International Journal of Essential Competencies in Education
https://journal-center.litpam.com/index.php/ijece/index

June 2025 Vol. 4, No. 1
e-ISSN: 2962-2131
pp. 54-69

Assessing Critical Thinking in Mathematics Education: A Systematic Review and Analysis Using the PRISMA Framework

Irham Azmi [1]*, Mohd Faizal Nizam Lee Abdullah [2], Zuraida Alwadood [3], Franklin P. Calaminos [4]

[1] Faculty of Applied Science and Engineering, Mandalika University of Education, Mataram, Indonesia.
[2] Department of Mathematics, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, Perak, Malaysia.
[3] Department of Mathematical Sciences, College of Computing, Informatics and Mathematics, Universiti Teknologi MARA (UiTM), Selangor, Malaysia.
[4] Education Department, Sorsogon State University - Magallanes Campus, Magallanes, Philippines.

*Correspondence: irhamazmi@undikma.ac.id

Article Info

Article History
Received: April 2024;
Revised: December 2024;
Published: June 2025

Keywords
Review literature;
Mathematics learning;
Assessment;
Critical thinking skills;
PRISMA Framework

Abstract

Developing critical thinking skills is paramount in the realm of mathematics education in today's era. It's crucial to monitor learners' advancements in critical thinking, as such insights are valuable for enhancing educational methodologies. Recognizing this, the availability of tools for evaluating critical thinking abilities is vital for fostering students' proficiency in these skills. This research sets out to review existing literature on the assessment of critical thinking capabilities within the sphere of mathematics. A thorough analysis was conducted on relevant scholarly articles focusing on the evaluation of critical thinking in mathematics education. The study adopted the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework, which includes the steps of identification, screening, eligibility, and inclusion. Literature from the SCOPUS database was meticulously reviewed, given its esteemed status in providing precise data for international research indexing. Data from the selected documents were visualized using VOSviewer software. Adhering to the PRISMA methodology, findings reveal the significance of critical thinking in the field of mathematics education and highlight various methods and tools that can assess these skills within mathematical contexts. Nonetheless, there remains a gap in consensus on the definition of critical thinking due to the diverse array of theories and perspectives, leading to variability in assessment standards, particularly in mathematics education. Thus, there is a pressing need to precisely define critical thinking within the mathematical domain and to develop accurate tools for its assessment.

https://doi.org/10.36312/ijece.v4i1.1858

How to Cite

Azmi, I., Abdullah, M. F. N. L., Alwadood, Z., & Calaminos, F. P. (2025). Assessing Critical Thinking in Mathematics Education: A Systematic Review and Analysis Using the PRISMA Framework. *International Journal of Essential Competencies in Education*, *4*(1), 54–69. https://doi.org/10.36312/ijece.v4i1.1858

INTRODUCTION

In recent years, the prominence of STEM (Science, Technology, Engineering, and Mathematics) education has grown significantly, driven by its essential role in cultivating

high-quality human capital (Montgomery & Fernández-Cárdenas, 2018). The emergence of novel occupational demands in the 21st century necessitates individuals equipped with adaptive, interdisciplinary, and independent thinking capabilities—competencies largely fostered through STEM-oriented curricula, particularly mathematics. As a discipline, mathematics has long been acknowledged not only for its instrumental value in technological and scientific development but also for its foundational status in cognitive training and systematic reasoning (Onion, 2004; Evendi & Verawati, 2021).

The current global educational climate increasingly emphasizes the integration of mathematical competence with critical thinking development, underscoring the responsibility of educational institutions to cultivate students' higher-order thinking skills (Evendi et al., 2022; Erikson & Erikson, 2019). Mathematical learning environments are not merely platforms for acquiring procedural fluency or arithmetic accuracy but are recognized as dynamic spaces for enhancing learners' capacity for abstraction, inference, and problem-solving. Nonetheless, despite its pedagogical potential, mathematics remains one of the most disliked subjects among students. This aversion stems from its perceived complexity and abstract nature, often resulting in diminished interest, low academic performance, and restricted opportunities for critical thinking growth (Salamah, 2020; Evendi et al., 2022).

Critical thinking, as a cognitive construct, involves deliberate, reasoned, and reflective judgment concerning what to believe or how to act. According to Ennis (2018), it encompasses active engagement with ideas through processes of interpretation, evaluation, and justification. In the domain of mathematics education, fostering these thinking dispositions is vital to achieving one of its central goals—developing students' reasoning capacities (Animasaun & Abegunrin, 2017). The National Council of Teachers of Mathematics (NCTM) asserts that reasoning and sense-making form the core of effective mathematics instruction, highlighting that mere memorization is insufficient for conceptual understanding. Instead, students must be empowered to construct logical arguments, analyze relationships, and justify conclusions, aligning closely with the foundational definitions of critical thinking proposed by Dewey (1933), Elder and Paul (2012), and Ennis (2018).

Promoting reasoning in mathematics instruction requires deliberate pedagogical planning, including the careful selection of tasks, the orchestration of classroom discourse, and the implementation of targeted assessments. NCTM (2000) emphasizes that effective instruction should provide learners with opportunities to explore problems from multiple perspectives and justify their strategies through logical explanation. Complementing this view, Maulyda (2020) underscores the importance of evaluating all educational processes against clearly articulated criteria and indicators. Only through such alignment can assessments serve as valid reflections of instructional success. It follows that the development of reasoning or critical thinking among students can only be effectively monitored through assessments designed to elicit and evaluate such competencies.

Assessments aligned with predetermined critical thinking criteria are essential for accurately tracking and enhancing students' reasoning processes. Williams and Lahman (2011) provide evidence that student engagement in structured online discussions correlates positively with critical thinking development, demonstrating the value of assessing interaction quality as part of critical thinking evaluation. Similarly, Tsai (2012) highlights that structured curricula that explicitly incentivize critical thinking contribute to significant improvements in students' reasoning abilities. These findings reinforce the necessity of implementing systematic, well-targeted assessment tools that not only diagnose students' current abilities but also inform instructional practices. Furthermore, as Hong et al. (2020)

assert, continuous evaluation of both student outcomes and pedagogical methodologies is vital to sustaining meaningful development in critical thinking.

Despite growing awareness of its importance, the implementation of valid and reliable assessments for critical thinking remains a complex and unresolved challenge (Suhirman & Prayogi, 2023). Teachers' competencies in administering these assessments—interpreting the results and adapting instruction accordingly—are critical to their success (Verawati et al., 2020). Moreover, assessments have the potential not only to measure thinking skills but also to shape them; in this sense, assessment functions as both a mirror and a motor for learning. Herpiana and Rosidin (2018) emphasize that well-designed assessments can catalyze students' intellectual growth by encouraging reflection, argumentation, and metacognition.

