



Development of Critical Thinking Skills Instruments : Cases for Essay Tests

Nurkhairo Hidayati^{1*}, Suryanti², Restesa Rahmayumita³, Selvira Aisya⁴

^{1*,2,4}Biology Education Department, ³Master Program in Science Education,
Universitas Riau, Indonesia.

*Corresponding Author. Email: khairbio@edu.uir.ac.id

Abstract: This research aims to develop a rubric used to assess students' critical thinking. This rubric refers to aspects of Ennis' critical thinking. The research method begins with determining the scope of the rubric so that an initial version of the rubric can be prepared. An initial version of the rubric received peer review. Next, the rubric was tested to assess critical thinking tests with essay-type questions. Data analysis was carried out using Cohen's kappa, while aspects of the critical thinking rubric were analyzed descriptively. The research results show that each aspect of critical thinking has been equipped with descriptors that can be used by raters. The level of reliability based on the kappa value of agreement is quite good, which means that different raters can use the rubric consistently. The value of agreement in the basic clarification aspect is 0.33 (fair), bases for a decision (0.571 in the moderate category), inference (0.455, moderate category), advanced clarification (0.250, fair category), supposition and integration (0.250, fair category), and strategies and tactics (0.182, poor category). Then, it can be concluded that the critical thinking rubric developed can be used to assess critical thinking abilities. As a recommendation, of course, this rubric cannot fulfill all types of critical thinking questions tested. It is also necessary to develop rubrics for question types other than essays to accommodate the wider need for critical thinking rubrics.

Article History

Received: 13-12-2024
Revised: 15-01-2024
Accepted: 05-02-2024
Published: 09-03-2024

Key Words:

Critical Thinking;
Rubric; Assessment.

How to Cite: Hidayati, N., Suryanti, S., Rahmayumita, R., & Aisya, S. (2024). Development of Critical Thinking Skills Instruments : Cases for Essay Tests. *Jurnal Kependidikan: Jurnal Hasil Penelitian dan Kajian Kepustakaan di Bidang Pendidikan, Pengajaran dan Pembelajaran*, 10(1), 77-88. doi:<https://doi.org/10.33394/jk.v10i1.10052>



<https://doi.org/10.33394/jk.v10i1.10052>

This is an open-access article under the [CC-BY-SA License](https://creativecommons.org/licenses/by-sa/4.0/).



Introduction

Critical thinking has been conceptualized by many experts by proposing various definitions. In general, it has been agreed that people with the ability to think will be able to dig up information, determine and carry out various evaluations of circumstances or conditions with logical and rational considerations when they want to solve problems or when making decisions (Ennis, 2011; Kong et al., 2014). Someone with critical thinking skills also tends to have an explorative and reflective mindset. Critical thinking is defined as an intellectual process of creating concepts, applying, analyzing, synthesizing, evaluating information and concluding and providing explanations about considerations based on judgment (Abrami et al., 2014) and the strength of arguments is an important part of the quality of critical thinking (Lin, 2013).

Critical thinking skills can be applied broadly to various scientific disciplines. This ability can also determine an academic career so that it becomes a consideration and concern in the curriculum (*Undergraduate Professional Education in Chemistry ACS Guidelines and Evaluation Procedures for Bachelor's Degree Programs*, 2015) and is a consideration when recruiting in the workplace (Gray & Koncz, 2018; Pearl et al., 2018). Critical thinking can be used as a reference for the success or failure of a program, including in the teaching and learning process. This ability can also be used as a benchmark for someone when they want



to enter the world of work and industry. Assessments carried out on critical thinking are also useful for educators to determine students' progress in participating in learning.

