Analysis of critical thinking in junior high school students through science learning in Indonesia: A systematic review

Arif Billah¹, Mohammad Masykuri¹, Sarwanto¹, Sajidan¹

¹Science Education Doctoral Program, Faculty of Teacher Training and Education, Universitas Sebelas Maret, Jl. Ir. Sutami No. 36A Surakarta, Indonesia

*Corresponding author: sajidan@fkip.uns.ac.id

Abstract. The purpose of this study was to see the results of the research analysis regarding the critical thinking of junior high school students through science lessons in Indonesia. This research was a systematic review of studies through the Scopus database. Search was done through nine stages with criteria at each predetermined stage, i.e.: stage I, initial search with the keywords “critical thinking” with the category “article title”; stage II, type of access “open access”; stage III, published in “2010 – 2018”; stage IV, “English” language; stage V, the type of source “journals and proceedings”; stage VI, keywords “critical thinking”; stage VII, “Indonesia” research location; stage VIII, field of study “Science”; and stage IX, “Junior High School” education level. Searching through these nine stages results in 13 articles. The focus of this study was to analyze critical thinking skills and the indicators used in these studies. Most studies reported that increasing the critical thinking skills of Junior High School students in Indonesia were still at a “moderate” level. This study recommends to do further research on experimentation-based learning with an environmental approach to enhance critical thinking skills.

Keywords: Critical thinking, Junior high school, Science learning, Systematic review

1. Introduction

Many 21st century learning frameworks stated that Critical Thinking (CT) skills are one of the thinking skills recommended for development in students, including: P21 [1], ATC21S [2], and UNESCO [3]. Learning in schools should be able to facilitate the development of these skills. In addition to students’ needs for their future, this skill can deliver students from Lower Order Thinking (LOT) to Higher Order Thinking (HOT) skills [4]; and support learning and innovation skills [1].

Critical thinking is a type of thinking activity with an emphasis on analysis and evaluation that aims to improve it [5]. In CT activities it requires the inclusion of self-direction, self-discipline, self-monitoring, and self-corrective thinking [6]. Critical thinking is thinking that has a purpose to prove a point, interpret what something means, and solve problems [7]; reflective and reasonable thinking that is focused on deciding what to believe or do [8].

Some research results have identified elements of student CT skills, including: analyzing arguments, claims, evidence, making inferences using inductive or deductive reasoning [8-10]; judging or evaluating [8,10]; making decisions or solving problems; asking and answering questions for clarification; defining terms [8]; identifying assumptions [8,9]; interpreting and explaining [10]; self-
regulation [11]. The purpose of this study was to see the results of the research analysis regarding the critical thinking of junior high school students through science lessons in Indonesia.

2. Research methodology

This research was conducted through a systematic review method. Literature was searched through the Scopus database. The systematic review research design was carried out through nine stages with the criteria for each stage that have been determined. In sequence, these criteria were: stage I, initial search with the keyword “critical thinking” with the category “article title”; stage II, type of access “open access”; stage III, published year “2010 – 2018”; stage IV, language “English”; stage V, source type “journals and proceedings”; stage VI, keyword “critical thinking”; stage VII, research location “Indonesia”; stage VIII, “natural science” field of study; and stage IX, “Junior High School” education level. Search time is held on 12 May 2019.

This nine-stage search resulted in 13 articles from 3945 articles on stage I. Furthermore, these 13 articles were tabulated based on: author, year, research method, instruments used to measure CT skills, respondents, codes, indicators of CT skills, theoretical bases, and results and findings. In detail, the flowchart of the systematic review stage in this study is shown in Figure 1. The focus of this study is to analyze CT skills and the indicators used in these studies. Several studies were rejected because they did not comply with the predetermined criteria.

![Figure 1. The systematic review process](image-url)
3. Results and discussion

3.1. Results
Relevant studies obtained in this study were 13 studies. These studies are shown in Table 1. Table 1 shows the authors, year, research methods, instruments used to measure CT skills, respondents, and codes. Encoding is done only to make it easier to connect and explain in the next section.