Nevertheless, a persistent issue is the ambiguity in distinguishing between assessments that truly measure critical thinking and those that assess general cognitive abilities or content knowledge (Verawati et al., 2020). In some instances, education systems may aim to promote critical thinking but rely on assessments misaligned with these goals, thus compromising the validity of the measurement (Sudrajat, 2018). Wiliam (2013) aptly notes that the quality and direction of learning are deeply influenced by the precision and relevance of the assessments used, highlighting the need for tools that are both targeted and pedagogically meaningful.

Numerous instruments have been developed to evaluate critical thinking skills. The Watson-Glaser Critical Thinking Appraisal (Watson & Glaser, 1980), originally designed for employment contexts, has been adapted for educational applications to assess logical reasoning and inference (O'Hare & McGuinness, 2015). The Ennis-Weir Critical Thinking Essay Test (Ennis & Weir, 1985), which measures students' written argumentation based on reading tasks, has also proven useful as a diagnostic tool and as a means of tracking learning effectiveness (Werner, 1991). Other notable instruments include the California Critical Thinking Skills Test (Facione, 1990), the California Critical Thinking Disposition Inventory (Facione et al., 1994), the Cornell Critical Thinking Test (Ennis et al., 1985), and the Halpern Critical Thinking Assessment (Halpern, 2010). These tools reflect a diversity of approaches, disciplinary backgrounds, and intended contexts of use.

However, the variety of available instruments also highlights the lack of consensus on what constitutes critical thinking and how it should be measured. This conceptual fragmentation has led to inconsistent definitions, divergent criteria, and measurement tools with varying levels of generalizability and validity (Verawati et al., 2020; Liu et al., 2014). According to Lai (2011), many existing assessments adopt overly broad approaches that fail to capture the specific cognitive demands of critical thinking. In mathematics education specifically, scholars have pointed out a dearth of assessment tools that accurately and consistently measure critical thinking within mathematical contexts (Faradillah & Adlina, 2021). The complexity of capturing abstract reasoning, logical consistency, and problem representation makes the measurement of critical thinking in mathematics particularly challenging (Quinn et al., 2020). Moreover, Leach et al. (2020) highlight concerns about the structural validity of many existing instruments, raising further doubts about their appropriateness for educational decision-making.

In mathematics learning, well-structured assessments not only diagnose student knowledge but also reveal underlying reasoning processes crucial for mathematical growth (Gultom et al., 2022). These processes are often tacit and complex, requiring assessment tools that go beyond correct answers to probe justification, strategy selection, and adaptive reasoning. Additionally, the development of numeracy skills—encompassing both critical and creative thinking—is indispensable for preparing students to meet the demands of modern education (Hidayah et al., 2021). As mathematics becomes increasingly integrated

with digital tools and interdisciplinary applications, assessments must evolve to capture the multifaceted nature of thinking and learning.

Given the complexity and interdisciplinary nature of critical thinking assessment, conducting comprehensive literature reviews is essential. Such reviews synthesize existing knowledge, map conceptual and methodological developments, and identify research gaps. Within mathematics education, these reviews serve to illuminate best practices, highlight effective instruments, and inform the development of new tools tailored to disciplinary demands. They are also invaluable in guiding teacher training, curriculum design, and policy decisions.

Therefore, this study aims to review the existing literature concerning the assessment of critical thinking skills specifically within the context of mathematics education. By systematically analyzing research findings and evaluating existing tools, the study seeks to offer a comprehensive foundation for future inquiry and practice. This includes identifying instruments that are valid, reliable, and appropriate for assessing critical thinking in mathematics; exploring their theoretical underpinnings and methodological robustness; and assessing their relevance to diverse student populations. In doing so, this review contributes to an ongoing effort to refine the conceptualization and measurement of critical thinking in mathematics, thereby enhancing instructional effectiveness and supporting student development in this essential domain.

METHODS

A thorough analysis of relevant literature concerning the assessment of critical thinking skills in mathematics education was undertaken. This research is structured as a literature review following the "Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)" framework (Page et al., 2021). The adoption of the PRISMA approach is motivated by its ability to synthesize current knowledge in the area of critical thinking skills within mathematics education, thereby highlighting areas ripe for future investigation. Additionally, it facilitates the identification of gaps and shortcomings in existing research that need to be addressed in subsequent studies.

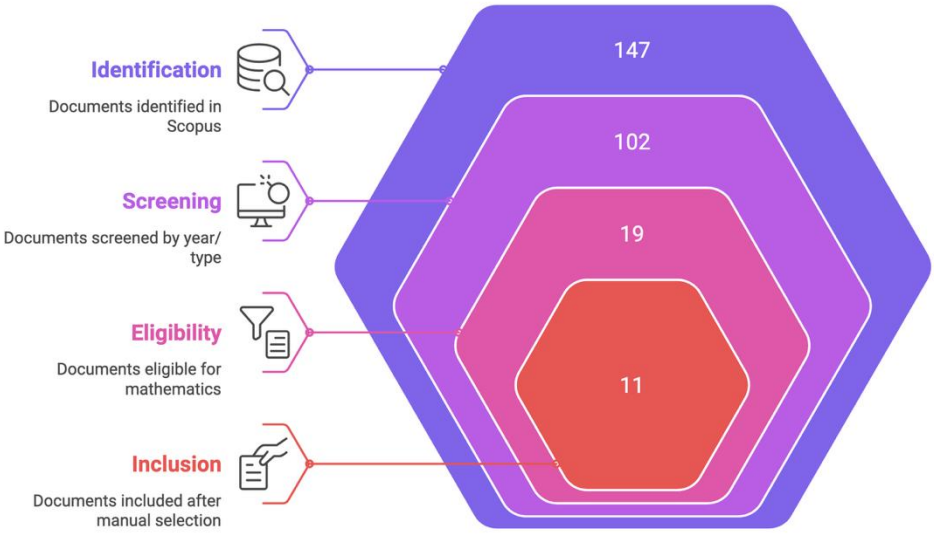


Figure 1. Application of the PRISMA framework for the review of literature documents focused on "assessment of critical thinking skills in mathematics."

The implementation of the PRISMA framework in this research involves a four-step process: "identification, screening, eligibility, and inclusion," utilizing the search terms "critical thinking skills assessment in mathematics." The PRISMA method in this study is as presented in Figure 1. This study relies on the SCOPUS database (https://www.scopus.com/) for its well-regarded accuracy in indexing data and its comprehensive features that enable detailed exploration of scholarly articles by various metrics, enhancing the quality assessment of publications under review.

The search within the SCOPUS database was conducted on December 31, 2023, employing the search string TITLE-ABS-KEY (critical AND thinking AND skills AND assessment AND in AND mathematics), yielding a total of one hundred and forty-seven documents encompassing all document types. The data from these documents were visualized using VOSviewer software, a tool designed for creating and visualizing bibliometric networks from the data collected. An initial screening filtered the documents by publication date, selecting those published within the last ten years (2013 to 2023), resulting in one hundred and twenty-one documents. A subsequent screening refined the selection to journal articles and conference papers, narrowing it down to one hundred and two documents. Out of these, nineteen were deemed relevant based on their specific focus on mathematics, with a final count of eleven documents chosen for review based on keyword relevance and citation analysis.

