Journal on Mathematics Education
Volume 15, No. 4, 2024, pp. 1311-1334

Curriculum and teacher assessment practices in mathematics learning: Alignment with higher order thinking skills in Indonesian secondary schools

Firdha Mahrifatul Zana[1], Cholis Sa'dijah[1,*], Susiswo[1], Lathiful Anwar[1], Hutkemri Zulnaidi[2]

[1]Mathematics Department, Universitas Negeri Malang, Malang, Indonesia
[2]Mathematics and Science Education Department, University of Malaya, Kuala Lumpur, Malaysia
*Correspondence: cholis.sadijah.fmipa@um.ac.id

Received: 22 June 2024 | Revised: 31 October 2024 | Accepted: 16 November 2024 | Published Online: 20 November 2024

Abstract

Higher-Order Thinking Skills (HOTS) are an essential element in education that must be integrated into curricula and classroom assessments. In Indonesia, educational initiatives have increasingly emphasized the incorporation of HOTS into both curriculum design and assessment practices. However, prior research has primarily focused on the challenges faced by teachers in developing HOTS-based assessments and aligning their teaching with curriculum demands. This study aims to investigate how the Indonesian mathematics curriculum integrates HOTS and evaluate the alignment between the curriculum objectives and teacher-developed assessments in fostering HOTS. The study employed a descriptive qualitative approach and was conducted in two Indonesian high schools, one located in an urban area and the other in a regional setting. A total of 15 mathematics teachers from grades ten, eleven, and twelve participated in the research. Data collection methods included focus group discussions, document analysis of mathematics assessments, and semi-structured interviews. The analysis employed Anderson and Krathwohl's Taxonomy to categorize cognitive levels. Findings reveal that the Indonesian Mathematics Curriculum predominantly emphasizes Low-Order Thinking Skills (LOTS), and teacher-developed assessments are largely aligned with these LOTS-focused objectives. Furthermore, even when curriculum indicators aim to target HOTS, teachers often struggle to design assessments that effectively evaluate students' higher-order cognitive abilities. These findings highlight a significant gap between curriculum goals and the practical implementation of HOTS in assessments. The results provide valuable insights for curriculum developers, suggesting the need for a curriculum redesign that places greater emphasis on HOTS. Additionally, the study underscores the importance of professional development initiatives to equip teachers with the skills necessary to design and implement HOTS-based assessments. This research contributes to advancing educational practices and policies that prioritize the integration of HOTS into teaching and assessment frameworks.

Keywords: Cognitive Level, Curriculum, Higher Order Thinking Skills, Mathematics, Teachers' Assessment

How to Cite: Zana, F. M., Sa'dijah, C., Susiswo, Anwar, L., & Zulnaidi, H. (2024). Curriculum and teacher assessment practices in mathematics learning: Alignment with higher order thinking skills in Indonesian secondary schools. *Journal on Mathematics Education, 15*(4), 1311–1334. https://doi.org/10.22342/jme.v15i4.pp1311-1334

The National Council of Teachers of Mathematics (NCTM) (1989; 2000; 2014) asserts that mathematics serves as a fundamental tool for teaching students to explore, connect, reason logically, solve problems, represent, and communicate mathematically, particularly in the context of rapid technological advancements. In parallel, numerous scholars emphasize the importance of fostering students' skills as a central objective in mathematics education. For instance, McHugh et al. (2021) highlight the significance of higher-level problem-solving abilities, content knowledge, and affective domains in mathematics, while Podkhodova et al. (2020) identify essential mathematical skills, including the ability to solve mathematical, instructional, and professional problems effectively. In accordance with these findings, mathematics educators utilize various instructional strategies, such as individual and group work, along with classroom assessments involving contextual mathematical problems. These assessments typically include tasks such as posing problems, prompting solutions, verifying answers, and encouraging the generation of new ideas (Shahrill et al., 2018; Sa'dijah et al., 2021). Additionally, peer assessment is implemented to enhance students' cognitive skills, fostering abilities in analysis, evaluation, problem-solving, and critical and creative thinking (Hadzhikoleva et al., 2019). Thus, mathematics plays a pivotal role in developing students' problem-solving and critical thinking capabilities through effective teaching practices, contextual problem-solving, and collaborative peer assessments within the classroom.

The 21st Century Learning represents a dynamic approach to education that emphasizes the integration of digital technology as a key component. This learning paradigm prioritizes the development of skills and knowledge relevant to the real world, positioning students as active and communicative learners (Istiyono, 2018; Syahida & Dewi, 2023). Amzaleg and Masry-Herzallah (2021) argue that teachers must provide cooperative education, foster creativity, offer motivation, and demonstrate empathy to cultivate reasoning, collaboration, and creative thinking skills among students. Similarly, Murtafiah et al. (2019) underscore the importance of decision-making skills, which involve generating ideas, clarifying ideas, and assessing the fairness of ideas, in the development of student thinking. These competencies are classified as HOTS by Anderson et al. (2001), a view also supported by Brookhart (2010), who emphasizes the centrality of reasoning, critical and creative thinking, collaborative thinking, and judgment in the development of high-level cognitive skills. The cultivation of HOTS in educational settings is believed to enhance students' quality and competence, thereby preparing them for the competitive demands of the 21st century (Thornhill-Miller et al., 2023). Moreover, HOTS-based learning in schools has the potential to produce students with critical insights who are capable of competing on an international scale (Jailani et al., 2017).

However, the implementation of HOTS-focused teaching remains a challenge for many educators (Thompson, 2008). Obstacles often arise due to various factors, including a lack of knowledge about HOTS (Marsh, 2016), insufficient preparation by teachers, and the influence of student-related factors (Musa & Yamat, 2021). Additionally, the comfort of teachers with conventional teaching methods (Seman et al., 2017), the heavy workload associated with classroom teaching (FitzPatrick & Schulz, 2015), and the belief that students will struggle with HOTS-based learning (Barak & Shakhman, 2008) contribute to the difficulties faced in integrating these skills into instruction. In rural and regional schools in Indonesia, these challenges are exacerbated by limited access to HOTS-focused training due to inadequate infrastructure, including insufficient internet access and substandard teaching environments (Ichwanto et al., 2024). As a result, many educators continue to rely on outdated teaching methods (Zulfikar, 2018), which further hampers their ability to effectively integrate HOTS into classroom assessment. Ultimately, developing HOTS is a critical necessity for all educators, as it is essential to the preparation of students for the complexities of modern global society.

HOTS have been widely acknowledged as essential across all educational levels. Previous research has emphasized the necessity of continuously teaching HOTS to students, as this will motivate them to become effective thinkers and problem-solvers in the real world. The findings of this study indicate that both teachers and students face significant challenges in the effective implementation of HOTS. For students, HOTS learning is often perceived as difficult to apply, as many students remain passive and require additional motivation to actively engage in the learning and discussion processes. Moreover, factors such as low reading interest, large class sizes, and varying student abilities further hinder the application of HOTS. From the teacher's perspective, difficulties arise in designing HOTS-based assessments, compounded by limited classroom time for instruction (Zana et al., 2022). The study by As'ari et al. (2017) also highlighted that prospective teacher candidates often lack proficiency in high-level thinking, with many future mathematics educators classified as non-critical thinkers. In contrast, the research by Lugosi and Uribe (2022) demonstrated that active teaching strategies—such as math projects, group work with discussion and feedback, student exploration, and interactive presentations—are effective in enhancing students' thinking skills. Additionally, Olivares et al. (2021), in their review of 131 articles published in reputable journals, highlighted the importance of embedding HOTS, such as problem-solving, into curriculum structures and characteristics, suggesting that the integration of HOTS in both teaching strategies and curriculum design is key to fostering students' higher-order cognitive abilities.

In education, the curriculum serves as the foundation for teachers in designing teaching strategies, learning activities, and assessment tasks. It is crucial for teachers to ensure that there is alignment between assessment, teaching, and learning activities with the intended outcomes outlined in the curriculum (Ali, 2018). While teachers bear significant responsibility for many instructional strategies and assessment decisions in the classroom, it is essential that explicit attention to higher-order thinking is incorporated into curriculum documents. This inclusion provides both guidance and consistency, benefiting students and teachers alike (Schulz & Fitzpatrick, 2016).

A developmental study by Fulmer and Polikoff (2014) introduced a method for analyzing the alignment between curriculum documents and teacher assessment practices. Consistent with this approach, previous research has emphasized that assessment items designed by teachers must be relevant to the desired outcomes specified in the curriculum (El-Hassan & Baassiri, 2019; Fulmer, 2011). This alignment ensures that students' learning experiences adhere to the curriculum, while the assessments effectively measure the students' learning progress. As noted by Suherman and Vidákovich (2022), assessment types may include both in-class methods, such as self-assessment, peer assessment, daily tests, and assignments, as well as out-of-class assessments like observations, practice tests, projects, portfolios, and performance evaluations (Sa'dijah et al., 2015). Therefore, it is imperative that the teacher's curriculum and assessment practices are mutually reinforcing to successfully achieve the learning objectives, particularly the development of students' HOTS.