| Authors (year)                  | Research methods                           | CT Instruments                  | Respondent                  | Code |
|--------------------------------|--------------------------------------------|---------------------------------|-----------------------------|------|
| Dhitareka, Firman, & Rusyati (2018) [12] | Quasi experimental research                 | Science Virtual Test (SVT) and Paper-Based Test (PBT) | 71 students of the 7th grade of one school | A    |
| Nurismawati, Sanjaya, & Rusyati (2018) [13] | Descriptive research                       | SVT                             | 168 students of the 8th grade of three schools | B    |
| Sya'bandari, Firman, & Rusyati (2018) [14] | Descriptive research                       | SVT                             | 117 students of the 7th grade of one school | C    |
| Fernandi, Firman, & Rusyati (2018) [15]   | Descriptive research                       | SVT                             | 110 students of the 9th grade of three schools | D    |
| Rosyidah, Firman, & Rusyati (2017) [16]   | Quasi experimental method with one group post-test only design | SVT and PBT                           | 40 students of the 8th grade of one school | E    |
| Maulida, Firman, & Rusyati (2017) [17]    | Descriptive research                       | SVT                             | 331 students of the 7th grade of five schools | F    |
| Akbar, Firman, & Rusyati (2017) [18]      | Descriptive research                       | SVT                             | 125 students of the 8th grade of three schools | G    |
| Budi, Sunarno, & Sugiyarto (2018) [19]     | True experimental pre-test and post-test of control group design (using IPA module approach SETS) | SVT                             | 64 students of the 7th grade of one school | H    |
| Fuad, Zubaidah, Mahanal, & Suarsini (2017) [20] | Quasi-experimental with pre-post nonequivalent control group design | Critical thinking assessments integrated essay tests [21] | 96 students of the 7th grade of three schools | I    |
| Haridza & Irving (2017) [22]              | Classroom action research with two cycles  | Observation checklist, rating scale, self, and peer assessment Written test | 32 students of the 8th grade of one school | J    |
| Karyadi, Ruyani, Sipriyadi, & Johan (2018) [23] | Research and Development                   | CT test, observation sheet of CT   | 30 students of the 7th grade of one school | K    |
| Hastuti, Nurohman, & Setianingsih (2018) [24] | Research and Development                   | CT test, observation sheet of CT   | 120 students of the 7th grade of one school | L    |
Table 2 shows indicators of CT skills, theoretical basis, and findings from studies

| Code | Indicators of CT Skills | Theoretical Bases | Findings |
|------|-------------------------|-------------------|----------|
| A    | Purpose, Question at issue, Assumption, Information, Concepts, Interpretation and inference, and Implication and consequences [26] | Inch, Warnick, & Endres | There were no significant differences between SVT and PBT on CT skills scores [12]. Through the Post Hoc LSD Test, there are significant differences between concrete-transitional groups and transitional-concrete groups on indicators: purpose, point of view, concepts, implication, and consequences [13]. Validation and reliability of the SVT-MH results obtained were accepted and stated sufficient to measure student CT skills [14]. Overall, there were no significant differences in the male and female groups for CT skills [15]. In most aspects of CT skills, SVT shows a higher score than PBT [16]. Improving student CT skills is generally categorized at a moderate level [17]. SVT development results can be used to measure students’ CT skills with good consistency [18]. Increasing CT skills is more effective using science modules with the SETS approach than learning with textbooks (print outs of Electronic School Books) [19]. The application of the Differentiated Science Inquiry (DSI) model combined with mind map has an impact on CT skills higher than DSI, and conventional [20]. |
| B    | Purpose, Question at issue, Assumption, Information, Concepts, Interpretation and inference, and Implication and consequences [26] | | |
| C    | Point of view, Information, Concepts, Interpretation and inference, and Implication and consequences [26] | Inch, Warnick, & Endres | |
| D    | Interpretation and inference, and Implication and consequences [26] | | |
| E    | Interpretation and inference, and Implication and consequences [26] | | |
| F    | Interpretation, Analysis, Explanation, Evaluation, Inference, and Self-regulation [7]. Focus, supporting reasons, reasoning, organization, conventions, and integration [27]. | Facione | |
| G    | Interpretation, Analysis, Explanation, Evaluation, Inference, and Self-regulation [7]. Focus, supporting reasons, reasoning, organization, conventions, and integration [27]. | Finken dan Ennis | |

