LITERATUR: Jurnal Bahasa dan Sastra
Vol 8 No 1 (2026) Page 30 – 47
e-ISSN: 1599445885/ ISSN

Cognitive Quality of Grade X Indonesian Textbook Assessments: Comparing the 2013 and Merdeka Curricula

Rio Sahputra [1], Amril Amir [2], Abdurahman[3]

[1]Universitas Negeri Padang, Indonesia (rio.sahputra.rs321@gmail.com)
[2]Universitas Negeri Padang, Indonesia (amril.amir@fbs.unp.ac.id)
[3]Universitas Negeri Padang, Indonesia (abdurahman@fbs.unp.ac.id)

ABSTRACT
The transition from the 2013 Curriculum (K13) to the Merdeka Curriculum is theoretically expected to improve the cognitive quality of assessment instruments in textbooks, particularly in terms of the proportion of higher-order thinking skills (HOTS) questions. However, whether this change is actually reflected in the assessment instruments of official textbooks has never been examined through a comparative study. This study aims to describe and compare the distribution of cognitive levels in assessment instruments in government-published 10th-grade Indonesian language textbooks between the K13 and the Merdeka Curriculum based on the Revised Bloom's Taxonomy (Anderson & Krathwohl, 2001). The study employed a qualitative approach using descriptive-comparative content analysis on 119 items (K13) and 60 items (Merdeka Curriculum). The results indicate that both textbooks remain dominated by LOTS items, with HOTS accounting for 42.0% in K13 and 40.0% in the Merdeka Curriculum. The most notable change is the shift in dominance from C1 to C2 in the Merdeka Curriculum, accompanied by a decrease in C5 and C6. These findings indicate that a curriculum change does not automatically increase the cognitive demands of assessment instruments in textbooks. This study provides practical implications for textbook developers, Indonesian language teachers, and education policymakers in designing assessment instruments that are more oriented toward the development of higher-order thinking skills.

Keywords: curriculum, HOTS, LOTS, Revised Bloom's Taxonomy, textbooks

Received: 07-02-2026 Revised:20-03-2026 Accepted:22-04-2026 Published:30-06-2026

Corresponding Author: Rio Sahputra (Email) : rio.sahputra.rs321@gmail.com

Published by Literatur (Jurnal Bahasa dan Sastra)
10.47766/literatur.v8i1.7632

INTRODUCTION

Textbooks are the primary learning tools that directly influence the quality of students' learning processes and outcomes. In the Indonesian education system, government-published textbooks hold a strategic position because they are used extensively at all levels of public and private schools, reaching millions of students with diverse backgrounds and learning conditions. According to BSNP standards, high-quality textbooks must meet four criteria: content, presentation, language, and graphics. Within the presentation component, assessment instruments in the form of exercises and evaluations are crucial elements that reflect the competency requirements to be achieved (Astari, 2022; Huda et al., 2021). Assessment tools in textbooks—whether in the form of practice questions, structured assignments, or end-of-chapter competency tests, serve a dual purpose: they measure competency achievement while also serving as stimuli that train and develop students' cognitive abilities (Huda et al., 2021; Saputro et al., 2021). The inclusion of assessment tools in textbooks is not merely an afterthought, but rather an indicator of the book's pedagogical orientation, whether it emphasizes memorization or critical thinking (Rosdiana et al., 2022; Saputro et al., 2021)In other words, the questions presented not only assess what has been learned, but also implicitly define the depth and complexity of the thinking expected of students.

To assess the cognitive quality of assessment instruments, Bloom's Taxonomy as revised by Anderson & Krathwohl (2001) is the most relevant framework. This taxonomy classifies cognitive processes into six levels: C1 Remembering, C2 Understanding, C3 Applying, C4 Analyzing, C5 Evaluating, and C6 Creating, with C1–C3 categorized as Lower Order Thinking Skills (LOTS) and C4–C6 as Higher Order Thinking Skills (HOTS) (Aini & Mukhlis, 2022). This revision transforms the knowledge dimension into nouns and the cognitive process dimension into verbs, making them more operational in designing and assessing learning objectives (Nafiati, 2021). The importance of HOTS is becoming increasingly evident in the context of international educational assessments: Indonesia's PISA scores in reading literacy and mathematics have consistently fallen below the OECD average, and one of the identified causes is students' inability to solve context-based problems that require higher-order thinking (OECD, 2022). At the national level, the Minimum Competency Assessment (AKM), which replaced the National Exam in 2021, explicitly requires reading literacy skills that go beyond merely understanding explicit information to include critically reflecting on, evaluating, and interpreting texts (Andrawana et al., 2022; Annam et al., 2023). If the assessment tools in textbooks only train students' ability to memorize and comprehend, then students will not be prepared to meet the far more demanding cognitive requirements of the AKM and

PISA (Andrawana et al., 2022; Izzah et al., 2025). It is the gap between national and international assessment requirements and the quality of cognitive exercises in textbooks that makes this study relevant and urgent.

A number of previous studies have consistently shown that the predominance of LOTS in Indonesian language textbook assessment instruments is a persistent and systemic problem, not merely a technical shortcoming in the design of questions in a single textbook or at a specific grade level. At the junior high school level, Alkarima et al., (2022) reported that the proportion of HOTS questions was only 30.0% in the 8th-grade textbook, while Permatasari (2021) found that the distribution was also uneven across chapters and was determined more by text genre than by a planned pedagogical design. At the high school level, Saputro et al., (2021) and Suvina & Ramly (2021) both identified the dominance of LOTS in 10th-grade textbooks, with Question C6 being almost nonexistent even though the curriculum formally requires critical thinking. This pattern even extends beyond subject boundaries: a study of English textbooks by Rustiyani et al., (2021) and Rinjaya & Halimi (2022) found a similar trend, indicating that the low proportion of HOTS questions is not a problem unique to the Indonesian language but rather a reflection of a broader culture of question design within the Indonesian textbook ecosystem. More crucially, this situation is directly correlated with Indonesia's low PISA scores, as assessment tools that focus solely on memorization and comprehension do not prepare students to tackle questions requiring complex reasoning (Rosdiana et al., 2022; Sutami et al., 2020). In other words, the issue of cognitive quality in textbooks is not merely a technical editorial matter, but a structural problem that has a tangible impact on the cognitive readiness of Indonesia's student generation