Teacher assessment serves several key functions in the learning process, acting as both a measure of student achievement and a tool for providing feedback to guide instructional improvements. Beyond these primary roles, assessments also help to evaluate students' progress after a series of lessons, determine the effectiveness of teaching methods, and classify students based on their skill levels (Murtafiah et al., 2020). Wilcox and Lanier (2000) argue that when assessments are used constructively and comprehensively, they can offer valuable insights into student understanding, including misconceptions and incomplete concepts. Furthermore, assessments enable teachers to gauge student progress towards the learning objectives outlined in the curriculum.

Given the crucial role that assessment plays in education, numerous studies have explored various aspects of assessment development. These studies range from investigating the types of assessments employed by teachers in the classroom (e.g., Berg, 2004; Berisha & Bytyqi, 2020) to the design of assessment instruments aimed at measuring students' skills (e.g., Adams & Wieman, 2011). Berisha and

Bytyqi (2020) found that teachers' assessment practices often failed to align with curriculum requirements. Specifically, they observed that teachers predominantly relied on closed, low-level cognitive tasks, which were presented in symbolic form, required simple operations, and lacked applicability. This finding contrasts with the results of Berg's (2004) study, where teachers employed a variety of formative assessments, summative assessments, and portfolio assessments, such as self-assessment, peer assessment, and lecturer assessment, to promote HOTS as outlined in the curriculum. From an assessment development perspective, Adams and Wieman (2011) demonstrated that assessments designed to integrate HOTS and align with curriculum standards have been proven effective in helping students develop thinking skills akin to those of experts.

In this study, the researchers assessed the extent to which mathematics teachers' assessments and the curriculum achievement indicators, particularly those outlined in the 2013 Curriculum (K-13) Revised Edition, targeted the development of HOTS in students. The research focused on the Indonesian mathematics curriculum for senior high schools, which encompasses basic competencies, instructional materials, and curriculum achievement indicators. These are further categorized into five main themes: Statistics and Probability, Algebra, Trigonometry, Geometry and Measurement, and Calculus. The assessments analyzed in this study were samples of both formative and summative assessments developed by teachers for students in grades ten, eleven, and twelve. The teachers provided the researchers with these assessments for analysis.

Summative assessments are defined as evaluations used at the end of a learning period to gauge students' mastery of the material, while formative assessments occur throughout the learning process, offering feedback to teachers for improving instructional strategies and advancing student progress (Retnawati et al., 2016). In addition to in-class assessments, mathematics teachers integrate assessments with classroom instruction, learning media, and teaching models to achieve curriculum objectives, including fostering HOTS development in students. Previous studies have emphasized that assessment serves as a critical link between the curriculum and the teaching process, underscoring the importance of aligning assessments with curriculum goals to ensure their effectiveness in supporting student learning outcomes (Madina & Kardena, 2021).

A comprehensive literature review has been conducted, highlighting significant research on the importance of integrating HOTS into curricula and classroom instruction, including studies from Indonesia. Numerous studies have examined the global implementation of the curriculum, revealing a growing trend toward incorporating HOTS into mathematics teaching (Anggraena et al., 2018; Grant & Gareis, 2013; Suyanto, 2018; Yeung, 2015). In Indonesia, the current curriculum, the "K-13 Revised Edition," emphasizes the integration of HOTS within the teaching and learning process (Pratiwi & Mustadi, 2021; Pulungan et al., 2021; Rizki et al., 2022). One notable effort in this direction has been the development of HOTS-based assessments for student worksheets, aligned with the basic competencies of the K-13 Revised Edition (Kusuma et al., 2017).

However, studies conducted in Indonesia have identified a gap between the curriculum's design and teachers' actual learning practices. Research on the implementation of the K-13 Revised Edition reveals that mathematics teachers face significant challenges when applying the curriculum in the classroom (Verdina et al., 2018). One manifestation of this gap is that teachers often struggle to develop assessments that accurately assess skills aligned with the curriculum's expectations (Suwarno et al., 2019; Zana et al., 2024b). Further studies have shown that teachers encounter obstacles in creating HOTS-based assessments and aligning them with the curriculum's desired outcomes (Dahlan et al., 2020; Zana et al., 2024a). Consequently, there remains a gap in research regarding an in-depth investigation into the skills targeted by the K-13 Revised Edition in Indonesia, and how Indonesian mathematics teachers design assessment items targeting HOTS cognitive levels, specifically in subjects such as Algebra, Statistics and Probability, Trigonometry, Geometry and Measurement, and Calculus. Finally, this study seeks to address this gap by examining the extent to which HOTS is integrated into the K-13 Revised Edition and exploring the alignment between mathematics teacher assessments and HOTS objectives in the context of Indonesia.

METHODS

Research Approach

This study aims to investigate whether the K-13 Revised Edition in Indonesia and the assessments developed by teachers, based on the curriculum's basic competencies, effectively support the development of students' HOTS. To achieve this, the researchers adopted a qualitative exploratory approach, which involves an in-depth examination and analysis of the skills targeted in the curriculum and teachers' abilities to design assessments (Creswell, 2014). In this research design, the researchers collected, explored, and analyzed two key elements: the curriculum achievement indicators and the assessments that teachers have developed based on their experiences teaching mathematics in schools.

Qualitative research offers a framework that addresses research questions related to the skills targeted by the curriculum and the teacher assessments that promote HOTS in students. This approach allows for a comprehensive investigation and deeper understanding of the phenomenon in question (Busetto et al., 2020). Furthermore, the findings of the study were presented in a narrative format, accurately reflecting the actual situation. The study specifically investigated the cognitive levels embedded in the curriculum and the teacher assessments across five main themes: Statistics and Probability, Algebra, Trigonometry, Geometry and Measurement, and Calculus.

The problem limitations in this study are established to ensure a focused approach, preventing any deviation or broadening of the research scope. This focus allows for a more structured discussion and helps achieve the study's objectives effectively. The key limitations of the study are as follows: the scope is restricted to cognitive thinking, mathematics teacher assessments, and the mathematics curriculum in Indonesia. The study specifically investigates the cognitive levels of the mathematics curriculum and teacher assessments in supporting students' HOTS, as well as the alignment between the cognitive levels in the curriculum and teacher assessments.

To collect relevant data, the research employs several methods, including the analysis of teacher assessment documents and the mathematics curriculum in Indonesia, Focus Group Discussions (FGD), and semi-structured interviews with selected participants. The data from the teacher assessment documents and curriculum were analyzed using Anderson and Krathwohl's Taxonomy, which categorizes cognitive levels. The researchers specifically chose this taxonomy over Bloom's Taxonomy, which is more widely known, because Anderson and Krathwohl's Taxonomy organizes cognitive levels into "verbs" rather than "nouns," as is the case with Bloom's Taxonomy (Darwazeh & Branch, 2015). This distinction is significant because the "verbs" in Anderson and Krathwohl's Taxonomy describe the cognitive processes targeted by teacher assessments and curriculum achievement indicators (Zana et al., 2024b). This approach aligns with the study's focus on analyzing the "verbs" found in both the assessments and curriculum achievement indicators.

After completing the data analysis, FGDs and semi-structured interviews were conducted with the research participants. The transcripts from the FGDs and interviews were systematically coded to identify recurring themes. Five researchers independently categorized the teacher assessments and curriculum achievement indicators into LOTS and HOTS levels. Any disagreements that arose during the categorization process were resolved through discussions among the four researchers, followed by consultation with an external examiner to ensure accuracy and consistency in the classification.

Research Instrument

The researchers utilized two types of instruments to collect data: main instruments and supporting instruments. The primary instruments in this study were the cognitive level indicator of Anderson and Krathwohl's Taxonomy, interview guidelines, and FGD guidelines. The cognitive level indicator from Anderson and Krathwohl's Taxonomy was employed to analyze and describe the cognitive skills targeted by the assessments and the curriculum. The FGD guidelines contained key questions concerning teachers' general perspectives on the implementation of the K-13 Revised Edition and their understanding of assessments. In contrast, the interview guidelines included more in-depth questions about the assessments created by the teachers and the challenges they faced in implementing HOTS-based assessments. Throughout the data collection process, the researchers maintained a collaborative relationship with the research participants, ensuring that data collection and interviews were adjusted to the contextual conditions within the participants' environment.

Supporting instruments included validation sheets for each main instrument, namely the validation sheet for the cognitive level indicator of Anderson and Krathwohl's Taxonomy, the validation sheet for the interview guidelines, and the validation sheet for the FGD guidelines. Prior to being used in the study, the instruments underwent validation to ensure their accuracy and appropriateness (Sudaryono et al., 2019). Initially, the instruments were validated by the research team, followed by an evaluation by one Mathematics Education Lecturer with over ten years of teaching experience and two Mathematics teacher practitioners with more than 15 years of experience. Several revisions were made to ensure the instruments' validity. The validity tests indicated that the cognitive level indicator of Anderson and Krathwohl's Taxonomy was adequately valid, and both the interview and FGD guidelines were valid for use in the study.

To minimize bias and enhance both internal and external validity, the researchers adhered to a semi-structured interview guide and recorded all interactions between the researchers and participants for later verification. The verification process included technique triangulation, where the researchers compared the assessment documents, curriculum objectives, FGD results, and teacher interviews to ensure the consistency and reliability of the data.

