Beta: Jurnal Tadris Matematika, 16(2) 2023: 85-97
DOI 10.20414/betajtm.v16i2.592

Research articles

Assessment of critical mathematical thinking skills: A literature review

Erpin Evendi[12], Heri Retnawati[3], Edi Istiyono[3]

Abstrak Studi ini bertujuan meninjau sejumlah literatur mengenai asesmen terhadap keterampilan berpikir kritis dalam pembelajaran matematika. Studi ini menggunakan metode *Preferred Reporting Items for Systematic reviews and Meta-Analyses* (PRISMA), yang meliputi empat tahap: *identification*, *screening*, *eligibility*, dan *inclusion*. Seleksi dokumen mengambil sumber dari database SCOPUS, yang merupakan salah satu sistem pengindeksan internasional paling akurat. Detail seleksi dokumen dilakukan sesuai urutan berikut: (1) *identification* dengan membatasi pencarian SCOPUS pada konteks penilaian berpikir kritis dalam pembelajaran matematika menghasilkan 147 dokumen; (2) *screening* pertama berdasarkan pembatasan rentang tahun (10 tahun terakhir: dari 2013 sampai 2023) menghasilkan 121 dokumen dan *screening* kedua berdasarkan tipe dokumen (*journal article* dan *conference paper*) menghasilkan 102 dokumen; (3) pada tahap *eligibility*, ditemukan 19 dokumen yang relevan dengan topik pembelajaran matematika; dan (4) pada tahap *inclusion*, ditemukan 11 dokumen yang diseleksi secara manual berdasarkan kata kunci dan artikel yang telah disitasi dan kemudian menjadi materi tinjauan dalam studi ini. Hasil studi ini menunjukkan pentingnya berpikir kritis dalam pembelajaran matematika, serta menyajikan berbagai pendekatan serta instrumen yang dapat digunakan untuk mengasesmen keterampilan berpikir kritis dalam konteks pembelajaran matematika. Namun, fakta tentang kekhasan dalam berpikir kritis masih belum terselesaikan karena banyaknya teori menurut sudut pandang yang berbeda, sehingga membuat standar pengukuran dan penilaian berpikir kritis dalam pembelajaran matematika pun menjadi berbeda-beda. Oleh karena itu, pendefinisian berpikir kritis pada konteks yang lebih spesifik, seperti pembelajaran matematika, perlu dilakukan, dan instrumen yang digunakan untuk mengasesmen keterampilan berpikir kritis dalam pembelajaran matematika perlu dikembangkan.

Kata kunci *Asesmen, Keterampilan berpikir kritis, Matematika*

Abstract This study aims to review several previous studies concerning the assessment of critical thinking skills in mathematics learning. This study used the method known as Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA), consisting of four stages: *identification*, *screening*, *eligibility*, and *inclusion*. SCOPUS database was used as the main source in the document selection since it is considered one of the most accurate international indexing systems. The detailed document selection was carried out in the following order: (1) the *identification* was performed by restricting the SCOPUS database search on the context of critical thinking assessment in mathematics learning and resulted in 147 documents; (2) the first *screening* based on a restriction of year range (the last 10 years: from 2013 to 2023) resulted in 121 documents and the second *screening* based on the types of documents (journal articles and conference papers) resulted in 102 documents; (3) in the *eligibility* stage, 19 documents were found to be relevant to the subject area of mathematics learning; and (4) in the *inclusion* stage, 11 documents were selected manually based on the keywords and citation and used as the review materials in this study. The existing studies reveal the importance of critical thinking in mathematics learning and present various approaches and instruments that can be adopted to assess the critical thinking skills in the context of mathematics learning. However, the

[1] Graduate School, Educational Research and Evaluation Program, Universitas Negeri Yogyakarta, Indonesia
erpinevendi.2023@student.uny.ac.id
[2] Department of Mathematics Education, Universitas Islam Negeri Mataram, Indonesia
[3] Department of Research and Evaluation in Education, Universitas Negeri Yogyakarta, Indonesia

facts about the peculiarities of critical thinking are still unresolved due to the abundance of theories according to different perspectives, thus causing the emergence of various measurement standards on critical thinking skills in mathematics learning. Therefore, the definition of critical thinking in a more specific context, such as mathematics learning, needs to be established, and the instruments used to assess the critical thinking skills in this specific context need to be developed.