A review of various literature has explained that one of the main obstacles in measuring critical thinking is that the operational definition of thinking has not been explained operationally, clearly and concretely (Liu et al., 2014). This is partly due to the lack of consensus regarding a definition or theoretical model of critical thinking over the years. Although there is debate, there are several things that are agreed upon. The concept of critical thinking concludes that there is significant agreement among researchers, educators, entrepreneurs, and policy-makers regarding the core cognitive components of critical thinking skills, which include analysis and interpretation, and judgment, evaluation, inference, and decision making (Shaw et al., 2019). In another important study conducted by Halpern in 2010, critical thinking was conceptualized as a set of cognitive skills that can facilitate problem-solving and increase the likelihood of desired outcomes. Key skills include reasoning, analysis, evaluation, judgment, and decision-making (Butler, 2012).

Measuring critical thinking abilities can be done using tests. Current tests available to assess critical thinking include the Critical Thinking Assessment Test (CAT) (Stein & Haynes, 2011), the California Critical Thinking Skills Test, and the Watson Glaser Critical Thinking Appraisal. These various tests use multiple-choice questions. Multiple-choice tests cannot always be used to provide feedback to students because students only choose answers from various alternative answers provided. This type of test can be accessed on the internet by first registering. As an alternative, an open-ended essay test can be a test to assess critical thinking skills. This form of test can encourage students to show responses or answers rather than just choosing answers. Essay tests have the potential to reveal students' abilities by reasoning, organizing, analyzing, synthesizing, and evaluating (Hart et al., 2021).

The use of essay tests requires an appropriate assessment rubric to truly be able to measure critical thinking abilities. Rubrics in the evaluation process are a very important measuring tool because they make it easier for assessors to give scores or write comments on student work. Apart from that, the use of rubrics also allows teachers to know the progress of individual students so that teachers can provide different treatment to each student according to their characteristics (Brookhart & Chen, 2015; Smit & Birri, 2014). One of the criteria for creating a rubric for assessing critical thinking skills is that teachers can use it when they want to provide feedback to students. Teachers can utilize this data to determine appropriate learning strategies to use in the teaching and learning process (Reynders et al., 2020). In the long term, it will be in accordance with the teacher's techniques to improve students' abilities. The usefulness of data from this assessment is not only for students but also for teachers so that both of these are learners. Teachers and students can use this data but for different purposes. Students need feedback from teachers that is obtained consistently. During learning, students can reflect on themselves by benchmarking their current performance and identifying what they can do to improve their respective abilities (Hattie & Gan, 2011).

The rubric developed refers to the critical thinking aspects of Ennis (2011). Ennis, in his writing, has explained that for the purposes of assessment and compiling assessment rubrics, criteria or aspects can be used as a basis for compiling them. These aspects are basic clarification, bases for a decision, inference, advanced clarification, supposition and integration, strategies and tactics. The development of this critical thinking rubric is also based on constructivism theory. Constructivism theory argues that students learn by constructing their own understanding of knowledge rather than deriving meaning from the teacher. This theory also considers how there is harmony between learning. Assignments, and assessments that can influence the knowledge and skills students develop. Students are more



likely to develop the desired knowledge and skills if there is alignment between the teacher's desired learning outcomes, the assignments given, and the assessment strategies the teacher uses (Biggs, 2014).

In accordance with the theoretical framework explained previously, the rubric used to assess students' critical thinking abilities must be adjusted to suit the expected results and be able to provide feedback to students and teachers about these abilities. Ennis (2011) determined the criteria that are used to measure in critical thinking. To make it easier for teachers to assess critical thinking skills, assessment rubrics need to be arranged in stages by providing descriptors and scores. The use of rubrics allows for focused feedback that is directly related to the tests that students have taken. This provides an opportunity to explicitly show students where and why they made mistakes, which provides an opportunity for them to correct their understanding (Golter et al., 2016).

Many researchers have developed critical thinking assessment rubrics (Anderson & Soden, 2001; Facione, 2000; *Undergraduate Professional Education in Chemistry ACS Guidelines and Evaluation Procedures for Bachelor's Degree Programs*, 2015; *VALUE Rubric Development Project*, 2019; *Vision for Science and Mathematics Education*, 2014). This rubric is mostly holistic in nature and does not refer one by one to aspects of critical thinking. A holistic rubric for assessing critical thinking implies the expectation that students may have good critical thinking skills. However, if we examine each aspect one by one, it is possible that students are not proficient in one aspect of critical thinking (Saxton et al., 2012). For example, a student who is proficient in basic clarification skills is not necessarily proficient in inference skills. As a result, generalizing students' critical thinking abilities to one holistic score represents a loss of diagnostic data that could be used for feedback.

