MALAYSIAN JOURNAL OF LEARNING AND INSTRUCTION
https://e-journal.uum.edu.my/index.php/mjli

How to cite this article:
Kania, N., & Kusumah, Y.S. (2025). The measurement of higher-order thinking skills: A systematic literature review. *Malaysian Journal of Learning and Instruction, 22*(1), 97-116. https://doi.org/10.32890/mjli2025.22.1.6

THE MEASUREMENT OF HIGHER-ORDER THINKING SKILLS: A SYSTEMATIC LITERATURE REVIEW

[1]Nia Kania & [2]Yaya Sukjaya Kusumah
[1,2]Faculty of Mathematics and Natural Sciences Education, Universitas Pendidikan Indonesia
[1]Mathematics Education Study Program, Universitas Majalengka, Indonesia

[1]Corresponding author: niakania@upi.edu

Received: 30/8/2023 Revised: 7/11/2023 Accepted: 7/9/2024 Published: 31/1/2025

ABSTRACT

Purpose - This article comprehensively examines the existing body of literature about the assessment tools employed for evaluating higher-order thinking skills (HOTS) in mathematics within the last five years. The study endeavors to offer a comprehensive overview of measuring instruments for HOTS in mathematics, utilizing precise and up-to-date global data.

Methodology - The research adheres to the PRISMA methodology, employing the systematic literature review (SLR) technique, encompassing ten distinct steps. By utilising the Scopus database, this review comprehensively analyses 18 scholarly articles written in English.

Findings - The results emphasize the significance of implementing a thorough evaluation approach that integrates standardized exams alongside other assessment methods to measure students' cognitive capacities in advanced mathematical reasoning precisely. The tendencies that have been observed provide valuable insights for the development of evaluation tools that can effectively measure effectiveness in many academic domains. Educators can augment instructional efficacy and foster heightened levels of cognitive reasoning in students by comprehending the equilibrium between evaluative and non-evaluative elements, while recognizing the significance of Bloom's taxonomy. Furthermore, this research presents opportunities for additional investigation, enabling scholars to delve into novel approaches and metrics, thus contributing to breakthroughs in education and improving student proficiencies in critical thinking and problem-solving.

Significance – This research provides the opportunity to enhance our comprehension of incorporating higher-order thinking skills in education, thereby promoting holistic cognitive growth.

Keywords: Cognitive domain of bloom's taxonomy, higher-order thinking skill, measurement, systematic review, thinking skill.

INTRODUCTION

The advancement of a country is inherently interconnected with the level of educational achievement among its populace. The claim presented is substantiated by many studies that underscore the significant importance attributed to formal and informal education (Cáceres-Reche et al., 2022; Susilo et al., 2023). These routes provide individuals with essential attributes (Slabko et al., 2019) and equip them with crucial competencies (Charungkaittikul et al., 2022). Cognitive skills and learning processes are interdependent (Kania & Juandi, 2023). The influence of mental talents, such as flexibility, problem-solving, processing speed, spatial orientation, and visuospatial acumen, on mathematical competency has been observed in previous research conducted (Mascia et al., 2018). Furthermore, developing higher-order thinking skills (HOTS) has become crucial in enhancing student academic performance and extracurricular accomplishments, a perspective strongly substantiated by empirical studies (Huang et al., 2023; Wahidin & Romli, 2020).

The demand for HOTS, which includes critical thinking, creativity, communication, and teamwork, is currently high (Husain, 2023; Zhou et al., 2023). Proficiency in HOTS is considered a crucial factor for mathematics student employability (Yigletu et al., 2023). This proficiency is also essential for addressing the increasing challenges in diverse industries (Boadu et al., 2020; Massetor et al., 2021). Therefore, it is evident that the importance of HOTS cannot be overstated since it plays a crucial role in developing a graduate's preparedness for the workforce.

Suanto et al. (2023) assert that HOTS have a significant impact on the process of education and learning. The cultivation of analytical thinking and problem-solving abilities, as highlighted (Wahono et al., 2020), is considered pivotal in achieving academic achievement (Gavronskaya et al., 2022; Phurikultong & Kantathanawat, 2022). The National Research Council (NRC) and the National Council of Teachers of Mathematics (NCTM) have emphasised the importance of HOTS in driving educational advancement, with a specific focus on their role in improving mathematics instruction (Abdurrahman et al., 2021). Research findings suggest that placing a higher emphasis on HOTS within the mathematics curriculum leads to favourable results, as seen by pupils exhibiting significant enhancements in their problem-solving and critical-thinking abilities (Devlin, 2011).

There is a need to expand research endeavors and explore literature about the assessment tools used to measure HOTS in mathematics education. It allows SLRs to map the condition of HOTS measuring instruments. A systematic review of HOTS-measuring mathematics learning studies published in the previous five years will provide a framework for future research. The SLRs which are currently being carried out have been about reviewing research subjects and studying the topics of research, types of instruments, characteristics of HOTS measuring instruments in mathematics learning, HOTS variables measured in mathematics lessons, research methods, and the outcomes of such research. In the context of the present article, the teaching of mathematics in schools is the focus of the current SLR.

NCTM (2021) describes the vital role for large-scale mathematics assessments in schools and districts as identifying the systemic deficiencies and successes in addressing the needs of students. Assessment is one of the fundamentals of a mathematics education, as it gives teachers and students essential data to support the possible advantages of studying the subject. The main goal of assessment is to improve learning during the learning process, as opposed to ranking, judging, or grading students (Popham, 2008). Student comprehension of mathematical ideas and skills should be continually developed throughout the teaching and learning process. The overhaul of the math curriculum emphasizes cognitive components in its teaching and learning evaluations that involve advanced thinking abilities. It is crucial to create mathematical problems that embody HOTS as provocations to enhance student cognitive capacities, since these problems enable students to straightaway apply HOTS. Additionally, everyone working in an educational setting needs HOTS.

Besides, there have been several systematic literature review studies on assessment. In their meta-analysis, Maison et al. (2022) investigated mathematical modelling. Nakakoji and Wilson (2020) and Wilson and Narasuman (2020) reviewed implementation and strategy as a higher-order thinking skill.

The present SLR which has included 18 articles after the final selection, used a different methodology that had a more focused objective. In addition, it had used a subjective study to cross-validate the results of the SLR (i.e., using experts' feedback). In light of the subjective survey results, the present SLR has helped in developing innovative concepts for future HOTS measurement.

METHODOLOGY

Research Design

The present study was aimed at investigating HOTS assessments thoroughly. It has used the Systematic Literature Review (SLR) technique to explore existing research on frameworks that included the following objectives: identifying, evaluating, and interpreting the development of HOTS (Ali et al., 2022; Budgen & Brereton, 2006). During the review phase of the SLR, there was stringent adherence to regulations (Budgen & Brereton, 2006; Khan et al., 2022). The study evaluated HOTS literature in mathematics education from 2018 to 2022. It had taken note of the advancements that had occurred throughout the past five years. Figure 1 illustrates the comprehensive research process undertaken in this SLR, which is divided into sections and subsections that rely on multiple sources (Brereton et al., 2007; Chambers et al., 2009; J. P. . Higgins & Green, 2006; S. Higgins, 2017; Paiva et al., 2014).

Figure 1

The Process of a Systematic Literature Review

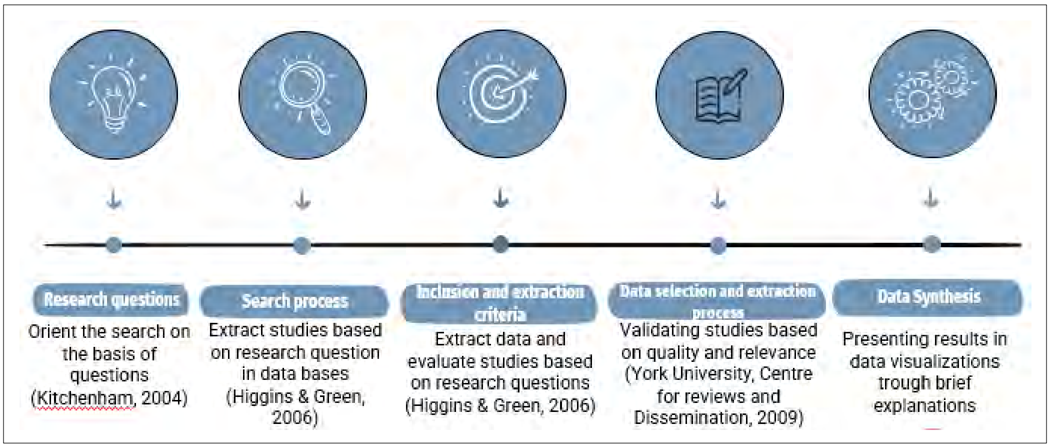


Figure 1 presents the process of a thorough and systematic analysis of the current body of literature. Similarly, this study has gathered and verified the pertinent information about the topic of concern. The graphic provides a comprehensive and logical representation of the research environment, highlighting important discoveries and concepts from a systematic literature study. The researcher's objective is to enhance the credibility and dependability of the research for the benefit of scholars, educators, and other stakeholders with a vested interest in the topic.