Keywords *Assessment, Critical thinking skills, Mathematics*

Introduction

The modern education system nowadays requires the achievement of competence in critical thinking skills, particularly through mathematics learning (Evendi et al., 2022). For this reason, educational institutions have a fundamental task of developing critical thinking skills in students (Erikson & Erikson, 2019). Mathematics has become one of the most important basic subjects in education because it has a very close connection with daily life, namely it aims to sharpen critical thinking skills towards complex problems. However, the reality is that many students still dislike mathematics because it has been considered a difficult subject (Evendi & Verawati, 2021). This causes students to be less interested and less motivated to study mathematics, eventually resulting in their learning outcome from mathematics still being low (Salamah, 2020) and their ability to think critically being less developed (Evendi et al., 2022).

Critical thinking is an intellectual process within the sphere of a person's cognitive dimension in active reasoning. In essence, critical thinking is a process of reasoning (Elder & Paul, 2008). Based on a definition widely adopted nowadays, critical thinking is a "reasonable and reflective thinking, that is focused on deciding what to believe or do" (Ennis, 2018). Thus, considering the importance of reasoning, the foremost expected outcome in all types of mathematics learning is thinking and reasoning skills (Animasaun & Abegunrin, 2017). It was explicitly stated in a framework of the National Council of Teachers of Mathematics (NCTM) that reasoning is the foundation of mathematics teaching, because it is not enough for students just to understand and remember facts, and the development of critical thinking skills is absolutely necessary for them to succeed in mathematics learning (National Council of Teachers of Mathematics, 2000). According to NCTM, mathematical interpretation involves logical conclusions based on evidence, and this is similar to the concept of critical thinking according to the perspective of other experts (e.g. Dewey, 1933; Elder & Paul, 2008; Ennis, 2018).

The focus on reasoning becomes important in the context of mathematics teaching in the classroom. In practice, it usually depends on the choice of tasks and the valuable learning experience in developing reasoning, including a classroom environment that supports it so that teachers can set up the learning discourse effectively and perform assessments accordingly to monitor the progress of students' reasoning (National Council of Teachers of Mathematics, 2000). Maulyda (2020) stated that each stage of the learning process should be assessed with the aim to measure the success rate of the learning process carried out, as well as the targeted goals. It was also stated in her study that the assessment at each stage should meet certain criteria, as well as the indicators specified as part of the reflection of the learning success. Finally, a student's progress in reasoning or critical thinking can be measured through an assessment of his or her critical level of thinking.

The presence of adequate evaluation instruments supports the learning implementation, but it also poses a challenge to the learning (Suhirman & Prayogi, 2023). Particularly, in achieving students' critical thinking competence, the teacher's ability to carry out the evaluation and utilize the evaluation instruments was highlighted as one of the most determining factors that sustainably support the success of critical thinking measurement (Verawati et al., 2020). Besides, these instruments can also be used to stimulate the development of critical thinking skills in students (Herpiana & Rosidin, 2018). Nonetheless, despite the continuity of training for teachers on critical thinking measurement, a weakness still lies precisely in the ambiguity of the evaluation instruments to measure students' success in critical thinking. As a consequence, it is impossible to distinguish whether teachers measure the success of their students in terms of critical thinking abilities or normal cognitive abilities (Verawati et al., 2020). In other cases, it was also revealed that although the learning process was aimed at training students to develop critical thinking skills, the teacher did not actually use instruments to measure it during the assessment (Sudrajat, 2018). This condition suggests that the impact of the learning process can only be measured through an assessment of the specific aimed learning outcome (Wiliam, 2013).

The assessment instruments to measure critical thinking skills have been extensively developed by many experts and researchers. The Watson Glaser Critical Thinking Appraisal (Watson & Glaser, 1980) was a critical thinking test originally developed to measure the critical thinking skills of workers for the selection and promotion of candidates, particularly in the field of employment or management (O'Hare & McGuinness, 2015). Later on, this test was also adopted in the context of education. Furthermore, the Ennis-Weir Critical Thinking Essay Test (Ennis & Weir, 1985) was used to measure the critical thinking of students based on reading materials in the form of essays. This test was developed in a specific format in the social psychology sciences. Based on the reviews that have been made, this test can be useful as an informal diagnostic instrument, a tool for evaluating learning effectiveness, and at the same time as a teaching material in the subject area of critical thinking (Werner, 1991). In addition, the California Critical Thinking Disposition Inventory (Facione et al., 1994) has been developed to measure the critical thinking disposition of nursing students. Several other tests have also been developed, such as the Cornell Critical Thinking Test (Ennis et al., 1985) and many others.

