SCIENTIAE EDUCATIA: JURNAL PENDIDIKAN SAINS

journal homepage: www.syekhnurjati.ac.id/jurnal/index.php/sceducatia

http://www.syekhnurjati.ac.id/jurnal/index.php/sceducatia/article/view/475

Development of Argumentation-Based Critical Thinking Skills Tests in Microbiology Laboratory

Evi Roviati[ab*], Ari Widodo[a], Widi Purwianingsih[a], Riandi[a]

[a] Prodi Pendidikan IPA, SPS, Universitas Pendidikan Indonesia, Bandung, Jawa Barat, 40154, Indonesia
[b] Prodi Pendidikan Biologi, Institut Agama Islam Negeri Syekh Nurjati, Cirebon 45135, Indonesia

[*]Corresponding author: Jl. Fatahillah No. 40, Kel. Watubelah, Kec. Sumber, Kab. Cirebon, Jawa Barat, 45100, Indonesia. E-mail addresses: evi1roviati@gmail.com

a r t i c l e   i n f o

a b s t r a c t

This study aims to develop an instrument used to measure students' argumentation-based critical thinking skills (CTS) in microbiology practicum activities. This study used a phased model for the development of tools, which consists of the stages of defining constructs and formulating objectives in the form of mapping aspects of the critical thinking skills and the microbiology concepts, formulating test item formats, constructing item questions, constructing scoring guidelines, evaluating by experts for content validation, pilot testing on students and analyzing the results. The test items were then analyzed to determine the validity, reliability, distinguishing features and degree of difficulty. The instrument produced in this study used 6 CTS indicators which were considered the most relevant to argumentation and laboratory activities, and consisted of 18 open-ended questions with 5 contexts. The results of expert validation show that the instrument is content valid and can be used for the next step. Furthermore, the results of the pilot test show that of the 18 questions that were tested, as many as 17 questions were declared valid and 1 question was corrected. Overall the test questions were declared reliable. Thus, the results of this study recommend the use of questions on this test in studies that measure argumentation-based critical thinking skills in microbiology laboratory.

2019 Scientiae Educatia: Jurnal Pendidikan Sains

1. Introduction

Critical thinking skills (CTS) as one of the 21st century skills which are claimed to be the main objective of science education need to be integrated in science education today, because the most important goal in science education is to develop students' thinking skills in a scientific context (Griffin et al., 2015; Bailin, 2002; Kemendikbud RI, 2013). CTS involves the ability to draw valid inferences, identify relationships, analyze opportunities, make predictions and logical decisions and solve complex problems (Facione, 2011). Skills in the CTS are linked to success in education in higher education, increased ability to make decisions by taking into account complex daily problems and participation as active and literate citizens in the era of democracy (Wright, 2011; Halpern, 2013).

Meanwhile, scientific argumentation in science education is critical in helping students develop scientific literacy (NRC, 2000; Cavagnetto, 2010). The ability to learn engaging in scientific argumentation becomes a challenge for students, such as the ability to test or

construct a claim, and then accept or reject an evidence and evaluate the explanation of the relationship with the evidence (Driver, Newton & Osborne, 2000). However, sometimes students do not use appropriate and sufficient evidence or do not try to provide an explanation of their choice in relation to the evidence in their arguments (Sadler, 2004). Therefore, the need to involve students in scientific argumentation is inevitable. However opportunities for students to engage in scientific argumentation-based science learning activities in a productive way are very rare (Simon, Erduran & Osborne, 2006).