Table 2. Indicators of CT skills, theoretical bases, and results and findings.
**Table 2:**

| Code | Indicators of CT Skills                                                                 | Theoretical Bases                  | Findings                                                                 |
|------|----------------------------------------------------------------------------------------|------------------------------------|--------------------------------------------------------------------------|
| J    | Identifying problem correctly, distinguish knowledge and opinion, providing possible solution, making decision and identify the impact of the implementation of their solution [22]. | Self-developed by the authors      | The use of the Problem Based Learning 4 Core Areas (PBL4C) model has been able to significantly improve CT skills [22]. |
| K    | Giving simple explanation, identification, translation, constructing argument, giving conclusion [23]. |                                    | Learning outside the classroom through the stages of Introduction, Exploration, and Interpretation (IEI) has a significant impact on improving CT skills [23]. |
| L    | Formulating questions, analysis (Not mentioned) discussion, problem-solving, decision making, and communication [24]. (Not mentioned) |                                    | The Worksheet Science Based on Inquiry Science Issues can potentially improve CT skills [24]. |
| M    | (Not mentioned)                                                                         |                                    | Problem-based learning can improve student CT skills. Students are required to think critically to solve problems [25]. |

**3.2. Discussion**

Table 2 shows that seven studies (codes A, B, C, D, E, F, and G) refer to Inch, Warnick, & Endres [26] for measuring CT skills. One study (code H) uses Fascione [7]. One study (code I) used Zubaidah, Corebima, & Mistianah [21] who adopted Finken and Ennis [27]. One study (code J) uses indicators that have been independently developed by the author, Haridza & Irving [22]. The other three studies (codes K, L, and M) do not mention the source of the theory of indicators of CT skills. This distribution is shown in Figure 2.
In Study A, the “purpose” indicator had the highest score and the “assumption” had the lowest score [12]. Based on the average score obtained by male and female students, the highest score was obtained for the “purpose” indicator and the lowest score obtained for the “assumption” indicator was reported by study D [15]. In Study E, the indicator “purpose” and “information” have the highest score. The lowest score is achieved in the “question at issue” indicator for SVT, and “assumption” for PBT [16]. Four studies (A, D, E, and F) have used Inch, Warnick, & Endres [26] theories to measure the CT skills of their students. Studies A, D, and E report the same thing, that students get the highest score for the indicator “purpose” and the lowest score for the indicator “assumption”.

Paul, Elder, & Bartell state that the aspect of “purpose” is important because the teacher will know that students are problematic or not in reasoning. Important in a learning to ensure students know the purpose, how to achieve goals, and meet their needs. The goals that students build must be realistic and significant [6]. Learning is designed so that it can guide students to ask: what do I want to achieve and what are my goals? [5]. In this case, most studies state junior high school students can apply it well.

Furthermore, the lowest indicator stated by most studies (A, D, and E) is “assumption”. Assumptions often work unknowingly [28]. All humans make assumptions in everyday life. This assumption guides people in acting and making decisions. This is shown by Paul & Elder in Figure 3. Learning should be able to facilitate students in understanding and connecting between “information”, “assumption”, and “inference” as shown in Figure 3. Indicators of “information” and “inference” be in conscious territory, while “assumption” is placed in the unconscious region [28]. Thus, CT learning should be able to put “assumption” as a step of thinking before drawing conclusions. This is important to guide and produce the right conclusions. Through design learning, the teacher is responsible for the students’ ability to carry out the “assumption” step. This is to ensure students use reason that is reasonable and based on correct information in assuming. The teacher can start by describing problems in daily life so that students demand to express their opinions about what alternative solutions to solve these problems. Then, students are asked to explain the reasons for the opinion.
Furthermore, several theories have put forward an assessment of the “assumption” indicator. One of the main principles of CT is the ability to identify and challenge assumptions [29]. Students’ abilities in CT skills can be viewed from aspects of recognizing and articulating assumptions. Clarity, relevance, and consistency in making assumptions have a profound effect on students’ CT abilities [6]. An important point in assuming is that assumptions are related to actions and decision making [30].

In contrast to the three studies above (A, D, and E), study F reported that the highest score was obtained for the indicator “assumption”. As for the “point of view” indicator, students get the lowest score [17]. “Points of view” is an element of the most challenging CT skills to master [28]. Unlike critical thinkers, ordinary thinkers tend to see things from their point of view. He did not think of the relevance of his thoughts on the problems at hand. The characteristics of critical thinkers are that they can place themselves to judge something. For example, when he read a book, he tried to put his point of view as the author of the book [28]. So, critical thinkers can be good listeners. Thus, learning to improve CT skills should be able to facilitate students to practice seeing problems from various perspectives. The “point of view” assessment can be focused on students in identification, relevance, strengths and weaknesses, and evaluation of their chosen perspectives [28].