Previous studies

There are various instruments for measuring critical thinking skills, mainly because of the different definitions of concepts or constructs built by experts pertaining to critical thinking itself (Ennis et al., 1985; Verawati et al., 2020). A previous study by Liu et al. (2014) stated an argument that the measurement of critical thinking still needs to be examined referring to its definition, degree of specificity, and generalization, as well as its practical impact. On the other hand, the assessment instruments for critical thinking have been well developed. However, although these instruments vary greatly, both in their purpose and format, the critical thinking assessments tend to be generalized (Lai, 2011). Specifically, in measuring the critical thinking skills in mathematics, previous studies (Faradillah & Adlina, 2021) revealed that there are currently no appropriate instruments to measure the critical thinking skills. The lack of agreement in the context of critical thinking and the best means to measure it is a problem to be solved (Quinn et al., 2020). Moreover, integrating the fundamental factor structure of the currently existing instruments is also a challenge to overcome (Leach et al., 2020).

Based on that background, a study in the form of literature review on the subject area of critical thinking skills is ultimately important to be carried out, particularly in the context of mathematics learning, as it can provide a strong foundation for understanding, developing, and

implementing more effective assessment practices in educational and mathematics learning contexts. By understanding the relevant previous studies, valid and reliable assessment instruments can be identified in measuring the students' critical thinking skills in mathematics learning. Therefore, this study specifically aims to perform a literature review on the subject area of critical thinking skills assessment in mathematics learning.

Methods

An extensive examination was conducted by selecting relevant empirical studies concerning the assessment of critical thinking skills in mathematics learning. This literature review adopted the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) method (Page et al., 2021). This method was chosen for two important reasons: 1) it can provide a synthesis of the state of knowledge in the field of critical thinking skills in mathematics learning so that the priorities of future studies can be formulated, and 2) it can identify problems in previous studies that need to be corrected in subsequent studies.

The most common PRISMA method comprises four stages, namely *identification, screening, eligibility,* and *inclusion*. The method in the document selection used the keyword "critical thinking skills assessment in mathematics," as shown in Figure 1.

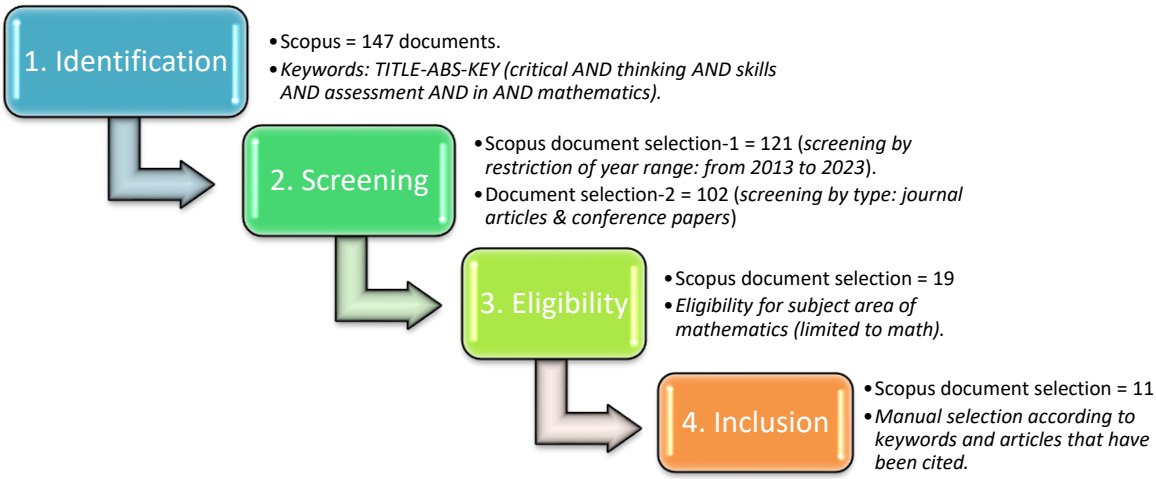

Figure 1. Four stages of document selection using PRISMA method in the literature review on the context of "critical thinking skills assessment in mathematics"

The document selection used SCOPUS database (https://www.scopus.com/), as it is considered one of the most accurate international data indexing systems. SCOPUS is equipped with complete features that allow anyone to browse any academic article based on several categories, such as by author, title, year, publisher, quote, or other metric data accurately and comprehensively.

The *identification* stage through SCOPUS database search was conducted on September 9, 2023. SCOPUS database was explored using the keyword TITLE-ABS-KEY (critical AND thinking AND skills AND assessment AND in AND mathematics) and resulted in 147 documents (all types). Next, a two-step *screening* stage was carried out to ensure that the documents selected were as relevant as possible to the subject area of this study. The first screening was carried out based on a restriction of year range (the last 10 years: from 2013 to 2023) and resulted in 121 documents. The second screening was carried out based on a restriction of the types of documents (article journals and conference papers) and resulted in 102

documents. Next, in the *eligibility* stage, 19 documents were found to be relevant to the subject area of mathematics learning. Lastly, in the *inclusion* stage, 11 documents were selected manually based on the keywords and citation and used as the review materials in this study.

