeared in the form of matching statements about the second poem with answers related to word connotations, diction effects, and sound effects explicitly reflected in the poet's lexical choices. The presence of literal questions in literary texts is not inherently problematic as long as they are balanced by higher-level questions. In this case, the balance between literal and inferential questions demonstrates a more proportional pattern compared with the nonfiction texts (Akbariski et al., 2024).

Inferential questions in literary texts generally required students to infer implied meanings. One representative example was a question asking students to identify the satire conveyed in the anecdote Indah pada Waktunya. To answer this question, students had to understand the social context underlying the anecdote and infer meanings not explicitly expressed in the text (Krismadayanti & Zainil, 2022). Inferential questions were also identified in Chapter III in the form of questions concerning social and educational values in the hikayat, values that were not directly stated but had to be inferred from the characters' actions and decisions within the story. In Chapter VI, all inferential questions were related to the interpretation of figurative meanings in poetry, including interpreting poetic lines as metaphors, explaining the meaning of the figurative phrase *detak hati malam*, and determining the main theme of the poem. The relatively high number of inferential questions in the poetry section strongly aligns with the

characteristics of poetry as a genre rich in condensed meaning and imagery, where comprehension cannot be achieved solely through literal reading (Sare et al., 2025).

Reorganization questions in literary texts totaled three items, all located in Chapters II and III. In Chapter II, reorganization questions appeared in the form of matching excerpts from stand-up comedy scripts with their structural components. In Chapter III, one question required students to arrange the correct sequence of plot events in the hikayat from four provided events, an activity involving the reorganization of explicitly presented information into a meaningful structure. Although limited in number, the presence of reorganization questions in literary texts indicates that the textbook pays attention to students' ability to actively understand and manage literary text structures rather than merely receive information passively (Wicaksana & Basuki, 2019).

Evaluation questions in literary texts were identified in two items, both found in Chapters II and III. The evaluation question in Chapter II required students to identify dialogue reflecting the author's opposition to certain societal habits presented in the comic, an activity requiring students to assess the author's stance based on linguistic choices. In Chapter III, students were asked to evaluate the educational values contained in the hikayat from several provided options, a cognitive demand requiring external evaluative standards derived from students' own experiences and value systems (Candra, 2014). The limited number of evaluation questions in literary texts, particularly in Chapter VI containing poetry, represents a notable weakness because poetry is one of the richest text types for developing evaluative questions. Students could, for instance, be asked to assess the effectiveness of the poet's diction, evaluate the relevance of the poem's message to contemporary social contexts, or judge the poet's success in conveying ideas and emotions.

Only one appreciation question was identified in the literary texts, namely in Chapter VI, where students were asked to explain the meaning of figurative expressions in the poem and respond to the aesthetic effects created by the poet's language choices (Sare et al., 2025). The minimal presence of appreciation questions in literary texts, particularly in hikayat and anecdotal texts, constitutes a significant weakness requiring attention. Amiruddin et al. (2025) emphasize that complex hierarchies of comprehension should be proportionally represented within assessment instruments rather than concentrated primarily at lower levels. Both hikayat texts, with their rich use of figurative language, similes, and hyperbole, and anecdotal texts, with their distinctive humor and social satire, actually provide substantial potential for the development of appreciation questions that require students to respond to linguistic beauty and the social criticism embedded within the texts.

Comparison of Level Distribution Between Literary and Nonfiction Texts

The comparison between the distribution of nonfiction and literary text questions reveals several noteworthy patterns. First, literary text questions demonstrate a more balanced distribution of Barrett's Taxonomy levels than nonfiction texts, particularly due to the greater presence of inferential and evaluative questions. The balance between literal and inferential questions in literary texts reflects the intrinsic characteristics of literature itself, which naturally contains layers of meaning that cannot be understood solely through surface-level reading (Damaiyanti & Mulyati, 2025). In contrast, the strong dominance of literal questions in nonfiction texts, particularly in Chapters I and V, indicates that the potential of factual texts to develop students' critical and evaluative thinking skills has not yet been optimally utilized in the