This diversity of student characteristics also occurs in high school students. The ability to provide answers when given a critical thinking test will vary from one student to another. This is where having a rubric that is able to assess high school students' answers plays an important role. Student characteristics start from the student's self, which is then influenced by the environment. Student characteristics influence the actions and decisions taken (Cooper & Brownell, 2020). In terms of responding to questions, it depends on their cognitive style (Kholid et al., 2020).

The purpose of this research was to develop a rubric that allows teachers to assess students' thinking skills in the classroom explicitly, which was modified from Ennis' critical thinking rubric. Ennis's critical thinking rubric is actually well designed. Given a well-designed rubric, there seems to be no reason to make modifications; However, the author believes that a more specific rubric, which uses language and examples from the biology discipline, would be more accessible and usable for biology teachers. Using this rubric also allows teachers to provide feedback to students. This rubric is specifically for assessing students' critical thinking based on students' answers when taking essay tests on biology material.

Research Method

Define The Scope of The Rubric

The rubric described here is intended to measure students' critical thinking abilities as a result of essay tests for biology learning. To develop this rubric, the definition of critical thinking and the aspects used refer to Ennis (2011). The development of this rubric was a modification of Ennis (2011) by maintaining the logical relationships of the original document and modifying it in a way that we believe builds content and expert validity for the purpose of assessing critical thinking.



Iterative Rubric Development

The initial step in developing this rubric is defining the scope. Through reviewing various literature, an initial version of the critical thinking rubric was prepared. Each aspect of critical thinking skills is assessed to determine appropriate descriptors and appropriate scoring. Next, the initial version is tested repeatedly. The first review was carried out by the researcher and colleagues in the research team. This initial review is intended to ensure that the rubric measures aspects of critical thinking abilities and is appropriate to the discipline of biology. This initial stage is also intended so that educators can understand each descriptor and can be applied when assessing student answers on essay tests.

This initial round of review is also intended to ensure that the rubric is ready to be tested in class. Next, each rubric is tested by checking students' answers after taking the essay test. The rubric is applied to student work that is checked by two different people on the same student's work. The scores given by each rater are used as a further review. Apart from that, researchers also held discussions with the raters if the scores given were different for the same answer. Feedback gathered from this follow-up review was used for further changes to the rubric. Feedback is collected in the form of discussion results and notes from the raters are used as a further review. Apart from that, researchers also held discussions with the raters. This follow-up review helps ensure that the rubric developed is truly capable of measuring critical thinking skills.

Data Sources and Data Analysis

The data sources used to develop the critical thinking rubric came from a literature review of how critical thinking is defined, feedback from validators, discussions regarding the suitability of the rubric to the essay tests that students took and the results of applying the rubric to samples of student work. Then the rubric developed is seen for its validity and reliability. Validity is demonstrated by the extent to which all the collected evidence supports the interpretation of the test scores intended for the proposed use. For the purposes of this research, we used two different types of validity, namely content validity and construct validity. Content validity takes into account the extent to which the rubric covers relevant aspects of each critical thinking ability. In this case, the definition of critical thinking skills and a review of the literature determine which categories are included in each rubric. The literature review is completed when the data is saturated or when no new aspects are discovered. Construct validity is the degree to which each rubric category accurately reflects the processes carried out by students. Evidence of construct validity is gathered through student test answers, and interviews with raters (Reynders et al., 2020).