In the next process, bibliometric analysis (Sarkingobir et al., 2023; Wirzal et al. 2022) was carried out using the PRISMA method on each document identified and selected. Through this method, each search result was carefully documented and curated in the form of a file [(.ris)/(.csv)] to keep a systematic record. In addition, the visual representation of the data was captured using the print-screen (prt-scr) of the SCOPUS database, with the aim to facilitate detailed analysis and constructive discussion. In comparison with several other related previous studies, the results of this study are a strong starting point for developing instruments to measure critical thinking skills, especially in mathematics learning.

Findings and Discussion: Theoretical review

Figure 2 (a, b, and c) presents the document identification results based on SCOPUS database search with the keyword "critical thinking skills assessment in mathematics" [TITLE-ABS-KEY (critical AND thinking AND skills AND assessment AND in AND Mathematics)]. In the *identification* stage, the document selection was not restricted based on year, subject areas, types of documents, publication stage, source title, keywords, source type, and others.

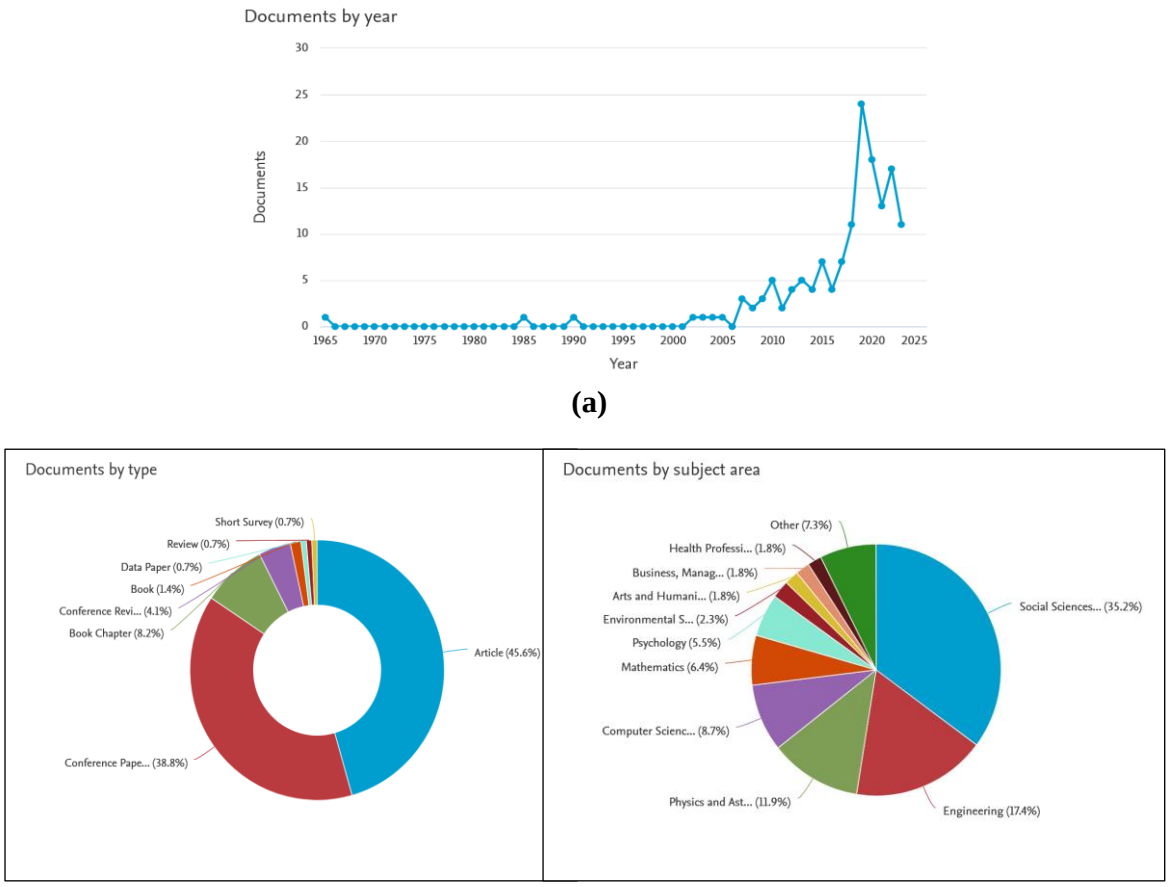

Figure 2. (a) Document identification results based on year (all years), (b) all types of document, and (c) all subject areas.