Fisher (2007) defines critical thinking as a skilled and active interpretation and evaluation of observation and communication, information and argumentation. Efforts to meet the challenges of developing CTS have been widely reported either in the form of teaching separately from regular learning (Ennis, 1993) or integrated in shared learning concepts of subjects (Niu, Behar-Horenstein & Garvan, 2013; Tiruneh, Verburgh & Elen, 2014). The implementation of CTS in learning activities on various subjects is expected to facilitate the acquisition of CTS that can be applied to thinking tasks and to everyday life (Lawson, 2004). However, the application of learning that facilitates CTS largely takes the general domain of skills in CTS, not specifically related to the potential of argumentation in CTS. Though the contribution of argumentation to the CTS has been discussed and suggested to be implemented in science education (Jimenez-Aleixandre & Puig, 2012; Facione, 1990). Argumentation has a significant contribution in developing CTS with unique characteristics, namely assessing the source of information, evaluating arguments and producing arguments and presenting them (Roviati & Widodo, 2019). Potential contributions of argumentation in science learning include supporting the development of critical thinking competencies through verification and reflection (Jimenez-Aleixandre & Erduran, 2007).

The characteristics of CTS which are related to scientific argumentation and used as indicators of skills tested in the development of this instrument are as follows: 1) assess the acceptability of information by considering the credibility of the source, evidence and claims; 2) identify the elements in the case being considered in the form of conclusions, reasons and assumptions; 3) assess / consider / evaluate the quality of arguments of various types, including whether the reasons, assumptions and evidence are acceptable; 4) produce arguments and present them; 5) develop and maintain a position on an issue by analyzing, evaluating and producing explanations; 6) plan experiments by evaluating experimental procedures and designs (Fisher, 2007; Ennis, 1993; Facione, 1990; Roviati & Widodo, 2019).

The development of CTS test instruments with domain specific concepts has been carried out by several studies (Putri, Isyono & Nurcahyanto, 2016; Nawawi & Wijayanti, 2018), but most are in the form of multiple choice tests or reasoned multiple choice. The CTS test instrument in the form of essays was developed by Amalia & Susilaningsih (2014), and Ritdamaya & Suhandi (2016) with specific domains on the chemical concepts of acids & bases and the physical concepts of temperature & heat matter, but have not been linked to the ability of scientific argumentation. Therefore, to meet the need for an instrument that measures argumentation-based CTS, it is necessary to study the development of an argumentation-based CTS measurement instrument. This study aims to produce instruments that measure CTS based on argumentation in microbiology laboratory activities.

2. Method

The stepwise model of instrument development was used in planning and developing the AB-CTS test instrument in this study with a descriptive cohort design.

2.1. Defining Constructs and Formulating Goals

The initial stage of developing an argumentation-based CTS (AB-CTS) test in this study is defining the CTS and selecting the targeted aspects of the CTS. The AB-CTS test aims to measure the CTS which focuses on scientific argumentation in the specific domain of microbiology laboratory activities. Therefore, it is necessary to identify specific characteristics of CTS which are relevant to scientific argumentation. The characteristics of the CTS are formulated with 6 aspects (indicators) based on the characteristics of the CTS delivered by Facione (1990), Fisher (2007) and Ennis (1993), which are selected and adjusted with scientific argumentation and laboratory activities for the purpose of developing this test and then used as a guide in creating items AB-CTS test questions as can be seen in the item grid in Table 1. While the concept domain used as context is the microbiology lab activities, which consist of the concepts of 1) aseptic work, 2) antimicrobial susceptibility, 3) microbes around us, 4) food microbes, and 5) microbiological testing of drinking water.

2.2. Formulating the Test Item Format

Most CTS tests currently use the multiple choice test format. Multiple choice tests are seen as less able to directly and efficiently measure CTS features such as drawing conclusions, analyzing arguments and solving problems systematically. Multiple choice tests can also cause bias because students might answer by guessing or coincidentally choose the right

answer. Experts usually recommend an essay format (open-ended) or a combination of multiple choice and essay. Therefore, this study uses the essay format (open-ended) to uncover the actual aspects of the AB-CTS that students have mastered (Ennis, 1993; Halpern, 2010; Norris, 1989). Based on this recommendation, it was decided that the CTS measurement based on argumentation in this study used essay items.

