Science Education and Application Journal (SEAJ)
Program Studi Pendidikan IPA
Universitas Islam Lamongan
http://jurnalpendidikan.unisla.ac.id/index.php/SEAJ

September, 2025. Vol. 7, No. 2,
p-ISSN: 2656-6672
e-ISSN: 2656-8365
pp. 210-220

A Systematic Review of Critical Thinking Assessment Instruments in Junior High School Science Education: Trends, Effectiveness, and Challenges

[1]Siti Shofa Assyifa'ul Qulbi Barid, [2]Arvidhea Safira Gunawan, [3]Firdha Yusmar, [4]Soekma Yeni Astuti, [5*]Sri Wahyuni
[1,2,3,5] Department of Science Education, Faculty of Education, Universitas Jember
[4] Department of Television and Film, Faculty of Humanities, Universitas Jember
Email Correspondence: sriwahyuni.fkip@unej.ac.id

Article Info
Article History
Received: September, 10th 2025
Revised: September, 23rd 2025
Accepted: September, 25th 2025
Published: September, 26th 2025
Keywords
Assessment, critical thinking, instruments, sciences

Abstract

A Systematic Review of Critical Thinking Assessment Instruments in Junior High School Science Education: Trends, Effectiveness, and Challenges. *This systematic literature review, conducted using the PRISMA method, synthesizes research on the development of critical thinking assessment instruments in junior high school science education between 2016 and 2025. Secondary data were obtained through the Publish or Perish application with sources from Google Scholar, keywords include science learning, assessment instruments, and critical thinking. By analyzing 7 peer-reviewed national journals, the study evaluates the validity, reliability, and effectiveness of these tools in enhancing students' critical thinking skills. The findings highlight diverse assessment models, their impact on student learning outcomes, and the challenges faced by educators in implementing these tools. This review offers recommendations for improving the design and application of critical thinking assessments in science curricula.*

Citations: Barid, S. S. A. Q., Gunawan, A. S., Yusmar, F., Astuti, S. Y., & Wahyuni, S. (2025). A systematic review of critical thinking assessment instruments in junior high school science education: Trends, effectiveness, and challenges. Science Education and Application Journal (SEAJ), 7(2), 211–221.

INTRODUCTION

The development of the times and technology requires Indonesian society to adapt by improving human resources, including education (Mardhiyah et al., 2021). Based on OECD data from the PISA survey, the literacy level of Indonesian students declined significantly in 2018 compared to 2015. Reading literacy scores fell from 397 to 371, mathematics literacy from 386 to 379, and science literacy from 403 to 396 (Yusmar & Fadilah, 2023). Literacy is not only the ability to read or gather information, but also the ability to understand, evaluate, and use information critically and reflectively. This skill is closely related to critical thinking. The decline in literacy scores indicates that even though access to education is becoming more widespread, the development of students' critical thinking skills still needs to be improved.

In the 21st century, mastery of soft skills has become one of the main goals of education in Indonesia. These skills consist of critical thinking, problem solving, creativity and innovation, digital literacy, experience-based learning, and media and information literacy (Mardhiyah et al., 2021). In the field of education, thinking skills are classified into higher-order thinking and lower-order thinking. Higher-order thinking includes critical thinking skills (Primayana, 2020). Critical thinking itself is the ability to process information systematically in order to arrive at valid facts and conclusions based on research, experience, and observation (Wahyuni et al., 2021). This skill is crucial in helping students face learning challenges, especially in science learning.

Science education has shifted from a model of knowledge transmission to an approach that focuses on developing higher-order cognitive skills. One crucial skill that is needed is critical thinking. The stages of critical thinking consist of clarification, data analysis, inference, evaluation of arguments, and formulation of solutions and conclusions (Ennis, 2011). Critical thinking skills in science learning focus on inquiry-based learning, where students can learn by asking questions, conducting investigations, and making discoveries (Vincent-Lancrin, 2021). Critical thinking skills are important for students because they enable them to view problems from a more scientific perspective (Zulfiana et al., 2023). Thus, when studying science, students no longer just memorize facts, but also think about how to implement facts as solutions to various challenges in science learning (García-Carmona, 2025).