The results in Figure 2 show that 147 documents were identified in a period of 1965 to 2023 based on search keywords. The types of documents identified were journal articles (45.60%), conference papers (38.80%), book chapters (8.20%), and others, such as conference reviews, books, data papers, reviews, and short surveys. Documents by subject areas identified were in areas related to social science (35.20%), engineering (17.40%), physics and astronomy (11.90%), and others. Meanwhile, the documents identified in the subject area of mathematics were only 6.40%. This suggests that there were not many studies in the SCOPUS database pertaining to critical thinking skills assessment in mathematics. The document distribution based on the subject areas provides an opportunity to investigate the development of critical thinking skills assessment instruments in the context of mathematics learning.

Subsequently, the screening process was carried out by restricting the year range and types of documents. This was to ensure that the documents analyzed or reviewed are the latest results of the study in the last 10 years and that the subject areas found in the journal articles and conference papers are as relevant as possible to the specific context of this study. The distribution of documents based on year restriction and types of documents is presented in Figure 3.

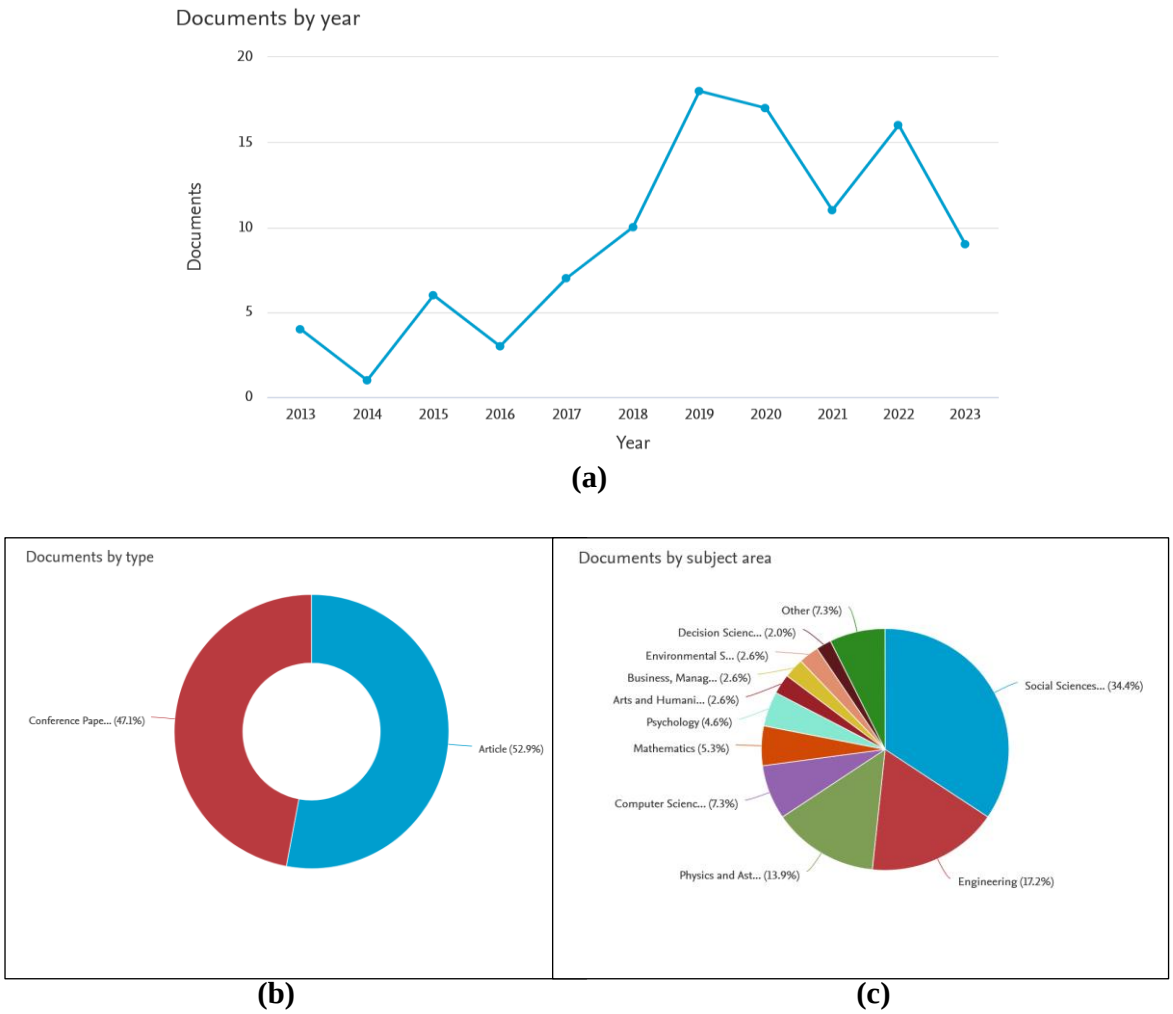

Figure 3. (a) Document screening results based on year range (the last 10 years: 2013–2023), (b) types of documents (journal articles and conference papers), and (c) all subject areas.