Indonesia has undergone several curriculum reforms in an effort to improve the quality of national education. The 2013 Curriculum (K13), as revised in 2017, is based on a scientific approach and authentic assessment grounded in strictly structured Core Competencies (KD) (Angga et al., 2022). However, the heavy administrative burden and difficulties in implementation on the ground led to the creation of the Merdeka Curriculum, which has been gradually implemented since the 2022–2023 school year. This curriculum shifts toward a Learning Outcomes (LO)-based paradigm, is oriented toward the Pancasila learner profile, and grants greater autonomy to teachers (Melani & Gani, 2023; Sumarsih et al., 2022). Philosophically, the Merdeka Curriculum positions assessment as an integral part of meaningful learning, shifting from the paradigm of "assessment of learning" toward "assessment for learning" and "assessment as learning," thereby theoretically opening up broader opportunities for the development of authentic higher-order thinking skills (HOTS) questions. However, this paradigm shift is not always immediately reflected in the

quality of the questions presented in published textbooks.

The most significant gap in the existing research landscape is the lack of comparative studies that directly compare the cognitive quality of textbook assessment instruments between the 2013 Curriculum and the Merdeka Curriculum in the subject of Indonesian Language. Most existing studies are descriptive in nature and analyze only one textbook within a single curriculum; the study Rohmah & Inderasari (2023) regarding the Cerdas Cergas Grade 10 textbook focuses only on literary competencies, while studies comparing the two curricula, such as Angga et al., (2022) and Febriana et al., (2022) focuses more on implementation and readability rather than the cognitive quality of the questions. As a result, a fundamental empirical question remains unanswered: does the Merdeka Curriculum truly produce assessment instruments that are cognitively more demanding and richer in higher-order thinking skills (HOTS) compared to the K13 curriculum? Given this gap, this study aims to map and compare the distribution of cognitive levels in the assessment instruments of 10th-grade Indonesian language textbooks between K13 and the Merdeka Curriculum based on Bloom's Revised Taxonomy by Anderson & Krathwohl (2001). The 10th grade was selected because it marks an important cognitive transition and is the first point of implementation for both curricula at the same educational level; the scope was limited to official government-published textbooks to ensure equal comparability. The research findings are expected to provide useful empirical evidence for textbook developers, teachers, and education policymakers.

Based on the background and research gaps outlined above, this study formulates three research questions: (1) what is the distribution of cognitive levels in the assessment instruments in the 10th-grade Indonesian Language textbook under the 2013 Curriculum based on the Revised Bloom's Taxonomy; (2) what is the distribution of cognitive levels in the assessment instruments in the Grade 10 Indonesian Language textbook under the Merdeka Curriculum based on the Revised Bloom's Taxonomy; and (3) how do the distributions of cognitive levels in the assessment instruments between the two textbooks compare, and what are the implications of these differences for the quality of cognitive assessment in Indonesian language learning.

RESEARCH METHOD

This study employs a qualitative approach using content analysis with a descriptive-comparative focus. Content analysis was chosen because it allows the researcher to systematically examine, categorize, and interpret the content of written texts based on predefined categories (Krippendorff, 2018). A descriptive approach was

used to describe the distribution of cognitive levels in each textbook, while a comparative approach was used to compare these distributions between the two textbooks.

The data sources for this study consist of two official government-published textbooks, namely: (1) Indonesian Language Textbook for Grade 10 of Senior High School/Madrasah Aliyah/Vocational High School/Madrasah Aliyah Keagamaan, 2013 Curriculum, Revised Edition, authored by Suherli, Maman Suryaman, Aji Septiaji, and Istiqomah (2017), published by the Ministry of Education and Culture in 2017; and (2) Indonesian Language Textbook for Grade 10 of Senior High Schools/Madrasah Aliyah/Vocational High Schools/Madrasah Aliyah Kejuruan (Revised Edition) under the Merdeka Curriculum, authors Fadillah Tri Aulia, Sefi Indra Gumilar, and Alvian Kurniawan (2023), published by the Ministry of Education, Culture, Research, and Technology in 2023. The units of analysis consist of all test items or task instructions in the "Tasks and Exercises" section of each chapter (K13 textbooks) and the "Competency Assessment" section of each chapter (Merdeka Curriculum textbooks). The selection of these units of analysis is based on the consideration that these two sections are the most representative and structured assessment instruments within each textbook.

The research instrument consists of a cognitive classification table based on the Revised Bloom's Taxonomy, containing seven columns: (a) question code number, (b) chapter and text topic, (c) question wording or instructions, (d) identified operational verbs (OVs), (e) cognitive levels C1–C6, (f) LOTS or HOTS category, and (g) indicator description. Coding was conducted in four stages. *First*, each item was read in its entirety to understand the cognitive orientation required. *Second*, OVs were identified from the item wording as the primary basis for classification; for example, the verbs "mention" and "identify" are coded as C1, "explain" and "describe" as C2, "apply" as C3, "analyze" and "compare" as C4, "evaluate" and "judge" as C5, and "create," "design," and "invent" as C6, referring to the operational KKO list by Anderson & Krathwohl (2001). *Third*, if the question wording does not contain an explicit KKO, inference is made based on two considerations: (1) the type of cognitive process inherently required by the task at hand, for example, a question asking students to construct an original argument is coded as C6 even though it does not explicitly use the verb "create"; and (2) the context of the stimulus text used, since certain text genres, such as anecdotes, structurally require the inference of implied meaning corresponding to the C4 cognitive process. *Fourth*, each item coded by the first coder is independently reviewed by a second coder without knowledge of the first coder's results, to ensure the objectivity of the classification. Items yielding coding discrepancies between the two coders are resolved through consensus discussion by

referring back to the KKO guidelines Anderson & Krathwohl (2001) until an agreement is reached. This coding procedure follows operational guidelines developed from the framework Anderson & Krathwohl (2001), as also applied in a similar study by Cressa & Mukhlis (2023) and Aini & Mukhlis (2022).