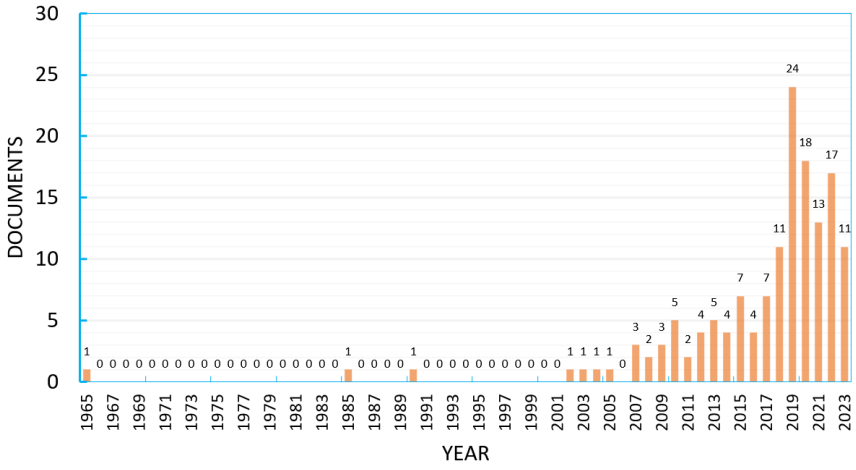
Following the PRISMA guidelines, a bibliometric analysis was conducted on each selected document (Sarkingobir et al., 2023; Wirzal et al., 2022). The process of documentation was meticulous, with search results saved in (.ris)/(.csv) formats to ensure organized record-keeping. Additionally, screenshots from the SCOPUS database were taken to visually document the data analysis process, aiding in thorough discussion and analysis. This literature review aims to yield insights into the theme of "critical thinking skills assessment in mathematics," offering a solid foundation for the development of assessment instruments tailored to mathematics education by cross-referencing with other pertinent literature.

RESULTS AND DISCUSSION

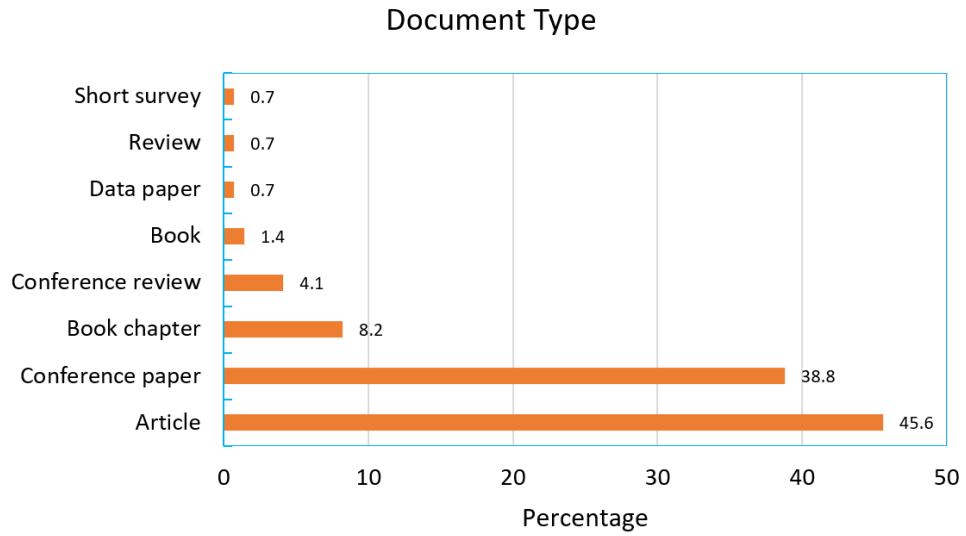
The outcomes of the document search within the SCOPUS database using the search terms "critical thinking skills assessment in mathematics" [TITLE-ABS-KEY (critical AND thinking AND skills AND assessment AND in AND mathematics)] are depicted in Figure 2. During this initial identification phase, documents were not filtered by any criteria such as publication year, subject area, type of document, stage of publication, title of source, keywords, type of source, or any other parameters.

The analysis depicted in Figure 2 reveals a comprehensive overview of the literature spanning from 1965 to 2023, with a total of one hundred and forty-seven documents identified using specific search keywords. This collection encompasses a diverse array of document types, including journal articles, which constitute 45.60% of the total, conference papers at 38.80%, book chapters making up 8.20%, and other forms of publications such as conference reviews, books, data papers, reviews, and short surveys contributing to the remainder.

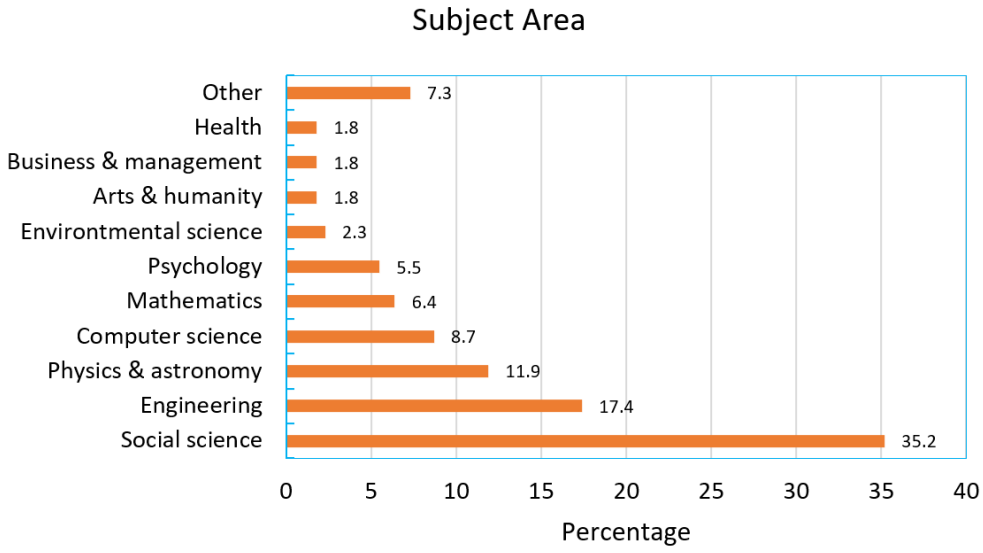
The subject areas covered by these documents are equally varied, with social science leading at 35.20%, followed by engineering at 17.40%, physics and astronomy at 11.90%, among others. Notably, the proportion of documents specifically related to mathematics stands at just 6.40%, indicating a relatively low volume of studies focused on the assessment of critical thinking skills within the field of mathematics as cataloged in the SCOPUS database. This finding underscores the limited research attention directed toward critical thinking skill assessment in mathematics, highlighting a significant gap in the literature.