Phase 1: Research questions

Strategically crafted inquiries are employed to efficiently accomplish the primary objective of comprehending the issue. This stage aims to discover certain literary traits that can offer valuable insights into the research subject as is presented in Table 1.

Table 1

Research Questions

Dimensions	Research Questions (RQ)
Core components: Measurement of HOTS	What type of Instruments were used? What were the variables of concern in the HOTS study? What were the indicators of the variables in the HOTS study?

The present SLR has assessed the effectiveness of HOTS by utilizing three primary inquiry points through its research questions (RQ). RQ 1 was aimed at determining the methodologies researchers had employed in evaluating cognitive talents at an advanced level. RQ2 explored the HOTS highlighted in the research. RQ3 was aimed at identifying the HOTS indicators used in previous studies. This could help facilitate the comprehension of the intricacy of high cognitive skills among educators and researchers, and underscore the significance of cognitive capabilities in mathematics.

Phase 2: The search process

The second part of the present SLR involved a comprehensive search of the SCOPUS database. To enhance search results, choosing the appropriate keywords for each search engine and utilizing the AND operator is necessary. It is recommended that one consults Table 2 to understand the critical terms used in this inquiry, which had significantly facilitated the meticulous analysis of the search method and findings.

Table 2

The Search String

Scopus
((“Higher-order Thinking Skills” OR “HOTS” OR “Higher Order Thinking Skills”) AND (“Measurement” OR “Instrument” OR “Test” OR “Assessment”) AND (“mathematics education” OR “mathematics” OR “mathematics teaching”))

The comparative advantage of Scopus compared to its competitors was a crucial aspect that had influenced the decision of this SLR to choose Scopus over the other available databases. Additionally, the research emphasized quality control methods and well-structured indexing systems that could enhance the advantages of utilizing Scopus as a research instrument.

This SLR deliberately concentrated its search efforts on the SCOPUS database, citing the extensive capabilities and enhanced dependability of the database in producing accurate outcomes. Scopus and similar databases could provide numerous benefits regarding extensive search functionality, heightened result reliability, and improved search capabilities (Gusenbauer & Haddaway, 2021). While considering the accessibility factor, it is worth mentioning that other prominent databases like the Web of Science (WoS) was not utilized in this study. In this instance, the authors could not utilize the WoS due to accessibility limitations, which had impacted the decision regarding its database selection.

Phase 3: Inclusion and extraction requirements

In this phase, methodological techniques were created to efficiently choose and evaluate pertinent studies to establish a comprehensive framework that included search criteria and procedures. The requirements encompassed a range of components, including identifying suitable keywords, choosing specific databases and search engines, applying inclusion and exclusion criteria, and developing a technique for organizing and evaluating the gathered information.

Table 3

The Criteria for Inclusion and Exclusion of Search Items

Criteria	Included	Excluded
Timeline of the publication	2018-2022	2017 and before
Type of document	Article	Books, chapter books, proceedings, etc
Type of source	journal	Non-journal
Language	English	Non-English
The topic of the research	Focus on mathematics education	Not focus on mathematics education

Moreover, these criteria offer explicit guidelines for extracting, analyzing, and synthesizing data, ensuring a uniform and transparent approach in research endeavors. By implementing rigorous search procedures and adhering to established policies, this SLR had been able to enhance its ability to locate, assess, and include scholarly publications of superior quality. These criteria have been utilized in the present SLR to acquire a comprehensive description of the equipment employed in HOTS. This will facilitate the development of appropriate assessment instruments for HOTS in the future.

Phase 4: Data selection and extraction process

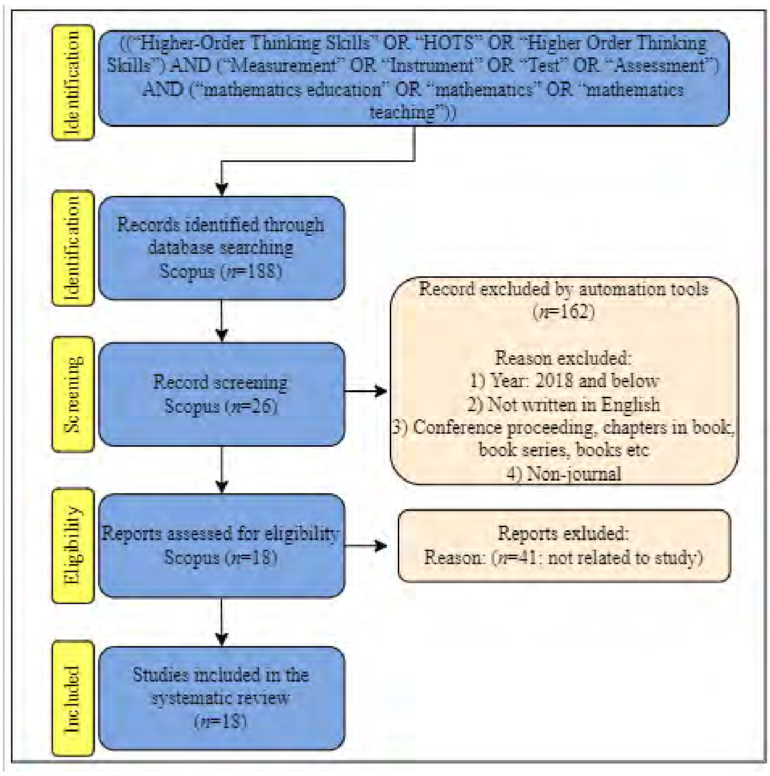
During the fourth phase, a thorough search for articles was conducted, and the relevant data was retrieved. Subsequently, Microsoft Excel databases were established to store the data. Several studies utilizing the HOTS measurement tool in mathematics education yielded the required data. Subsequently, 188 papers that were freely available for access were evaluated regarding their quality and emphasis on HOTS. Verifying the document's title and abstract is the initial stage. During the eligibility examination, 162 papers were rejected due to their lack of relevance. If the title and abstract were insufficient, the methodology, results, and discussion sections were thoroughly scrutinized. Eight publications were removed due to their lack of emphasis on the relevant evaluation strategies for HOTS in mathematics education. Over the past five years, research has directed the search for assessment methodologies that evaluate HOTS in mathematics education. This search has been aimed at exploring the originality of the research within a specific framework. Based on the procedures above, 18 articles were further examined to assess their quality. Figure 2 illustrates the application of PRISMA in establishing these boundaries.

Phase 5: Data synthesis

The primary goal in the fifth phase of the categorization procedure was to pinpoint relevant research in the SCOPUS database that explicitly focused on HOTS in mathematics education. To guarantee precision and comprehensiveness, each publication's abstract, keywords, and title underwent a meticulous verification process. Figure 2 depicts the application of the Systematic Review and Meta-study (PRISMA) protocol for conducting an exhaustive study of every research from 2018 to 2022. This assessment was then utilized to amalgamate data from chosen studies. The findings have been utilized to provide methodical reports and observations.

Figure 2

A Flow Chart Illustrating the Detailed Application of PRISMA 2022 to Studies that were Published Between 2018 and 2022



Note. the spelling of Excluded in the Chart

A SLR is an essential undertaking that involves formulating acceptable research inquiries and employing a systematic and transparent technique to locate, choose, assess, and acquire data from prior relevant research for further analysis. The PICO Bilondi et al. (2024) framework, which encompasses the factors of population, intervention, comparison, and outcome, aids researchers in creating pertinent and suitable questions for systematic reviews. The present SLR has incorporated the following three essential components, namely identification, screening, and feasibility evaluation, into its systematic search technique of the literature review, using these notions as a theoretical framework.