Data validity was ensured through an inter-rater reliability procedure involving two independent coders. Inter-coder agreement was calculated using Cohen's Kappa coefficient (Krippendorff, 2018). The Kappa values obtained were 0.78 (for the K13 textbook) and 0.81 (for the Merdeka Curriculum textbook), both of which fall within the range of *substantial agreement* to *almost perfect agreement* and are considered reliable for content analysis research. Items that resulted in coding discrepancies between the two *raters* were resolved through consensus discussions until agreement was reached. Data analysis was conducted through the following stages: (1) initial coding, (2) validity testing, (3) consensus, (4) tabulation of frequencies and percentages of C1–C6 per chapter and per textbook, and (5) comparison and qualitative interpretation of the distributions found.

RESULT AND DISCUSSION

Result

1. Distribution of Cognitive Levels in K13 Textbooks

Based on an analysis of 119 questions spread across 8 chapters of the 10th-grade Indonesian Language textbook for the 2013 Curriculum, the distribution of cognitive levels is as shown in Table 1 below.

Table 1. Distribution of Cognitive Levels in the K13 Textbook Assessment Instrument by Chapter

Chapter / Text	C1	C2	C3	C4	C5	C6	Total	LOTS (%)	HOTS (%)
Chapter 1 (LHO)	9	7	5	7	4	1	33	63,6	36,4
Chapter II (Exposition)	3	3	3	3	1	2	15	60,0	40,0
Chapter III (Anecdotes)	3	2	0	4	1	1	11	45,5	54,5
Chapter IV (Legend)	4	2	1	3	0	2	12	58,3	41,7
Chapter V (Negotiation)	4	2	2	3	1	1	13	61,5	38,5
Chapter VI (Debate)	2	3	2	3	1	1	12	58,3	41,7
Chapter VII (Biography)	3	2	0	2	2	1	10	50,0	50,0
Chapter VIII (Poetry)	6	2	2	5	2	3	20	50,0	50,0
TOTAL	34	23	12	30	12	12	119	58,0	42,0
Percentage (%)	28,6	19,3	10,1	22,7	10,1	9,2	100%	-	-

Overall, the K13 textbooks show that Level C1 (28.6%) is the most frequently appearing level, followed by C4 (22.7%), C2 (19.3%), and C3, C5, and C6 at 10.1%, 10.1%, and 9.2%, respectively. The proportion of LOTS (C1–C3) reached 58.0%, while HOTS (C4–C6) accounted for 42.0%. The distribution across chapter levels showed significant variation: Chapter III (Anecdotes) had the highest proportion of HOTS at 54.5%, followed by Chapter VII (Biographies) and Chapter VIII (Poetry) at 50.0% each.

Conversely, Chapter I (Observation Report Text) has the lowest HOTS proportion, at 36.4%. C6-level questions (creative) were identified in all chapters, although in small numbers.

2. Distribution of Cognitive Levels in Merdeka Curriculum Textbooks

An analysis of the 60 items on the Competency Test, which are distributed across the six chapters of the Grade 10 Indonesian Language textbook under the Merdeka Curriculum, yielded the distribution of cognitive levels presented in Table 2 below.

Table 2. Distribution of Cognitive Levels in the Merdeka Curriculum Textbook Assessment Instrument by Chapter

Chapter / Text	C1	C2	C3	C4	C5	C6	Total	LOTS (%)	HOTS (%)
Chapter 1 (LHO)	3	3	1	1	1	0	9	77,8	22,2
Chapter II (Anecdotes)	1	1	0	6	0	2	10	20,0	80,0
Chapter III (Legend)	1	4	1	2	0	1	9	66,7	33,3
Chapter IV (Negotiation)	2	5	1	2	0	0	10	80,0	20,0
Chapter V (Biography)	2	4	1	1	2	0	10	70,0	30,0
Chapter VI (Poetry)	1	5	0	6	0	0	12	50,0	50,0
TOTAL	10	22	4	18	3	3	60	60,0	40,0
Percentage (%)	16,7	36,7	6,7	30,0	5,0	5,0	100%	-	-

The Merdeka Curriculum textbooks show a dominance of C2 (36.7%), which is significantly higher than in the K13 textbooks, followed by C4 (30.0%), C1 (16.7%), C3 (6.7%), C5 (5.0%), and C6 (5.0%). The proportion of LOTS reaches 60.0% and HOTS 40.0%. At the chapter level, Chapter II (Anecdotes) has the highest HOTS at 80.0%, followed by Chapter VI (Poetry) at 50.0%. Conversely, Chapter IV (Negotiation) has the lowest HOTS, at 20.0%. It is important to note that C6 questions (creating) were found in only two chapters, namely Chapter II and Chapter III, while the other four chapters did not have any C6 questions at all.

3. Comparison of the Distribution of Cognitive Levels in the 2013 Curriculum and the Merdeka Curriculum

A comprehensive comparison of the distribution of cognitive levels between the two textbooks is presented in Table 3 below.