The reliability of the rubric developed is carried out through interrater agreement. Interrater agreement was chosen as a measure of interrater reliability because of its common use in rubric development projects (Saxton et al., 2012). This process is carried out by looking at the percentage at which two raters agree on a rating or differ by one level (that is, they give adjacent ratings to the same test answers). The threshold for agreement, based on the number of possible levels of performance for each aspect in the rubric, was set at 80%. To calculate critical thinking rubric agreement, two raters discussed the scoring descriptors for each rubric and then independently scored students' critical thinking answers. Interrater agreement data were analyzed using Cohen's Kappa. The research stages can be seen more simply in Figure 1

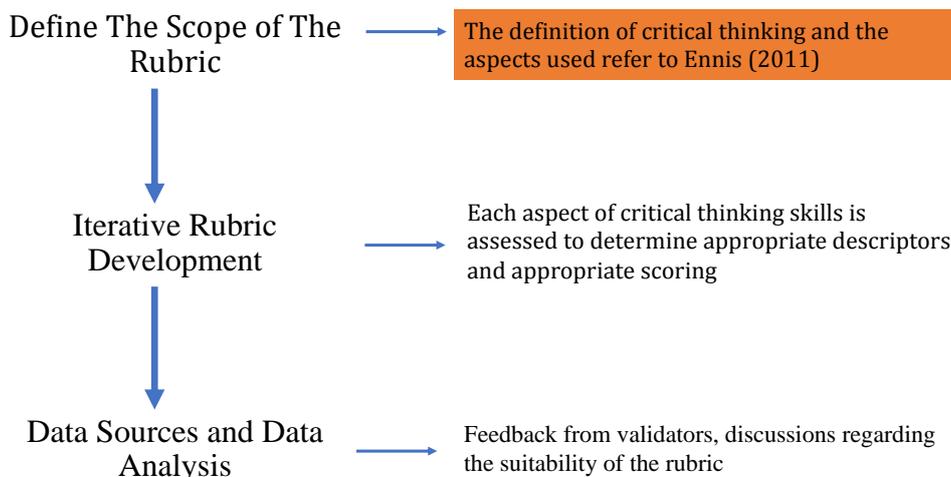


Figure 1. Research Process Steps

Results and Discussion

The rubric developed to assess critical thinking focused on the essay test. Aspects of critical thinking that are measured include basic clarification, bases for a decision, inference, advanced clarification, supposition and integration, strategies and tactics (Ennis, 2011). An overview of examples of essay test questions and targeted critical thinking aspects can be seen in Table 1.

Table 1. Examples of Questions and Critical Thinking Aspects Measured

Example Questions	Aspects of Critical Thinking
Gea went to school without breakfast because she was in a rush to attend lectures. Today lectures lasted long enough that Gea felt hungry. When we are hungry, we hear a distinctive sound from our stomach or what is known as "rumbling". Based on the problem in this question, formulate one question! Then from the question, determine the possible answer criteria accompanied by examples!	Basic clarification
Which blood groups in the ABO and Rh blood group systems can be safely transfused into someone with blood type A+? Write your analysis!	Bases for a decision
Astronauts who travel for long periods in space in a zero-gravity environment can experience muscle and bone changes. Give your argument to reach a conclusion regarding the statement in this question!	Inference
The term diabetes in ordinary people is always associated with diabetes. This type of diabetes is called diabetes mellitus. However, there is another diabetes term that is less known to the public, namely diabetes insipidus. What do you know about diabetes insipidus? In both diabetes (mellitus and insipidus) kidney function is affected. Why can urine frequency be a symptom of these two types of diabetes?	Advanced clarification
As an SPG (sales promotion girl), Kinara has to stand all day to promote the products she sells. After standing all day, he felt his legs swelling, but when he woke up in the morning the swelling was no longer there. How can this happen? Give an explanation!	Supposition and integration
The kidney is the first organ to be successfully transplanted. A donor can live a normal life with only one kidney so someone can donate their kidney to another person. Why can we still live with one kidney? As a student who has studied the structure and function of the kidneys, design a healthy lifestyle that can be lived by people who only live with one kidney so that! Write down the reasons for each rule you make!	Strategies and tactic

Critical thinking assessment using the developed rubric. This rubric was finalized after obtaining feedback from the raters who tested the rubric. This rubric has 6 aspects, and each aspect has a rating scale equipped with a descriptor. The critical thinking assessment rubric can be seen in Table 2.