Table 1. Description of the outcome targeted in AB-CTS test test instrument

Characteristics of AB-CTS	Item	Amount
1. Assess the acceptability of information by considering the credibility of the source, the evidence and its claims.	1.a.; 4.c.; 5.a.	3
2. Identify the elements in the case being considered in the form of conclusions, reasons and assumptions.	1.b.; 2.c.; 4.a.	3
3. Assess / consider / evaluate the quality of arguments of various types, including whether the reasons, assumptions and evidence are acceptable.	2.d.; 3.b.; 3.c.	3
4. Generate arguments and present them.	1.c.; 3.d.; 5.b.	3
5. Develop and maintain a position on an issue by analyzing, evaluating and generating explanations.	3.a.; 3.e.; 5.c.	3
6. Plan experiments by evaluating experimental procedures and designs	2.a.; 2.b.; 4.b.	3
Total		18

2.3. Constructing Test Items

The construction of test items that revealed the special domain of CTS based on argumentation in this study developed through repetitive improvement. Initially 5 question items were arranged with 5 contexts on different concepts. Each item was reviewed and discussed by the researchers to follow the test criteria that reveal the desired performance of the CTS and the clarity of the questions to be understood by students. Next, seven question items were added to better accommodate the CTS aspects measured in this study so they could be represented. Through discussion and revision, each item was developed to meet all the desired criteria. Then 6 more question items were added, so that in the end 18 questions were obtained. The eighteen question items were presented in the form of 5 question numbers according to the context, each of which contained 3 to 5 questions that represent the specified indicators. Each of 6 CTS indicator was represented by 3 questions spread in all 5 question contexts. The distribution of the question contex and indicators that being assessed in each item of the test can be seen in Table 2. The discussion and revision process continued until all the question items are considered sufficient to meet the required requirements.

Table 2. Distribution of the question contex and AB-CTS indicators

No.	Contex	Question sub no.	Indicators being assessed
1.	aseptic work	1.a.	1 (Assess the acceptability of information by considering the credibility of the source, the evidence and its claims.)
		1.b.	2 (Identify the elements in the case being considered in the form of conclusions, reasons and assumptions.)
		1.c.	4 (Generate arguments and present them.)
2.	antimicrobial susceptibility	2.a.	6 (Plan experiments by evaluating experimental procedures and designs)
		2.b.	6 (Plan experiments by evaluating experimental procedures and designs)
		2.c	2 (Identify the elements in the case being considered in the form of conclusions, reasons and assumptions.)
		2.d.	3 (Assess / consider / evaluate the quality of arguments of various types, including whether the reasons, assumptions and evidence are acceptable.)
3.	microbes around us	3.a.	5 (Develop and maintain a position on an issue by analyzing, evaluating and generating explanations)
		3.b.	3 (Assess / consider / evaluate the quality of arguments of various types, including whether the reasons, assumptions and evidence are acceptable.)
		3.c.	3 (Assess / consider / evaluate the quality of arguments of various types, including whether the reasons, assumptions and evidence are acceptable.)
		3.d.	4 (Generate arguments and present them.)
		3.e.	5 (Develop and maintain a position on an issue by analyzing, evaluating and generating explanations)
4.	food microbes	4.a.	2 (Identify the elements in the case being considered in the form of conclusions, reasons and assumptions.)
		4.b.	6 (Plan experiments by evaluating experimental procedures and designs)
		4.c.	1 (Assess the acceptability of information by considering the credibility of the source, the evidence and its claims.)
5.	microbiological testing of drinking water	5.a.	1 (Assess the acceptability of information by considering the credibility of the source, the evidence and its claims.)
		5.b.	4 (Generate arguments and present them.)
		5.c.	5 (Develop and maintain a position on an issue by analyzing, evaluating and generating explanations)

2.4. Creating Scoring Guidelines

In line with the preparation of item questions, answer keys and scoring guidelines for each question item were also made and reviewed by researchers. The answer keys were arranged according to the objectives of each test item and the types of answers expected. Assessment guidelines prepared as a guide to provide a consistent score.