(a)



(b)



(c)

Figure 2. Document identification results based on (a) all years, (b) all types, and (c) all subject areas

The underrepresentation of mathematics in the context of critical thinking skills assessment suggests a ripe area for future research, especially in the development of specialized assessment instruments tailored for mathematics learning. The distribution of documents across various subject areas, with a scant percentage dedicated to mathematics, points to the necessity for a more focused investigation into how critical thinking skills are evaluated within this discipline. This gap presents an opportunity for researchers and educators alike to contribute to the expansion of this crucial area of study. By developing and refining assessment tools that accurately measure critical thinking skills in mathematics, the academic community can better understand and enhance the ways in which these skills are cultivated and evaluated in educational settings. This expansion is not only vital for advancing the field of mathematics education but also for ensuring that students are equipped with the critical thinking abilities essential for success in a rapidly evolving and increasingly complex world.

The resulting data from the identified documents is then visualized using VOSviewer software. The results are presented in Figure 3 and Figure 4. The Figure 3 presented is a bibliometric visualization from VOSviewer, showcasing the interconnectivity and thematic clusters within the body of research concerning the assessment of critical thinking skills in mathematics. The various terms represented as nodes suggest key focus areas within this academic discourse. Large nodes such as "students," "critical thinking," and "education" indicate these are central concepts in the literature, with a high frequency of occurrence. The clusters, delineated by different colors, group together terms that are often referenced together in the literature, implying a thematic relationship. For example, "critical thinking" and "higher-order thinking skills" are close to each other, suggesting that research on critical thinking in mathematics often involves discussions on advanced cognitive processes.

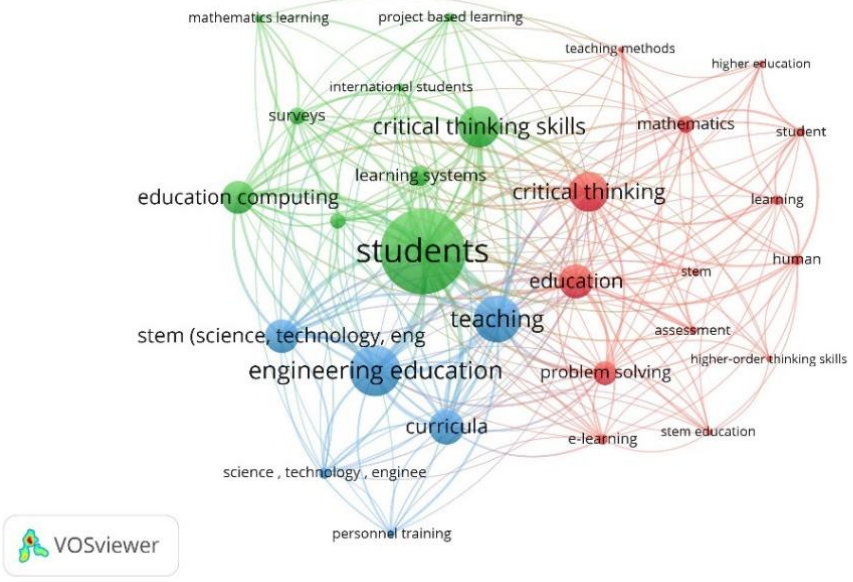


Figure 3. Visualization of networks using VOSviewer software

The lines connecting the nodes (Figure 3) represent the strength and frequency of the relationships between terms across the literature, with thicker lines indicating a stronger or more commonly cited connection. The network map illustrates the interdisciplinary nature of the research, bringing together domains such as "education computing," "engineering education," and "e-learning," and highlighting their relevance to the study of critical thinking in mathematics. The presence of terms like "problem solving" and "learning systems" suggests

a focus on practical applications and methodologies in teaching and learning environments. This visualization allowing researcher to navigate through complex thematic territories, identify central and peripheral research themes, and explore potential gaps or new directions for further investigation in the domain of mathematics education and critical thinking assessment.

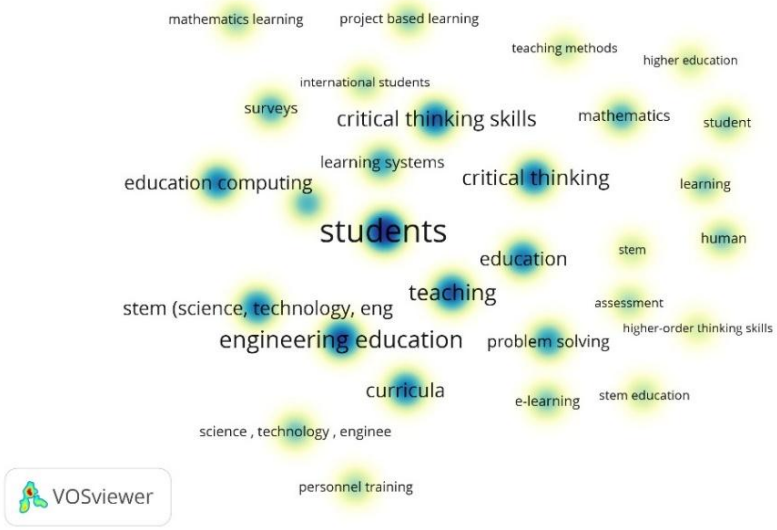


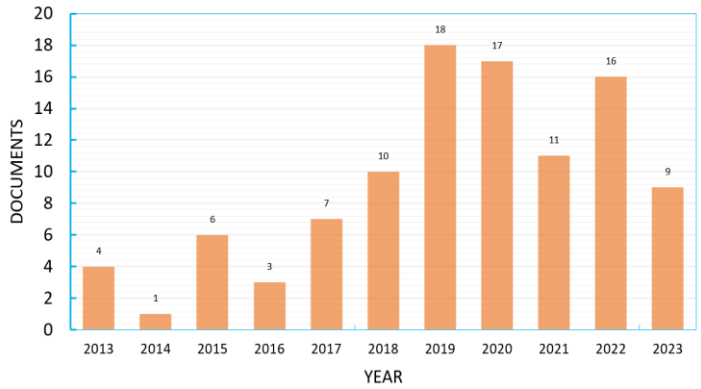
Figure 4. Visualization of density using VOSviewer software

The Figure 4 is another visualization from VOSviewer, depicting a density view which highlights the most prominent and densely connected terms within the same body of research on the assessment of critical thinking skills in mathematics education. In this density map, "students," "education," "teaching," and "critical thinking" are among the most prominent terms, indicating that they are not only frequent but also heavily interconnected within the research. This suggests that studies focusing on critical thinking in mathematics education are particularly concerned with how students are taught and how their critical thinking skills are developed within educational settings. The blue to green gradient areas represent less dense but still relevant terms, which may indicate emerging areas of research or supporting concepts that provide additional context to the central themes. The density visualization complements the network map by providing a heat-mapped overview of the research landscape, allowing for quick identification of the most intensively researched areas and offering a visual summary of the field's thematic concentrations.