Table 3. Comparison of Cognitive Level Distributions in the 2013 Curriculum and the Merdeka Curriculum

Cognitive Level	K13 (n=119)	K13 (%)	KM (n=60)	KM (%)	Difference (%)
C1: Remembering	34	28,6	10	16,7	-11,9
C2: Understanding	23	19,3	22	36,7	+17,4
C3: Applying	12	10,1	4	6,7	-3,4
C4: Analyzing	27	22,7	18	30,0	+7,3
C5: Evaluate	12	10,1	3	5,0	-5,1

C6: Creating	11	9,2	3	5,0	-4,2
LOTS (C1+C2+C3)	69	58,0	36	60,0	+2,0
HOTS (C4+C5+C6)	50	42,0	24	40,0	-2,0

The data in Table 3 show that the transition from K13 to the Merdeka Curriculum did not result in an increase in the proportion of HOTS; on the contrary, HOTS decreased slightly from 42.0% to 40.0%. The most notable changes were a 17.4 percentage point increase in the proportion of C2 and an 11.9 percentage point decrease in C1. The proportion of C4 increased by 7.3 points, but this was offset by decreases in C5 (-5.1 points) and C6 (-4.2 points), so that, overall, there was no increase in HOTS.

Discussion

1. LOTS Dominance and Low Cognitive Pattern Persistence

The finding that K13 textbooks are dominated by LOTS questions (58.0%) is consistent with a number of previous studies that analyzed the cognitive quality of questions in Indonesian language textbooks. Saputro et al., (2021) found that the Grade 10 Indonesian Language textbook based on the K13 curriculum tends to emphasize LOTS-based thinking processes in its learning objectives, even though the curriculum calls for critical thinking skills. Alkarima et al., (2022) More specifically, it shows that HOTS questions in the 8th-grade junior high school Indonesian language textbook account for only 30.0%, far below the ideal threshold of 50.0% recommended by the Ministry of Education and Culture. A similar pattern was also found by Permatasari (2021) In the 8th-grade junior high school Indonesian language textbook under the 2013 Curriculum, LOTS questions account for over 60.0% of the content. The predominance of LOTS in the textbook's assessment instruments is not merely a technical issue of question design; rather, it reflects the textbook's pedagogical orientation, which still prioritizes the mastery of facts and concepts as the primary learning objective.

The fact that even textbooks under the Merdeka Curriculum still feature a higher proportion of LOTS (60.0%) than HOTS (40.0%) indicates the presence of structural inertia in the culture of textbook question development in Indonesia, an inertia that cannot be overcome simply by changing the curriculum's terminology. This inertia has been shown to transcend grade levels and subject areas: a study of elementary school textbooks Fadhilah et al., (2024); Izzah et al., (2025) as well as thematic textbooks Andrawana et al., (2022) consistently found that the dominance of LOTS persists despite differences in curricular contexts, and that the distribution of HOTS is sporadic and uneven across chapters. This cross-grade-level pattern reveals something more fundamental: the issue lies not in which curriculum is currently in

effect, but in the absence of binding minimum cognitive standards in the process of writing and evaluating textbooks themselves. The implications are serious: students who are constantly exposed to LOTS-based materials will not have sufficient learning experiences to tackle the AKM and PISA assessments, both of which require context-based complex reasoning skills, far beyond merely memorizing and understanding explicit information.

The persistence of LOTS's dominance must be understood within a broader framework than simply the issue of question-writing quality. Evidence from various studies points to a single root cause: the absence of systemic mechanisms to ensure adequate cognitive diversity in the process of writing and publishing textbooks. Nurlela et al., (2026) found that textbook authors generally lack practical technical guidance on the distribution of cognitive levels, so the development of test items is driven more by intuition and habit than by taxonomy-based design. Rahim et al., (2024) reinforces this finding by showing that even a change in curriculum is not automatically accompanied by training for authors in the design of HOTS-based assessments, so that question patterns tend to carry over from one curriculum to another with only cosmetic changes. Together, these two findings explain why the data in this study show similar patterns: without systemic interventions at the level of writing guidelines and the textbook evaluation process, no curriculum change will result in substantive changes in the cognitive quality of its assessment instruments.

2. The Shift from C1 to C2: A Surface Change, not a Substantial One

The most striking change between the K13 and the Merdeka Curriculum is the decrease in the proportion of C1 from 28.6% to 16.7% and the increase in C2 from 19.3% to 36.7%. At first glance, the decline in C1 appears to be progress because questions that previously required only the recall of facts have shifted to the comprehension level. However, this shift actually harbors an epistemological pitfall that warrants careful attention: the dominance of C2 in the Merdeka Curriculum indicates that the orientation of the questions still relies on the comprehension of explicit information, rather than on the ability to analyze, evaluate, or create. This phenomenon is not new in the literature, Cressa & Mukhlis (2023) found that C2 questions often appear to be HOTS because they use the question words "why" or "how," when in fact they merely require an understanding of information that is already explicitly available in the text. More importantly, this type of categorization error is not the fault of individual test writers, but rather a reflection of the lack of rigorous taxonomic verification in the textbook editorial process. In other words, the shift from C1 to C2 in the Merdeka Curriculum is better understood as a modernization of question packaging rather than a genuine increase in cognitive

demands, a change that is visible on the surface but does not touch the substance.

The dominance of C2 in the Merdeka Curriculum textbook is, in fact, inextricably linked to the characteristics of the question formats used: multiple-choice, true-false, and matching questions dominate the Competency Assessment section, and these formats structurally create a cognitive ceiling that limits thinking demands to no higher than the C3 level. This is not merely a theoretical assumption, Huda et al., (2021) and Rustiyani et al., (2021) independently found that multiple-choice questions, although claimed by their creators to be analytical, in practice only require the identification of answers already provided in the text, rather than an actual analytical process. This reveals a design paradox: the chosen question format is inherently incompatible with the cognitive demands it aims to achieve. Authentic implementation of HOTS requires essay, open-ended, or problem-based formats that provide students with the space to analyze, evaluate, and create productively (Gunartha, 2024; Himawan & Nurgiyantoro, 2022). As long as the question format remains unchanged, the shift from C1 to C2 in the Merdeka Curriculum will amount to nothing more than a cosmetic (regulatory) change, an illusion of cognitive progress that fails to address the root of the problem.