2.5. Expert Validation

Three lecturers with expertise in each field of microbiology, scientific argumentation and biology education in Universitas Pendidikan Indonesia were asked to review and pass judgment on the 18 items developed. The main purpose of the preparation of the AB-CTS test instrument was explained to the three experts and then they were asked to assess the

suitability of the test items with the CTS indicators, answer keys and assessment guidelines. Specifically, these experts were asked to rate with the following criteria: 1) the suitability of the questions with the CTS indicators; 2) the suitability of the question with the microbiology concept being tested; 3) accuracy of science content on questions and answer keys; 4) the correct use of words and terms or language; 5) Questions do not lead to multiple interpretations; and 6) the appropriateness and relevance of the assessment criteria and scores with questions and answers. All three experts agreed that most items about AB-CTS were appropriate and relevant to measure aspects of the CTS that were targeted in the context of microbiology lab activities. The results of the validation of the three experts can be seen in Table 3. The experts also provided useful feedback on a number of item items that were considered to require revision. In accordance with their suggestions and comments, the necessary revisions were also made.

Table 3. Results of expert validation of AB-CTS test instruments

| No. | Aspects / Indicators | Expert Validator | | | Average |
		1	2	3	
1	The suitability of the questions with indicators of critical thinking skills	2	2	3	2.33
2	The suitability of the question with the concept being tested	3	3	3	3
3	Accuracy of scientific content on questions and answers	3	3	3	3
4	Accuracy in using words and terms or languages	3	3	3	3
5	The problem does not lead to double disclaimer	3	3	3	3
6	Suitability and relevance of assessment criteria and scores with questions and answers	3	2	3	2.67

Note: 3= good; 2= adequate; 1= deficient

2.6. Pilot Testing

The pilot testing performed to the 5th semester biology education students (N = 35) with an average age of 21 years in one of the universities in Cirebon Indonesia. The student participants consisted of 32 female and 3 male students. These students were taking microbiology courses and having an argumentation-based microbiology lab. The students had never participated in the AB-CTS test instrument before. The subject of the trial was chosen because their campus has a microbiology laboratory which is sufficiently representative for the implementation of microbiology laboratory courses and allows for conducting argumentation-based inquiry laboratory activities.

2.7. Analysis of Trial Results

The results of trials conducted on student participants were then analyzed to obtain data on validity, reliability, distinguishing features and degree of difficulty of the test. The validity of

the test instrument is obtained by analyzing the test results using the Pearson product Moment formula. While reliability is analyzed using the Cronbach's Alpha formula (Norris, 1989).

3. Result and Discussion

3.1. Test Instrument Results

The AB-CTS test instrument produced from this study consisted of 18 items of test questions divided into 5 context numbers. An example of the generated question items can be seen in Figure 1. In the next section, the results of the AB-CTS test analysis will be explained including the results of the item validity, reliability, discrimination and difficulty level of the test results.

The AB-CTS test questions consisted of 5 questions, each consisting of 3 to 5 questions, as described in Table 2. Problem no. 1 has 3 questions in the context of microbiology topics about working aseptically in the microbiology lab, while question no. 2 with 4 questions about the context of antimicrobial material, question no. 3 with 5 questions about the microbial context around us, question number 4 with 3 questions about the context of food fermentation and question no. 5 with 3 questions about the context of microbiological testing of drinking water quality. The context of the question is about the activities of the microbiology lab and the question is adjusted to the indicator of argumentation-based critical thinking skills.

3.2. Validity and Reliability

The results of the validity test show that 17 of the 18 questions tested obtained a significant correlation value of validity. Thus only 1 problem was declared invalid. The invalid questions are then revised so that they can be used in further tests. The results of the validity test can be seen in **Table 4**.

Meanwhile, the results of the reliability test showed Cronbach's alpha coefficient of 0.71 which meant the AB-CTS test questions were reliable. Thus, this problem can be said to be reliable and can be used to measure CTS based on student argumentation on the topic of microbiology laboratory.