Phase 6: Data extraction and analysis

The third phase was extracting high-quality data from previously evaluated publications, with a specific emphasis on crucial elements present in each work, namely the abstract, findings, and discussion sections. The data that has been extracted was then arranged in Table 4, improving its accessibility and enabling the application of comprehensive analytical methods. After the extraction of the data, a thorough analysis was conducted. The rationale behind opting for an SLR stems from its ability to incorporate several research methodologies, with qualitative synthesis being the preferred analytical method. Thematic analysis is a highly regarded method for a qualitative synthesis that is beneficial in examining data obtained from many research methodologies. This approach facilitates the identification of patterns by recognizing similarities and correlations among the collected observations. Within various analytical methods, thematic analysis emerges as a particularly productive instrument for qualitative synthesis. Subsequently, codes were developed to analyze the content of the 18 articles.

RESULTS

General characteristics of the research

After evaluating their eligibility, 18 papers were chosen to be analyzed systematically. This study employed three research questions to analyze the specific scholarly papers directly.

Table 4

Summary of Studies Included in the Systematic Literature Review

Authors and Year	RQ1	RQ2
Abdullah (2020)	Cognitive Domain of Bloom's Taxonomy	Tests, Rubrics, and Interviews
Yayuk & Husamah (2020)	Problem-solving Skills	Test, Interview, and Observation
Koçak (2020)	Problem-solving Skills	Tests, general evaluations, rating scales, and rubrics
Ibrahim et al. (2020)	Cognitive Domain of Bloom's Taxonomy	Test
Azid et al. (2022)	Cognitive Domain of Bloom's Taxonomy	Tests and interviews
Ismail et al. (2022)	Critical thinking skills	Questionnaire
Abdullah et al. (2019)	Cognitive Domain of Bloom's Taxonomy	Interviews
Ansari et al. (2021)	Solving the problem	Tests and interviews
Moyo et al. (2022)	Cognitive Domain of Bloom's Taxonomy	Tests and interviews
Hidajat (2021)	Creative thinking skills	Tests, observations, and interviews.
Ramlee et al. (2019)	Cognitive Domain of Bloom's Taxonomy	Test
Julius et al. (2018)	Cognitive Domain of Bloom's Taxonomy	Interview, rubric, and tests
MD-Ali & Kim (2018)	Cognitive Domain of Bloom's Taxonomy	Questionnaire, interviews, video recording, observation, and field notes
Tanujaya et al. (2021)	Cognitive Domain of Bloom's Taxonomy	Interview
Ismail et al. (2019)	Critical thinking skills	Questionnaires
Abdurrahman et al. (2021)	Cognitive Domain of Bloom's Taxonomy	Questions and worksheets
Tajudin et al. (2018)	Six Principles of HOTS in Teaching and Learning	Survey

This section offers a concise overview of the critical research characteristics addressed in this study, such as the year the study was published, the types of HOTS in mathematics education, and the research tools used. Table 4 provides a concise overview of the components of the literature review. This compilation of articles all dealt with the researches which were conducted in 10 different nations. It included the analysis and comparison of certain scientific papers, using specific inclusion criteria, to examine HOTS variables, measurement tools, and indicators. The subsequent subsections will delve into each of the topics covered.

RQ1: What type of instruments were used?

A comprehensive examination of the analyzed data was conducted to identify an appropriate topic for this study. This technique generates three distinct themes, specifically as follows: (1) test instruments, (2) non-test instruments, and (3) test and non-test instruments. The distinction between tests and assessments primarily resides in their extent, goals, and technique. Tests are specialized instruments utilized to assess specific qualities or competencies and are predominantly quantitative (Yeager et al., 2015). On the other hand, a non-test is a more comprehensive procedure that incorporates several qualitative and quantitative data sources to generate an expert evaluation of an individual's capabilities, performance, or growth (Zumbo & Hubley, 2016).

Furthermore, the results derived from these overarching themes were further analyzed into sub-subthemes. As a consequence of this methodology, three distinct categories of instruments were identified. The research topic presented involves a comprehensive and thorough examination of many themes and types of instruments since they were directly relevant to the subject of study.

Table 5

Types of Instruments Used in Previous Research

RQ1	Test	Non-Test	Tests and non-tests
What types of instruments are used?	Ibrahim et al. (2020a); Ramlee et al. (2019)	Ismail et al. (2019); Tajudin et al., (2018); Tanujaya et al. (2021)	Abdullah (2020); Abdullah (2021); Ansari et al. (2021); Azid et al. (2022); Hidajat (2021); Julius et al. (2018); Koçak (2020); MD-Ali & Kim (2018); Yayuk & Husamah (2020)

RQ2: What HOTS Variables of Mathematics Education were Used in the Articles?

Reviewing the empirical literature on the HOTS construct in mathematics education, it can be seen that that only a few studies have specifically examined how to test students' mathematical proficiency. Remarkably, many of these publications primarily focused on assessing HOTS as a single construct, and have neglected to thoroughly investigate its numerous characteristics and possible intricacies within mathematical cognition.

Table 6

Variables of HOTS Used in Previous Studies

RQ2: HOTS Variables of Mathematics Education	Critical thinking skills	Problem-solving skills	Creative thinking skills	Cognitive Domain of Bloom's Taxonomy	Principles of HOTS
Studies	Erdogan (2019); Ismail et al. (2019); Ismail et al. (2022)	Abdullah et al. (2019); Ansari et al. (2021); Koçak (2020); Yayuk & Husamah (2020)	Hidajat (2021)	Abdullah (2020); Azid et al. (2022); Ibrahim et al. (2020); Julius et al. (2018); MD-Ali & Kim (2018); Ramlee et al. (2019); Tanujaya et al. (2021)	Tajudin et al. (2018)

Despite the extensive body of research on HOTS in mathematics education, it is apparent that there is a notable deficiency in conducting thorough investigations and assessments of mathematical proficiency as a fundamental element of HOTS. The findings demonstrate the need for additional studies and the need to develop evaluation tools to gauge the various aspects of mathematical thinking covered by the HOTS framework. Scholars can learn more about the importance of advanced mathematical cognition in the classroom by filling this gap and further explore the complexities of mathematical aptitude in HOTS.

RQ3: What were the indicators of HOTS variables used in the articles?

In RQ3, indicators associated with the HOTS variables were investigated. The significance of employing instruments to evaluate the efficacy of an activity should not be overlooked. These profound discoveries can significantly aid educators and curriculum designers in modifying teaching techniques and learning experiences that foster the development of HOTS. The insufficient consideration given to assessing students' higher-order cognitive skills in mathematics education raises questions about how educators and researchers utilize dependable evaluation instruments to measure the complexities of HOTS accurately.

Table 7 provides a comprehensive overview of the various indicators employed in different research studies, making it a valuable resource for comprehending the intricate properties of the variables of HOTS. Every indicator represents a distinct facet of HOTS, and collectively they contribute to a more thorough comprehension of this fundamental educational idea. This information will enhance the advancement of future research endeavors and educational activities. In addition, educators can use this resource to identify specific areas of focus when developing instructional methods that aim to promote HOTS in their pupils. As a result, this adds to the enhancement and effectiveness of instructional strategies.

Table 7

Indicators of HOTS Variables Used in Previous Studies

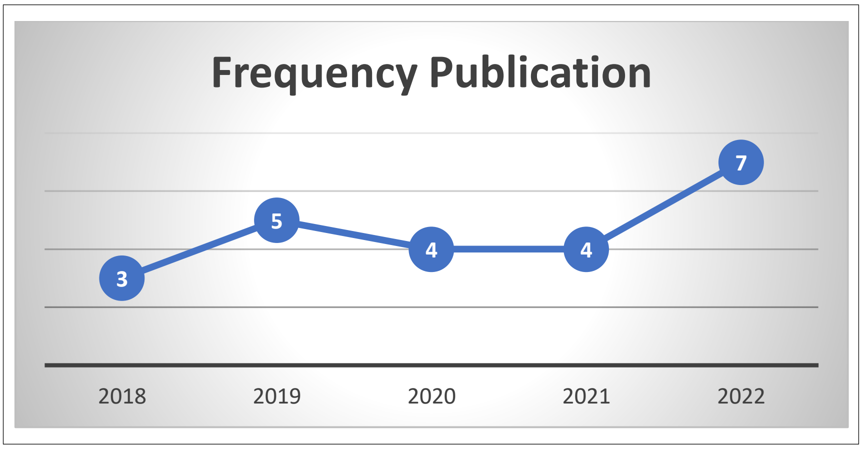
RQ3	HOTS *Variables*	Author	Indicator of HOTS
What were the indicators of HOTS variable?	Critical thinking skills	Erdogan (2019)	Utilization of thinking, arguing, and questioning, as well as enhancing abilities in problem-solving
		Ismail et al. (2019).	Reasoning, decision-making, and problem-solving
		Ismail et al. (2022).	Perception and readiness
	Problem-solving Skills	Ansari et al. (2021); Yayuk & Husamah (2020)	The process involves identifying patterns, establishing links between ideas, and integrating these two aspects to develop linkages between the concepts.
		Koçak (2020)	Open-ended items
		Abdullah et al. (2019); Yayuk & Husamah (2020)	(Model of Polya) Understanding Planning Implementation Final answer
	Creative thinking skills	Hidajat (2021)	The individual intends to engage in problem-solving activities by identifying the fundamental aspects of the issue at hand. S/he aims to generate novel ideas and propose alternative strategies for resolving the problem. Additionally, S/he plans to integrate prior concepts with the inquiries raised by the issue. Ultimately, S/he will implement plans by generating diverse creative solutions.