Research Participants

Participants in this study were mathematics teachers selected from two schools with an accreditation rating of 'A'. The schools were located in both districts and cities, with the aim of broadening the subject and research data. The schools included in the study were senior high schools from representative districts and cities that gave permission and were willing to participate. This study is limited to the selected schools based on these criteria and is confined to the districts and cities involved. Ethical considerations in the study included participant volunteerism, the agreement to collect assessment documents and curriculum materials, maintaining the anonymity of participants, and the communication of results where triangulation of sources was employed.

The two schools were located in Malang City and Nganjuk District, East Java, Indonesia. Data collection took place from July 2022 to 2023, involving a total of fifteen teachers. Nine teachers (60%)

were from high schools in Nganjuk District and six teachers (40%) were from high schools in Malang City. The distribution of teachers was as follows: two teachers from Nganjuk taught tenth grade, two taught eleventh grade, and five taught twelfth grades. In Malang, two teachers taught tenth grade, one taught eleventh grade, and three taught twelfth grades.

The fifteen teachers involved in the study had an average of 21 years of teaching experience. Of these, eight teachers held undergraduate degrees, while seven had completed postgraduate degrees in Mathematics or Mathematics Education (see Figure 1). There was variation in the teachers' familiarity with HOTS; nine teachers had attended professional development workshops on the implementation of the K-13 Revised Edition and the integration of HOTS.

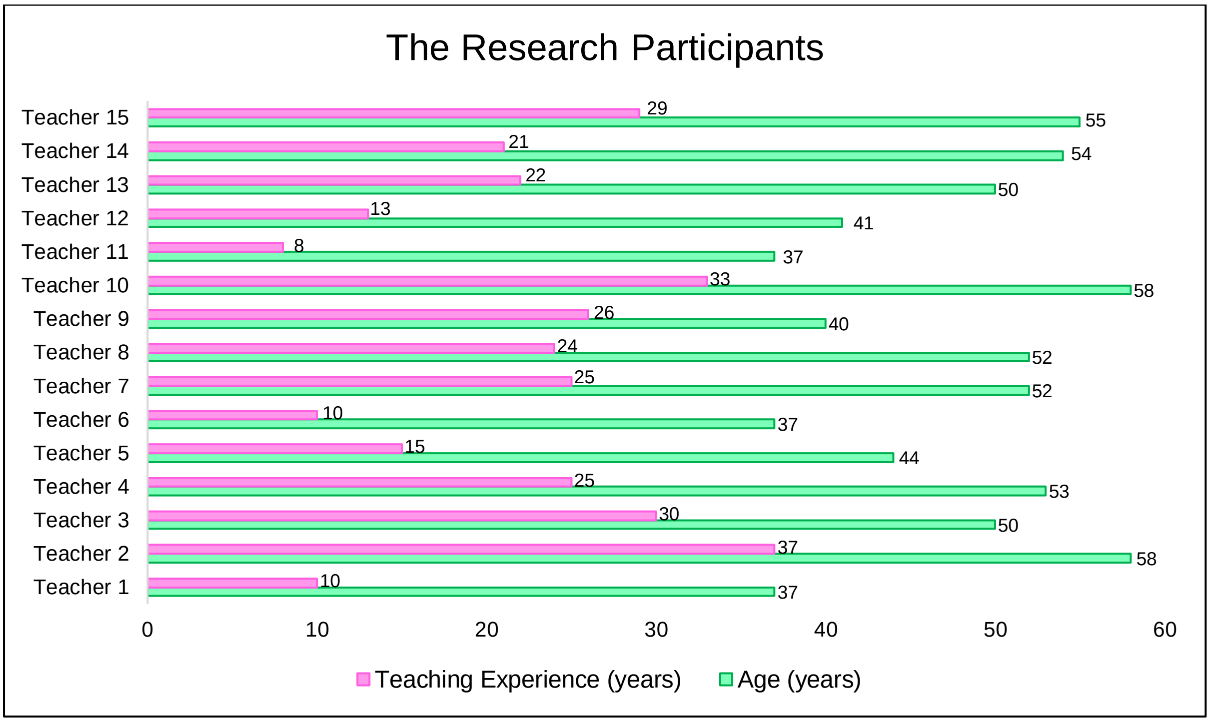


Figure 1. Research participants

Data Collection

Data collection was conducted using a combination of FGDs, document analysis, and semi-structured interviews with each research participant (Creswell, 2012). The first FGD aimed to explore teachers' general perspectives on the implementation and content of the mathematics curriculum in Indonesia, as well as their knowledge and understanding of assessment practices. FGDs were held at each participating school and lasted 50–60 minutes, allowing researchers to gather insights from the participants collectively.

Following the FGDs, researchers collected documents prepared by the participants, including curriculum achievement indicators and assessments developed by the teachers based on these indicators. These assessments were used during the classroom learning process to evaluate student progress. Subsequently, researchers conducted semi-structured interviews with individual participants. The interviews, guided by a predefined set of questions (Byrne, 2001), were tailored to the specific needs of the research process and lasted approximately 30–45 minutes per participant.

During the data collection phase, participants were asked to submit curriculum indicators and assessment items via email within 1–8 weeks following the FGDs. Each participant provided 5–20 assessment items, which included both formative and summative assessments. For participants who delayed submission, researchers followed up via email, phone calls, and WhatsApp messages to ensure timely data collection. Submission delays were tolerated for up to 14 days without impacting the validity of the collected data.

To verify the accuracy and consistency of the data, individual interviews were conducted with each participant to cross-check the validity of their submitted assessments through audio-video interviews. The validity of the research data was further strengthened through triangulation, which involved comparing teacher assessments with interview transcripts to ensure alignment and reliability (Creswell, 2012).

Data Analysis

The data obtained in this study were analyzed in three key phases: reduction, presentation, and conclusion (Miles et al., 2018). In reduction phase, the researchers focused on summarizing and distilling the interview transcripts and other collected data to identify key aspects relevant to the research questions. The data reduction process involved selecting the most important information and looking for emerging themes and patterns. The researchers grouped the curriculum achievement indicators and teacher assessments into five main themes: Algebra, Trigonometry, Geometry and Measurement, Calculus, and Statistics and Probability. Within each theme, the researchers then examined the cognitive levels targeted by the curriculum indicators and teacher assessments, using Anderson and Krathwohl's Taxonomy as a framework to categorize the cognitive skills being assessed. This allowed for a clearer understanding of the cognitive demands present in the curriculum and assessments. Additionally, the interview transcripts were reduced by focusing on key points related to the teachers' views on assessment practices and their alignment with the curriculum.

Furthermore, in the presentation phase, the researchers presented the analyzed data in the form of descriptive text, complemented by documentation such as photographs and interview transcripts. This presentation aimed to illustrate the cognitive levels found in both the mathematics curriculum and the formative and summative assessments used by the teachers. The researchers ensured that the assessments used in the study had undergone quality testing, as the teachers explained during the FGDs. Teachers reported that all assessments were first evaluated by the school's Question Inspection Team and subjected to peer evaluations before being used in the classroom. This process ensured that the assessments were of high quality and fit for use in the study. The descriptive text, therefore, included not only the cognitive levels of the assessments but also the teachers' feedback and validation processes that ensured the quality of the materials.

Finally, the researchers drew conclusions based on the analysis and synthesis of the data in the conclusion phase. The conclusions were aimed at answering the main research questions, focusing on the alignment between the cognitive levels in the curriculum and teacher assessments, and evaluating whether the assessments effectively supported the development of students' HOTS. The conclusions also considered how teachers' experiences and practices influenced their ability to design assessments aligned with HOTS objectives and curriculum requirements. These conclusions provided insights into the effectiveness of the K-13 Revised Edition in fostering HOTS through teacher assessments and highlighted areas for improvement.

In data reduction, the analysis process involved reviewing the operational verbs used in both teacher assessments and curriculum achievement indicators and cross-referencing them with Anderson and Krathwohl's cognitive Taxonomy. This cross-referencing refers to the process of comparing operational verbs with cognitive verbs at six levels in Anderson and Krathwohl's Taxonomy. For instance, in the cognitive level analysis of curriculum achievement indicators, the teacher compiled the indicator as follows: "Formulate linear equations and/or inequalities of one variable that contain absolute values." It was found that two researchers categorized "formulating" as HOTS and the other as LOTS, both of whom reviewed the accompanying assessments to ensure consistency. Next, an example of the teacher assessment analysis. Given a teacher assessment item as follows: "The curve $f(x) = x^3 + bx^2 + cx - 6$ passes through the point $(2, -4)$ and has a maximum turning point $(1, -2)$. Prove that the following statements are true or false. Give your reasons! (a) Value of $a = 1$, (b) Value of $b = 6$, (c) Sum of values of $a + b + c = -4$, (d) Minimum turning point of $f(x)$ curve is $(3, -6)$". During the cognitive level analysis process of the assessment, three researchers categorized the operational verb "proving" as LOTS and the others as HOTS. Therefore, a re-analysis was carried out among all researchers by reviewing the completion steps of the assessment items. Then, the results were consulted with the external examiner until an agreement was reached.