Sample question item

Context:
Antiseptics and disinfectants are distinguished based on their use. The choice of antiseptic material for the body surface, one of which is based on its nature that does not cause irritation. However, the selection of antiseptics and disinfectants to kill certain types of microbes requires careful testing. In a lab, students test the effectiveness of three kinds of antiseptics against test bacteria *E. coli* and *S. aureus* and put them together in a design experiment, with the results of the diameter of inhibition zone as follows:

Table 1. Data on diameter of antiseptic inhibition zones against test bacteria.

Antiseptic type	*Escherichia coli*			*Staphylococcus aureus*		
	U1	U2	U3	U1	U2	U3
Antiseptic A	2.3 cm	2.5 cm	2.0 cm	1.0 cm	1.3 cm	1.2 cm
Antiseptic B	1.9 cm	2.0 cm	2.5 cm	1.6 cm	1.5 cm	1.7 cm
Antiseptic C	3.8 cm	3.8 cm	3.8 cm	4.0 cm	3.6 cm	3.9 cm

Ket: U= repetition

Question: From the inhibitory zone measurement data, students draw 2 conclusions:
        1) the most effective antiseptic is C, and
        2) both test bacteria are sensitive to all three antiseptics.
        In your opinion, is that conclusion correct? Why?

Aspects of the AB-CTS indicator being tested:
        Identify the elements in the case being considered in the form of conclusions, reasons and assumptions.

Answer key:
        The student's conclusion is correct, with reason:
        1) The largest average diameter of inhibition zone is in antiseptic C,
        2) the most effective antiseptic is seen from the diameter of the largest inhibitory zone,
        3) all three antiseptics show inhibition zones of more than 1 cm in both test bacteria, which means they are in the sensitive category.

Scoring guide:
        Answer correctly and give 2 / more reasons (score 3)
        Answer correctly, and give 1 good reason (score 2)
        Answering correctly, but giving a wrong reason or not giving a reason (score 1)
        Giving incorrect answers or not answering (score 0)

Figure 1. Sample of question items and its scoring guide

3.3. Item Difficulty Level and Discrimination

To further strengthen the results of the development of this instrument, a level of difficulty and discrimination test was carried out. Difficulty and discrimination test results can be seen in Table 3. Difficulty level test results showed that 10 items were easy, 7 items were medium questions and 1 item was very difficult question. Although the ideal conditions of the level of difficulty of the problem depend on the purpose of linking the questions, but it can be seen here that most item items are easy and medium difficulty. There were no very easy or difficult

questions. There was only 1 problem that is classified as very difficult and this problem was also invalid. Thus the item was revised for being used later.

The discrimination test explains how well a question item can distinguish between students with different levels of ability. The results of the discrimination test on the development of this instrument showed that all the items were quite good in distinguishing students at different levels of ability.

Table 4. The results of the analysis of the test items of pilot testing of the AB-CTS test instrument

No.	Question	Corelation	Validity	Reli-ability	Discrimi-nation	Difficulty	Conclusion
1.	a.	0.371	Sig	0.71	14.81	Easy	Item used
	b.	0.406	Sig		33.33	Medium	Item used
	c.	0.649	Hi. Sig		59.26	Medium	Item used
2.	a.	0.506	Sig		11.11	Medium	Item used
	b.	0.521	Sig		29.63	Easy	Item used
	c.	0.447	Sig		11.11	Easy	Item used
	d.	0.320	Sig		18.52	Easy	Item used
3.	a.	0.611	Hi. Sig		40.00	Easy	Item used
	b.	0.527	Sig		29.63	Easy	Item used
	c.	0.415	Sig		14.81	Easy	Item used
	d.	0.382	Sig		14.81	Medium	Item used
	e.	0.425	Sig		40.74	Medium	Item used
4.	a.	0.686	Hi. Sig		40.74	Easy	Item used
	b.	0.149	-		7.41	Hi. difficult	Item revised
	c.	0.365	Sig		11.11	Easy	Item used
5.	a.	0.597	Hi. Sig		62.96	Medium	Item used
	b.	0.369	Sig		22.22	Medium	Item used
	c.	0.404	Sig		25.93	Easy	Item used

3.4. Discussion

Because of the importance of developing students' CTS and the ability of scientific argumentation that continues to grow, researchers and practitioners should have a valid and reliable test instruments to evaluate the effectiveness of various learning development efforts aimed at developing both skills. This study argues that an accurate and comprehensive assessment should emphasize the two dimensions of the CTS, the specific and general domain. With the known lack of CTS tests in the specific domain in science learning, a test that can evaluate the CTS elements based on argumentation in the microbiology laboratory concept has been developed and validated.