Data of the H study show that the highest average score is achieved for the indicator “evaluation” [19]. This study chose Facione as the basis of his theory. To improve ability in “evaluation”, the guiding questions needed include: credibility of claims built, why we should think that we can trust someone’s words, the strength of the argument, whether we have the right facts, and the level of our beliefs about our conclusion [7]. Students from the H study were reported to be successful in the “evaluation” indicator.

However, for the “inference” indicator, the H study reported that the students got a low score. Efforts to increase the ability of “inference” are recommended by Facione to refer to the sub-skills, namely: evidence query; alternative conjecture; draw conclusions using; inductive or deductive reasoning [7]. Some guidance questions that can be used as a basis for applying “inference” skills, among others, refer to: (1) students’ memories of knowledge that has been obtained. So, the questions can guide students in drawing conclusions, (2) and can get rid of something that is not related; (3) information from the evidence obtained; (4) selection of assumptions and impacts; (5) additional information needed to resolve questions; and (6) the consequences of each choice [7].

The results of the J study revealed that the increase in “identifying correct problems” indicators was the most significant compared to other indicators. While the lowest increase is obtained for indicators of “distinguish knowledge and opinion” [22]. In the K study, the highest score was shown in the “identify” and “translation” indicators. This means that students can identify, classify, and interpret well the role of each component of the ecosystem. Both abilities are related to students’ concrete abilities. Students work based on facts and observation data, have not been able to think abstractly. However, student achievement for the other three indicators is low. The three indicators are “giving simple explanation”, “constructing argument”, and “giving conclusion”. The lowest score is indicated by the indicator “giving conclusion” [23]. The data from the L study revealed that the “problem solving” indicator had the highest score, while “communication” had the lowest score [24]. Studies I, C, G, and M do not include detailed data from CT skills indicators.

Most studies report that increasing the CT skills of junior high school students in Indonesia is still at the “moderate” level. This leads to the conclusion that learning efforts are still needed to improve CT skills. The development of the ability of CT skills in middle-aged children (around 12-15 years) is very feasible. Ornstein & Hunkins equate the development of CT skills as concrete and formal operations (Piaget). This is because the stages of cognitive development are related to intellectual potential and environmental experience [31]. For this reason, environmental-based experimentation [32][33] can be an alternative solution for developing student CT skills.

Besides the CT skills indicators, other variables found in these studies include: gender, learning model, and geographical location of the school (rural and urban). Study D reports that overall, there were no significant differences between the male and female groups [15]. However, Study I reported...
that female students had higher CT skills than male students [20]. One finding revealed that to improve CT skills, the learning model was more influential than gender [20]. On the geographical variable, the K study reported that there was no difference between groups of urban and rural students in CT skills [23].

4. Conclusion and Suggestion
Relevant studies obtained in this study were 13 studies. Seven studies refer to Inch, Warnick, & Endres. Facione, Finken and Ennis, and Haridza & Irving each have been used by one study. The other three studies do not mention the source of the theory of CT skill indicators. For a theory based on Inch, Warnick, & Endres it was reported that: (1) Three studies obtained the highest score for the indicator “purpose” and the lowest score for the indicator “assumption”; (2) One study reported the highest score for “assumption” and the lowest for “point of view”. Through learning that has been designed, the teacher should be responsible for ensuring students use reason that is reasonable and based on correct information to take assumptions. For the basic theory that uses Facione, it was reported that one study obtained the highest average score achieved for the indicator “evaluation” and the lowest score for “inference”. Some guidance questions that can be used as a basis for applying the “inference” skills, among others, refer to: (1) students’ memories of the knowledge that has been obtained. So, the questions can guide students in drawing conclusions, (2) and can get rid of something that is not related; (3) information from the evidence obtained; (4) selection of assumptions and impacts; (5) additional information needed to resolve questions; and (6) the consequences of each choice. Other studies based on self-developed indicators report that there is a significant increase in indicators: “identifying problems correctly” (one study); “Identify” and “translation” (one study); indicator of “problem solving” (one study). The lowest score is reported for the indicator: “distinguish knowledge and opinion” (one study); “Giving conclusion” (one study); and “communication” (one study). Most studies report that the CT skills of junior high school students in Indonesia are still at a “moderate” level. Based on the literature review in the previous section, this study recommends that the need for experimentation-based learning with an environmental approach. This type of learning is expected to be further investigated so that it can be an alternative solution for developing student CT skills. Another recommendation from this literature review is the need for further research on the influence of gender, learning models, and the geographical location of schools (rural or urban) on CT skills.