3. The Position of C4 as the Peak of HOTS: Gaps at C5 and C6

In both the K13 and the Merdeka Curriculum, Level C4 (analyzing) accounts for the largest share of HOTS, at 22.7% in the K13 and 30.0% in the Merdeka Curriculum. This finding is not an anomaly, but rather a confirmation of a pattern that has been repeatedly reported in the literature. Rohmah & Inderasari (2023) and Suvina & Ramly (2021) found that C4 represents the highest level achieved in Indonesian language textbook assessment instruments, while C5 and C6 are almost entirely absent. The consistency of this pattern across grade levels, publishers, and curricula suggests that C4 has become a sort of de facto ceiling in the design of Indonesian textbook questions, a limit that was never consciously set but is systematically never exceeded. This is a serious issue because it is precisely C5 and C6—evaluating and creating, that represent the cognitive levels most relevant to the demands of real life, the workplace, and 21st-century competencies. Students who are never trained to critically evaluate arguments or independently produce new work within the context of formal learning will face a significant cognitive deficit when confronted with real-world challenges.

The decline in C5 from 10.1% (K13) to 5.0% (Merdeka Curriculum) and in C6 from 9.2% to 5.0% is the most critical finding in this study, and at the same time the most paradoxical. The Merdeka Curriculum rhetorically positions critical, creative, and reflective thinking skills as the core elements of the Pancasila learner profile

(Annam et al., 2023), yet the data actually shows that questions requiring critical evaluation (C5) and creativity (C6) are less common in the Merdeka Curriculum textbooks than in the K13 textbooks. This paradox is not an isolated phenomenon, Febriyani & Mu'arifah (2024) and Komariah et al., (2023) consistently found that Level C6 is the least frequently appearing level in Indonesian language assessment instruments across various levels; its distribution is uneven across chapters; and this shortcoming persists without systematic improvement despite ongoing changes in curriculum policy. What makes this finding even more concerning is its implication for the coherence between curriculum aspirations and the reality of assessment instruments: a curriculum that is philosophically the most ambitious in terms of creativity and critical thinking actually produces textbooks with the lowest proportions of C5 and C6 questions. This is not merely a technical inconsistency; it is a systemic gap between what the curriculum promises students and what is actually taught to them every day in the classroom.

The absence of C6 questions in four of the six chapters of the Merdeka Curriculum textbook indicates that the ability to think creatively and productively, which is the philosophical essence of the Merdeka Curriculum, has not yet been consistently reflected in its assessment instruments. Moreover, the simultaneous decline in C5 and C6 highlights the gap between policy aspirations and implementation realities: while the curriculum sets a profile for graduates to be creative and critical, the instruments used to develop and measure these abilities have actually regressed. Ariyati et al., (2025) emphasizes that without adequate C5 and C6 questions, students have no opportunity to practice making argument-based critical judgments or producing meaningful new work within the context of formal learning. This situation must be addressed immediately through stricter revisions to the guidelines for developing textbook questions, rather than simply waiting for the next curriculum revision cycle.

4. Inter-chapter Variation: Uneven Cognitive Distribution

Analysis at the chapter level reveals significant variations in distribution, both in the K13 and the Merdeka Curriculum, and these variations themselves constitute findings that require critical interpretation. In the K13 textbook, Chapter III (Anecdotes) has the highest proportion of HOTS (54.5%) because the anecdote genre inherently demands analysis of implied meaning and identification of social criticism; these cognitive qualities emerge not because they were designed to, but because they are dictated by the genre's own characteristics. Conversely, Chapter I (Observation Report Text) actually has the lowest HOTS (36.4%), even though observation-based scientific texts are actually highly suitable for C4–C6 questions, such as comparing

two observation texts, evaluating data accuracy, or designing a new observation framework. This contrast confirms the argument Permatasari (2021) that the distribution of HOTS in textbooks is largely determined by the inherent characteristics of the text genres, rather than by a deliberate pedagogical design. In other words, the level of HOTS in each chapter does not reflect the instructional design awareness of the textbook authors, but is rather a byproduct of genre characteristics—a situation that should not be accepted as the standard in the development of high-quality textbooks.

In the Merdeka Curriculum textbook, Chapter II (Anecdotes) stands out with an 80.0% HOTS score, an achievement worthy of recognition, yet one that actually highlights a larger problem. The strength of Chapter II is supported by a question format that consistently encourages dialogue inference, analysis of social satire, and the creation of comic strips, all of which demand high-level cognitive processes. However, in contrast to Chapter IV (Negotiation), which stands at only 20.0%, and the other four chapters, all of which are below 55.0%, this reveals that the impressive achievement of this single chapter is the exception, not the norm. A range from 80.0% to 20.0% within the same textbook is not merely a reasonable variation but a clear indicator of the absence of systematic cognitive distribution planning. Andrawana et al., (2022) dan Ariyati et al., (2025) consistently asserts that such inconsistencies reflect a lack of rigorous editorial oversight in the design of textbook assessment instruments. More importantly, this inconsistency has tangible pedagogical consequences: students who sequentially progress through chapters with a high proportion of LOTS will lack experience in higher-order thinking before suddenly encountering a highly demanding chapter, creating an unbridgeable cognitive leap that is counterproductive to the development of thinking competencies, which should occur gradually and in a structured manner.

This unplanned variation across chapters has a serious impact on the cumulative effects of learning, which are often overlooked. The unevenness of cognitive development is not merely a quantitative issue, but one of sequencing: how the sequence of thinking experiences students encounter over the course of a school year cumulatively shapes, or, conversely, hinders, the development of their cognitive competencies (Himawan & Nurgiyantoro, 2022). Rahim et al., (2024) emphasizes that the design of textbook assessment instruments should be progressive, with the proportion of higher-order thinking skills (HOTS) gradually increasing from chapter to chapter as students' competencies develop, rather than fluctuating unpredictably. Together, Himawan & Nurgiyantoro (2022) and Rahim et al., (2024) This indicates that the quality of textbook assessments cannot be evaluated solely based on the aggregate proportion of HOTS, but also on the consistency and progression of their distribution

across chapters. The findings of this study, where HOTS variation across chapters is extremely high and lacks a clear pattern, indicate that this dimension of sequentiality has not been considered at all in the process of textbook development and evaluation by the National Textbook Center.