Analysis of qualitative and quantitative data as a whole provides sufficient evidence that the AB-CTS test at an early stage can be the basis for measuring student CTS in the context of microbiology laboratory. The argumentation-based critical thinking skills test instrument (AB-CTS test) is used to capture the data of critical thinking skills based on student

argumentation before and after taking part in an argumentation based microbiology laboratory courses. The development of the AB-CTS test instrument refers to the indicators of critical thinking skills formulated and selected from the characteristics of critical thinkers put forward by Ennis (1993), Fisher (2007) and Facione (1990) and formulated into 6 indicators as shown in table 1. adjusted to the laboratory activities and argumentation are learned and trained on the program implemented.

The results of the expert validation showed that most aspects of the assessment of the AB-CTS test instrument showed good criteria, and only a small proportion showed sufficient criteria and none were included in the poor category. Aspects that were still included were sufficient, such as the appropriateness of questions with indicators of critical thinking skills and the appropriateness and relevance of assessment criteria and scores with questions and answers, then revised.

The procedure described in this study for developing and validating AB-CTS test items is in line with the suggested guidelines for preparing essay tests and other performance tests according to Adam & Wieman, (2011), Benjamin et al., (2017) and Tiruneh et al., (2017). Although following the guidelines for the development of existing research, this research proposes an assessment framework that encourages the measurement of CTS based on argumentation in specific domains. It is hoped that the AB-CTS test can be used for evaluation of learning and research. The development and validation of this instrument is the first attempt to meet the need for a AB-CTS test instrument, which is expected to be able to demonstrate an approach that can be applied to developing and validating the CTS test in other domains and other fields.

4. Conclusions and Implications

Argument-based critical thinking skills (AB-CTS) test instrument that has been developed and validated in this study consists of 18 item items divided into 5 contexts. The instrument was developed based on CTS indicators relating to scientific argumentation in the context of microbiology laboratory. The results of expert validation and trial analysis showed that the instruments are valid and can be used to measure relevant capabilities. The results of this study are recommended for the research and learning of microbiology laboratory courses based on argumentation.

Acknowledgements

References

Adams W. K. & Wieman, C. E. (2011). Development and validation of instruments to measure learning of expert like thinking. *International Journal of Science Education*, *33*(9), 1289-1312. doi:10.1080/09500693.2010.512369.

Amalia, N. F. & Susilaningsih, E. (2014). Pengembangan instrumen penilaian keterampilan berpikir kritis siswa SMA pada materi asam basa. *Jurnal Inovasi Pendidikan Kimia, 8*(2), 80-89.

Bailin, S. (2002). Critical thinking and science education. *Science & Education*, *11*(4), 361-375.

Benjamin, T. E., Marks, B., Demetrikopoulos, M. K., Rose, J., Pollard, E., Thomas, A., & Muldrow, L. L. (2017). Development and validation of scientific literacy scale for college preparedness in STEM with freshmen from diverse institutions. *International Journal of Science and Mathematics Education*, *15*(4), 607-623.

Cavagnetto, A. R. (2010). Argument to foster scientific literacy: A review of argument interventions in K–12 science contexts. *Review of Educational Research*, *80*(3), 336-371.

Driver, R., Newton P. & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education, 84*(3), 287-312.

Ennis, R. H. (1993). Critical thinking assessment. *Theory into practice*, *32*(3), 179-186.

Facione, P. A. (1990). *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction*. California Academic Press, California.

Facione, P. A. (2011). Critical thinking: What it is and why it counts. *Insight assessment*, *2007*(1), 1-23.