Table 2. Critical Thinking Rubric

Aspect	Descriptor	Score
Basic Clarification	Formulate the question correctly. Identify criteria for appropriately considering possible answers. Complete the explanation or answer with examples	4
	Formulate the question correctly. Identify criteria for appropriately considering possible answers. Cannot complete explanation with examples	3
	Formulate questions correctly but the criteria used to provide explanations are not appropriate	2
	Formulates questions correctly but is unable to consider possible answers correctly and is not accompanied by examples	1
	Unable to formulate questions	0
	Bases for a decision	Determine whether the information is correct or incorrect by paying attention to the credibility of the source. Utilize data in making decisions or completing answers based on appropriate references
Determining whether information is correct or incorrect without paying attention to the credibility of the source. Utilize data in making decisions or providing answers		3
Determine whether information is correct or incorrect by paying attention to the credibility of the source. Not utilizing data in making decisions or providing answers		2
Determining the truth of information without paying attention to the credibility of the source and not using data in making decisions or providing answers		1
Unable to determine exact information		0
Inference		Make and determine conclusions correctly based on background facts. Putting forward reasons that are general to specific or vice versa towards drawing conclusions
	Make and determine conclusions correctly based on background facts. State the reasons for drawing conclusions but do not do it sequentially	3
	Make and determine conclusions correctly based on background facts. Not stating reasons for drawing conclusions	2
	Making and determining conclusions but they are not accurate because they are not based on background facts	1
	Can't make a conclusion	0
	Advanced clarification	Identify terms correctly. Define the term or consider a definition. Constructing arguments by providing answers in accordance with existing theories or facts
Identify terms correctly. Explain further terms accurately and correctly. The answer does not match existing theory or facts		3
Identify terms correctly. Explains the term further but is less precise		2
Identify and address intentional untruths but cannot explain further		1
Unable to provide further explanation		0
Supposition and integration	Consider and find the reasons for the questions asked correctly. Integrate theories or other facts in making assumptions	4
	Consider and find the reasons for the questions asked correctly. Making assumptions can integrate other theories or facts but is not precise	3
	Consider and find the reasons for the questions asked correctly. Unable to integrate theories or other facts in making assumptions	2
	Consider and find the reasons for the questions asked but are not quite right	1
Strategies and tactic	Unable to make assumptions or correct answers	0
	Formulate alternative solutions appropriately. Consider actions based on appropriate criteria	4
	Formulate alternative solutions appropriately. Considering action but the criteria used are inappropriate	3
	Formulate alternative solutions appropriately. Unable to consider actions based on appropriate criteria	2

Aspect	Descriptor	Score
	Formulate a solution but it is not precise	1
	Unable to formulate a solution	0

This critical thinking rubric consists of six aspects modified from Ennis (2011). In the first aspect, basic clarification, critical thinking assessment is focused on students' abilities in formulating problems and determining appropriate criteria for solving these problems. This aspect requires students' ability to provide clarification and explanation of existing problems by providing arguments and examples. When this critical thinking rubric was initially developed, it was not equipped with examples, then after receiving input from the raters, examples were added to the descriptor. An argument depends on its meaning and how others respond to it. Someone with good critical thinking skills will be trained to anticipate the failure of their arguments as a consequence of other people's objections, or in other words, they will already be able to imagine their own potential objections (Kuhn, 2019).

The bases for a decision aspect are viewed from students' ability to determine correct or incorrect information by utilizing data. Students who are able to determine whether information is correct or incorrect based on data will get the highest score, namely 4. The scoring results given by the raters show that students have not achieved the highest score. As a result of interviews with raters, determining scores for this aspect refers to the use of data or information when providing answers to questions. For example, students use information about blood type to determine whether someone can become a blood donor, but they have not stated whether the source used is credible or not. Based on these considerations, two raters agreed to give a score of 3 (Figure 2).