The first screening based on a restriction of year range resulted in 121 documents and the second screening based on the types of documents resulted in 102 documents in the form of journal articles (52.90%) and conference papers (47.10%). Next, an eligibility examination was conducted on 102 documents, based on the specific subject area of mathematics learning. As a result, 19 documents in the form of journal articles and conference papers were found. Finally, a manual selection was performed based on the keywords and articles that had been cited. This was the final stage of document selection based on the PRISMA method. This process resulted in 11 final documents in the form of journal article documents and conference papers and used as the review materials in this study, as described in Table 1.

Table 1. Documents in the form of journal articles and conference papers to be reviewed based on the PRISMA method

No.	Title	Author	Source	Evaluation tool developed
1.	Embedding Sustainable Mathematics Higher Education in the Fourth Industial Revolution Era Post-COVID-19: Exploring Technology-Based Teaching Methods	(Naidoo & Reddy, 2023)	Sustainability, MDPI	Explore the experiences, views, implications, and suggestions of participants for technology-based teaching methods for mathematics (problem-solving, critical thinking and analytical skills) in postgraduate students and teachers of mathematics schools in KwaZulu-Natal, South Africa
2.	Assessing Students' Critical Thinking Skills Viewed from Cognitive Style: Study on Implementation of Problem-Based E-Learning Model in Mathematics Course	(Evendi et al., 2022)	EJMSTE, 18(7), em2129	Reviewing cognitive styles by applying a problem-based e-learning model (e-PBL) effectively in improving critical thinking skills of students at mathematics courses in Mataram, Indonesia
3.	Mathematics Instruction to Promote Mathematics Higher-Order Thinking Skills of Students in Indonesia: Moving Forward	(Tanujaya et al., 2021)	TEM Journal	Qualitative research to review literature related to HOTS mathematics courses in Indonesian students
4.	Visual Literacy Intervention for Improving Undergraduate Student Critical Thinking of Global Sustainability Issues	(Krejci et al., 2020)	Sustainability, MDPI	Implementation of visual intelligence education interventions in courses can support the development of advanced metacognitive skills in students with the benefit of improved academic success and critical thinking outside the classroom in education in the United States
5.	STEM-Project Based Learning Integration On Bioacoustics Worksheet to	(Widiyawati et al., 2020)	ACM Int' conf.	Develop a work sheet of bioacoustic students based on STEM-projects to enhance

No.	Title	Author	Source	Evaluation tool developed
	Enhance Critical Thinking Skills			critical thinking skills in Semarang, Indonesia
6.	Development of Assessment Tools of Critical Thinking in Mathematics in the Context of HOTS	(Tanjung et al., 2020)	Adv. in Math.	Develop tools to evaluate critical thinking skills in mathematics learning, in the context of higher thinking skills (henceforth, HOTS) students in Malang, Indonesia
7.	Pathways to Effective K-12 STEM Programs.	(Kimmel et al., 2014)	Frontiers in Edu. Conf.	Develop factual, procedural, conceptual, and meta knowledge of students and teachers, as well as essential skills in problem-solving, analytical, critical thinking, teamwork, and communication on the pre-technical curriculum, Integration of Engineering into Mathematics and Science at the New Jersey Institute of Technology (NJIT)
8.	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	(Sontgerath & Meadows, 2018)	CoNECD Conf. 2018	Measures mathematical science interests in higher education, science identities, and the four learning skills of the 21st century: critical thinking, persistence, relationships with peers, and relationships to adults of the Worcester Polytechnic Institute (WPI)
9.	An Early Mathematical Patterning Assessment: identifiying young Australian Indigenous children's patterning skills	(Papic, 2015)	Math. Edu. Res. J.	Development of the Early Mathematical Patterning Assessment (EMPA) to identify the mathematical thinking of children aged four to five before entering formal school in Australia.
10.	Assessment Techniques and Students' Higher-Order Thinking Skills.	(Abosalem, 2016)	Int. J. Sec. Edu.	Assessment of Higher Order Critical Thinking on mathematics test characteristics made by teachers at Bloom's taxonomic level in the UAE
11.	Assessing Conceptual Understanding in Mathematics.	(De Zeeuw et al., 2013)	Frontiers in Edu. Conf.	Develop NetLogo to measure the ability of students in the United States on calculus subject in applying conceptual understanding (critical thinking, modeling, and content application) of mathematics concepts to scientific phenomena.