Fisher, A. (2007). Critical thinking: An introduction. Cambridge University Press.

Griffin, P., Care, E., & Harding, S. M. (2015). Task characteristics and calibration. In *Assessment and Teaching of 21st Century Skills* (pp. 133-178). Springer, Dordrecht.

Halpern, D.F. (2010). The Halpern critical thinking assessment: Manual. Modling, Austria: Schuhfried GmbH.

Halpern, D.F. (2013). Thought and knowledge: An introduction to critical thinking. New York: Psychology Press

Jimenez-Aleixandre M.P. & Puig. B. (2012). Argumentation, evidence evaluation and critical thinking." (Chapter 66) in *Second International Handbook of Science Education* Ed. by B.J. Fraser et al. Springer International Handbooks of Education 24.

Jiménez-Aleixandre, M. P. & Erduran, S. (2007). Argumentation in Science Education Perspectives from Classroom-Based Research. Chapter 1. *Argumentation in Science Education: An Overview*. Springer Netherlands. doi: 10.1007/978-1-4020-6670-2.

Kemendikbud, R. I. (2013). Bahan pelatihan kurikulum 2013. Jakarta: Kemdikbud.

Lawson, A. (2004). The nature and development of scientific reasoning: A synthetic view. *International Journal of Science and Mathematics Education, 2*(3), 307-338. doi:10.1007/s10763-004-3224-2.

National Research Council, (2000). *Inquiry and The National Science Education Standards: A Guide for Teaching and Learning*. National Academic Press

Nawawi, S. & Wijayanti, T. F. (2018). Pengembangan asesmen biologi berbasis keterampilan berpikir kritis terintegrasi nilai Islam. *Jurnal Inovasi Pendidikan IPA, 4*(2), 136-148.

Niu, L., Behar-Horenstein, L. S. & Garvan, C. W. (2013). Do instructional interventions influence college students' critical thinking skills? a meta-analysis. *Educational Research Review*, 9, 114–128. doi:10. 1016/j.edurev.2012.12.002

Norris, S.P. (1989). Can we test validly for critical thinking? *Educational Researcher, 18*(9), 21-26. doi:10. 2307/1176715.

Pascarella, E. T. & Terenzini, P. T. (2005). How college affects students: A third decade of research (Vol. 2). San Francisco, CA: Jossey-Bass.

Putri, F.S. Istiyono E. & Nurcahyanto E. (2016). Pengembangan instrumen tes keterampilan berpikir kritis dalam bentuk pilihan ganda beralasan (politomus) di DIY. *Unnes Physical Education Journal, 5*(2), 76-84.

Ritdamaya, D. & Suhandi, A. (2016). Konstruksi instrumen tes keterampilan berpikir kritis terkait materi suhu dan kalor. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, *l*(2), 87-96.

Roviati, E. & Widodo, A. (2019). Kontribusi argumentasi ilmiah dalam pengembangan keterampilan berpikir kritis. *Titian Ilmu: Jurnal Ilmiah Multi Sciences, 11*(2), 56-66.

Sadler, T. D. (2004). Informal reasoning regarding socioscientific issues: A critical review of research. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *41*(5), 513-536.

Simon, S., Erduran, S., & Osborne, J. (2006). Learning to teach argumentation: Research and development in the science classroom. *International journal of science education*, *28*(2-3), 235-260.

Tiruneh, D. T., DeCock, M., Weldeslassie, A. G., Elen, J., and Janssen, R. (2017). Measuring critical thinking in physics: development and validation of a critical thinking test in electricity and magnetism. *International Journal of Science and Mathematics Education*, *15*(4), 663-682.

Tiruneh, D. T., Verburgh, A. & Elen, J. (2014). Effectiveness of critical thinking instruction in higher education: A systematic review of intervention studies. *Higher Education Studies, 4*(1), 1-17. doi:10. 5539/hes.v4n1p1.

Wright, G.B. (2011). Student-centered learning in higher education. *International Journal of Teaching and Higher Education, 23*(3), 92-97