construction of assessment items (Kusumawardani et al., 2024). Second, appreciation questions were found more frequently in nonfiction texts, specifically one question in the biography section of Chapter V, than in literary texts on an individual chapter basis, although Chapter VI poetry still contributed one appreciation question overall. This phenomenon is somewhat paradoxical because appreciation is generally more strongly associated with literary texts than with nonfiction texts (Azizah et al., 2023). Third, no appreciation questions were identified in the hikayat texts (Chapter III) or anecdotal texts (Chapter II), despite the fact that both text types contain rich values, stylistic features, and aesthetic dimensions that provide considerable potential for the development of appreciation-based questions.

These findings are consistent with Harsiati's (2018) argument that high-quality reading comprehension questions should measure not only the ability to retrieve information but also the ability to develop opinions and evaluate reading materials. In the context of literary texts, the abilities to develop opinions and evaluate texts are closely related to literary appreciation, which should represent the highest level of literary reading assessment (Azizah et al., 2023). Similarly, Harsiati and Priyatni (2017), in their analysis of PISA reading literacy items from 2000 to 2009, found that PISA questions were dominated by higher-order thinking skills involving interpretation, reflection, and evaluation. In other words, international standards position inferential and higher-level comprehension as the core of reading assessment, a standard that has not yet been fully achieved by the textbook examined in this study, either in literary or nonfiction texts (Harsiati & Priyatni, 2017). This condition is also relevant to the findings of Sultan et al. (2023), who reported that reading comprehension questions in Indonesian language textbooks have not yet fully supported students' readiness to face international literacy assessments emphasizing higher-order thinking skills.

The imbalance in the distribution of Barrett's Taxonomy levels identified in this study also has direct implications for the quality of students' learning experiences in the classroom. If the questions provided in textbooks, as the primary assessment source used by teachers, are dominated by literal-level items, students will indirectly become more accustomed to answering questions that merely require retrieving explicitly stated information. Such conditions risk fostering shallow and non-critical reading habits, which may ultimately contribute to the low overall reading literacy abilities of students (Himawan et al., 2025). Therefore, improving the distribution of questions based on Barrett's Taxonomy is not merely a technical matter of assessment design but is directly related to the development of students' critical thinking abilities, which constitute a central objective of Indonesian language learning within the Merdeka Curriculum (Akbariski et al., 2024).

On the other hand, one positive finding worth highlighting is the considerable variety of question formats used across both literary and nonfiction texts, including multiple-choice, true-false, matching, fill-in-the-blank, and essay questions. Wicaksana and Basuki (2019) noted that variation in question formats constitutes one of the strengths of textbook-based assessment instruments in training diverse aspects of reading skills. This variation is important because each question format possesses its own advantages in measuring reading ability. Multiple-choice questions are efficient for assessing broad comprehension, essay questions provide opportunities for deeper and more personal responses, and matching questions are effective for measuring the ability to organize and connect information (Porhiel, 2004). This positive variation in question formats may serve as an important foundation for the future development of more balanced assessment instruments for both literary and nonfiction texts (Setyoningrum et al., 2025). The

primary challenge moving forward lies in improving the cognitive quality of the questions so that the existing diversity of formats can be optimized to reach evaluation and appreciation levels, which remain minimally represented at present.

CONCLUSION

Based on the analysis of 31 reading comprehension test items contained in the competency test sections of the Grade X Indonesian Language textbook under the Merdeka Curriculum (Revised Edition, 2023), several conclusions can be drawn. First, reading comprehension questions overall were dominated by literal and inferential comprehension levels, while evaluation and appreciation questions remained very limited in number. Second, nonfiction text questions were more heavily concentrated at the literal level, whereas literary text questions demonstrated a more varied distribution of Barrett's Taxonomy levels. Third, appreciation questions were almost entirely absent from hikayat and anecdotal texts. Fourth, Chapter IV (Negotiation) did not contain any reading comprehension questions at all.