RQ3	HOTS *Variables*	Author	Indicator of HOTS
What were the indicators of HOTS variable?	Cognitive Domain of Bloom's Taxonomy	Abdullah (2021); Julius et al. (2018); MD-Ali & Kim (2018); Moyo et al. (2022); Ramlee et al. (2019).	Applying, analyzing, evaluating, and creating
		Tanujaya et al. (2021)	Analyzing, evaluating, and creating
		Azid et al. (2022).	Analysis, synthesis, and evaluation
		Ibrahim et al. (2020)	Skills, ideas, and information for problem-solving purposes to improve one's decision-making abilities.
	Principles of HOTS	Tajudin et al. (2018)	Examples of practices that can be included in this category encompass the optimal implementation of assessment for learning, the determination of HOTS learning outcomes, the utilization of HOTS questioning strategies, the integration of information, communication, and technology, the engagement in active learning, and the cultivation of habits of mind.

After analyzing the search results, it has become apparent that a few researchers have been actively pursuing studies about measuring instruments for HOTS. The lack of research in this area highlights the need for increased focus and resources dedicated to creating and verifying dependable tools for evaluating HOTS in the context of mathematics education. Comprehensive and rigorously validated assessment instruments are essential in education as they offer educators unique insights into the cognitive capacities associated with higher-order thinking. By promoting increased research in this area and facilitating cooperation between researchers and educators, it is feasible to progress the field of HOTS assessment and make valuable contributions to improving mathematics education. This, in turn will empower students with the cognitive abilities necessary for success in their academic pursuits and beyond.

Figure 2

Frequency of Article Publication on Instruments for the Past Five Years



Educators need to possess a greater comprehension of high-level cognitive processes (Harpster, 1999; Thompson, 2008). HOTS is a set of mental processes (Lewis & Smith, 1993; Resnick, 2016) that involve complex evaluation and analysis of complex situations based on many criteria, or the identification of possible solutions in difficult situations. The task of instructing students in higher-order thinking has significant difficulties due to the necessity of the teacher possessing creative abilities (Alhassora et al., 2017; Henningsen & Stein, 1997; Nisa et al., 2023; Thompson, 2008), as well as the prerequisite of students possessing strong foundational abilities (Kania et al., 2023). However, educators still strive to create assessment instruments that effectively evaluate students' higher-order thinking skills (Rintayati, 2020). Teachers' inability to incorporate HOTS) can be attributed to their limited training and lack of exposure to HOTS. It is recommended that extra courses or specialized workshops be provided alongside teaching courses in order to equip future educators with the necessary information and skills to effectively incorporate HOTS into their instruction (Ahmat et al., 2017). Fewer than ten scholarly studies were carried out on this topic yearly, which is a matter of concern.

DISCUSSION

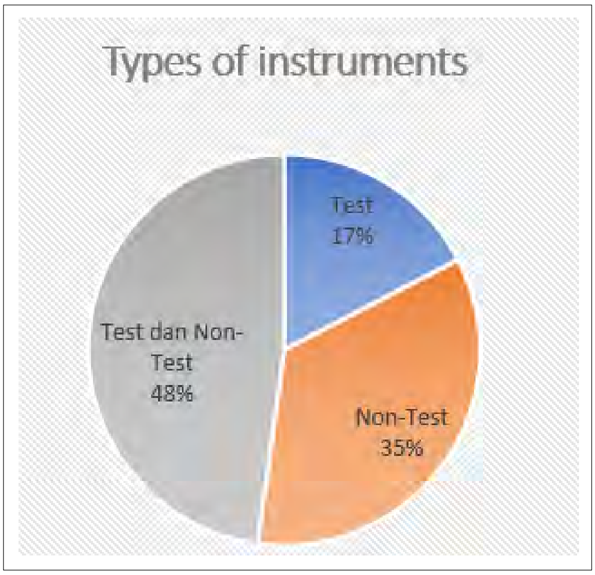
Types of HOTS Measuring Instruments

Numerous instruments measuring a broad range of HOTS, have been devised to assess cognitive abilities that extend beyond the mere acquisition of fundamental knowledge. These instruments have been designed to facilitate the investigation of complex cognitive processes such as critical thinking, problem-solving, creativity, and analytical skills. Educators and researchers are presented with a wide array of HOTS measuring tools, enabling them to choose the most suitable method for assessing specific cognitive abilities and enhancing their understanding of students' higher-order thinking capabilities.

There are three types of HOTS test instruments, and they are as follows: (1) test instruments (Erdogan, 2019; Ibrahim et al., 2020a; Ramlee et al., 2019); (2) non-test instruments (Abdullah et al., et al., 2019; Ismail et al., 2019a, 2022b; Newton et al., 2022; Tajudin et al., 2018; Tanujaya et al., 2021); and (3) test and non-test instruments (Abdullah, 2020; Abdurrahman et al., 2021; Ansari et al., 2021; Azid et al., 2022; Hidajat, 2021; Julius et al., 2018; Koçak, 2020; MD-Ali & Kim, 2018; Yayuk & Husamah, 2020).

Figure 3

Types of HOTS Measuring Instruments



Researchers have created various HOTS measuring instruments and methodologies to meet assessment needs. Some HOTS test instruments have been formulated as multiple-choice or essay-based assessments, and the studies conducted by Erdogan (2019) and Ibrahim et al. (2020) have contributed to this area. HOTS test instruments can be developed as multiple choice (Setiawan et al., 2021; Suprapto, 2020) or essay. It has been shown to accurately identify students' HOTS (Maryani et al., 2021). Several commonly used tools for testing HOTS include essay questions, case studies, project-based exams, and performance tasks. Essay prompts necessitate students engaging in the analysis and synthesis of information to formulate well-considered responses, demonstrating their proficiency in critical thinking and compelling articulation of ideas.

On the other hand, non-test instruments have also been employed to evaluate HOTS. These include interview-based assessments (Abdullah et al., 2019; Tanujaya et al., 2021) and questionnaires utilized in studies by Ismail et al. (Ismail et al., 2019, 2022). Surveys have also been employed as another non-test instrument (Tajudin et al., 2018). Non-test instruments facilitate the assessment of HOTS by allowing students to apply mathematical concepts in real-world situations and exhibit a profound understanding of the subject matter. Non-test instruments are of significant importance in evaluating and enhancing students' cognitive skills at a higher level of thinking within the domain of mathematical education.

The variety in the design of non-test instruments utilized in mathematics instruction is contingent upon the particular objectives and learning context. Non-test instruments encompass various assessment methodologies that differ from conventional examinations. This academic instrument includes a variety of research projects, oral presentations, compiled portfolios, collaborative group discussions, in-depth case studies, and written assignments. Research endeavors allow students to explore a particular mathematical topic, formulate an experimental approach, and subsequently assess and analyze the findings acquired. Presentations allow students to disseminate their knowledge and skills in specific mathematical topics to their peers within the classroom or group setting.

Case studies encompass authentic situations that compel students to employ their knowledge to resolve intricate issues, showcasing their aptitude for problem-solving within pragmatic settings. Project-based evaluations evaluate students' capacity for creativity by assigning them the responsibility of generating creative solutions or manifestations about a specific subject matter. In contrast, performance tasks require students to exhibit their advanced cognitive capabilities by engaging in practical activities or simulations, exemplifying their capacity to use acquired knowledge and skills in real-world scenarios.