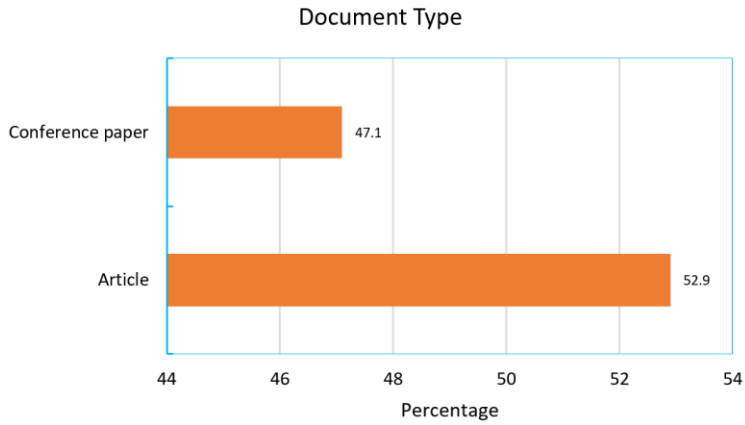
Subsequently, the selection process involves implementing criteria based on the publication year and type of document to guarantee that the materials reviewed in this research represent the most recent findings from the past decade. This approach ensures that the documents considered, specifically journal articles and conference papers, are directly relevant to the study's theme. The categorization of documents following these specified restrictions, including the timeframe and nature of the publications, is illustrated in Figure 5.

The initial phase of document screening, focusing on the publication timeframe of the last decade (2013 to 2023), successfully identified one hundred and twenty-one documents. A subsequent review, narrowing the scope to specific types of documents, yielded one hundred and two documents, categorically split into journal articles (52.90%) and conference papers (47.10%). This meticulous process ensures that the selection is finely tuned to the most recent and relevant forms of scholarly communication within the stipulated period. Following this, a detailed eligibility examination was conducted on the one hundred and two documents to

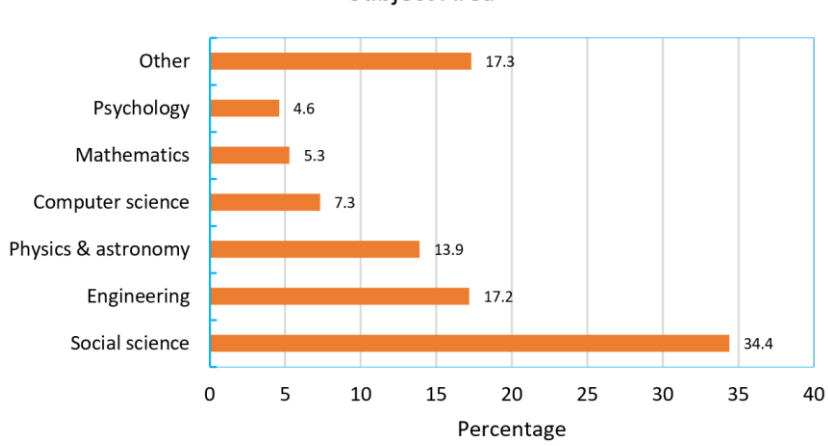
align closely with the study's thematic focus, particularly in the context of mathematics education. This scrutiny led to the identification of nineteen documents, encompassing both journal articles and conference papers, that specifically addressed the targeted area of mathematics. The culmination of this rigorous selection process involved a manual review based on relevance to the chosen keywords and the frequency of citations, ultimately distilling the pool to eleven documents earmarked for in-depth review in this research. This meticulous curation, guided by the principles of the PRISMA method, culminates in the final compilation of documents poised for analysis.



(a)



(b)



(c)

Figure 5. Results from the document filtering process were determined by criteria including: (a) a publication date within the last decade (2013-2023), (b) the specific types of documents, namely articles and conference papers, and (c) encompassing all subject areas.

The literature review conducted on the eleven documents incorporates a series of empirical studies that offer significant insights into the role of mathematics in fostering problem-solving, critical thinking, and analytical skills within educational settings. Highlighted within these discussions is the work of Naidoo and Reddy (2023), who underscore the critical role of technology-enhanced mathematics teaching strategies. These strategies serve as pivotal scaffolds in teaching, learning, and assessment processes, emphasizing the necessity for these elements to be authentic and directly applicable to students' experiences. Furthermore, the study by Evendi et al. (2022) places a strong emphasis on the integration of mathematics teaching within real-world contexts to bolster critical thinking, a competency deemed indispensable for the 21st century. Through their research, Evendi et al. (2022) implemented problem-based learning (PBL) in a remote learning environment, aiming to evaluate students' critical thinking abilities in the context of mathematics education. Their methodology incorporates a rigorous framework for critical thinking derived from seminal theories (e.g., Dewey, 1933; Elder & Paul, 2012; Ennis, 2018), which delineates critical thinking indicators such as analysis, inference, evaluation, and decision-making skills, encapsulated in an essay test format employed in their study.

In contrast, the study conducted by Tanujaya et al. (2021) points out the infrequent explicit programming of higher order thinking skills (HOTS), including critical thinking, in mathematics education by school teachers. Their research delineates critical thinking skills as cognitive abilities characterized by the ability to analyze and evaluate information. Tanujaya et al. (2021) subtly explore the assessment of HOTS based on criteria established by international educational standards such as the Program for International Student Assessment (PISA) and the Trends in Mathematics and Science Study (TIMSS). Furthermore, the development of a HOTS instrument in mathematics, as discussed by Samritin & Suryanto (2016), takes the form of an essay test designed around cognitive aspect indicators: connections, problem-solving, and mathematical reasoning.

Broadening the scope to a more generalized understanding, critical thinking is often intertwined with reasoning capabilities (Ennis, 2015). The research by Krejci et al. (2020) delves into critical thinking as a complex cognitive construct involving multiple interconnected components, including interpretation, explanation, reasoning, evaluation, synthesis, reflection, judgment, metacognition, and self-regulation. These components are similarly examined in the work of Spector and Ma (2019). Studies within the STEM fields also endeavor to quantify critical thinking through specific indicators aligned with Ennis (2011), as demonstrated by the research conducted by Widiyawati et al., (2020). Moreover, the exploration of mathematical reasoning by Papic (2015) through the "Early Mathematical Patterning Assessment (EMPA)" instrument illustrates the dual purpose of this tool in enhancing children's mathematical reasoning and their capacity to conceptualize abstract mathematical ideas and relationships. This approach not only advances our understanding of mathematical reasoning but also emphasizes the importance of developing tools that facilitate the conceptualization of mathematical concepts in young learners.