In this study, after reducing the data from the interviews and FGDs, the researchers moved on to the presentation phase, where the data were presented in the form of descriptive text. This descriptive presentation was supported by graphic charts, tables, and interview transcripts from the research participants. The descriptive text included detailed analyses of the cognitive levels targeted by both the mathematics curriculum and the teacher assessments. These analyses provided insights into how well the curriculum and the assessments aligned with the goals of fostering HOTS among students.

The final step in the data analysis process was drawing conclusions. This phase involved continuous verification of the data throughout the research process, ensuring that the conclusions drawn were based on thorough analysis and interpretation. The researchers carefully examined the research data from the senior high school mathematics teachers, focusing on the cognitive levels present in both the curriculum achievement indicators and the teacher assessments. Finally, the researchers used this analysis to answer the two main research questions.

RESULTS AND DISCUSSION

In this section, the researchers present a discussion of two key findings: the skills outlined in the indicators of the Indonesian Mathematics curriculum and the extent to which teachers' assessments evaluate students' HOTS. Initially, the researchers conducted an analysis of the cognitive levels associated with the teacher's indicators, based on the basic competencies outlined in the mathematics curriculum, which were organized into five distinct themes. Subsequently, the cognitive levels of the assessments were examined and the skills assessed during the evaluation process were categorized. Furthermore, the results of this study are categorized into two primary findings: the mathematics curriculum and teachers' formative and summative assessments.

Mathematics Curriculum

The outcomes of the K-13 Revised Edition in Indonesia encompass four key competencies: spiritual attitude competence, social attitude competence, knowledge competence, and skills competence. This study specifically focused on the learning objectives related to KI-3 (knowledge competence) and KI-4 (skills competence). These competencies for the tenth, eleventh, and twelfth grades were then categorized into five thematic areas. The researchers systematically analyzed the material within each basic competency and classified it under the themes of Statistics and Probability, Algebra, Trigonometry, Geometry, Measurement, or Calculus. Consequently, there are ten basic competencies for grade ten, ten for grade eleven, and four for grade twelve.

Following this classification, the researchers examined and categorized the cognitive levels of each competency achievement indicator formulated by the teachers based on these basic competencies. For this categorization, the Anderson and Krathwohl Taxonomy of Cognitive Levels was applied. A total of 220 indicators were collected. Initially, the material or content within these indicators was categorized according to the five themes: algebra, trigonometry, geometry and measurement, calculus, and statistics and probability. Furthermore, each indicator was assessed and classified as either LOTS (remembering, understanding, and applying) or HOTS (analyzing, evaluating, and creating), as outlined by Anderson et al. (2001).

Moreover, the researchers aligned the operational verbs used in the indicators with those listed in the Anderson et al. (2001) and Krathwohl Taxonomy of Cognitive Levels. In cases where an operational verb from the assessment did not correspond to any verb in the Taxonomy's list, the researchers conducted cross-referencing. This process involved identifying synonyms for the operational verbs in the assessment and comparing them with the cognitive levels defined in Anderson and Krathwohl's Taxonomy. The researchers then thoroughly reviewed the accompanying assessment to ensure consistency. This categorization process required careful and detailed analysis, as the mere use of identical operational verbs does not necessarily imply that the cognitive level of the curriculum attainment indicators corresponds directly to that of Anderson et al. (2001) and Krathwohl's Taxonomy.

The researchers engaged in a rigorous process of analyzing, matching, and categorizing the cognitive levels of the curriculum achievement indicators. If disagreements arose regarding the categorization of an indicator, a re-analysis was conducted until consensus was achieved on its cognitive level. The results of this re-analysis were then reviewed by an external examiner for final validation.

For instance, consider an indicator in which the teacher provides the equation of a curve and a point through which the curve passes, along with the equation of a line parallel to the curve's tangent. The task required students to determine the various equations of tangents to the curve. According to the Anderson et al. (2001) and Krathwohl Taxonomy, the operational verb "to determine" is typically classified under LOTS, specifically within the "applying" level. However, following cross-referencing with the Taxonomy and the corresponding assessment, it was determined that the indicator aligned with the "creating" level. This conclusion was based on the nature of the open-ended problem, where students were required to formulate different tangent equations based on given conditions, rather than simply applying a predefined formula. As a result, each student's response would vary from others, indicating a higher level of cognitive engagement and creativity.

After analyzing the 220 competency achievement indicators in the curriculum, the researchers found that the cognitive level of "applying" was most prevalent within the Algebra theme, with 51 indicators. This predominance can be attributed to the fact that the essential competencies emphasized in the tenth and eleventh grades were largely focused on the Algebra theme. Additionally, the researchers identified that Algebra contained the highest number of indicators targeting HOTS, particularly the "creating" level, with four indicators, compared to other themes. However, within Algebra, the number of indicators at the lower cognitive levels surpassed those targeting higher cognitive levels.

A similar trend was observed in the Statistics and Probability theme, where the majority of indicators focused on "applying" skills, with 23 indicators identified. Only seven indicators addressed HOTS at the "analyzing" level, and no indicators were found targeting the "evaluating" or "creating" levels in this theme. A comparable pattern was found in the Trigonometry, Geometry and Measurement, and

Calculus themes as well.

These findings align with previous research, which underscores the importance of integrating HOTS into the curriculum, as teachers are consistently expected to meet curriculum objectives (FitzPatrick & Schulz, 2015). Additionally, studies have shown that teachers who are adept at analyzing and interpreting the competencies outlined in the curriculum, and subsequently applying them in the learning process, can provide clearer student achievement goals in each class, thereby enhancing student performance (Troia et al., 2018). The categorization of each competency achievement indicator is presented in Table 1.

Table 1. Categories of competency achievement indicators in mathematics curriculum

Theme	LOTS			HOTS		
	Remembering	**Understanding**	**Applying**	**Analyzing**	**Evaluating**	**Creating**
Algebra	2	7	51	16	4	4
Trigonometry	0	0	9	2	0	0
Geometry and Measurement	0	1	20	5	1	1
Calculus	0	0	43	17	4	1
Statistics and Probability	0	2	23	7	0	0
Total	**2**	**10**	**146**	**47**	**9**	**6**

The explanation of Table 1 highlights the differences in the cognitive level of indicators across each theme, categorized as either LOTS or HOTS. The research findings revealed that nearly half of the total competency achievement indicators were classified at the LOTS level, accounting for 71.82% (158 out of 220 indicators). The Algebra theme contained the highest number of LOTS indicators, with 60 indicators (27.27%), while the Trigonometry theme had the fewest, with only nine indicators (4.20%). Other themes at the LOTS level included Calculus with 43 indicators (19.60%), Statistics and Probability with 24 indicators (11.10%), and Geometry and Measurement with 21 indicators (9.65%).

Conversely, the researchers identified only 62 indicators (28.18%) at the HOTS level out of the 220 total indicators. The fewest HOTS indicators were found in the Trigonometry theme, with just two indicators, while the Algebra theme had the highest number, with 24 indicators. The remaining HOTS indicators were distributed across the other themes as follows: 22 indicators in Calculus, seven in Geometry and Measurement, and seven in Statistics and Probability. These findings are consistent with previous research, which suggests that the competencies embedded in curricula tend to be more concentrated at lower cognitive levels, with little to no emphasis on HOTS (El-Hassan & Baassiri, 2019).

The results of this study indicate that the majority of the Competency Achievement Indicators used by teachers are classified at the LOTS level, with fewer indicators at the HOTS level. This finding is further supported by interviews with one of the research participants, Teacher 3, who noted that many teachers still face challenges in developing competency achievement indicators from the mathematics curriculum based on cognitive levels. This is reflected in Teacher 3's statement, in which they expressed the view that "applying" is already considered a high cognitive level. The following interview transcript illustrates Teacher 3's perspective.

Transcript Dialog between Researcher (R) and Teacher 3 (T3)

R : *"When compiling competency achievement indicators, in your opinion, what skills are included in HOTS?"*
T3 : *"When students analyze, they continue to apply."*
R : *"Okay. So, when students apply a procedure, can they be said to have HOTS?"*
T3 : "*Yes, they can, in my opinion. It is because students don't just remember a material"*

The researchers also identified a limited number of indicators targeting the HOTS level. One example of an indicator at the HOTS level comes from Teacher 5, who developed an indicator based on the following basic competency: "Constructing a system of three-variable linear equations from contextual problems." This basic competency is part of the tenth-grade curriculum. The indicator, which focuses on the topic of a Three-Variable Linear Equation System, was classified under the Algebraic theme due to its alignment with the relevant basic competencies. In the indicator created by Teacher 5, students are presented with a contextual problem related to a system of three-variable linear equations, requiring them to analyze the mathematical model of the problem. The following is an example of a competency achievement indicator at the HOTS level.