Moreover, there are combined HOTS measuring instruments that incorporate multiple assessment methods. For example, some researchers have used a combination of tests and interviews, as seen in studies by Ansari et al. (2021a) and Azid et al. (2022a). Others have utilized tests, interviews, and observations, as evidenced in the works of Hidajat (2021) and Yayuk & Husamah (2020). Additionally, the study by Julius et al. (2018) recommended the use of a test, a rubric, and interviews, while the research by Abdurrahman et al. (2021) used questions and worksheets. These comprehensive and varied HOTS measuring instruments allow researchers and educators to employ various assessment approaches to accurately gauge and foster higher-order thinking skills among students.

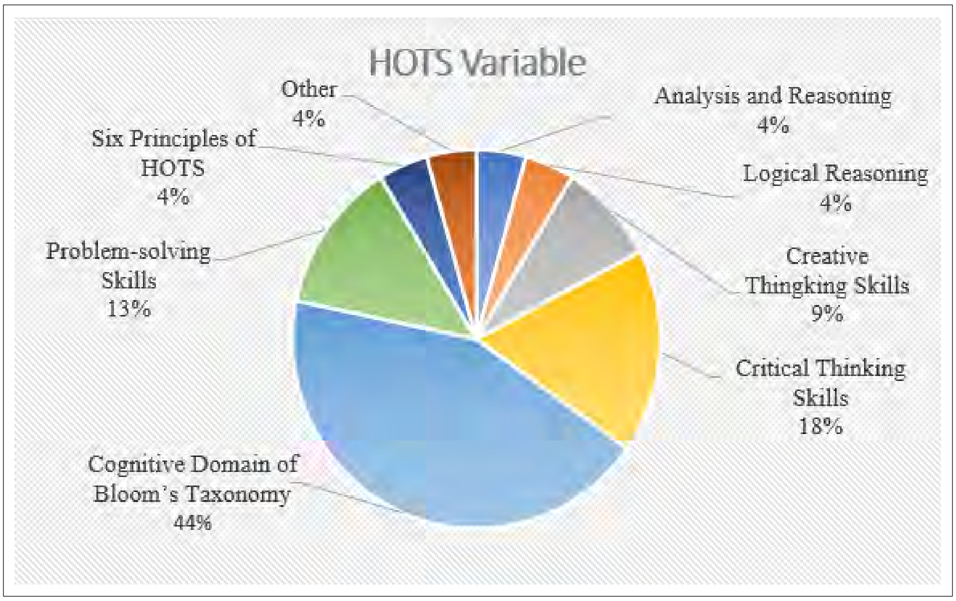
HOTS Variables in Research

The significance of an instrument in assessing the level of success achieved in an activity cannot be overstated. The analysis of articles published in the last five years reveals a notable lack of emphasis on the evaluation of measuring instruments in mathematics education. This phenomenon is visually depicted in Figure 2. Remarkably, the annual frequency of research conducted to examine or evaluate the precision of these measuring instruments has consistently remained low, with less than ten investigations being undertaken per year. The findings above indicate a worrisome pattern in research on the assessment tools used for HOTS, as this aspect has not received the necessary attention and recognition. Since this area is not getting enough attention, it raises questions about how well teachers and researchers are using reliable assessment tools to capture the complexity of HOTS, despite how important it is to accurately measure students' higher-order cognitive skills in mathematics education.

The present SLR has clearly shown that there is not much research in the area of HOTs assessment tools; more time, effort, and money need to be put into creating and validating robust assessment tools that can test students' ability to use HOTS in math class. Educators who are provided with extensive and rigorously validated measurement tools are empowered to personalize educational strategies that can foster critical thinking, problem-solving, and analytical abilities. These instruments are of utmost importance in furnishing educators with valuable insights into their students' cognitive ability at a higher level of thinking. The advancement of HOTS measurement and its potential contribution to the improvement of mathematics education can be facilitated by the promotion of further research in this area and the cultivation of collaborative efforts between researchers and educators.

Figure 4

Distribution of HOTS Variables in Research Articles



Note correction: Analysis and Reasoning in Figure 4

Figure 4 shows that several experts have defined HOT in terms of several thinking skills, including the most frequently used the cognitive domain of Bloom's taxonomy (Conklin, 2012; Krulik & Rudnick, 1999; Yuliati & Lestari, 2018). HOT was considered critical thinking and creative thinking by Presseisen (1985). Brookhart (2010) stated that HOT was problem-solving. King et al. (2013) explained that HOT stood for logical, reflective, and metacognitive thinking. Presseisen (1985) was of the view that HOT was decision-making. Figure 4 presents a visually compelling representation of the distribution of research about the HOTS variable. The current state of research on HOTS indicates a very narrow focus, with most studies concentrating on the cognitive domain of Bloom's Taxonomy. This observation suggests significant potential for future research and development in several domains of HOTS.

In contrast, the Critical Thinking Skills category has garnered considerable scholarly attention, as evidenced by 18 percent of research on this subject. This observation highlights the widespread recognition among scholars of the need to cultivate critical thinking abilities in educational settings. Nevertheless, further investigation is still required into several dimensions of HOTS, including creative thinking skills, logical reasoning, and analytical reasoning abilities.

The limited allocation of 4 percent of research efforts towards investigating the Six Principles of HOTS implies a significant gap in the current body of research about HOTS. Consequently, additional exploration and investigation of this idea within the realm of the HOTS study is warranted. In general, Figure 4 underscores the potential for further advancement and progress in research on higher-order thinking skills. Conducting additional research on several facets of HOTS will contribute to the advancement of one's comprehension regarding strategies to enhance higher-order thinking skills within the field of education.

CONCLUSION

The significance of an instrument in assessing the level of success achieved in an activity cannot be overstated. The analysis of articles released over the last five years reveals a notable lack of focus on the assessment of measuring instruments within the realm of mathematics education. This phenomenon is visually depicted in Figure 2. Remarkably, the annual quantity of research conducted to examine or evaluate the precision of these measuring instruments has always remained limited, with fewer than ten investigations being completed every year. The findings above indicate a worrisome pattern, wherein the significance of research about measuring instruments for HOTS has not been adequately emphasized. The fact that it has not received much recognition is proof of this inclination. The insufficient level of engagement within this domain raises concerns regarding the degree to which educators and researchers are employing valid measuring tools to capture the complexities of HOTS, despite the crucial importance of accurately evaluating students' higher-order cognitive abilities in mathematics education.

Upon examining the search results, it has become evident that only a few scholars were researching assessment instruments associated with HOTS. Because there was not a lot of research in this area, more time, effort, and money need to be put into creating and validating robust assessment tools that can measure how well students use HOTS in mathematics class. Educators with extensive and rigorously validated measuring tools are empowered to tailor instructional tactics that foster critical thinking, problem-solving, and analytical abilities. These tools are essential in providing valuable information to educators about the cognitive skills of their pupils at a higher level of thinking. The encouragement of additional research in this area and the development of collaborative efforts between researchers and educators can help advance HOTS measurement and improve mathematics education. Ultimately, this will give pupils the cognitive abilities essential for success in their educational endeavors and beyond.

Recommendations

The present SLR has broader implications that go beyond the boundaries of its stated aims, thereby offering significance for future research undertakings and educational practices. The observed patterns and trends offer valuable insights for designing suitable assessment instruments for HOTS in mathematics and other academic subjects. Educators should optimize their educational strategies by developing a comprehensive comprehension of the dynamic relationship between evaluative and non-evaluative components and recognizing the significance of Bloom's Taxonomy, empowering students to nurture and demonstrate their higher-order cognitive abilities more efficiently. Moreover, this study presents potential avenues for future exploration within the HOTS domain. As the expansion of HOTS evaluation develops, it is recommended that prospective scholars seize the opportunity to investigate innovative ways, study supplemental indications, and untangle the subtleties of higher-order cognitive processes. Ongoing research efforts have the potential to drive advancements in educational methodologies through the improvement of teaching strategies and the cultivation of critical thinking and problem-solving abilities in students. By utilizing the comprehensive foundation established in this inquiry, scholars can develop their capacity to generate significant contributions in the pursuit of a comprehensive understanding of sophisticated cognitive processes. This, in turn enables the integration of the relevant HOTs processes within educational settings, thereby fostering the intellectual development of students and nurturing their aptitude in the realm of critical thinking.