5. The Impact of Curriculum Changes on the Quality of Assessment

The findings of this study provide strong empirical evidence that curriculum changes, despite bringing about significant philosophical and structural reforms, do not automatically lead to improvements in the cognitive quality of assessment instruments in textbooks. Evidence from various implementation studies points to the same conclusion: the gap between curricular aspirations and assessment realities arises because policy changes are not accompanied by changes in capacity at the implementation level. Angga et al., (2022) indicates that the implementation of the new curriculum requires profound changes at the level of instructional design and assessment, not merely structural changes and changes in terminology. Sumarsih et al., (2022) and Melani & Gani (2023) reinforce this argument by showing that even teachers at pilot schools, who were among the first to implement the Merdeka Curriculum, still face real difficulties in designing HOTS-based assessments, and the inertia of conventional practices persists due to a lack of adequate technical support. If teachers who are directly dealing with the new curriculum are still stuck in old patterns, it is not surprising that textbooks produced within the same ecosystem also maintain a cognitive orientation that is not much different

From the perspective of textbook development, these findings indicate that improvements in the cognitive quality of assessment instruments will not be realized without systemic interventions at the level of publishing standards. The convergence between empirical recommendations and national policy directions is actually strong enough to serve as a basis for action: Rinjaya & Halimi (2022) recommend a minimum HOTS proportion of 50.0%, The Directorate of Secondary School Development (2017) emphasizes that assessments in senior high schools should focus more on HOTS, and The Center for Educational Assessment (2019) underscores the importance of instruments based on the abilities to analyze, evaluate, and create. What is missing is not an awareness of the importance of HOTS, but rather the mechanisms for enforcing cognitive standards within the textbook publishing process itself. Therefore, the practical implications of this research for the Textbook Center include three interrelated steps: first, establishing a minimum HOTS proportion of 50.0% and the presence of C5 and C6 questions in every chapter as mandatory eligibility criteria, on par with existing content and language criteria; second, requiring verification of cognitive distribution based on the Revised Bloom's Taxonomy as an integral part of

the manuscript review process before the book is published; and third, providing operational guidelines for developing taxonomy-based questions that can be directly used by textbook authors, rather than merely normative policy documents.

The implications for Indonesian language teachers are no less fundamental, and in fact are the most urgent, since it is the teachers who deal directly with the limitations of textbooks every day in the classroom. Gunartha (2024) and Himawan & Nurgiyantoro (2022) Together, these findings indicate that a teacher's ability to independently develop HOTS questions is not merely a matter of technical skill, but rather a reflection of deep pedagogical awareness regarding the purpose of assessment in language learning. Teachers who possess this awareness will not rely entirely on textbook questions, but will actively supplement and go beyond their limitations. Concretely, these implications require teachers to: (1) supplement textbook questions with self-created questions that intentionally target C5 and C6, for example, by asking students to evaluate the author's word choices or create a new text based on the genre studied; (2) routinely use assessment rubrics based on the Revised Bloom's Taxonomy in daily assessment planning, not just in end-of-semester exams; and (3) document and share self-created HOTS questions through the Indonesian Language Teachers' Working Group (MGMP) forum as a step toward sustained collective strengthening. These three steps will not be realized without systemic support in the form of structured, ongoing training, ideally organized by the Directorate of Teacher and Education Personnel (GTK) or provincial education offices within the framework of teacher professional development. Ultimately, the cognitive quality of assessment instruments in Indonesian language learning will not improve substantially merely through curriculum changes or the publication of new textbooks, but rather through simultaneous changes at three levels: textbook publication standards, teachers' capacity as assessment developers, and policy commitment that goes beyond mere curriculum rhetoric.

CONCLUSION

This study yielded three main conclusions. *First*, the 10th-grade Indonesian language textbook for the 2013 Curriculum shows a distribution of cognitive levels dominated by C1 (28.6%) and a ratio of 58.0% LOTS to 42.0% HOTS. Level C4 dominates the HOTS section, while C6 (creating) is the least common level, although it is found in all chapters. *Second*, the Grade 10 Indonesian Language textbook under the Merdeka Curriculum shows a distribution dominated by C2 (36.7%) and a ratio of 60.0% LOTS to 40.0% HOTS. The most significant changes are the shift in dominance from C1 to C2, as well as the increase in the proportion of C4, though accompanied by a decrease in C5 and C6. *Third*, a comparison between the two textbooks shows that

the transition from the 2013 Curriculum to the Merdeka Curriculum did not result in an aggregate increase in cognitive demands on the textbook's assessment instruments; the proportion of HOTS actually decreased slightly from 42.0% to 40.0%. These findings indicate that the curriculum change has not substantially altered the cognitive orientation of questions in official textbooks, and the changes that occurred are more a redistribution across LOTS levels than an actual increase in cognitive demands.

Based on these findings, several recommendations can be made. For developers and textbook authors, it is necessary to plan for a more explicit and even distribution of cognitive levels across each chapter, with a minimum target of 50.0% for Higher-Order Thinking Skills (HOTS), as encouraged by the Ministry of Education and Culture's policy, Direktorat Pembinaan Sekolah Menengah Atas; Pusat Penilaian Pendidikan accompanied by the consistent inclusion of C5 and C6 questions in every chapter. The guidelines for developing assessments based on the Revised Bloom's Taxonomy should be made an integral part of the textbook writing and evaluation process by the Textbook Center. For Indonesian language teachers, these findings underscore the importance of not relying entirely on questions in textbooks, but rather developing independent assessment instruments that are richer in HOTS questions, particularly at levels C5 and C6. For future researchers, similar studies could be expanded to other grade levels, private publishers, or focused on more specific language skills such as writing or speaking. The use of two *raters* and the Cohen's Kappa test should be maintained as a standard for data validity in similar content analysis studies.