Students' critical thinking skills in solving various challenges in science learning can be measured using assessment instruments. These instruments play an important role in describing the extent to which students in science learning can analyze, evaluate, and solve problems based on scientific concepts. According to (Sugiarti et al., 2017) Assessment instruments are useful for measuring student achievement and the development of critical thinking skills. Assessment instruments can take the form of written tests, essays, portfolios, or interactive digital devices that can systematically measure students' thinking outcomes. The effectiveness of these assessment instruments must also depend on their validity, reliability, and fairness in the context of science learning.

Although critical thinking skills have become an important part of 21st-century learning objectives, especially in science subjects, various studies show that the assessment instruments currently in use still have fundamental weaknesses. One notable weakness is that many instruments still assess only the final result in the form of students' answers, without recording the thought process behind those answers. For example, in physics wave material, the validity indicators for measuring answers to this material can be categorized as high. However, assessment instruments do not yet have a mechanism to directly assess the stages of students' critical thinking when constructing problem-solving processes (Lina & Desnita, 2022). In addition, the assessment instruments that are often used are still frequently found in the form of written tests, such as multiple-choice tests and essay tests. These types of tests are not very contextual and rarely integrate real-life situations, and are not problem-based. In fact, critical thinking in science is closely related to the ability to solve problems contextually, design experiments, and evaluate observation results (Maghfiroh et al., 2023). Assessments are also still paper-based and do not optimize the use of technology. In fact, digital media has great potential to provide interactive simulations and process-based assessments (Berlian et al., 2025).

The development of critical thinking skills assessment instruments in science education at the junior high school level still faces many limitations. Therefore, a study using a Systematic Literature Review approach with the PRISMA method is needed. This research is an important step in identifying trends in the use of critical thinking assessment instruments in recent years, evaluating the effectiveness of these instruments in measuring various aspects of critical thinking skills, and summarizing the challenges and obstacles encountered in their development and application. The period from 2016 to 2025 was chosen to examine the development of critical thinking assessment instruments because this period covers the academic response to the decline in Indonesia's PISA survey results in 2018. This analysis is expected to capture the latest innovations and trends in efforts to improve the quality of critical thinking skills assessment instruments in junior high school science education.

METHODS

This study uses a Systematic Literature Review (SLR) approach to comprehensively examine various scientific articles discussing the development of critical thinking assessment instruments in science learning at the junior high school level. In the process of synthesizing the data obtained, this study applies descriptive narrative analysis techniques, namely by sorting and grouping articles based on similarities in topic, research focus, and findings. Although it did not use formal coding techniques such as thematic analysis, the analysis process was still carried out systematically by examining the content of the articles in depth, then summarizing the findings narratively to identify trends, effectiveness, and challenges in the development of critical thinking assessment instruments.

In the data collection process, this study applied inclusion and exclusion criteria to ensure the relevance and quality of the articles analyzed. The inclusion criteria included: (1) articles indexed in the Google Scholar database, (2) articles published between 2016 and 2025, and (3) articles that specifically discussed the development of critical thinking assessment instruments in science education at the junior high school level. Meanwhile, the exclusion criteria include articles that do not meet the inclusion criteria, articles that discuss topics outside the focus of the study, and assessments at other education levels or fields of study other than science. The process of identifying and selecting articles can be carried out systematically and efficiently by utilizing Publish or Perish software. This application is used to collect scientific articles from the Google Scholar database based on predetermined keywords, as well as facilitating initial screening based on publication year and topic relevance. The use of Publish or Perish can strengthen the validity and reliability of the literature selection process in studies through a more transparent approach.

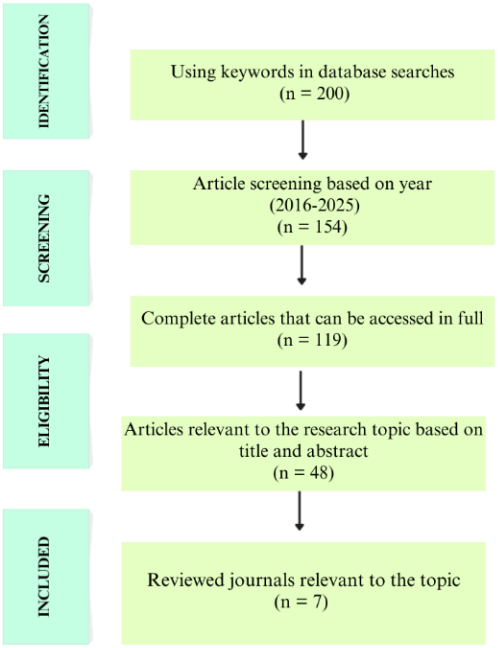

Figure 1. PRISMA Diagram

The PRISMA diagram in Figure 1 shows the article selection process using Publish or Perish with the Google Scholar database, which began based on the keywords used in this study. These keywords consisted of science learning, assessment tools, and critical thinking. Finally, 200 articles were obtained and selected based on the year. The year range used in this study was from 2016 to 2025. This year range was chosen to examine the development of critical thinking assessment instruments because this period covers the academic response to