Limitations

Though there were significant findings obtained from the present SLR, it is imperative to also acknowledge its many limitations. The study's utilization of 18 scholarly papers may not provide a complete representation of the extensive range of assessment methods for HOTS in mathematics education. Although there were considerable efforts made to select pertinent and dependable sources, the restricted scope of the literature study might impede the applicability of the present findings. Moreover, limiting the SLR’s attention solely to English-language articles risks introducing linguistic and cultural biases into the research, which might have resulted in neglecting many other different viewpoints and cultural contexts. Therefore, it is recommended to use caution in the interpretation of the data and not generalize beyond the particular contexts studied. Furthermore, engaging in further studies that span a broader array of sources has the ability to augment one’s comprehension of this subject matter more thoroughly.

In addition, the research carried out in the present SLR demonstrates an inclination towards employing Bloom's Taxonomy as the principal framework for evaluating HOTS in the domain of mathematics. This inclination has the ability to generate an inherent bias towards the paradigm outlined earlier. While recognizing the presence of diverse theoretical frameworks and models about higher-order thinking, it is essential to realize that the study has encountered certain limitations in exploring them.

The present SLR has substantially contributed to assessing higher-order cognitive skills within mathematical education. Nevertheless, it is imperative to recognize the intrinsic constraints of the SLR study to prevent generating baseless extrapolations. To further one’s understanding of the intricate characteristics of higher-order thinking and its evaluation in diverse educational contexts, it is crucial that future research endeavors to prioritize the resolution of these constraints.

ACKNOWLEDGEMENT

This research was made possible by the Lembaga Pengelola Dana Pendidikan, the Indonesian Endowment Fund for Education (LPDP Indonesia).

REFERENCES

Abdullah, A. H. (2017). Mathematics teachers’ level of knowledge and practice on the implementation of higher-order thinking skills (HOTS). *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(1), 3–17. https://doi.org/10.5539/ass.v11n21p133

Abdullah, A. H. (2020). Using active learning with smart board to enhance primary school students’ higher order thinking skills in data handling. *Universal Journal of Educational Research*, *8*(10), 4421–4432. https://doi.org/10.13189/ujer.2020.081009

Abdullah, A. H. (2021). Does the Use of Smart Board Increase Students’ Higher Order Thinking Skills (HOTS)? *IEEE Access*, *9*, 1833–1854. https://doi.org/10.1109/ACCESS.2020.3042832

Abdullah, A. H., Fadil, S. S., Abd Rahman, S. N. S., Tahir, L. M., & Hamzah, M. H. (2019). Emerging patterns and problems of higher-order thinking skills (HOTS) mathematical problem-solving in the form-three assessment (PT3). *South African Journal of Education*, *39*(2). https://doi.org/10.15700/saje.v39n2a1552

Abdurrahman, M. S., Halim, A. A., & Sharifah, O. (2021). Improving polytechnic students’ high-order-thinking-skills through inquiry-based learning in mathematics classroom. *International Journal of Evaluation and Research in Education*, *10*(3), 976–983. https://doi.org/10.11591/IJERE.V10I3.21771

Ahmat, N., Azmee, N. A., Mohamed, N. H., Zamzamir, Z., Zahari, N. S., Shafie, S., Mohamed, N. A., & Shah, R. N. F. A. R. M. (2017). Knowledge, Skills and Attitude of Pre-Service Mathematics Teachers Towards Higher-Order Thinking Skills. *International Journal of Educational Methodology*, *8*(4), 795–804. https://doi.org/https://doi.org/10.12973/ijem.8.4.795

Alhadi, A., Dr Tom, B., & Yacine, R. (2025). Enhancing asset management: Integrating digital twins for continuous permitting and compliance - A systematic literature review. *Journal of Building Engineering*, *99*, 111515. https://doi.org/10.1016/j.jobe.2024.111515

Alhassora, N. S. A., Abu, M. S., & Abdullah, A. H. (2017). Hindering factors in mastering higher-order thinkings skills: Application of Rasch measurement model. *Man in India*, *97*(19), 275–280.

Ali, N., Ullah, S., & Khan, D. (2022). Interactive laboratories for science education: A subjective study and systematic literature review. *Multimodal Technologies and Interaction*, *6*(10). https://doi.org/10.3390/mti6100085

Ansari, B. I., Saleh, M., Nurhaidah, & Taufiq. (2021). Exploring students' learning strategies and self-regulated learning in solving mathematical higher-order thinking problems. *European Journal of Educational Research*, *10*(2), 743–756. https://doi.org/10.12973/eu-jer.10.2.743

Azid, N., Ali, R. M., El Khuluqo, I., Purwanto, S. E., & Susanti, E. N. (2022). Higher order thinking skills, school-based assessment and students' mathematics achievement: understanding teachers' thoughts. *International Journal of Evaluation and Research in Education*, *11*(1), 290–302. https://doi.org/10.11591/ijere.v11i1.22030

Bilondi, H. T., Valipour, H., Khoshro, S., Jamilian, P., Ostadrahimi, A., & Zarezadeh, M. (2024). Corrigendum to the effect of caffeine supplementation on muscular strength and endurance: A meta-analysis of meta-analyses" [Heliyon *10*(15), August 2024, e35025]. *Heliyon*, *10*(23), e40064. https://doi.org/10.1016/j.heliyon.2024.e40064

Boadu, E. F., Wang, C. C., & Sunindijo, R. Y. (2020). Characteristics of the construction industry in developing countries and its implications for health and safety: An exploratory study in ghana. *International Journal of Environmental Research and Public Health*, *17*(11), 1–21. https://doi.org/10.3390/ijerph17114110

Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M. (2007). Lessons from applying the systematic literature review process within the software engineering domain. *Journal of Systems and Software*, *80*(4), 571–583. https://doi.org/10.1016/j.jss.2006.07.009

Bronkhorst, H., Roorda, G., Suhre, C., & Goedhart, M. (2022). Students' use of formalisations for improved logical reasoning. *Research in Mathematics Education*, *0*(0), 1–33. https://doi.org/10.1080/14794802.2021.1991463

Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. Library of Congress Catalog.

Buang, Z., Mohamad, M. M., Ahmad, A., & Yuniarti, N. (2020). The earnings and employment of community colleges' graduates: occupational field and gender analysis. *Jurnal Pendidikan Teknologi Dan Kejuruan*, *26*(1), 11–17. https://doi.org/10.21831/jptk.v26i1.29750

Budgen, D., & Brereton, P. (2006). Performing systematic literature reviews in software engineering. *Proceedings of the 28th International Conference on Software Engineering*, 1051–1052. https://doi.org/10.1145/1134285.1134500

Cáceres-Reche, M. P., Tallón-Rosales, S., Navas-Parejo, M. R., & De la Cruz-Campos, J. C. (2022). Influence of sociodemographic factors and knowledge in pedagogy on the labor market insertion of education science professionals. *Education Sciences*, *12*(3). https://doi.org/10.3390/educsci12030200

Chambers, D., Rodgers, M., & Woolacott, N. (2009). Not only randomized controlled trials, but also case series should be considered in systematic reviews of rapidly developing technologies. *Journal of Clinical Epidemiology*, *62*(12), 1253-1260.e4. https://doi.org/10.1016/j.jclinepi.2008.12.010

Charungkaittikul, S., Pathumcharoenwattana, W., & Kovitya, M. (2022). A study on the required work competencies of graduates in the non-formal education program. *Kasetsart Journal of Social Sciences*, *43*(4), 825–832. https://doi.org/10.34044/j.kjss.2022.43.4.04

Conklin, W. (2012). *Strategies for Developing Higher-Order Thinking Skills, Grades 6-12: Grades 6-12*. Shell Educational Publishing.Inc.

Devlin, K. (2011). *Mathematics Education for a New Era: Video Games as a Medium for Learning*. A K Peters, Ltd.