A literature review was carried out on 11 documents as displayed in Table 1, in which several relevant empirical studies also contributed to important discussions based on the findings emphasized in these studies. Mathematics is an important subject in education, which provides a foundation for problem solving, critical thinking, and analytical skills (Naidoo & Reddy, 2023). A qualitative study by Naidoo and Reddy (2023) emphasized the importance of technology-based mathematics teaching methods as scaffolding to ensure better outcomes in teaching, learning, and assessment. The problem solving, critical thinking, and analytical skills must be authentic (connected to the physical world in the surrounding environment) and can be applied to students. Evendi et al. (2022) emphasized the importance of mathematics teaching in authentic contexts to support critical thinking as an essential skill in the 21st century. In their study, the problem-based learning (PBL) method was applied remotely in the teaching to assess the students' critical thinking skills in mathematics learning. More intensive critical thinking standards were adopted in several studies based on previous theories (e.g. Dewey, 1933; Elder & Paul, 2008; Ennis, 2018) that suggested detailed fragments on critical thinking indicators, such as analyzing, inferencing, evaluating, and decision making skills (Prayogi et al., 2018). These four critical thinking skills indicators were used by Evendi et al. (2022) in their study and formatted in an essay-shaped test.

Tanujaya et al. (2021) highlighted mathematics learning using HOTS (higher order thinking skills) instruments, including critical thinking aspect, that was rarely adopted explicitly by school teachers. Critical thinking skills refer to cognitive abilities based on two main indicators: analyzing and evaluating skills (Tanujaya et al., 2021). The study implicitly discussed how HOTS is organized based on the standards set by the Programme for International Student Assessment (PISA) and Trends in Mathematics and Science Study (TIMSS). In another study (Samritin & Suryanto, 2016), HOTS instruments in mathematics were developed in the form of tests (test essays) based on several indicators pertaining to cognitive aspects: creating connections, problem solving, and mathematical reasoning.

In a more general context, reasoning was identified through critical thinking (Ennis, 2015). A study by Krejci et al. (2020) investigated critical thinking as a cognitive construction in relation to its interrelated elements: interpretation, explanation, reasoning, evaluation, synthesis, reflection, judgment, metacognition, and self-regulation. These indicators were also discussed in a study by Spector and Ma (2019). The study areas related to STEM also measured critical thinking on the corresponding indicators according to Ennis (2011), such as a study by Widiyawati et al. (2020). Moreover, the context of mathematical reasoning was also investigated by Papic (2015) using an instrument he developed known as "Early Mathematical Patterning Assessment (EMPA)". Based on his study, the instrument was useful to identify the development of children's mathematical reasoning and their ability to abstract mathematical ideas and relationships.

Assessment of critical thinking skills in mathematics learning is essential to be developed for measuring the extent of students' success in critical thinking (Tanjung et al., 2020). The study developed an instrument for assessing critical thinking skills in mathematics learning in the context of HOTS. In another study, Krathwohl (2002) precisely developed an instrument developed based on Bloom's cognitive theory that has been revised in three domains: analyze (C4), evaluate (C5), and create (C6). In interdisciplinary areas, such as STEM, critical thinking was influenced by project-based learning (Sontgerath & Meadows, 2018). Likewise, a study by Sontgereth and Meadow (2018) used the critical-thinking indicators "Holistic Student

Assessment (HAS)" developed by the PEAR (Partnerships in Education and Resilience) Institute. These indicators comprised examination of information, exploration of ideas, and independent thinking. Meanwhile, in another study, the area of critical thinking measurement was included in the sphere of higher-order cognition abilities, involving applying skills, synthesis, and evaluation (Kimmel et al., 2014). A study by Abosalem (2016) emphasized the four levels of HOTS, namely application, analysis, synthesis, and evaluation, as a domain in critical thinking skills. It was also stated in his study that the assessment techniques to measure critical thinking in mathematics learning can be conducted through both traditional assessments (paper and pencil tests) and performance (practice) assessments.