the decline in the 2018 PISA survey results in Indonesia. After this selection, 154 articles remained. At the eligibility stage, the articles were checked for accessibility, resulting in 119 articles that were fully accessible. Then, the titles and abstracts were reviewed, and 48 articles that met the research focus were obtained. After passing the quality and relevance evaluation stage, seven articles were finally selected for in-depth analysis and used as the basis for this research synthesis.

Data analysis in this study involved a series of systematic steps to transform raw data into meaningful and interpretable information. The analysis was conducted after collecting data from various journals related to the development of assessment instruments for students' critical thinking in science learning, obtained through a literature review. A qualitative approach was used to produce descriptive data in the form of written statements. Research results are written systematically and comprehensively. In addition, there are data visualization steps in the form of data presentation, which consists of a brief descriptive narrative, tables, and the relationship between subjects and their types. After that, conclusions are drawn based on the research findings.

RESULTS AND DISCUSSION

The results of the literature review analysis are listed in Table 1 below.

Author and Year	Object	Lessons	Method	Results
Baharizki, S., Sabtiawan, W.B., and Widodo, W (2021)	Critical Thinking	Additives and Addictives	Research and Development (R&D)	The developed instrument consists of 12 essay questions and is valid (included in the excellent category), reliable (reability value of 0.889), and valid for use. Therefore, the instrument is recommended as a reference for science teachers to measure the critical thinking skills of junior high school students.
Wijaya, U.R.B., Sumarni, W., Haryani, S. (2016)	Critical Thinking	Chemistry	ADDIE Model	Research shows that SETS-based critical thinking instruments on buffer solution material are valid, reliable (0.816) and effective to use. A total of 17 questions were declared valid with good discrimination power and moderate to difficult levels of difficulty. The instruments were able to objectively distinguish students' critical thinking

Author and Year	Object	Lessons	Method	Results
				abilities, but only 11% of students reached the critical category. Student responses were positive, even though the questions were considered quite difficult.
Haryanti., Susongko, P., and Arfiani, Y (2024)	Critical Thinking	Science	Research and Development (R&D)	This instrument can map students' critical thinking skills with an average score of 46.8% (sufficient category). The analysis indicator is in the high category (66.54%), while interpretation, evaluation, and inference are still low, and explanation is in the sufficient category. These findings confirm that although the instrument is suitable for use, these skills need to be improved through more innovative learning strategies.
Trimawati K., Tjandrakirana., and Raharjo (2020)	Critical Thinking	Integrated Sciences	Research and Development (R&D) with modifications to the Dick & Carey model	The PjBL-based assessment instrument for the excretory system material is suitable for use because it meets the criteria of validity (highly valid category), reliability, and practicality. The instrument can improve critical thinking skills to 87.76 and creative thinking skills to 84.85, with both N-gain scores in the high category.
Sudirman, Kistiono, Akhsan, H., and Ariska, M (2020)	Critical Thinking	Concept of electricity	Research Development	This study produced an instrument for assessing science knowledge, attitudes, and skills. Of the 25 knowledge questions, 20 were declared valid (validity score of 86) with a

Author and Year	Object	Lessons	Method	Results
				reliability of 0.55. Teachers also rated the instrument as easy to use because it was in line with the curriculum (score of 4.3).
Dharmawati., Rahayu, S., and Mahanal, S (2016)	Critical Thinking	Interaction between living things and the environment	Research and Development model according to Borg & Gall	The instrument developed using the Borg & Gall model was validated by experts with a validity percentage of 88.35% (highly valid category). The readability analysis of the questions was in the very good category (93.51%), with a reliability score (0.951) in the very good category.
Saputra, I.G.E., Jampel, I.N., and Parwata, I.G.L.A (2022).	Critical Thinking	Vibrations and Waves	R&D with 4D models	Validity testing was conducted using the Content Validity Ratio (CVR), which showed that all items met the validity criteria. Meanwhile, reliability testing showed a result of 0.941. The instrument has good discriminating power, appropriate difficulty, and, based on the results of the Graded Response Model analysis, shows adequate quality.