REFERENCES

Aini, S. W., & Mukhlis, M. (2022). Analisis taksonomi pada soal Asesmen Kompetensi Minimum literasi membaca di SMK Negeri 3 Pekanbaru. *Diglosia: Jurnal Kajian Bahasa, Sastra, Dan Pengajarannya*, *5*(4), 933–948. https://doi.org/10.30872/diglosia.v5i4.513

Alkarima, O., Sumarwati, S., & Suryanto, E. (2022). Muatan keterampilan berpikir tingkat tinggi pada buku pelajaran Bahasa Indonesia SMP Kelas VIII. *GERAM: Gerakan Aktif Menulis*, *10*(1), 55–67. https://doi.org/10.25299/geram.2022.vol10(1).9021

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's educational objectives*. Longman.

Andrawana, T., Halidjah, S., & Suparjan. (2022). Analisis konten Higher Order Thinking Skills (HOTS) dalam buku teks tematik siswa Kelas V Tema 1. *PALAPA: Jurnal Studi Keislaman Dan Ilmu Pendidikan*, *10*(1), 61–79. https://doi.org/10.36088/palapa.v10i1.1668

Angga, A., Suryana, C., Nurwahidah, I., Hernawan, A. H., & Prihantini, P. (2022). Komparasi implementasi Kurikulum 2013 dan Kurikulum Merdeka di Sekolah Dasar Kabupaten Garut. *Jurnal Basicedu*, *6*(4), 5877–5889. https://doi.org/10.31004/basicedu.v6i4.3149

Annam, F. K., Lestari, M. I., Okvisari, R., Hasanah, T. L., & Handayani, V. (2023). Pengembangan keterampilan berpikir kritis dalam penerapan Kurikulum Merdeka di Sekolah Dasar. *Jurnal Pendidikan Guru Sekolah Dasar*, *1*(2), 1. https://doi.org/10.47134/pgsd.v1i2.204

Ariyati, I., Rustam, & Gati Ningsih, A. (2025). Analisis butir soal dalam buku teks Bahasa Indonesia SMP Kelas IX berdasarkan HOTS. *Jurnal Pendidikan Bahasa & Sastra Indonesia*, *13*(2), 1–10. https://doi.org/10.32682/sastranesia.v13e3/26/

Astari, T. (2022). Buku teks dalam implementasi Kurikulum Merdeka di Sekolah Dasar. *Madako Elementary School*, *1*(2), 163–175. https://doi.org/10.56630/mes.v1i2

Aulia, F. T., Gumilar, S. I., & Kurniawan, A. (2023). *Bahasa Indonesia untuk SMA/MA/SMK/MAK kelas X (Edisi revisi)*. Pusat Kurikulum dan Perbukuan.

Cressa, J., & Mukhlis, M. (2023). Level kognitif Taksonomi Bloom pada soal mata pelajaran Bahasa Indonesia. *J-LELC: Journal of Language Education, Linguistics, and Culture*, *3*(1), 55–64. https://doi.org/10.25299/j-lelc.2023.12094

Direktorat Pembinaan Sekolah Menengah Atas. (2017). *Panduan Penilaian oleh Pendidik dan Satuan Pendidikan Sekolah Menengah Atas*. Direktorat Pembinaan SMA, Direktorat Jenderal Pendidikan Dasar dan Menengah, Kementerian Pendidikan dan Kebudayaan. https://repositori.kemendikdasmen.go.id/18053/1/Panduan_Penilaian_SMA_Final_02082017.pdf

Fadhilah, H., Ratyasha, S., Donda, T., & Giwangsa, S. F. (2024). Analisis distribusi soal Higher Order Thinking Skills (HOTS) dan Lower Order Thinking Skills (LOTS) dalam buku ajar Bahasa Indonesia. *Indonesian Journal of Action Research*, *3*(1), 25–31. https://doi.org/10.14421/ijar.2024.31-03

Febriana, I., Wulandari, A. N., & Sari, Y. (2022). Keterbacaan buku teks Kurikulum Merdeka Bahasa Indonesia Kelas 7 dengan Grafik Fry. *Basastra: Jurnal Bahasa Dan Sastra Indonesia*, *11*(2), 174–184. https://doi.org/10.24114/bss.v11i2.38197

Febriyani, H., & Mu'arifah, S. (2024). Muatan soal LOTS dan HOTS kompetensi sastra buku teks Bahasa Indonesia Kelas VII terbitan Erlangga. *ANUFA: Jurnal Bahasa, Sastra, Dan Budaya*, *2*(1), 1–9. https://doi.org/10.63629/anufa.v2i1.66

Gunartha, I. W. (2024). Pengembangan penilaian berorientasi HOTS: Upaya peningkatan kemampuan berpikir kritis siswa di era global abad ke-21. *Widyadari*, *25*(1), 133–147. https://doi.org/10.59672/widyadari.v25i1.3660

Himawan, R., & Nurgiyantoro, B. (2022). Analisis butir soal latihan penilaian akhir semester ganjil mata pelajaran Bahasa Indonesia Kelas VIII SMPN 1 Bambanglipuro Bantul. *KEMBARA: Jurnal Keilmuan Bahasa, Sastra, Dan Pengajarannya*, *8*(1), 160–180. https://doi.org/10.22219/kembara.v8i1.20530

Huda, M., Purnomo, E., Anggraini, D., & Prameswari, D. H. (2021). Higher Order Thinking Skills (HOTS) dalam materi dan soal pada buku pelajaran Bahasa Indonesia SMA terbitan Kemendikbud RI. *PRASI: Jurnal Bahasa, Seni, Dan Pengajarannya*, *16*(2), 128–143. https://doi.org/10.23887/prasi.v16i02.40671

Izzah, A. N., Azizah, S. S. L., Ekawati, R., & Rufiana, I. S. (2025). Eksplorasi analisis butir soal Bahasa Indonesia di Kelas II Sekolah Dasar berdasarkan Taksonomi Bloom. *Jurnal Pendidikan Dan Pembelajaran Indonesia (JPPI)*, *5*(2), 1031–1043. https://doi.org/10.53299/jppi.v5i2.1064

Komariah, I., Hamdani, A., Hamdani, N. A., Suherman, E., & Kartini, A. (2023). Analisis butir soal ujian madrasah Bahasa Indonesia berdasarkan tingkat berpikir siswa di KKM 1 Garut. *Al Qodiri: Jurnal Pendidikan, Sosial Dan Keagamaan*, *21*(1), 298–306. https://doi.org/10.53515/qodiri.2023.21.1.298-306

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology* (4th ed.). SAGE Publications.