Di, W., Danxia, X., & Chun, L. (2019). The effects of learner factors on higher-order thinking in the smart classroom environment. *Journal of Computers in Education*, *6*(4), 483–498. https://doi.org/10.1007/s40692-019-00146-4

Dipitso, P. (2021). Work-integrated learning for mining engineering training and the employability nexus in traditional research universities. *Journal of Comparative & International Higher Education*, *12*(6S1), 122–136. https://doi.org/10.32674/jcihe.v12i6s1.2192

Erdogan, F. (2019). Effect of cooperative learning supported by reflective thinking activities on students' critical thinking skills. *Eurasian Journal of Educational Research*, *2019*(80), 89–112. https://doi.org/10.14689/ejer.2019.80.5

Gavronskaya, Y., Larchenkova, L., Berestova, A., Latysheva, V., & Smirnov, S. (2022). The development of critical thinking skills in mobile learning: fact-checking and getting rid of cognitive distortions. *International Journal of Cognitive Research in Science, Engineering and Education (IJCRSEE)*, *10*(2), 51–68. https://doi.org/10.23947/2334-8496-2022-10-2-51-68

Gusenbauer, M., & Haddaway, N. R. (2021). What every researcher should know about searching – clarified concepts, search advice, and an agenda to improve finding in academia. *Research Synthesis Methods*, *12*(2), 136–147. https://doi.org/10.1002/jrsm.1457

Hadiastriani, Y. (2022). Bibliometric Analysis for Development of Analytical Thinking Based Assessment. *European Journal of Research and Reflection in educational sciences, 10*(2), 52–64. http://www.idpublications.org/wp-content/uploads/2022/12/Full-Paper-BIBLIOMETRIC-ANALYSIS-FOR-DEVELOPMENT-OF-ANALYTICAL-THINKING-BASED-ASSESSMENT.pdf

Harpster, R. A. (1999). How to take the out of FMEAs. *Measuring Business Excellence*, *3*(3), 20–24. https://doi.org/10.1108/eb025573

Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, *28*(5), 524–549. https://doi.org/10.2307/749690

Herayanti, L., Widodo, W., Susantini, E., & Gunawan, G. (2020). The effectiveness of blended learning model based on inquiry collaborative tutorial toward students' problem-solving skills in physics. *Journal for the Education of Gifted Young Scientists*, *8*(3), 959–972. https://doi.org/10.17478/JEGYS.675819

Hidajat, F. A. (2021). Students creative thinking profile as a high order thinking in the improvement of mathematics learning. *European Journal of Educational Research*, *10*(4), 1907–1918. https://doi.org/https://doi.org/10.12973/eu-jer.10.3.1247

Higgins, J. P. ., & Green, S. (2006). *Cochrane handbook for systematic reviews of interventions 4.2.6. September*, 49–55.

Higgins, S. (2017). Managing Higher-Order Thinking Skills. In *Managing Academic Libraries* (pp. 29–40). Elsevier. https://doi.org/10.1016/b978-1-84334-621-0.00004-2

Huang, Y. M., Silitonga, L. M., Murti, A. T., & Wu, T. T. (2023). Learner engagement in a business simulation game: impact on higher-order thinking skills. *Journal of Educational Computing Research*, *61*(1), 96–126. https://doi.org/10.1177/07356331221106918

Husain, F. N. (2023). Impact of multiple intelligences and 21st century skills on future work force. *International Education Studies*, *16*(3), 16. https://doi.org/10.5539/ies.v16n3p16

Ibrahim, N. N., Ayub, A. F. M., & Yunus, A. S. M. (2020). Impact of higher order thinking skills (hots) module based on the cognitive apprenticeship model (CAM) on student's performance. *International Journal of Learning, Teaching and Educational Research*, *19*(7), 246–262. https://doi.org/10.26803/ijlter.19.7.14

Indah, P. (2020). Development of HOTS (High Order Thinking Skill) oriented learning through discovery learning model to increase the critical thinking skill of high school students. *International Journal of Chemistry Education Research*, *3*(3). https://doi.org/10.20885/ijcer.vol4.iss1.art4

Ioannou, A., & Makridou, E. (2018). Exploring the potentials of educational robotics in the development of computational thinking: A summary of current research and practical proposal for future work. *Education and Information Technologies*, *23*(6), 2531–2544. https://doi.org/10.1007/s10639-018-9729-z

Ismail, S. N., Muhammad, S., Kanesan, A. G., & Ali, R. M. (2019). The influence of teachers' perception and readiness towards the implementation of critical thinking skills (CTS) practice in mathematics. *International Journal of Instruction*, *12*(2), 337–352. https://doi.org/10.29333/iji.2019.12222a

Ismail, S. N., Muhammad, S., Omar, M. N., & Shanmugam, S. K. S. (2022). The practice of critical thinking skills in teaching mathematics: Teachers' perception and readiness. *Malaysian Journal of Learning and Instruction*, *19*(1), 1–30. https://doi.org/10.32890/mjli2022.19.1

Jaenudin, R., Chotimah, U., Farida, F., & Syarifuddin, S. (2020). Student development zone: higher order thinking skills (hots) in critical thinking orientation. *International Journal of Multicultural and Multireligious Understanding*, *7*(9), 11. https://doi.org/10.18415/ijmmu.v7i9.1884

Julius, E., Mun, S. H., Abdullah, A. H., Mokhtar, M., & Suhairom, N. (2018). Using digital smart board to overcome higher order thinking skills learning difficulties in data handling among primary school students. *International Journal of Interactive Mobile Technologies*, *12*(7), 43–59. https://doi.org/10.3991/ijim.v12i7.9644

Kania, N., Fitriani, C., & Bonyah, E. (2023). Analysis of students' critical thinking skills based on prior knowledge mathematics. *International Journal of Contemporary Studies in Education (IJ-CSE)*, *2*(1), 49–58. https://doi.org/10.56855/ijcse.v2i1.248

Kania, N., & Juandi, D. (2023). *Does self-concept affect mathematics learning achievement ? 17*(3), 455–461. https://doi.org/10.11591/edulearn.v17i3.20554

Khan, D., Plopski, A., Fujimoto, Y., Kanbara, M., Jabeen, G., Zhang, Y. J., Zhang, X., & Kato, H. (2022). Surface Remeshing: A systematic literature review of methods and research directions. *IEEE Transactions on Visualization and Computer Graphics*, *28*(3), 1680–1713. https://doi.org/10.1109/TVCG.2020.3016645.

Koçak, D. (2020). Investigation of rater tendencies and reliability in different assessment methods with many facet rasch model. *International Electronic Journal of Elementary Education*, *12*(4), 349–358. https://doi.org/10.26822/iejee.2020459464

Krulik, S., & Rudnick, J. A. (1999). *Innovative Tasks to Improve Critical and Creative Thinking Skills. in L. Stiff & F. Curcio (Eds.), Developing Mathematical Reasoning in Grades K-12*. National Council of Teachers of Mathematics.

Lewis, A., & Smith, D. (1993). Defining higher order thinking. *Theory Into Practice*, *32*(3), 131–137. https://doi.org/10.1080/00405849309543588

Maison, Hidayat, M., Kurniawan, D. A., Yolviansyah, F., Sandra, R. O., & Iqbal, M. (2022). How critical thinking skills influence misconception in electric field. *International Journal of Educational Methodology*, *8*(2), 377–390.

Marchuk, L., & Yatsyna, O. (2020). Practices of informal education as a resource for self-realization of self-referential identities in a pandemic and post-pandemic period. *Postmodern Openings*, *11*(2), 99–104. https://doi.org/10.18662/po/11.2/162

Maryani, I., Prasetyo, Z. K., Wilujeng, I., Purwanti, S., & Fitrianawati, M. (2021). HOTs multiple choice and essay questions: a validated instrument to measure higher-order thinking skills of prospective teachers. *Journal of Turkish Science Education*, *18*(4), 674–690. https://doi.org/10.36681/tused.2021.97

Mascia, M. L., Perrone, S., Bardi, D., Agus, M., Penna, M. P., & Lucangeli, D. (2018). Digital life, mathematical skills and cognitive processes. *Proceedings of the 15th International Conference on Cognition and Exploratory Learning in the Digital Age, CELDA 2018*, *Celda*, 371–374.

Massetor, N. A., Rozali, M. Z., & Marian, M. F. (2021). K-workers practice in increasing graduate employability to overcome the challenges of the fourth industrial revolution. *Research and Innovation in Technical and Vocational Education and Training*, *1*(1), 224–231. https://publisher.uthm.edu.my/periodicals/index.php/ritvet/article/view/261

MD-Ali, R., & Kim, K. M. (2018). Geogebra in learning of mathematics towards supporting "stem" education. *Journal of Social Sciences Research*, *2018*(Special Issue 6), 776–782. https://doi.org/10.32861/jssr.spi6.776.782

Moyo, S. E., Combrinck, C., & Van Staden, S. (2022). Evaluating the impact of formative assessment intervention and experiences of the standard 4 teachers in teaching higher-order-thinking skills in mathematics. *Frontiers in Education*, *7*. https://doi.org/10.3389/feduc.2022.771437

Nakakoji, Y., & Wilson, R. (2020). Interdisciplinary learning in mathematics and science: transfer of learning for 21st century problem solving at university. *Journal of Intelligence*, *8*(3), 1–23. https://doi.org/10.3390/jintelligence8030032

NCTM. (2021). The Effective and Appropriate Use of Large-Scale Assessments in Mathematics Education to Guide Systemic Improvement and Equitable Student Learning Large-Scale Assessments Can Support Systemic Improvement Large-Scale Assessment Must Always Be Viewed Throug. *National Council of Teachers of Mathematics*, *March*.