The implications of this study indicate the need for improvement in the development of reading comprehension assessment instruments in Indonesian language textbooks. Curriculum developers and textbook authors need to deliberately design questions that are more proportionally distributed across all levels of Barrett's Taxonomy by increasing the proportion of evaluation and appreciation questions so that students are more effectively trained to think critically and to appreciate texts in a deeper and more reflective manner.

ACKNOWLEDGE

If any, thanks are addressed to official institutions or individuals who have provided funding or have made other contributions to the research. Acknowledgments are accompanied by a research contract number.

REFERENCES

- Akbariski, H. S., Harsiati, T., & Basuki, I. (2024). Assessment of Critical Reading Ability Using Indonesian Language Textbooks for the 2013 Curriculum and the Merdeka Curriculum. *Jurnal Ilmu Pendidikan*, 30(1), 19-28. <https://doi.org/10.17977/um048v30i1p19-28>
- Ambika, L. G. S., Reddy, R. M. K., Raj, S. N., & Pasha, S. P. V. (2024). Quality textbook assessment system. *International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)*, 4(2), 8-15. <https://doi.org/10.48175/IJARSCT-18102>
- Amiruddin, N., Basuki, I. A., & Harsiati, T. (2025). Analisis Instrumen Butir Soal Asesmen Menyimak pada Latihan Uji Kemahiran Berbahasa Indonesia. *ESTETIK: Jurnal Bahasa Indonesia*, 8(2), 281-302.
- Azizah, N., Karsono, K., & Kamsiyati, S. (2023). Perbandingan kompetensi literal dan apresiatif siswa kelas IV SD dalam memahami isi teks fiksi. <https://doi.org/10.20961/ddi.v11i3.76415>
- Barrett, T. C. (1968). Taxonomy of cognitive and affective dimensions of reading comprehension. In T. Clymer, What is 'reading?': Some current concepts. In H. M. Robinson (Ed.), *Innovation and change in reading instruction, Sixty-seventh Yearbook of the National Society for the Study of Education, Part II* (pp. 17-23). University of Chicago Press.
- Burns, P. C., Roe, B. D., & Ross, E. P. (1996). *Teaching reading in today's elementary schools* (6th ed.). Houghton Mifflin.
- Cahyanto, Z. D., Basuki, I. A., & Martutik, M. (2016). Pengembangan Instrumen Asesmen Membaca Kritis untuk Siswa SMA/SMK Kelas X dan XI. *Jurnal Pendidikan: Teori, Penelitian, dan Pengembangan*, 1(5), 780-784.