Melani, A., & Gani, E. (2023). Implementasi Kurikulum Merdeka dalam pembelajaran Bahasa Indonesia di SMP Negeri 16 Padang. *Educaniora: Journal of Education and Humanities*, *1*(2), 23–32. https://doi.org/10.59687/educaniora.v1i2.28

Nafiati, D. A. (2021). Revisi Taksonomi Bloom: Kognitif, afektif, dan psikomotorik. *Humanika: Kajian Ilmiah Mata Kuliah Umum*, *21*(2), 151–172. https://doi.org/10.21831/hum.v21i2.29252

Nurlela, R., Kuswari, U., & Nugraha, H. S. (2026). Analisis HOTS buku teks Rancage Diajar Basa Sunda Kelas VII SMP Kurikulum Merdeka. *Jurnal Ilmiah Global Education*, *7*(1). https://doi.org/10.55681/jige.v7i1.4867

OECD. (2022). *Indonesia Student Performance (PISA 2022)*. OECD: Education GPS The World of Education at Your Finger Tips. https://gpseducation.oecd.org/CountryProfile?primaryCountry=IDN&treshold=10&topic=PI

Permatasari, I. (2021). Distribusi pertanyaan Higher Order Thinking Skill (HOTS) pada buku teks Bahasa Indonesia SMP/MTs Kelas VIII Kurikulum 2013. *Nuances of Indonesian Language*, *2*(1), 1–7. https://doi.org/10.51817/nila.v2i1.60

Pusat Penilaian Pendidikan, B. P. dan P. (2019). *Panduan Penulisan Soal HOTS: Higher Order Thinking Skills*. Pusat Penilaian Pendidikan, Badan Penelitian dan Pengembangan, Kementerian Pendidikan dan Kebudayaan.

Rahim, A., Aziz, A., Rahim, I., Herdiani, R., Ambarita, J., & Nurhikmah. (2024). Integrating HOTS in Indonesian high school textbooks: A comparative study of the 2013 and Merdeka Curriculum. *Lingue: Jurnal Bahasa, Budaya, Dan Sastra*, *7*(1). https://doi.org/10.33477/lingue.v7i1.9426

Rinjaya, D., & Halimi, S. S. (2022). An evaluation of Indonesian textbooks using revised Bloom's taxonomy. *OKARA: Jurnal Bahasa Dan Sastra*, *16*(1), 1–18. https://doi.org/10.19105/ojbs.v16i1.5813

Rohmah, N., & Inderasari, E. (2023). Muatan HOTS pada soal kompetensi bersastra

buku teks Cerdas Cergas Kelas X. *ANUFA: Jurnal Bahasa, Sastra, Dan Budaya*, *1*(1), 11–18. https://doi.org/10.63629/anufa.v1i1.19

Rosdiana, R., Budiana, S., Mahajani, T., & Talitha, S. (2022). Penerapan HOTS pada soal-soal buku teks pelajaran Bahasa Indonesia tingkat SMA. *Aksara: Jurnal Ilmu Pendidikan Nonformal*, *8*(2), 1065–1074. https://doi.org/10.37905/aksara.8.2.1065-1074.2022

Rustiyani, R., Sofyan, D., & Syafryadin, S. (2021). Levels of cognitive domain of tasks in English textbooks for senior high school: A revised Bloom's taxonomy analyses. *English Education: Jurnal Tadris Bahasa Inggris*, *14*(2), 280–293. https://doi.org/10.24042/ee-jtbi.v14i2.9960

Saputro, D., Sabardila, A., Prayitno, H. J., & Markhamah, M. (2021). Integrasi keterampilan berpikir kritis dalam buku teks Bahasa Indonesia Kurikulum 2013 berperspektif HOTS. *Diglosia: Jurnal Kajian Bahasa, Sastra, Dan Pengajarannya*, *4*(3), 365–374. https://doi.org/10.30872/diglosia.v4i3.168

Suherli, Suryaman, M., Septiaji, A., & Istiqomah. (2017). *Bahasa Indonesia untuk SMA/MA/SMK/MAK kelas X* (Edisi revi). Kementerian Pendidikan dan Kebudayaan.

Sumarsih, I., Marliyani, T., Hadiyansah, Y., Hernawan, A. H., & Prihantini, P. (2022). Analisis implementasi kurikulum merdeka di sekolah penggerak sekolah dasar. *Jurnal Basicedu*, *6*(5), 8248–8258. https://doi.org/10.31004/basicedu.v6i5.3216

Sutami, S., Hudiyono, Y., & Ilyas, M. (2020). Pengembangan instrumen asesmen Higher Order Thinking Skills (HOTS) pada mata pelajaran Bahasa Indonesia SMA dan SMK. *Diglosia: Jurnal Kajian Bahasa, Sastra, Dan Pengajarannya*, *3*(1), 102–113. https://doi.org/10.30872/diglosia.v3i1.24

Suvina, N., & Ramly, R. (2021). Analisis pertanyaan HOTS buku teks mata pelajaran Bahasa Indonesia SMK/MAK Kelas X terbitan Erlangga. *INDONESIA: Jurnal Pembelajaran Bahasa Dan Sastra Indonesia*, *2*(1). https://doi.org/10.59562/indonesia.v2i1.19292