Newton, D., Wang, Y., & Newton, L. (2022). 'Allowing them to dream': Fostering creativity in mathematics undergraduates. *Journal of Further and Higher Education*, *46*(10), 1334–1346. https://doi.org/10.1080/0309877X.2022.2075719

Nieminen, J. H., & Atjonen, P. (2022). The assessment culture of mathematics in Finland: A student perspective. *Research in Mathematics Education*, *0*(0), 1–20. https://doi.org/10.1080/14794802.2022.2045626

Nisa', K., Nasrullah, A., Hidayat, A., Mahuda, I., & Bhat, I. A. (2023). *Problem-Based Learning in Improving Problem-Solving Ability and Interest in Learning Mathematics: An Empirical Study*. *1*, 206–217.

Paiva, A., Craveiro, R., Aroso, I., Martins, M., Reis, R. L., & Duarte, A. R. C. (2014). Natural deep eutectic solvents - Solvents for the 21st century. *ACS Sustainable Chemistry and Engineering*, *2*(5), 1063–1071. https://doi.org/10.1021/sc500096j

Phurikultong, N., & Kantathanawat, T. (2022). Flipping the undergraduate classroom to develop student analytical thinking skills. *Emerging Science Journal*, *6*(4), 739–757. https://doi.org/10.28991/ESJ-2022-06-04-06

Pollarolo, E., Størksen, I., Skarstein, T. H., & Kucirkova, N. (2022). Children's critical thinking skills: Perceptions of Norwegian early childhood educators. *European Early Childhood Education Research Journal*, *May*, 1–13. https://doi.org/10.1080/1350293X.2022.2081349

Popham, W. (2008). *Classroom assessment: What teachers need to know (5th ed)*. Allyn and Bacon.

Prakash, O., & Kumar, K. (2016). Higher education: Informal learning for unorganized sector. *International Journal of Applied Science and Engineering*, *4*(2), 103. https://doi.org/10.5958/2322-0465.2016.00012.5

Presseisen, B. . (1985). *Thinking Skills Throughout The Curriculum*: *A conceptual Design.* Research for Bette*r* Schools

Ramlee, N., Rosli, M. S., & Saleh, N. S. (2019). Mathematical HOTS cultivation via online learning environment and 5E inquiry model: Cognitive impact and the learning activities. *International Journal of Emerging Technologies in Learning*, *14*(24), 140–151. https://doi.org/10.3991/ijet.v14i24.12071

Resnick, I. (2016). Developmental growth trajectories in understanding of fraction magnitude from fourth through sixth grade. *Developmental Psychology*, *52*(5), 746–757. https://doi.org/10.1037/dev0000102

Rintayati, P. (2020). Development of two-tier multiple choice test to assess indonesian elementary students' higher-order thinking skills. *International Journal of Instruction*, *14*(1), 555–566. https://doi.org/10.29333/IJI.2021.14133A

Setiawan, J., Sudrajat, A. A., Aman, A., & Kumalasari, D. (2021). Development of higher order thinking skill assessment instruments in learning Indonesian history. In *International Journal of Evaluation and Research in Education (IJERE)* (Vol. 10, Issue 2, p. 545). Institute of Advanced Engineering and Science. https://doi.org/10.11591/ijere.v10i2.20796

Sidiq, Y., Ishartono, N., Desstya, A., Prayitno, H. J., Anif, S., & Hidayat, M. L. (2021). Improving elementary school students' critical thinking skill in science through hots-based science questions: A quasi-experimental study. *Jurnal Pendidikan IPA Indonesia*, *10*(3), 378–386. https://doi.org/10.15294/JPII.V10I3.30891

Suanto, E., Maat, S. M., & Zakaria, E. (2023). The effectiveness of the implementation of three dimensions geometry kara module on higher order thinking skills (HOTS) and motivation. *International Journal of Instruction*, *16*(3), 95–116. https://doi.org/10.29333/iji.2023.1636a

Sukholova, M. A. (2020). Formal and non-formal education in Ukraine in the context of new socio-cultural challenges. *Innovative Solution in Modern Science*, *7*(43), 61. https://doi.org/10.26886/2414-634x.7(43)2020.5

Suprapto, E. (2020). The analysis of instrument quality to measure the students' higher order thinking skill in physics learning. *Journal of Turkish Science Education*, *17*(4), 520–527. https://doi.org/10.36681/tused.2020.42

Susilo, F. J., Usodo, B., & Sari, D. (2023). The profile of high-order thinking skills of junior high school students. *International Journal of Mathematics and Mathematics Education*, *1*(1), 77–82. https://doi.org/https://doi.org/10.56855//ijmme.v1i1.266

Tajudin, N. M., Puteh, M., & Adnan, M. (2018). Guiding principles to foster higher order thinking skills in teaching and learning of mathematics. *International Journal of Engineering and Technology(UAE)*, *7*(4), 195–199. https://doi.org/10.14419/ijet.v7i4.15.21445

Tambunan, H., & Naibaho, T. (2019). Performance of mathematics teachers to build students' high order thinking skills (HOTS). *Journal of Education and Learning (EduLearn)*, *13*(1), 111–117. https://doi.org/10.11591/edulearn.v13i1.11218

Tanujaya, B., Prahmana, R. C. I., & Mumu, J. (2021). Mathematics instruction to promote mathematics higher-order thinking skills of students in Indonesia: Moving forward. *TEM Journal*, *10*(4), 1945–1954. https://doi.org/10.18421/TEM104-60

Thompson. (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

Wahidin, D., & Romli, L. A. M. (2020). Students critical thinking development in the national sciences and mathematics competition in Indonesia: A descriptive study. *Jurnal Pendidikan IPA Indonesia*, *9*(1), 106–115. https://doi.org/10.15294/jpii.v9i1.22240

Wahono, B., Lin, P.-L., & Chang, C.-Y. (2020). Evidence of STEM enactment effectiveness in Asian student learning outcomes. *International Journal of STEM Education*, *7*(1). https://doi.org/10.1186/s40594-020-00236-1

Wilson, D. M., & Narasuman, S. (2020). Investigating teachers' implementation and strategies on higher order thinking skills in school based assessment instruments. *Asian Journal of University Education*, *16*(1), 70–84. https://doi.org/10.24191/ajue.v16i1.8991

Yaacob, Y., Atika, K., & Zawawi, M. (2021). *An Automated Symbolic Package to Enhance Higher Order Thinking Skills (HOTS): Critical Thinking*. 2–3.

Yayuk, E., & Husamah, H. (2020). The difficulties of prospective elementary school teachers in item problem solving for mathematics: Polya's steps. *Journal for the Education of Gifted Young Scientists*, *8*(1), 361–378. https://doi.org/10.17478/jegys.665833

Yeager, J., Sommer, L., Faughnan, J., & Geerkens, C. (2015). Prototype Modeling vs. Item Response Theory – A Paradigm Shift for Measurement Professionals. *The Qualitative Report*. https://doi.org/10.46743/2160-3715/2013.1455

Yigletu, A., Michael, K., & Atnafu, M. (2023). The effect of assessment for learning on pre-service mathematics teachers' higher-order thinking skills in algebra. *Arhiv Za Farmaciju*, *7*(1), 187–202. https://doi.org/10.33902/JPR.202317679

Yuliati, S. R., & Lestari, I. (2018). Higher-order thinking skill (HOTS) Analysis of Students in Solving HOTS Question in Higher Education. In *Perspektif Ilmu Pendidikan* (Vol. 32, Issue 2, pp. 181–188). Universitas Negeri Jakarta. https://doi.org/10.21009/pip.322.10

Zhou, Y., Gan, L., Chen, J., Wijaya, T. T., & Li, Y. (2023). Development and validation of a higher-order thinking skills assessment scale for pre-service teachers. *Thinking Skills and Creativity*, *48*, 101272. https://doi.org/10.1016/j.tsc.2023.101272

Zumbo, B. D., & Hubley, A. M. (2016). Bringing consequences and side effects of testing and assessment to the foreground. *Assessment in Education: Principles, Policy & Practice*, *23*(2), 299–303. https://doi.org/10.1080/0969594X.2016.1141169