Research conducted by (Baharizki et al. (2021) successfully developed an instrument for assessing critical thinking skills on the subject of additives and addictive substances, which was found to be valid and reliable. However, this study still has several limitations that open up opportunities for further research. The first limitation is that this study still uses a small-scale trial involving 26 students. This results in the generalization of the results still being relatively low. Another limitation is that the instrument is in the form of essay questions, so it does not cover other forms of assessment, such as performance-based assessment or computer-assisted adaptive tests. The practicality of the instrument has also not been tested in depth, for example, in terms of the time required to complete it, its ease of use by teachers, and student responses. This study also only uses Ennis's indicators of critical thinking, so there is still an opportunity to integrate other, more comprehensive frameworks. For example, Suputra et al. (2023) research uses Facione's more comprehensive critical thinking framework, through six

main skills, namely interpretation, analysis, inference, evaluation, explanation, and self-regulation. This framework does not assess students' cognitive abilities in answering questions, but also emphasizes metacognitive dimensions such as self-regulation.

The research conducted by Wijaya et al., (2017) has limitations in that although the instrument can distinguish skill levels, the results show that only 11% of students reached the critical category. Therefore, it is necessary to integrate the instrument with innovative learning strategies so that it has a greater impact on improving student abilities. The limited sample size of 37 students also resulted in low external validity and generalizability. The instrument is still in the form of essay questions that do not utilize adaptive digital assessment. Similar to Baharizki's research, this study also uses Ennis indicators. Therefore, it can be further developed by enlarging the scale of the trial, integrating assessment technology, and using a more comprehensive framework.

The research by Haryanti et al., (2024) developed a critical thinking assessment instrument based on the Facione framework. The instrument consists of 10 multiple-choice questions analyzed using the Rasch model and was found to be valid and reliable. The use of Rasch analysis has advantages and provides a profile of students' critical thinking skills on the five Facione indicators. The limitations of this study consist of the form of questions, which are only multiple-choice without any variation in authentic assessment. In the future, it is recommended that the instrument use open-ended questions, project-based tasks, or performance-based assessments, which are better able to capture the quality of students' arguments.

Research conducted by Trimawati et al., (2020) shows that the development of an instrument in the form of an essay test can improve students' critical thinking skills from low to higher levels. The sample size for this study was 90 students. The limitation of this study is that the assessment instrument is still in the form of an essay test. This instrument requires a long time to complete and is not efficient for large-scale evaluation. Another limitation is that student response to learning is only moderate (50-60%). This indicates a challenge in familiarizing students with the PjBL model, which requires more time, collaborative skills, and independent learning. In the context of PjBL, teachers play an important role as task designers who adapt instruments to learning objectives. Teachers also act as facilitators of the process, monitoring student work, providing direction, and recording progress. Teachers also act as assessors, using rubrics to assess products and processes, while providing feedback.

The knowledge assessment instrument used in the study by Sudirman et al. (2020) consisted of 25 multiple-choice questions and produced 20 valid items with a reliability coefficient of 0.55. The attitude instrument, which included indicators of honesty, discipline, cooperation, and responsibility, obtained an average validity score of 84, while the skills instrument, in the form of a performance rubric (preparing tools, assembling, using, analyzing, and concluding), obtained an average validity score of 86. In terms of practicality, teachers gave a very good rating (score of 4.3), assessing the instruments as easy to understand, systematic, and relevant to the requirements of the 2013 Curriculum. The limitations of this study are evident in the relatively low reliability of the knowledge questions and the practicality that was only assessed from the teachers' perceptions, without testing time efficiency or inter-rater consistency. Thus, although this instrument has been proven to be valid and practical, further research is needed to improve reliability and explore aspects of usability from the

teachers' perspective so that the instrument can be more sustainable in junior high school science learning practices.