Indicator:
"Given a contextual problem related to a system of three-variable linear equations, students can analyze the mathematical model of the problem" (Teacher 5)

This indicator demonstrates that the material used, the Three-Variable Linear Equation System, aligns with the Basic Competency from the Mathematics Curriculum. The indicator specifically asks students to analyze contextual problems and subsequently construct a mathematical model to represent the problem. As such, this indicator is classified as targeting the HOTS level, specifically the "analyzing" level, since it requires students to engage in the process of analysis by examining the available information and selecting the relevant data necessary for constructing a mathematical model and forming a system of equations.

Previous research supports this classification, as mathematical modeling activities necessitate students' ability to analyze and interpret the relationships between the various elements of a problem (Viseu et al., 2020). Furthermore, mathematical modeling involves solving real-world problems and interpreting the results within the context of mathematical concepts. The use of the Three-Variable Linear Equation System in this HOTS indicator is further supported by Teacher 5's statements, as reflected in the following interview transcript.

Transcript Dialog between Researcher (R) and Teacher 5 (T5)

R : *"In your opinion, what is material in your indicator? Please explain."*
T5 : *"The material is the system of linear equations of three variables and according to their basic competence."*
R : *"Then, can you explain, what level cognitive target of your indicator?"*
T3 : *(Reading the indicator). "Given a contextual problem related to a system of three-variable linear equations, students can analyze the mathematical model of the problem. The indicator uses operational verbs and asks students to analyze so that they are classified as HOTS".*

Teachers' Formative and Summative Assessments

Other findings pertain to the formative and summative assessments developed by the research participants based on the competency achievement indicators. On average, the fifteen teachers required between 2 to 3 weeks to prepare formative assessments, while summative assessments took approximately 1 to 3 months to prepare. These assessments were used to evaluate student learning outcomes in the classroom and to measure both knowledge and skill competencies.

Additionally, it was found that the research participants were able to construct one or more assessments from a single indicator. As a result, the number of competency achievement indicators and their associated assessments varied. For instance, Teacher 8 was able to create two distinct assessments from one indicator, which involved determining the image of a point and a line after a transformation. Then, the assessment is (1) If the line $3x - 2y = 6$ is translated to $T\ (2{,}3)$, then the image is…..; and (2) The point $(2, -9)$ is rotated $240°$ counterclockwise with the center of rotation $P\ (-3, -6)$ then it is followed by a reflection on the $x - axis$. The image from point T is ……. Therefore, the assessments collected from all participants totaled 223 assessments.

First, the researchers categorized each assessment according to the five themes of competency achievement indicators. Next, they analyzed each assessment in detail to determine the cognitive level it targeted. Like the process used for categorizing cognitive levels in competency achievement indicators, assessments were classified into LOTS (remembering, understanding, and applying) or HOTS (analyzing, evaluating, and creating), based on Anderson et al. (2001) and Krathwohl's Taxonomy. However, there were differences in the cognitive level analysis process for indicators versus assessments. While the categorization of indicators relied on operational verbs and the intended purpose of the indicators, the categorization of assessments also considered definitions, operational verbs, and the measurement objectives outlined in the teacher's assessments.

For example, Teacher 1's assessment, which asks, "The sum of the first five terms of a geometric series is 93. If the ratio of the series is 2, find the product of the second and sixth terms!" was classified at the "applying" level. This classification is based on two factors, such as the assessment requires students to apply routine procedures related to geometric series formulas (which corresponds to the definition of applying) and the operational verb "to determine" is used. Since students are required to use a specific procedure, this demonstrates that they are engaged in the "applying" stage (Bloom, 1956).

The findings show that the assessments across the five themes predominantly measured students' skills at the "applying" level, with 186 out of 223 assessments falling into this category. The Algebra theme contained the most assessments targeting the "applying" level, with 65 items, while the fewest were found in the Trigonometry theme, with only 8 items. Other themes included 27 items in Geometry and Measurement, 59 in Calculus, and 27 in Statistics and Probability.

In contrast, few assessments targeted HOTS. At the analyzing level, only 22 items assessed students' analytical skills, with the majority found in the Algebra theme (10 items). At the evaluating level, only nine items assessed students' evaluation skills. Furthermore, creative thinking skills were almost entirely absent, with the researchers identifying only one assessment at the "creating" level, which was in the Calculus theme.

These findings support previous research indicating that teachers often do not fully understand the concept of HOTS and experience difficulties in preparing high-level assessments (Widana, 2020). As a result, teachers are unable to develop assessments that assess students' skills in higher-order thinking, including analyzing, evaluating, and creating (Barak & Shakhman, 2008). The categorization of each teacher's assessment is presented in Table 2.

Table 2. Categories of teachers' assessments

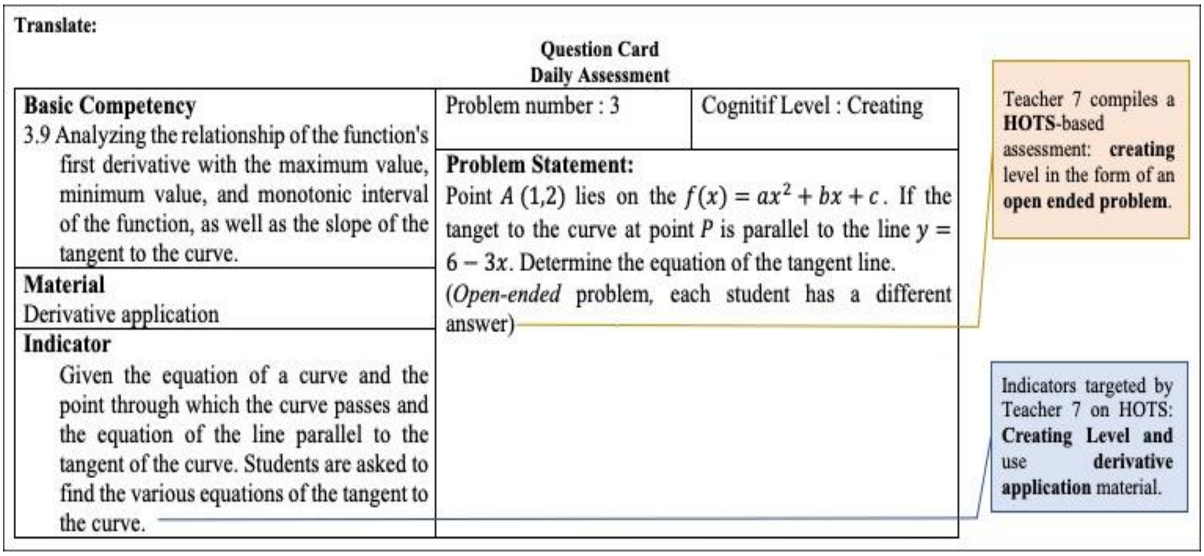
Theme	LOTS			HOTS		
	Remembering	Understanding	Applying	Analyzing	Evaluating	Creating
Algebra	0	3	65	10	4	0
Trigonometry	0	0	8	3	0	0
Geometry and Measurement	0	0	27	1	1	0
Calculus	0	0	59	6	4	1
Statistics and Probability	0	2	27	2	0	0
Total	**0**	**5**	**186**	**22**	**9**	**1**

Almost all assessments focused on evaluating students' LOTS, with 191 out of 223 assessments (85.65%) targeting this level. The majority of LOTS assessments were found in the Algebra theme, with 68 items (30.50%), while the fewest were in Trigonometry, with only 8 items (3.60%). The remaining LOTS assessments were distributed across the other themes: 27 items (12.10%) in Geometry and Measurement, 59 items (26.55%) in Calculus, and 29 items (13%) in Statistics and Probability.

Conversely, only 32 assessments (14.35%) measured students' HOTS. The largest proportion of HOTS assessments was found in Algebra, with 14 items (6.27%), followed by 11 items (4.93%) in Calculus, 3 items (1.35%) in Trigonometry, 2 items (0.90%) in Geometry and Measurement, and 2 items (0.90%) in Statistics and Probability. These findings suggest that teachers, when designing assessments, primarily focus on evaluating LOTS. This is consistent with previous research, which indicates that many assessments and textbooks predominantly feature low-level questions, with fewer questions requiring students to engage in HOTS (Davies et al., 2022). When teachers rely on routine questions in their assessments, students may be unprepared for HOTS-based assessments and may struggle with non-routine problems (Firmansyah et al., 2022; Abdullah et al., 2015). This highlights a gap in teachers' ability to develop assessments that effectively target higher cognitive skills such as analyzing, evaluating, and creating, despite the curriculum's emphasis on HOTS in learning. Therefore, it is crucial to implement professional development programs for teachers that focus on bridging this gap and providing them with practical strategies for creating HOTS-oriented assessments.

Our findings also revealed that HOTS assessments were rare, with only one assessment item at the "creating" level. This assessment, developed by Teacher 7, involved the application of the derivative of a function to find the slope of a tangent line to a curve. This assessment, classified under the Calculus theme and targeting the "creating" level of HOTS, serves as an example of how higher-level cognitive skills can be assessed. Teacher 7's assessment preparation is shown in Figure 2.

Figure 2 illustrates that Teacher 7 developed an indicator aligned with the basic competencies related to the first derivative of a function. Based on the curriculum achievement indicators, Teacher 7 created an open-ended assessment as follows.