The development of assessments for mathematical critical thinking skills is recognized as a crucial tool for gauging students' success in engaging in critical thought (Tanjung et al., 2020). The literature reveals that there have been efforts to create assessment tools specifically tailored to evaluate mathematical critical thinking within the context of Higher Order Thinking Skills (HOTS). For instance, the study by Tanjung et al. (2020) highlights the construction of an instrument grounded in the revised Bloom's cognitive theory (Krathwohl, 2002), covering three domains: analyze (C4), evaluate (C5), and create (C6). This emphasis on cognitive processes underscores the importance of developing comprehensive assessment

tools that can accurately measure the depth and breadth of students' critical thinking in mathematics. In interdisciplinary areas such as STEM, critical thinking is notably enhanced by project-based learning approaches (Sontgerath & Meadows, 2018). The research conducted by Sontgerath and Meadows (2018) adopted critical thinking indicators from the "Holistic Student Assessment (HAS)" developed by the PEAR Institute (Partnerships in Education and Resilience), which include the examination of information, exploration of ideas, and independent thought. These indicators provide a nuanced framework for assessing critical thinking, emphasizing the integration of knowledge exploration and the autonomy of thought in learning processes.

Furthermore, in other research, the measurement of critical thinking falls within the realm of higher-order cognition, encompassing skills such as application, synthesis, and evaluation (Kimmel et al., 2014). The study by Abosalem (2016) delineates four levels of HOTS, namely application, analysis, synthesis, and evaluation, which serve as domains within critical thinking skills. These levels offer a structured approach to identifying indicators that can effectively measure critical thinking in mathematics education. The assessment techniques employed to evaluate critical thinking in mathematics range from traditional assessments (paper and pencil tests) to performance assessments (Abosalem, 2016). This diversity in assessment methods reflects the multifaceted nature of critical thinking and the need for varied approaches to accurately assess these skills in educational settings. By employing a broad range of assessment tools, educators can obtain a more comprehensive understanding of students' abilities to apply critical thinking in mathematical contexts, thereby enhancing instructional strategies to better support the development of these essential skills.

Contemporary mathematics education emphasizes the enhancement of students' conceptual abilities across three key domains: critical thinking, modeling, and the application of mathematical concepts. Given the dynamic and interactive nature of mathematics, which inherently fosters higher-order thinking and complex problem-solving skills, there exists a pronounced need for assessment methods that specifically target the development of students' conceptual understanding. The research conducted by De Zeeuw et al. (2013) critically evaluates the effectiveness of current assessment tools in capturing the depth of students' mathematical conceptual skills. They argue that most existing instruments fall short in this regard and propose the utilization of “NetLogo Hotlink Replay” software as a novel approach for evaluating students' conceptual prowess in mathematics. However, their research stops short of providing an in-depth exploration of how the software functions or its application in educational settings, leaving a gap in understanding its potential benefits and limitations.

The scholarly discourse surrounding mathematics education consistently highlights the critical role of thinking skills and the diverse methodologies and tools available for assessing these competencies within the mathematical context. Yet, there persists a notable ambiguity regarding the nature of critical thinking itself, attributed to the myriad of theoretical perspectives that offer varying definitions and criteria for its assessment. This ambiguity is compounded by the challenges outlined in research by Verawati et al. (2020), which include the difficulty in categorizing critical thinking as either a general cognitive skill or one that is subject-specific, as well as the complexities involved in measuring the transferability of critical thinking skills across different subject areas. Such transferability is crucial for understanding how critical thinking developed in mathematics can be applied in other contexts, yet the specificity of knowledge required for critical thinking in mathematics poses significant challenges for cross-disciplinary application. These issues underscore the necessity for a more

refined understanding and evaluation of critical thinking within the sphere of mathematics education.

Therefore, there is a pressing need to define critical thinking more precisely within the context of mathematics and to devise assessment tools specifically designed to evaluate these skills. The development of such instruments would not only facilitate a deeper understanding of students' critical thinking abilities in mathematical scenarios but also enhance the ability of educators to tailor instructional strategies that effectively nurture these competencies. As the field of mathematics education continues to evolve, addressing these challenges will be essential for fostering an educational environment that truly enhances students' ability to think critically, model complex problems, and apply mathematical concepts in diverse and meaningful ways. This endeavor requires a concerted effort from educators, researchers, and curriculum developers to bridge the gap between theoretical conceptualizations of critical thinking and practical assessment methodologies in mathematics education.

CONCLUSION

The comprehensive literature review conducted using the PRISMA method has provided valuable insights into the theme of assessing critical thinking skills in mathematics education. By synthesizing the findings of current literature studies obtained from the extensive SCOPUS database alongside insights from other empirical research, it becomes evident that critical thinking holds a paramount role in the realm of mathematics learning. The reviewed studies collectively emphasize the significance of nurturing critical thinking abilities among students, as it is closely intertwined with their success in mathematics education. These findings reaffirm the notion that critical thinking represents a fundamental cognitive skill that underpins the development of problem-solving capabilities, analytical reasoning, and effective decision-making in mathematical contexts.

However, despite the consensus on the importance of critical thinking in mathematics education, the literature also highlights a significant challenge: the lack of a universally accepted definition and standardized measurement framework for critical thinking, particularly within the domain of mathematics. This issue stems from the multitude of theories and perspectives on critical thinking, each offering unique viewpoints and criteria for assessment. Consequently, the absence of a clear and specific definition of critical thinking in the context of mathematics education has led to variations in the standards used to measure and assess this skill. These variations pose a substantial obstacle in ensuring consistency and accuracy in evaluating students' critical thinking abilities in mathematics.

LIMITATION

One limitation of this review is the reliance on literature available up to the year 2023. The field of mathematics education is continuously evolving, and new assessment tools and methodologies may have emerged since then. Additionally, the review focused on documents available in the SCOPUS database, which may not encompass all relevant literature on this topic. The limited representation of mathematics-focused documents within the database highlights the need for a more extensive exploration of critical thinking assessment tools in mathematics education. Furthermore, the review acknowledges that the selected documents encompass a wide range of subject areas and interdisciplinary approaches, which may introduce variability in the assessment methodologies discussed.

RECOMMENDATION

For future research in the domain of assessing critical thinking skills in mathematics education, it is recommended to prioritize the development of a universally accepted

definition of critical thinking within the context of mathematics. Additionally, researchers should focus on creating standardized assessment criteria or frameworks that can be applied across various mathematical topics and grade levels. Longitudinal studies tracking the progression of critical thinking skills in students, cross-cultural investigations, and interdisciplinary collaborations with fields like psychology and cognitive science are essential for a holistic understanding of critical thinking in mathematics. Moreover, exploring the integration of technology, ensuring assessment validity and reliability, and addressing the needs of diverse student populations should be central themes in future research endeavors. Finally, conducting meta-analyses and systematic reviews can help consolidate existing knowledge and guide evidence-based practices in mathematics education.