The study conducted by Dharmawati et al., (2016) used instruments developed through the Borg and Gall model and involved 34 students in a limited trial. Meanwhile, in the field stage, 75 students were involved, bringing the total number of students involved in this study to 109. Of the total 38 questions developed, 18 multiple-choice questions and 9 valid questions were obtained and could be used to measure critical thinking skills. The limitations of this study can be seen from the fact that the questions still use multiple-choice and essay formats and do not utilize digital or performance-based assessment formats that can measure critical thinking skills more authentically. Examples of digital assessment tools include Computer-Based Test platforms, Nearpod, Learning Management Systems (LMS), Google Forms, and interactive quiz applications. In a study Ratnady et al., (2024) a CBT-based science learning module on the respiratory system showed very valid results (93.75%), was practical (89.41%), and effectively improved the critical thinking skills of junior high school students. Meanwhile, the use of Nearpod, which is web-based interactive learning, showed validity (81.7%) and practicality (97.66%). Nearpod allows essay-format answers and supports HOTS assessment (Maghfiroh et al., 2023).

The results of Saputra et al., (2022) research shows that the development of assessment instruments for vibration and wave material was successfully carried out using the Research and Development model. This instrument was designed based on critical thinking indicators that refer to Bloom's taxonomy, specifically the dimensions of analysis, evaluation, and synthesis. The instrument items were developed from 12 essay questions, and based on the validity test results through CVR, all items met the validity criteria. The reliability test also showed that the internal consistency was very high, so that the instrument could be used to assess students' critical thinking skills. Item analysis through field trials involved 90 students. After being analyzed using the Graded Response Model, the quality of the instrument can be categorized as adequate in measuring critical thinking skills in the subject of waves and vibrations. In this study, the perspective of teachers as users of the instrument has not been explored in depth. This limitation concerns the aspects of usability, ease of application in the classroom, and adaptation of the instrument to the teaching style of teachers. These aspects can affect the success of the instrument's implementation. Based on research conducted by Hartono et al., (2022) the solution for teachers to get involved and understand the assessment instrument well is to prepare the material thoroughly. For example, by analyzing the basic competencies of HOTS questions, compiling question grids, formulating interesting and contextual stimuli, writing questions in accordance with the question grids, and creating scoring guidelines.

Critical thinking refers to the ability to think at a higher level (higher-order thinking skills), which is important in science learning (Widyapuraya et al., 2023). Critical thinking skills include examining information, analyzing opinions, drawing logical conclusions, and solving problems based on scientific evidence. Teaching and learning science in junior high school requires critical thinking skills to understand scientific concepts, connect various natural phenomena, and develop a scientific attitude (Cahyani et al., 2025). There are several indicators of critical thinking skills, including: identifying problems, collecting and reviewing information, analyzing assumptions and arguments, concluding, and being able to provide alternative solutions. The measurement of critical thinking skills must meet several main criteria, namely validity, reliability, objectivity, and practicality. The Rasch model from one of

the articles reviewed can provide an objective picture of students' results in critical thinking and detect questions that are not in line with students' abilities (Fauzi et al., 2022).

Assessment tools often used to measure critical thinking skills include open-ended tests, which require students to provide scientific reasoning and explain their thought processes; context-based instruments that relate science concepts to real life; and assessment rubrics that evaluate projects based on critical thinking indicators. The development of science assessment instruments is able to consider indicators of critical thinking. The above studies indicate that developed assessment instruments have been found to perform well in measuring the critical thinking skills of students. Nevertheless, there are numerous problems in the implementation of assessment instruments to measure critical thinking in schools, including lack of understanding of how to design and use critical thinking assessment instruments by teachers, lack of time, which frequently leads to the in-depth critical thinking assessment instrument not being used, lack of training and resources in the instrument development process, that is, development in line with student characteristics and the curriculum.

CONCLUSION

Based on several articles that have been analyzed, it was found that the development of critical thinking assessment instruments in junior high school science learning generally succeeded in compiling valid and reliable instruments, both in essay and multiple-choice questions. However, most instruments still do not utilize digital technology, such as CBT or other interactive platforms. Other consistent challenges that emerged were small test sample sizes, monotonous question formats, minimal integration of metacognitive dimensions, and a lack of exploration of teachers' perspectives as users of the instruments. Research instruments can be developed in the future by integrating digital technology and adaptive assessment, varying question formats, and applying a more comprehensive critical thinking framework. Teachers also need to be actively involved to ensure that the instruments are more practical, easy to use, and relevant to teaching styles, so that they can effectively improve the critical thinking skills of junior high school students in science learning.

SUGGESTION

Further research can be conducted by integrating digital technology and artificial intelligence into critical thinking instruments to make them more adaptive and interactive. In addition, teacher training is also needed to improve understanding, skills in using instruments, and adaptation according to teaching styles. Further review research can also expand the scope of literature through international databases such as Scopus and Web of Science, thereby providing more comprehensive and relevant references to improve critical thinking skills.