Translate:

Question Card
Daily Assessment

Basic Competency 3.9 Analyzing the relationship of the function's first derivative with the maximum value, minimum value, and monotonic interval of the function, as well as the slope of the tangent to the curve.	Problem number : 3	Cognitif Level : Creating
Material Derivative application	Problem Statement: Point A (1,2) lies on the $f(x) = ax^2 + bx + c$. If the tanget to the curve at point P is parallel to the line $y = 6 - 3x$. Determine the equation of the tangent line. (*Open-ended* problem, each student has a different answer)	
Indicator Given the equation of a curve and the point through which the curve passes and the equation of the line parallel to the tangent of the curve. Students are asked to find the various equations of the tangent to the curve.		

Figure 2. Teachers' assessment at HOTS: Creating level

Indicator:

"Given the equation of a curve and the point through which the curve passes and the equation of the line parallel to the tangent of the curve. Students are asked to find the various equation of the tangent line to the curve"

Problem Statement:

"Point $A(1,2)$ lies on the $f(x) = ax^2 + bx + c$. If the tangent to the curve at point P is parallel to the line $y = 6 - 3x$. Determine the equation of the tangent line. (*Open-ended problem*, each student has a different answer)" **(Teacher 7)**

The mathematics assessment developed by Teacher 7 is an open-ended problem. Open-ended problems are designed to encourage diverse thinking and emphasize creativity, as they allow for multiple possible answers, thereby highlighting the originality of students' responses (Lin & Lien, 2013). Teacher 7's assessment focuses on the derivative of algebraic functions, with the highest degree being two. The assessment discusses the concept of the derivative of a function, namely the tangent line at a certain point, for example, $A\ (x_1, y_1)$. The concept of the derivative function used in this assessment is the gradient of the tangent line at point $A\ (x_1, y_1)$, which becomes the limit of the line's gradient at point $A\ (x_1, y_1)$ if the limit exists. On the other hand, the assessment also requires students to make connections between the concept of derivatives of algebraic functions and the concept of straight-line equations, specifically the equations of two parallel straight lines. This assessment falls under the HOTS level, specifically at the "creating" level. Teacher 7's assessment challenges students to generate multiple possible forms of the equation of a tangent to a curve. Before attempting the assessment, students must first correlate the concept of the derivative of an algebraic function—namely, the gradient of a line—with the concept of straight-line equations.

In the assessment, Teacher 7 stated point $A\ (x_1, y_1)$ lies on the curve $f(x) = ax^2 + bx + c$. Therefore, point $A\ (1, 2)$ where $x = 1$ and $y = f\ (x) = 2$ satisfies the equation $f(x) = ax^2 + bx + c$. Furthermore, the assessment also informs that the tangent to the curve is at point P where the point P is parallel to a line $y = 6 - 3x$. Based on this known information, students use their previous concepts related to straight-line equations, namely, if two lines are parallel to each other, they have the same gradient. For example, mp is the gradient of the tangent to the curve at point P, and my is the gradient of the line $y = 6 - 3x$. Students use the straight-line concept, namely: $y = mx + c$, where "m" is the gradient of the line. Students will get that $m_y = -3$. Thus, students get the gradient of the tangent to the curve at point P, namely $m_P = m_y = -3$. Based on the calculations, students will get two equations, namely equations (i) $2ax + b = -3$ and (ii) $a + b + c = 2$. In this section, the creativity of students' answers will appear. Students are required to think openly by choosing any value a, b, c, or x. In other words, students are creative in making their own various possible algebraic functions $f(x) = ax^2 + bx + c$ or point $P(x, y)$ that crosses the curve. After students determine the various possibilities $f(x)$ or point $P(x, y)$, students can create various equations of the tangent to the curve at that point.

Furthermore, providing open-ended problems to students can enhance their ability to find, acquire knowledge, and solve problems using diverse strategies, thereby fostering the development of creative thinking skills (Shahrill et al., 2018; Marsitin et al., 2022). This approach enables students to view problems from multiple perspectives, facilitating the generation of creative solutions (Hwang et al., 2007; Ahamad et al., 2018). Similar studies support the notion that open-ended problems offer students the opportunity to explore and discover various alternative solutions, which can stimulate their creativity (Brookhart, 2010; Munroe, 2015; Nadjafikhah et al., 2012). However, in practice, many teachers are still unable to develop assessments that cultivate higher-order thinking, despite the importance of teaching HOTS to students (Zana et al., 2022).

CONCLUSION

The findings of this study reveal two significant conclusions. First, the competency achievement indicators outlined in the K-13 Revised Edition curriculum predominantly target LOTS, including remembering, understanding, and applying, rather than HOTS. Among these, the skill of applying is the most frequently emphasized in the Indonesian mathematics curriculum. Given that the curriculum serves as a foundation for teachers in designing assessments, it is essential for the competency indicators to focus more on fostering HOTS. Second, the study highlights that the assessments developed by teachers are similarly aligned with the curriculum's emphasis on LOTS, particularly in areas such as Algebra, Statistics and Probability, Trigonometry, Geometry and Measurement, and Calculus. These assessments primarily evaluate students' applying skills, with limited focus on higher-order cognitive processes, especially the creating level. This underscores the rarity of assessments designed by mathematics teachers that aim to develop students' HOTS.

This research, however, has several limitations. First, it exclusively utilized the K-13 Revised Edition curriculum document. Employing other curriculum documents in Indonesia, such as the Merdeka Curriculum, could yield different findings concerning the cognitive level distribution. Additionally, the cognitive level analysis in this study was based on Anderson and Krathwohl's Taxonomy, which emphasizes thinking processes through verbs. An alternative approach using Bloom's Taxonomy, which focuses on thinking outcomes expressed as nouns, could result in different interpretations of cognitive level distributions in assessments and curricula. Furthermore, the study's sample size was limited to 15 teachers, which restricts the generalizability of the findings to the broader population of mathematics educators. Finally, the qualitative nature of the research, coupled with the reliance on self-reported assessment documents and interview responses, presents potential biases. These biases were mitigated through triangulation techniques, wherein researchers cross-verified data from multiple sources using different methods. For instance, semi-structured interviews with participants regarding the cognitive levels of their assessments were corroborated with the actual assessment documents provided by the participants.

Furthermore, the limitations encountered in this study offer valuable insights for future research, which could further refine the investigation of cognitive levels in teacher assessments. Future studies should consider incorporating classroom observations or field notes to verify the alignment of assessments with teaching practices and the intended goals of the curriculum. By addressing these limitations, subsequent research could also explore in greater depth the cognitive skills targeted in other Indonesian curriculum documents, utilizing thinking taxonomies such as Bloom's Taxonomy across a wider array of teacher assessments. This should not be confined solely to formative or summative assessments but could also encompass portfolio assessments, peer assessments, and other forms of evaluation.

Finally, the findings of this research have important implications for both teachers and curriculum developers. Teachers should take proactive steps to prepare assessments that promote HOTS, enabling students to engage with high-level cognitive tasks such as analyzing, evaluating, and creating. These efforts may include incorporating more open-ended questions, using contextual or real-world problems that require critical analysis and evaluation, and engaging in professional development programs to enhance their ability to design assessments that foster HOTS and align with the curriculum's objectives. Additionally, the results of this study provide valuable insights for teachers and prospective teachers on the application of thinking taxonomies in assessment design to meet the desired learning outcomes. For curriculum developers, the findings suggest that the current curriculum achievement indicators focus insufficiently on HOTS. Therefore, a review or enhancement of the curriculum is necessary to better support the development of students' HOTS.

Acknowledgments

This research was funded by the Program Magister menuju Doktor Sarjana Unggul (PMDSU), in collaboration with the Direktorat Riset, Teknologi, dan Pengabdian Masyarakat (DRTPM), under contract number 036/E5/PG.02.00.PL/2023, and Universitas Negeri Malang, under contract number 3.4.52/UN32.20.1/LT/2023.

Declarations

Author Contribution	:	FMZ: Conceptualization, Writing – original draft, Data curation, Methodology, Investigation. CS: Resources, Investigation, Writing – original draft, Supervision. S: Writing – original draft, Supervision, Investigation. LA: Formal analysis, Methodology, Visualization. HZ: Methodology, Validation, Writing – review & editing
Funding Statement	:	This research was funded by the Directorate of Research, Technology and Community Service for supporting and funding this research.
Conflict of Interest	:	The authors have no conflict of interest to declare.
Additional Information	:	Additional information is available for this paper.