Author Contributions

The authors have sufficiently contributed to the study, and have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Acknowledgement

We would like to express our heartfelt gratitude to all the researchers, scholars, and educators whose work has contributed to the body of knowledge on critical thinking in mathematics education. Their valuable insights and dedicated efforts have been instrumental in shaping this literature review.

Declaration of Interest

The authors declare no conflict of interest.

REFERENCES

Abosalem, Y. (2016). Assessment Techniques and Students' Higher-Order Thinking Skills. *International Journal of Secondary Education*, *4*(1), 1. https://doi.org/10.11648/j.ijsedu.20160401.11

Animasaun, I. L., & Abegunrin, O. A. (2017). Gender difference, self-efficacy, active learning strategies and academic achievement of undergraduate students in the Department of Mathematical Sciences, Federal University of Technology, Akure, Nigeria. *International Journal of Teaching and Case Studies*, *8*(4), 255. https://doi.org/10.1504/IJTCS.2017.088929

De Zeeuw, A., Craig, T., & You, H. S. (2013). Assessing conceptual understanding in mathematics. *2013 IEEE Frontiers in Education Conference (FIE)*, 1742–1744. https://doi.org/10.1109/FIE.2013.6685135

Dewey, J. (1933). *How We Think: A Restatement of the Relation of Reflective Thinking to the Educative Process*. D.C. Heath & Co Publishers.

Elder, L., & Paul, R. (2012). *The thinker's guide to intellectual standards: The words that name them and the criteria that define them*. Foundation for Critical Thinking Press.

Ennis, R. (2018). Critical Thinking Across the Curriculum: A Vision. *Topoi*, *37*(1), Article 1.

Ennis, R. H. (2011). The nature of critical thinking: An outline of critical thinking dispositions and abilities. *Inquiry: Critical Thinking Across the Disciplines*, *26*(2), Article 2. https://doi.org/10.5840/inquiryctnews201126214

Ennis, R. H. (2015). Critical Thinking: A Streamlined Conception. In M. Davies & R. Barnett (Eds.), *The Palgrave Handbook of Critical Thinking in Higher Education* (pp. 31–47). Palgrave Macmillan US. https://doi.org/10.1057/9781137378057_2

Ennis, R. H., Millman, J., & Tomko, T. N. (1985). *Cornell Critical Thinking Tests*. Midwest Publications.

Ennis, R. H., & Weir, E. (1985). *The Ennis-Weir Critical Thinking Essay Test*. Midwest Publication.

Erikson, M. G., & Erikson, M. (2019). Learning outcomes and critical thinking – good intentions in conflict. *Studies in Higher Education*, *44*(12), Article 12. https://doi.org/10.1080/03075079.2018.1486813

Evendi, E., Kusaeri, A. K. A., Pardi, M. H. H., Sucipto, L., Bayani, F., & Prayogi, S. (2022). Assessing students' critical thinking skills viewed from cognitive style: Study on implementation of problem-based e-learning model in mathematics courses. *Eurasia Journal of Mathematics, Science and Technology Education*, *18*(7), em2129. https://doi.org/10.29333/ejmste/12161

Evendi, E., & Verawati, N. N. S. P. (2021). Evaluation of Student Learning Outcomes in Problem-Based Learning: Study of Its Implementation and Reflection of Successful Factors. *Jurnal Penelitian Pendidikan IPA*, *7*(SpecialIssue), 69–76. https://doi.org/10.29303/jppipa.v7iSpecialIssue.1099

Facione, N. C., Facione, P. A., & Sanchez, C. A. (1994). Critical Thinking Disposition as a Measure of Competent Clinical Judgment: The Development of the California Critical Thinking Disposition Inventory. *Journal of Nursing Education*, *33*(8), 345–350. https://doi.org/10.3928/0148-4834-19941001-05

Facione, P. A. (1990). *The California Critical Thinking Skills Test—College Level. Technical Report #1. Experimental Validation and Content Validity*. California Academic Press, 217 La Cruz, Millbrae, CA 94030. https://eric.ed.gov/?id=ED327549

Faradillah, A., & Adlina, S. (2021). Validity of critical thinking skills instrument on prospective Mathematics teachers. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *25*(2). https://doi.org/10.21831/pep.v25i2.40662

Gultom, C., Triyanto, T., & Saputro, D. (2022). Students' mathematical reasoning skills in solving mathematical problems. *Jurnal Pendidikan Indonesia, 11*(3), 542-551. https://doi.org/10.23887/jpiundiksha.v11i3.42073

Halpern, D. F. (2010). *Halpern Critical Thinking Assessment Manual*. Schuhfried GmbH.

Herpiana, R., & Rosidin, U. (2018). Development of instrument for assessing students' critical and creative thinking ability. *Journal of Physics: Conference Series*, *948*, 012054. https://doi.org/10.1088/1742-6596/948/1/012054

Hidayah, I., Kusmayadı, T., & Fitriana, L. (2021). Minimum competency assessment (akm): an effort to photograph numeracy. *Journal of Mathematics and Mathematics Education, 11*(1), 14. https://doi.org/10.20961/jmme.v11i1.52742

Hong, Z., Yuan, R., & He, X. (2020). Investigating university efl teachers' perceptions of critical thinking and its teaching: voices from china. *The Asia-Pacific Education Researcher, 29*(5), 483-493. https://doi.org/10.1007/s40299-020-00500-6

Kimmel, H. S., Burr-Alexander, L. E., Hirsch, L., Rockland, R. H., Carpinelli, J. D., & Aloia, M. (2014). Pathways to effective K-12 STEM programs. *2014 IEEE Frontiers in Education Conference (FIE) Proceedings*, 1–6. https://doi.org/10.1109/FIE.2014.7044362

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, *41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Krejci, S. E., Ramroop-Butts, S., Torres, H. N., & Isokpehi, R. D. (2020). Visual Literacy Intervention for Improving Undergraduate Student Critical Thinking of Global Sustainability Issues. *Sustainability*, *12*(23), 10209. https://doi.org/10.3390/su122310209

Lai, E. R. (2011). *Critical thinking: A literature review* (pp. 1–50). https://images.pearsonassessments.com/images/tmrs/CriticalThinkingReviewFINAL.pdf

Leach, S. M., Immekus, J. C., French, B. F., & Hand, B. (2020). The factorial validity of the Cornell Critical Thinking Tests: A multi-analytic approach. *Thinking Skills and Creativity*, *37*, 100676. https://doi.org/10.1016/j.tsc.2020.100676

Liu, O. L., Frankel, L., & Roohr, K. C. (2014). Assessing Critical Thinking in Higher Education: Current State and Directions for Next-Generation Assessment: Assessing Critical Thinking in Higher Education. *ETS Research Report Series*, *2014*(1), 1–23. https://doi.org/10.1002/ets2.12009

Maulyda, M. A. (2020). *Mathematics Learning Paradigm based on NCTM*. CV IRDH.