ACKNOWLEDGEMENT

We, the Community Empowerment Partnership Team (PKM), would like to thank DPPM for the PKM Grant for Fiscal Year 2025 that has been given to us.

REFERENCES

Baharizki, S., Sabtiawan, W. B., & Widodo, W. (2021). Pengembangan Instrumen Penilaian Kemampuan Berpikir Kritis Pada Materi Zat Aditif dan Adiktif. *PENSA: E-JURNAL PENDIDIKAN SAINS*, *9*(3), 304–308. https://ejournal.unesa.ac.id/index.php/pensa/article/view/38778

Berlian, Z., Isnaini, M., Farwati, R., Metafisika, K., Velina, Y., & Fitriyanni, V. (2025). Paper vs. Digital Assessments: Evaluating Critical Thinking on Ecological Issues in Indonesian

Madrasahs. *Journal of Applied Science, Engineering, Technology, and Education*, *7*(1), 41–55. https://doi.org/10.35877/454RI.ASCI3785

Cahyani, D. K., Nugroho, A. S., Nizaruddin, & Hayat, M. S. (2025). Analisis Kemampuan Literasi Sains dan Berpikir Kritis Siswa SMP pada Pembelajaran IPA. *PENDIPA Journal of Science Education*, *8*(3), 593–600. https://doi.org/10.33369/pendipa.8.3.593-600

Dharmawati, D., Rahayu, S., & Mahanal, S. (2016). Pengembangan Instrumen Asesmen Berpikir Kritis untuk Siswa SMP Kelas VII pada Materi Interaksi Makhluk Hidup dengan Lingkungan. *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan*, *1*, 1598–1606.

Ennis, R. H. (2011). An outline of critical thinking dispositions and abilities. *University of Illinois*, *2*(4), 1–8.

Fauzi, A. A., Susongko, P., & Hayati, M. N. (2022). Tes Kemampuan Berpikir Kritis pada Pembelajaran IPA di SMP Berbasis Model Rasch. *PSEJ (Pancasakti Science Education Journal)*, *7*(1), 59–67. https://doi.org/10.24905/psej.v7i1.146

García-Carmona, A. (2025). Scientific Thinking and Critical Thinking in Science Education : Two Distinct but Symbiotically Related Intellectual Processes. *Science and Education*, *34*(1), 227–245. https://doi.org/10.1007/S11191-023-00460-5/TABLES/5

Hartono, H., Makkulau, M., & Samparadja, H. (2022). Upaya Mengatasi Kesulitan Guru dalam Menyusun Instrumen Tes Berpikir Tingkat Tinggi. *Jurnal Pembelajaran Berpikir Matematika*, *7*(1). https://doi.org/10.33772/JPBM.V7I1.25665

Haryanti, H., Susongko, P., & Arfiani, Y. (2024). Pengembangan Instrumen Penilaian Berpikir Kritis Menurut FACIONE Pada Pembelajaran IPA Di Sekolah Menengah Pertama Berbasis Model Rasch. *PSEJ (Pancasakti Science Education Journal)*, *9*(1), 78–87. https://doi.org/10.24905/PSEJ.V9I2.207

Lina, E. Y., & Desnita, D. (2022). Validity and Reliability of Critical Thinking Instruments to Measure the Effectiveness of Context-Based Physics E-Module on Wave Materials. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, *8*(1), 57–64. https://doi.org/10.21009/1.08106

Maghfiroh, R., Wahyuni, S., & Ridlo, Z. R. (2023). The Development of Critical Thinking Skills Assessment Instrument Based on Nearpod in Junior High School Science Learning. *Prisma Sains : Jurnal Pengkajian Ilmu Dan Pembelajaran Matematika Dan IPA IKIP Mataram*, *11*(3), 706–718. https://doi.org/10.33394/J-PS.V11I3.7876

Mardhiyah, R. ., Aldriani, S. N. ., Chitta, F., & Zulfikar, M. R. (2021). Pentingnya Keterampilan Belajar di Abad 21 sebagai Tuntutan dalam Pengembangan Sumber Daya Manusia. *Lectura : Jurnal Pendidikan*, *12*(1), 29–40. https://doi.org/10.31849/LECTURA.V12I1.5813

Primayana, K. . (2020). Menciptakan Pembelajaran Berbasis Pemecahan Masalah Dengan Berorientasi Pembentukan Karakter Untuk Mencapai Tujuan Higher Order Thingking Skilss (HOTS) Pada Anak Sekolah Dasar. *Purwadita: Jurnal Agama Dan Budaya*, *3*(2), 85–92.