REFERENCES

Abdullah, A. H., Abidin, N. L. Z., & Ali, M. (2015). Analysis of students' errors in solving Higher Order Thinking Skills (HOTS) problems for the topic of fraction. *Asian Social Science, 11*(21), 133–142. https://doi.org/10.5539/ass.v11n21p133

Adams, W. K., & Wieman, C. E. (2011). Development and validation of instruments to measure learning of expert-like thinking. *International Journal of Science Education, 33*(9), 1289–1312. https://doi.org/10.1080/09500693.2010.512369

Ahamad, S. N. S. H., Li, H. C., Shahrill, M., & Prahmana, R. C. I. (2018). Implementation of problem-based learning in geometry lessons. *Journal of Physics: Conference Series, 943*(1), 012008. http://dx.doi.org/10.1088/1742-6596/943/1/012008

Ali, L. (2018). The design of curriculum, assessment and evaluation in higher education with constructive alignment. *Journal of Education and E-Learning Research, 5*(1), 72–78. https://doi.org/10.20448/journal.509.2018.51.72.78

Amzaleg, M., & Masry-Herzallah, A. (2021). Cultural dimensions and skills in the 21st century: the Israeli education system as a case study. *Pedagogy, Culture and Society, 30*(5), 765–785. https://doi.org/10.1080/14681366.2021.1873170

Anderson, L. W., Krathwohl Peter W Airasian, D. R., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *Taxonomy for Assessing a Revision of Bloom's Taxonomy of Educational Objectives (Abridged)*. Addison Wesley Longman, Inc.

Anggraena, Y., Abdulhak, I., & Rusman, R. (2018). The development of mathematics curriculum to increase the higher order thinking skills in the 21st century era. *Proceedings of the 1st International Conference on Science and Technology for an Internet of Things*. https://doi.org/10.4108/eai.19-10-2018.2281302

As'ari, A. R., Mahmudi, A., & Nuerlaelah, E. (2017). Our prospective mathematic teachers are not critical thinkers yet. *Journal on Mathematics Education, 8*(2), 145–156. https://doi.org/10.22342/jme.8.2.3961.145-156

Barak, M., & Shakhman, L. (2008). Fostering higher-order thinking in science class: Teachers' reflections. *Teachers and Teaching: Theory and Practice, 14*(3), 191–208. https://doi.org/10.1080/13540600802006079

Berg, G. Van Den. (2004). The use of assessment in the development of higher-order thinking skills. *Africa Education Review, 1*(2), 279–294. https://doi.org/10.1080/18146620408566285

Berisha, V., & Bytyqi, R. (2020). Types of mathematical tasks used in secondary classroom instruction. *International Journal of Evaluation and Research in Education, 9*(3), 751–758. https://doi.org/10.11591/ijere.v9i3.20617

Bloom, B. S. (1956). Taxonomy of educational objectives: the classification of educational goals; Handbook I: Cognitive domain. In M. D. Engelhart, E. J. Furst, W. H. Hill, & D. R. Krathwohl (Eds). *Taxonomy of educational objectives: the classification of educational goals; Handbook I: Cognitive domain*. David McKay.

Brookhart, S. M. (2010). *How to Assess Higher-Order Thinking Skills in Your Classroom*. ASCD.

Busetto, L., Wick, W., & Gumbinger, C. (2020). How to use and assess qualitative research methods. *Neurological Research and Practice*, *2*(1), 14. https://doi.org/10.1186/s42466-020-00059-z

Byrne, M. (2001). Interviewing as a data collection method. *AORN Journal, 74*(2), 233–235. https://doi.org/10.1016/S0001-2092(06)61533-0

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Sage.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage.

Dahlan, D., Permana, L., & Oktariani, M. (2020). Teacher's competence and difficulties in constructing hots instruments in economics subject. *Cakrawala Pendidikan, 39*(1), 111–119. https://doi.org/10.21831/cp.v39i1.28869

Darwazeh, A. N., & Branch, R. M. (2015). A revision to the revised Bloom's taxonomy. *The Annual Proceedings, 2*(19), 2330–2335. https://doi.org/10.35445/alishlah.v15i2.2823

Davies, D. J., McLean, P. F., Kemp, P. R., Liddle, A. D., Morrell, M. J., Halse, O., Martin, N. M., & Sam, A. H. (2022). Assessment of factual recall and higher-order cognitive domains in an open-book medical school examination. *Advances in Health Sciences Education, 27*(1), 147–165. https://doi.org/10.1007/s10459-021-10076-5

El-Hassan, K., & Baassiri, M. (2019). The alignment between the fourth grade Lebanese science curriculum and classroom assessments. *Education 3-13, 47*(5), 612–624. https://doi.org/10.1080/03004279.2018.1515236

Firmansyah, F. F., Sa'dijah, C., Subanji, S., & Qohar, A. (2022). Characterizations of students' metacognition in solving geometry problems through positioning group work. *TEM Journal, 11*(3), 1391–1398. https://doi.org/10.18421/TEM113-50

FitzPatrick, B., & Schulz, H. (2015). Do curriculum outcomes and assessment activities in science encourage higher order thinking? *Canadian Journal of Science, Mathematics and Technology Education, 15*(2), 136–154. https://doi.org/10.1080/14926156.2015.1014074

Fulmer, G. W. (2011). Estimating critical values for strength of alignment among curriculum, assessments, and instruction. *Journal of Educational and Behavioral Statistics, 36*(3), 381–402. https://doi.org/10.3102/1076998610381397

Fulmer, G. W., & Polikoff, M. S. (2014). Tests of alignment among assessment, standards, and instruction using generalized linear model regression. *Educational Assessment, Evaluation and Accountability, 26*(3), 225–240. https://doi.org/10.1007/s11092-014-9196-z

Grant, L. W., & Gareis, C. R. (2013). Teacher-made assessments: How to connect curriculum, instruction, and student learning. In *Teacher-Made Assessments: How to Connect Curriculum, Instruction, and Student Learning* (2nd ed.). Eye on Education. https://doi.org/10.4324/9781315855240

Hadzhikoleva, S., Hadzhikolev, E., & Kasakliev, N. (2019). Using peer assessment to enhance higher order thinking skills. *TEM Journal, 8*(1), 242–247. https://doi.org/10.18421/TEM81-34

Hwang, W.-Y., Chen, N.-S., Dung, J.-J., & Yang, Y.-L. (2007). Multiple representation skills and creativity effects on mathematical problem solving using a multimedia whiteboard system. *Journal of*

Educational Technology & Society, 10(2), 191–212. http://www.jstor.org/stable/jeductechsoci.10.2.191

Ichwanto, M. A., Al-ansyorie, M. M., Isnandar, I., Alfianto, I., & Weifen, D. (2024). Development of TVET Teacher Professionalism in Indonesian Rural Area. *Proceedings of 5th Vocational Education International Conference*. https://doi.org/10.2991/978-2-38476-198-2_32

Istiyono, E. (2018). IT-based HOTS assessment on physics learning as the 21st century demand at senior high schools: Expectation and reality. *AIP Conference Proceedings*, *2014*(1), 020014. https://doi.org/10.1063/1.5054418

Jailani, J., Sugiman, S., & Apino, E. (2017). Implementing the problem-based learning in order to improve the students' HOTS and characters. *Jurnal Riset Pendidikan Matematika, 4*(2), 247–259. https://doi.org/10.21831/jrpm.v4i2.17674

Kusuma, D. M., Rosidin, U., & Suyatna, A. (2017). The development of higher order thinking skill (hots) instrument assessment in physics study. *IOSR Journal of Research & Method in Education (IOSR-JRME), 7*(1), 2320–2737. https://doi.org/10.9790/7388-070103XXXX

Lin, W. L., & Lien, Y. W. (2013). Exploration of the relationships between retrieval-induced forgetting effects with open-ended versus closed-ended creative problem solving. *Thinking Skills and Creativity, 10*, 40–49. https://doi.org/10.1016/j.tsc.2013.05.004

Lugosi, E., & Uribe, G. (2022). Active learning strategies with positive effects on students' achievements in undergraduate mathematics education. *International Journal of Mathematical Education in Science and Technology, 53*(2), 403–424. https://doi.org/10.1080/0020739X.2020.1773555

Madina, S., & Kardena, A. (2021). The difficulties in implementing scientific approach based on 2013 curriculum: A case study on one english teacher. *Elsya: Journal of English Language Studies, 3*(2), 126–133. https://doi.org/10.31849/elsya.v3i2.5794

Marsh, C. (2016). *Marsh's becoming a teacher* (N. Poole, K. Millar, & J. Tyrrell, Eds.; 6th ed.). Pearson Australia.

Marsitin, R., Sa'dijah, C., Susiswo, S., & Chandra, T. D. (2022). Creative mathematical reasoning process of climber students in solving higher order thinking skills geometry problems. *TEM Journal, 11*(4), 1877–1886. https://doi.org/10.18421/TEM114-56

McHugh, L., Kelly, A. M., Fisher, J. H., & David Burghardt, M. (2021). Graphing as a means to improve middle school science learning and mathematics-related affective domains. *Research in Science Education, 51*(2), 301–323. https://doi.org/10.1007/s11165-018-9796-6

Miles, M. B., Huberman, M. A., & Saldana, J. (2018). *Qualitative data analysis: A methods sourcebook* (4th ed.). SAGE Publications.