Montgomery, C., & Fernández-Cárdenas, J. M. (2018). Teaching STEM education through dialogue and transformative learning: Global significance and local interactions in Mexico and the UK. *Journal of Education for Teaching*, *44*(1), 2–13. https://doi.org/10.1080/02607476.2018.1422606

Naidoo, J., & Reddy, S. (2023). Embedding Sustainable Mathematics Higher Education in the Fourth Industrial Revolution Era Post-COVID-19: Exploring Technology-Based Teaching Methods. *Sustainability*, *15*(12), 9692. https://doi.org/10.3390/su15129692

National Council of Teachers of Mathematics. (2000). *Principles, Standards, and Expectations—National Council of Teachers of Mathematics*. https://www.nctm.org/Standards-and-Positions/Principles-and-Standards/Principles,-Standards,-and-Expectations/

O'Hare, L., & McGuinness, C. (2015). The validity of critical thinking tests for predicting degree performance: A longitudinal study. *International Journal of Educational Research*, *72*, 162–172. https://doi.org/10.1016/j.ijer.2015.06.004

Onion, A. J. (2004). What use is maths to me? A report on the outcomes from student focus groups. *Teaching Mathematics and Its Applications*, *23*(4), 189–194. https://doi.org/10.1093/teamat/23.4.189

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Systematic Reviews*, *10*(1), 89. https://doi.org/10.1186/s13643-021-01626-4

Papic, M. (2015). An Early Mathematical Patterning Assessment: Identifying young Australian Indigenous children's patterning skills. *Mathematics Education Research Journal*, *27*(4), 519–534. https://doi.org/10.1007/s13394-015-0149-8

Prayogi, S., Yuanita, L., & Wasis. (2018). Critical Inquiry Based Learning: A Model of Learning to Promote Critical Thinking Among Prospective Teachers of Physic. *Journal of Turkish Science Education*, *15*(1), Article 1.

Quinn, S., Hogan, M., Dwyer, C., Finn, P., & Fogarty, E. (2020). Development and Validation of the Student-Educator Negotiated Critical Thinking Dispositions Scale (SENCTDS). *Thinking Skills and Creativity*, *38*, 100710. https://doi.org/10.1016/j.tsc.2020.100710

Salamah, D. P. (2020). Error Analysis based on Newman Error Analysis on Compound Event Probability Material Viewed from Gender and Self Confidence. *JPMI (Jurnal Pembelajaran Matematika Inovatif)*, *3*(4), 273-284. https://doi.org/10.22460/jpmi.v3i4.p%25p

Samritin, S., & Suryanto, S. (2016). Developing an assessment instrument of junior high school students' higher order thinking skills in mathematics. *Research and Evaluation in Education*, *2*(1), 92. https://doi.org/10.21831/reid.v2i1.8268

Sarkingobir, Y., Egbebi, L. F., & Awofala, A. O. A. (2023). Bibliometric Analysis of the Thinking Styles in Math and Its' Implication on Science Learning. *International Journal of Essential Competencies in Education*, *2*(1), 75–87. https://doi.org/10.36312/ijece.v2i1.1391

Sontgerath, S., & Meadows, R. N. (2018). A comparison of changes in science interest and identity and 21st century learning skills in a mixed-gender and single-gender robotics program for elementary/ middle school youth. *CoNECD 2018 - Collaborative Network for Engineering and Computing Diversity Conference*, *164476*, 1–12.

Spector, J. M., & Ma, S. (2019). Inquiry and critical thinking skills for the next generation: From artificial intelligence back to human intelligence. *Smart Learning Environments*, *6*(1), 8, s40561-019-0088-z. https://doi.org/10.1186/s40561-019-0088-z

Sudrajat, A. K. (2018). Development of Formative Assessment on Circulatory System Material to Measure Students' Critical Thinking Abilities. *Journal of Educational Research*, *18*(3), 243–251. https://doi.org/10.17509/jpp.v18i3.15291

Suhirman, S., & Prayogi, S. (2023). Overcoming Challenges in STEM Education: A Literature Review That Leads to Effective Pedagogy in STEM Learning. *Jurnal Penelitian Pendidikan IPA*, *9*(8), 432–443. https://doi.org/10.29303/jppipa.v9i8.4715

Tanjung, H. S., Nababan, S. A., Sa'dijah, C., & Subanji. (2020). Development of Assessment Tools of Critical Thinking in Mathematics in the Context of HOTS. *Advances in Mathematics: Scientific Journal*, *9*(10), 8659–8667. https://doi.org/10.37418/amsj.9.10.91

Tanujaya, B., Indra Prahmana, R. C., & Mumu, J. (2021). Mathematics Instruction to Promote Mathematics Higher-Order Thinking Skills of Students in Indonesia: Moving Forward. *TEM Journal*, 1945–1954. https://doi.org/10.18421/TEM104-60

Tsai, K. (2012). Dance with critical thinking and creative thinking in the classroom. *Journal of Sociological Research, 3*(2). https://doi.org/10.5296/jsr.v3i2.2323

Verawati, N. N. S. P., Prayogi, S., Yusup, M. Y., & Taha, H. (2020). Development of the Test Instrument for Measuring Students' Critical Thinking Abilities on Fluid Material. *Prisma Sains*, *8*(1), 46. https://doi.org/10.33394/j-ps.v8i1.2487

Watson, G., & Glaser, E. M. (1980). *Watson-Glaser Critical Thinking Appraisal, forms A and B manual*. The Psychological Corporation.

Werner, P. H. (1991). The Ennis-Weir Critical Thinking Essay Test: An Instrument for Testing and Teaching (Test Review). *Journal of Reading*, *34*(6), 494–495.

Widiyawati, Y., Nurwahidah, I., Sari, D. S., Masykuri, M., & Budiyanto, C. W. (2020). STEM-Project Based Learning Integration On Bioacoustics Worksheet To Enhance Critical Thinking Skills. *Proceedings of the 4th International Conference on Learning Innovation and Quality Education*, 1–7. https://doi.org/10.1145/3452144.3452279

Williams, L. & Lahman, M. (2011). Online discussion, student engagement, and critical thinking. *Journal of Political Science Education, 7*(2), 143-162. https://doi.org/10.1080/15512169.2011.564919

Wiliam, D. (2013). Assessment: The Bridge between Teaching and Learning. *Voices from the Middle*, *21*(22), 15–20.

Wirzal, M. D. H., Nordin, N. A. H. M., Bustam, M. A., & Joselevich, M. (2022). Bibliometric Analysis of Research on Scientific Literacy between 2018 and 2022: Science Education Subject. *International Journal of Essential Competencies in Education*, *1*(2), 69–83. https://doi.org/10.36312/ijece.v1i2.1070