- Candra, I. (2014). The classification of reading comprehension questions in senior high school textbook entitled English using Barrett's taxonomy.
- Damaiyanti, M. F., & Mulyati, Y. (2025). Between Imagination and Information: Exploring the Assessment Practices of Literary and Non-Literary Texts on Reading Skills. <https://doi.org/10.26714/lensa.15.2.2025.194-219>
- Dewi, A. K. (2023). Karakteristik Kemampuan Membaca Pemahaman dan Literasi Siswa Tingkat Menengah. *Jurnal Pembelajaran Bahasa dan Sastra Indonesia*, UPT Publikasi Ilmiah Universitas Negeri Malang.
- Fauni, A. M., Rini, S., Setyaningrum, D. R., & Huda, M. (2025). Kelayakan Materi dan ilustrasi dalam buku Pelajaran bahasa Indonesia Edisi Revisi Kelas X. *Buletin Pengembangan Perangkat Pembelajaran*. <https://doi.org/10.23917/bppp.v7i1.11632>
- Gasong, D. (2016). Developing Interpersonal and Intrapersonal Competence in Learning Literature Appreciation. *Journal of Education and Practice*, 7(11), 24-28.
- Harsiati, T., & Priyatni, E. T. (2017). Karakteristik tes literasi membaca pada programme for international student assesment (PISA). 1(2), 1-11. <https://doi.org/10.17977/UM008V1I22017P001>
- Himawan, R., Rahayu, T., Alfian, M. A., & Hermawan, M. A. (2025). Analysis of the Quality of Assessment Questions of Standardized Regional Assessment (ASPD) of Junior High School Reading Literacy: A Review Based on Barrett's Taxonomy Cognitive Levels. *Jurnal Paedagogy*, 12(1), 208. <https://doi.org/10.33394/jp.v12i1.13693>
- Huang, R., Tlili, A., Zhang, X., Sun, T., Wang, J., & Sharma, R. (2022). A Comprehensive Framework for Comparing Textbooks: Insights from the Literature and Experts. *Sustainability*, 14(11), 6940. <https://doi.org/10.3390/su14116940>
- Kalhor, I. A., Rabbi, F., & Haram, B. U. (2025). A Comprehensive Review of Reading Comprehension Assessment Methods and Cognitive Levels. 4(3), 1747-1763. <https://doi.org/10.63056/acad.004.03.0488>
- Krismadayanti, A., & Zainil, Y. (2022). The level of the students' reading comprehension analyzed by using Barrett taxonomy. *Journal of Cultura and Lingua*, 3(1), 39-48. <https://doi.org/10.37301/culingua.v3i1.110>
- Kuleshova, P. N. (2025). A typology of tasks of increased complexity for younger students. <https://doi.org/10.31483/r-152744>
- Kusumawardani, A. R., Harsiati, T., & Basuki, I. A. (2024). Karakteristik Asesmen Membaca Nonsastra dalam Buku Bahasa Indonesia Kurikulum Merdeka. *Jurnal Onoma: Pendidikan, Bahasa, dan Sastra*, 10(3), 3283-3291.
- Porhiel, S. (2004). Linguistic controls that are used to assess a reader's understanding. *Journal of Pragmatics*, 36(6), 1009-1035. <https://doi.org/10.1016/J.PRAGMA.2003.09.004>
- Riry, J. G. I., & Binnendyk, S. (2025). Analyzing Reading Comprehension Questions in English Textbooks Using Barrett's Taxonomy. *Matai*, 5(2), 144-154. <https://doi.org/10.30598/matai.v5i2.18343>
- Sare, M. G. P. A., Nurbaya, S., & Kusmiatun, A. (2025). Kesesuaian Level Soal Membaca Pemahaman dalam Buku Bahasa Indonesia Kelas VIII dengan Tujuan Pembelajaran Kurikulum Merdeka. *Jurnal Onoma: Pendidikan, Bahasa, Dan Sastra*, 11(4), 4237-4253. <https://doi.org/10.30605/onoma.v11i4.6979>
- Setyoningrum, R. A., Basuki, I. A., & Harsiati, T. (2025). Asesmen membaca teks nonsastra dalam buku bahasa Indonesia kurikulum merdeka kelas VIII. <https://doi.org/10.59672/stilistika.v14i1.4867>
- Sultan, S., Rapi, Muh., Basri, M., Mangila, B. B., & Rahmat, W. (2023). Texts and Tasks in Indonesian Language Textbooks: Do They Support Indonesian Students in the International Reading Literacy Test? <https://doi.org/10.26858/ijole.v7i3.56017>
- Supriyono, S. (2024). Research methodology: qualitative approach. <https://doi.org/10.13140/rg.2.2.26723.57121>

- Tabatabaee-Yazdi, M. (2020). Hierarchical Diagnostic Classification Modeling of Reading Comprehension. SAGE Open, 10(2). <https://doi.org/10.1177/2158244020931068>
- Uysal, P. K. (2025). The Use of Taxonomies in Constructing Text-Based Reading Comprehension Questions: Bloom and Barrett Taxonomies. E- Kafkas Eğitim Araştırmaları Dergisi. <https://doi.org/10.30900/kafkasegt.1716522>
- Wicaksana, S., & Basuki, I. A. (2019). Analisis Soal Latihan Keterampilan Membaca Dalam Bahan Ajar BIPA Sahabatku Indonesia Tingkat A1. Basindo: Jurnal Kajian Bahasa, Sastra Indonesia, dan Pembelajarannya, 3(2), 261-270.