Munroe, L. (2015). The open-ended approach framework. *European Journal of Educational Research, 4*(3), 97–104. https://doi.org/10.12973/eu-jer.4.3.97

Murtafiah, W., Sa'dijah, C., Chandra, T. D., & Susiswo, S. (2019). Decision making of the winner of the national student creativity program in designing ICT-based learning media. *TEM Journal, 8*(3), 1039–1045. https://doi.org/10.18421/TEM83-49

Murtafiah, W., Sa'dijah, C., Chandra, T. D., & Susiswo. (2020). Exploring the types of problems task by mathematics teacher to develop students' HOTS. *AIP Conference Proceedings, 2215*(1), 1–7. https://doi.org/10.1063/5.0000656

Musa, H. M., & Yamat, H. (2021). Exploring primary school teachers' challenges in incorporating higher order thinking skills (hots) in English classroom. *International Journal of Academic Research in Progressive Education and Development, 10*(2), 665–1045. https://doi.org/10.6007/ijarped/v10-i2/9970

Nadjafikhah, M., Yaftian, N., & Bakhshalizadeh, S. (2012). Mathematical creativity: Some definitions and characteristics. *Procedia - Social and Behavioral Sciences, 31*, 285–291. https://doi.org/10.1016/j.sbspro.2011.12.056

National Council of Teachers of Mathematics (NCTM). (1989). *Curriculum and evaluation standards for school mathematics*. NCTM,Inc.

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and Standards for School Mathematics*. NCTM,Inc.

National Council of Teachers of Mathematics (NCTM). (2014). *Principles to actions: Ensuring mathematics success for all*. NCTM,Inc.

Olivares, D., Lupiáñez, J. L., & Segovia, I. (2021). Roles and characteristics of problem solving in the mathematics curriculum: A review. *International Journal of Mathematical Education in Science and Technology, 52*(7), 1079–1096. https://doi.org/10.1080/0020739X.2020.1738579

Podkhodova, N., Snegurova, V., Stefanova, N., Triapitsyna, A., & Pisareva, S. (2020). Assessment of mathematics teachers' professional competence. *Journal on Mathematics Education, 11*(3), 477–500. https://doi.org/10.22342/jme.11.3.11848.477-500

Pratiwi, N., & Mustadi, A. (2021). Hots-based learning in 2013 curriculum: Is it suitable?. *JPI (Jurnal Pendidikan Indonesia), 10*(1), 128–135. https://doi.org/10.23887/jpi-undiksha.v10i1.22781

Pulungan, M., Toybah, T., & Suganda, V. A. (2021). Development of HOTS-based 2013 curriculum assessment instruments in elementary school. *Journal of Teaching and Learning in Elementary Education (JTLEE), 4*(1), 51–64. https://doi.org/10.33578/jtlee.v4i1.7858

Retnawati, H., Hadi, S., & Chandra Nugraha, A. (2016). Vocational high school teachers' difficulties in implementing the assessment in curriculum 2013 in Yogyakarta province of Indonesia. *International Journal of Instruction, 9*(1), 33–48. https://doi.org/10.12973/iji.2016.914a

Rizki, R. N., Mustadi, A., & Wangid, M. N. (2022). Evaluation of the implementation of assessment in higher order thinking skills oriented learning 2013 curriculum in elementary schools. *AL-ISHLAH: Jurnal Pendidikan, 14*(3), 4363–4370. https://doi.org/10.35445/alishlah.v14i3.757

Sa'dijah, C., Murtafiah, W., Anwar, L., Nurhakiki, R., & Cahyowati, E. T. D. (2021). Teaching higher-order thinking skills in mathematics classrooms: Gender differences. *Journal on Mathematics Education, 12*(1), 159–180. https://doi.org/10.22342/jme.12.1.13087.159-180

Sa'dijah, C., Qohar, Abd., & Anwar, L. (2015). *Asesmen dalam pembelajaran matematika yang mendukung sikap, pengetahuan, dan keterampilan matematika* [Assessment in mathematics learning that supports mathematical attitudes, knowledge, and skills]. UM Penerbit & Percetakan.

Schulz, H., & Fitzpatrick, B. (2016). Teachers' understandings of critical and higher order thinking and what this means for their teaching and assessments. *Alberta Journal of Educational Research, 62*(1), 61–86. https://doi.org/10.11575/ajer.v62i1.56168

Seman, S. C., Yusoff, W. M. W., & Embong, R. (2017). Teachers challenges in teaching and learning for higher order thinking skills (hots) in primary school. *International Journal of Asian Social Science, 7*(7), 534–545. https://doi.org/10.18488/journal.1.2017.77.534.545

Shahrill, M., Putri, R. I. I., Zulkardi, & Prahmana, R. C. I. (2018). Processes involved in solving mathematical problems. *AIP Conference Proceedings, 1952*(1), 020019. https://doi.org/10.1063/1.5031981

Sudaryono, Rahardja, U., Aini, Q., Isma Graha, Y., & Lutfiani, N. (2019). Validity of test instruments. *Journal of Physics: Conference Series, 1364*(1), 012050. https://doi.org/10.1088/1742-6596/1364/1/012050

Suherman, S., & Vidákovich, T. (2022). Assessment of mathematical creative thinking: A systematic review. *Thinking Skills and Creativity, 44*, 1–13. https://doi.org/10.1016/j.tsc.2022.101019

Suwarno, R., Randall, E. V., & Hite, J. M. (2019). Alignment of classroom instruction with Indonesian national standards. *Indonesian Research Journal in Education |IRJE|, 3*(1), 6–28. https://doi.org/10.22437/irje.v3i1.5611

Suyanto, S. (2018). The implementation of the scientific approach through 5Ms of the revised curriculum 2013 in Indonesia. *Cakrawala Pendidikan, 37*(1), 22–29. https://doi.org/10.21831/cp.v37i1.18719

Syahida, H., & Dewi, L. (2023). Identifying 21st-century skills in the "Kurikulum Merdeka" at the elementary level numeracy aspect: A literature review. *International Conference on Elementary Education, 5*(1), 308–314. http://proceedings.upi.edu/index.php/icee/article/view/3119

Thompson, T. (2008). Mathematics teachers' interpretation of higher-order thinking in Bloom's taxonomy. *International Electronic Journal of Mathematics Education, 3*(2), 96–109. https://doi.org/10.29333/iejme/221

Thornhill-Miller, B., Camarda, A., Mercier, M., Burkhardt, J. M., Morisseau, T., Bourgeois-Bougrine, S., Vinchon, F., El-Hayek, S., Augereau-Landais, M., Mourey, F., Feybesse, C., Sundquist, D., & Lubart, T. (2023). Creativity, critical thinking, communication, and collaboration: Assessment, certification, and promotion of 21st century skills for the future of work and education. *Journal of Intelligence, 11*(3), 1–32. https://doi.org/10.3390/jintelligence11030054

Troia, G. A., Olinghouse, N. G., Zhang, M., Wilson, J., Stewart, K. A., Mo, Y., & Hawkins, L. (2018). Content and alignment of state writing standards and assessments as predictors of student writing achievement: An analysis of 2007 National Assessment of Educational Progress data. *Reading and Writing, 31*(4), 835–864. https://doi.org/10.1007/s11145-017-9816-3

Verdina, R., Gani, A., & Sulastri. (2018). Improving students' higher order thinking skills in thermochemistry concept using worksheets based on 2013 curriculum. *Journal of Physics: Conference Series, 1088*(1), 012105. https://doi.org/10.1088/1742-6596/1088/1/012105

Viseu, F., Martins, P. M., & Leite, L. (2020). Prospective primary school teachers' activities when dealing with mathematics modelling tasks. *Journal on Mathematics Education, 11*(2), 301–318. https://doi.org/10.22342/jme.11.2.7946.301-318

Widana, W. I. (2020). The effect of digital literacy on the ability of teachers to develop hots-based assessment. *Journal of Physics: Conference Series, 1503*(1), 012045. https://doi.org/10.1088/1742-6596/1503/1/012045

Wilcox, S. K., & Lanier, P. E. (2000). *Using assessment to reshape mathematics teaching: A casebook for teachers and teacher educators, curriculum and staff development specialists*. Routledge.

Yeung, S. yin S. (2015). Conception of teaching higher order thinking: Perspectives of Chinese teachers in Hong Kong. *Curriculum Journal, 26*(4), 553–578. https://doi.org/10.1080/09585176.2015.1053818

Zana, F. M., Sa'dijah, C., & Susiswo, S. (2022). LOTS to HOTS: How do mathematics teachers improve students' higher-order thinking skills in the class? *International Journal of Trends in Mathematics Education Research, 5*(3), 251–260. https://doi.org/doi.org/10.33122/ijtmer.v5i3.143

Zana, F. M., Sa'dijah, C., & Susiswo. (2024a). Mathematics teachers' obstacles in aligning cognitive levels of the assessment and curriculum to develop HOTS for high school students. *AIP Conference Proceedings, 3049*(1), 030026. https://doi.org/10.1063/5.0195402

Zana, F. M., Sa'dijah, C., & Susiswo. (2024b). The cognitive alignment of mathematics teachers' assessments and its curriculum. *International Journal of Evaluation and Research in Education, 13*(3), 1561–1575. https://doi.org/10.11591/ijere.v13i3.26814

Zulfikar, T. (2018). The making of Indonesian education: An overview on empowering Indonesian teachers. *Journal of Indonesian Social Sciences and Humanities, 2*, 13–39. https://doi.org/10.14203/jissh.v2i0.19