References
[1] P21, P21 Framework Definitions 2009 (Massachusetts: Partnership for 21st Century Skills)
[2] Griffin P, McGaw B and Care E (eds) 2012 Assessment and Teaching of 21st Century Skills (Dordrecht: Springer)
[3] Learning Metrics Task Force (LTMF) 2013 Toward universal learning: Recommendations from the Learning Metrics Task Force (Montreal and Washington, D. C.: UNESCO Institute for Statistics and Center for Universal Education at the Brookings Institution).
[4] Miri B, David B and Uri Z 2007 Res. Sci. Educ. 37 353–369
[5] Paul R and Elder L 2007 The Miniature Guide to Critical Thinking Concepts & Tool (California: Found. Crit. Think.)
[6] Paul R, Elder L and Bartell T 1997 California Teacher Preparation for Instruction in Critical Thinking: Research Findings and Policy Recommendations (California: California Commission on Teacher Credentialing)
[7] Facione P A 2015 Critical Thinking: What It Is and Why It Counts (California: Insight Assess)
[8] Ennis R H 1985 Educ. Leadersh. 43 44–48
[9] Paul R W 1992 Community Coll. 77 3–24
[10] Facione P A 1990 Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction (California: The California Academic Press)
[11] Facione P A 2011 Critical Thinking: What It Is and Why It Counts (California: California Academic Press)
[12] Dhitareka P, Firman H and Rusyati L 2018 J. Phys. Conf. Ser. 1013 012070
[13] Nurismawati R, Sanjaya Y and Rusyati L 2018 J. Phys. Conf. Ser. 1013 012071
[14] Sya’Bandari Y, Firman H and Rusyati L 2018 J. Phys. Conf. Ser. 1013 012067
[15] Fernandi R, Firman H and Rusyati L 2018 J. Phys. Conf. Ser. 1013 012066
[16] Rosyidah T, Firman H and Rusyati L 2017 J. Phys. Conf. Ser. 812 012028
[17] Maulida N, Firman H and Rusyati L 2017 J. Phys. Conf. Ser. 812 012058
[18] Akbar M, Firman H and Rusyati L 2017 J. Phys. Conf. Ser. 812 012106
[19] Budi A, Sunarno W and Sugiyarto 2018 J. Phys. Conf. Ser. 1022 012015
[20] Fuad N, Zubaidah S, Mahanal S and Suarsini E 2017 Int. J. Instr. 10 pp 101–116
[21] Zubaidah S, Corebima A and Mistianah 2015 Pros. Symp. Biol. Educ. pp 200–213
[22] Haridza R and Irving K 2017 J. Phys. Conf. Ser. 812 012081
[23] Karyadi B, Ruyani A, Sipriyadi and Johan H 2018 J. Phys. Conf. Ser. 1116 052036
[24] Hastuti P, Nurohman S and Setianingsih W 2018 J. Phys. Conf. Ser. 1097 012004
[25] Aswan D, Lufri L and Sumarmin R 2018 IOP Conf. Ser. Mater. Sci. Eng. 335 012128
[26] Inch E S, Warnick B and Endres D 2006 Critical Thinking and Communication: The Use of Reason in Argument, 5th ed. (Boston: Pearson/Allyn & Bacon)
[27] Finken M and Ennis R H 1993 Illinois Critical Thinking Essay Test (Illinois: University of Illinois)
[28] Paul R W and Elder L 2002 Critical Thinking: Tools for Taking Charge of Your Professional and Personal Life (New Jersey: Financial Times Prentice Hall)
[29] Brookfield S 1987 Developing Critical Thinkers (Milton Keynes: Open University Press)
[30] Gabbitas B W 2009 Critical Thinking and Analyzing Assumptions in Instructional Technology (Utah: Brigham Young University)
[31] Ornstein A C and Hunkins F P 2004 Curriculum: Foundations, Principles and Issues, 4th ed. (Boston: Allyn & Bacon)
[32] Slavin R 2006 Educational Psychology: Theory and Practice (Michigan: Pearson)
[33] Karyadi B, Susanta A, Winari E W, Ekaputri R Z and Enersi D 2018 J. Phys. Conf. Ser. 1013 012074