The modern mathematics learning studies focus on student conceptual skills covering three areas, namely critical thinking, modeling, and content application. Considering that mathematics learning is an active process that promotes high-level thinking and problem-solving, an assessment that focuses on the growth of conceptual understanding is required (De Zeeuw et al., 2013). The study acknowledged that the existing assessment instruments currently lead to a lack of measurement of student's conceptual mathematics skills. In the study, they recommended the use of "NetLogo Hotlink Replay" software as an instrument to match students' conceptual skills in mathematics. However, they did not provide a detailed description regarding this software.

This literature review of relevant empirical studies reveals the importance of critical thinking skills in mathematics learning and the various approaches and instruments that can be adopted to assess the skills in the context of mathematics education. However, we argue that the facts about the peculiarities of critical thinking are still unresolved because of the abundance of theories from different perspectives, thus causing the emergence of different standards of measurement and assessment of critical thinking skills. This fact was also revealed in a previous study (Verawati et al., 2020), based on two grounds: First, the type of conclusion is still unclear to the extent that researchers cannot agree whether critical thinking is a general or specific subject (specifically according to the subject matter of learning). Second, it is difficult to assess the transfer of critical thinking to another context, because transfer to another context can be different from the specificity of knowledge in critical thinking. Besides, as mentioned before, the theories and perspectives built by previous researchers and experts to identify critical thinking skills were quite different from each other. Therefore, it is going to be difficult in practice to describe the separate cases of critical thinking in different contexts in order to determine the appropriate measurement method.

Conclusion

A study in the form of literature review has been conducted using the PRISMA method concerning the assessment of critical thinking skills in the subject area of mathematics learning. Comparing the findings of current studies from the SCOPUS database along with other empirical studies, it can be concluded that the existing studies reveal the importance of critical thinking skills in mathematics learning and suggest various approaches and instruments that can be adopted to measure critical thinking skills in the context of mathematics education. However, the facts about the peculiarities of critical thinking are still unresolved due to the abundance of theories according to different perspectives, thus causing the emergence of different measurement standards on critical thinking skills in mathematics learning. Therefore, the definition of critical thinking in a more specific context, such as mathematics learning, needs to

be established, and the instruments used to measure critical thinking skills in this context need to be developed.

References

Abosalem, Y. (2016). Assessment techniques and students' higher-order thinking skills. *International Journal of Secondary Education*, *4*(1), 1-11. https://doi.org/10.11648/j.ijsedu.20160401.11

Anggraini, S., Setyaningrum, W., Retnawati, H., & Marsigit. (2020). How to improve critical thinking skills and spatial reasoning with augmented reality in mathematics learning? *Journal of Physics: Conference Series*, *1581*(1), 012066. https://doi.org/10.1088/1742-6596/1581/1/012066

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*. Boston, MA: D.C. Heath & Co Publishers.

Elder, L., & Paul, R. (2008). *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. H. (2018). Critical thinking across the curriculum: A vision. *Topoi*, *37*(1), 165-184. https://doi.org/10.1007/s11245-016-9401-4

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), 4. 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*(Special Issue), 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

Faradillah, A., & Adlina, S. (2021). Validity of critical thinking skills instrument on prospective Mathematics teachers. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *25*(2), 126-137. https://doi.org/10.21831/pep.v25i2.40662

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

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://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=b42cffa5a2ad63a31fcf99869e7cb8ef72b44374

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). Paradigma pembelajaran matematika berbasis NCTM. CV IRDH Malang.

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

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), 43-56. http://dx.doi.org/10.12973/tused.10220a

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). Analisis kesalahan berdasarkan newman error analysis terhadap materi peluang kejadian majemuk ditinjau dari gender dan self confidence pada siswa elas XII SMK di bandung barat. *JPMI (Jurnal Pembelajaran Matematika Inovatif)*, *3*(4), Article 4. 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), 92-107. 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. https://api.semanticscholar.org/CorpusID:55421583

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*(8), 1-11. https://doi.org/10.1186/s40561-019-0088-z

Sudrajat, A. K. (2018). Pengembangan asesmen formatif pada materi sistem sirkulasi untuk mengukur kemampuan berpikir kritis siswa kelas XI SMA laboratorium UM. *Jurnal Penelitian Pendidikan, 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*, *10*(4), 1945–1954. https://doi.org/10.18421/TEM104-60

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 : Jurnal Pengkajian Ilmu Dan Pembelajaran Matematika Dan IPA IKIP Mataram*, *8*(1), 46-56. 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 New York.

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. https://api.semanticscholar.org/CorpusID:140788435

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*, *137*, 1–7. https://doi.org/10.1145/3452144.3452279

Wiliam, D. (2013). Assessment: The bridge between teaching and learning. *Voices from the Middle*, *21*(2), 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