Ratnady, H. D., Tindangen, M., Maasawet, E. T., Hakim, A., Hudiyono, Y., & Hardoko, A. (2024). Development of Science Learning Modules with Computer-Based Test (CBT) Assessments to Train Critical Thinking Skills of Students at SMP Negeri 7 Bontang. Jurnal Penelitian Pendidikan IPA, 10(9), 6564-6573. *Jurnal Penelitian Pendidikan IPA*, *10*(9), 6564–6573.

Saputra, I. G. E., Jampel, I. N., & Parwata, I. G. L. A. (2022). Pengembangan Instrumen Penilaian Kemampuan Berpikir Kritis Peserta Didik Kelas VIII SMP pada Materi Getaran dan Gelombang. *Jurnal Pendidikan Dan Pembelajaran Sains Indonesia (JPPSI)*, *5*(2), 154–164.

Sudirman, S., Kistiono, K., Akhsan, H., & Ariska, M. (2020). Pengembangan instrumen

penilaian pengetahuan, sikap dan keterampilan IPA berbasis berpikir kritis pada konsep listrik siswa SMP. Jurnal Inovasi Dan Pembelajaran Fisika, 7(1), 28-40. *Jurnal Inovasi Dan Pembelajaran Fisika*, *7*(1), 28–40.

Sugiarti, T., Kaniawati, I., & Aviyanti, L. (2017). Development of assessment instrument of critical thinking in physics at senior high school. *Journal of Physics: Conference Series*, *12*(1), 12018.

Suputra, G. K. A., Rafika, R., Palliwi, E. D., Nuraini, N., & Khairunnisa, K. (2023). Examining the Validity of Inquiry-Based Biology Learning Tools in Fostering Critical Thinking Ability. Jurnal Penelitian Pendidikan IPA, 9(6), 4863-4868. *Jurnal Penelitian Pendidikan*, *9*(6), 4863–4868.

Trimawati, K., Kirana, T., & Raharjo, R. (2020). Pengembangan instrumen penilaian ipa terpadu dalam pembelajaran model project based learning (pjbl) untuk meningkatkan kemampuan berpikir kritis dan kreatif siswa smp. *Quantum: Jurnal Inovasi Pendidikan Sains*, *11*(1), 36.

Vincent-Lancrin, S. (2021). Fostering Students' Creativity and Critical Thinking in Science Education. *Education in the 21st Century: STEM, Creativity and Critical Thinking*, 29–47. https://doi.org/10.1007/978-3-030-85300-6_3

Wahyuni, S., Rizki, L. A. K., Budiarso, A. S., Putra, P. D. A., & Narulita, E. (2021). The development of e-student worksheet on environmental pollution to improve critical thinking skills of Junior High School Students. *Jurnal Penelitian Pendidikan IPA*, *7*(4), 723–728.

Widyapuraya, N. W., Suryana, A. L., Suyanta, S., & Wilujeng, I. (2023). Profil Keterampilan Berpikir Kritis Siswa SMP Negeri 1 Juwangi pada Pembelajaran IPA. *Jurnal Penelitian Pendidikan IPA*, *9*(3), 1368–1374.

Wijaya, U. R. B., Sumarni, W., & Haryani, S. (2017). Pengembangan instrumen penilaian berpikir kritis pada pembelajaran kimia berpendekatan SETS (Science, environment, technology, and society). *Chemistry in Education*, *6*(2), 35–41.

Yusmar, F., & Fadilah, R. E. (2023). Analisis Rendahnya Literasi Sains Peserta Didik Indonesia: Hasil PISA dan Faktor Penyebab. *LENSA (Lentera Sains): Jurnal Pendidikan IPA*, *13*(1), 11–19. https://doi.org/10.24929/LENSA.V13I1.283

Zulfiana, S., Gunamantha, I. M., & Putrayasa, I. B. (2023). Pengembangan Instrumen Kemampuan Berpikir Tingkat Tinggi dan Literasi Sains Pada Pembelajaran IPA Kelas V SD. *PENDASI Jurnal Pendidikan Dasar Indonesia*, *7*(1), 13–24. https://doi.org/10.23887/JURNAL_PENDAS.V7I1.1384