<p>2) menurut saya golongan darah yang dapat ditransfusi pada golongan darah A adalah golongan A dan O, Hal ini dikarenakan tidak semua darah dapat cocok. A akan cocok dengan A karena sejenis, dan O dapat mentransfuskikan ke A karena O bersifat netral jadi dapat diterima oleh semua jenis golongan darah. namun O tidak dapat menerima semua dari jenis golongan darah.</p>	<p>Translate: In my opinion, the blood types that can be transfused to blood group A+ are blood groups A+ and O. This is because not all blood can be compatible with A+ because it is similar and O can be transfused to A+ and) because it is neutral or can be accepted by all types of blood groups, however O cannot accept all blood types.</p>
<p>2) menurut saya golongan darah yang dapat ditransfusi pada golongan darah A adalah golongan A dan O, Hal ini dikarenakan tidak semua darah dapat cocok. A akan cocok dengan A karena sejenis, dan O dapat mentransfuskikan ke A karena O bersifat netral jadi dapat diterima oleh semua jenis golongan darah. namun O tidak dapat menerima semua dari jenis golongan darah.</p>	

Figure 2. Raters Scoring Results on the Bases for a Decision Aspect

When making decisions, students need to evaluate the information provided to relate it or see its relevance to the problem that needs to be solved. Students should ideally show evidence by identifying what information is relevant or not relevant, and showing why the information is relevant (Reynders et al., 2020). An important fundamental characteristic of critical thinking is the ability to search for an evidentiary basis to justify a point of view. Critical

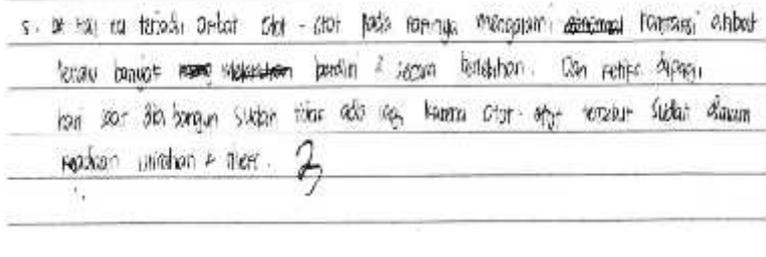


thinking is the first line of defense when information is not always reliable because it guides a person to hold beliefs that are consistent with the available evidence. Moreover, currently, objective facts and evidence are less powerful in shaping public opinion than personal beliefs, anecdotes and popular views (Cooke, 2017; Ku et al., 2019).

The next aspect of critical thinking that will be discussed is making conclusions. Making conclusions is given a score based on students' ability to make conclusions by stating facts or reasons stated sequentially. There is no need to revise the discussions held with the rater regarding this aspect because the descriptors presented already contain complete information in determining the score. For example, the highest score of 4 is given when participants systematically state reasons before making a decision. Students can express reasons appropriately when they have a lot of information, one of which is obtained from reading. By reading literature, they learn how to examine different points of view, find differences and similarities as well as variant interpretations and determine correct information (Din, 2020). Reading by exploring various information allows students to find cause and effect phenomena in a reading so that students are able to make conclusions (Khonamri & Karimabadi, 2015).

Advanced clarification is the fourth aspect discussed. Scoring using a critical thinking rubric in this aspect refers to students' ability to construct arguments by providing appropriate definitions and considering various facts. Argumentation improves decision-making and problem-solving through critical thinking processes. By providing arguments, students are able to provide clarification (Bezanilla et al., 2019). Reviews from the raters suggested that the description on score 3 be added so that the difference was directly visible from score 2. Improvements were made by adding details on score 3, namely the answer did not match existing theory or facts. Improvements to this rubric make it easier for raters to determine scores because sometimes students are able to identify terms correctly, but they do not match existing theory or facts. Sometimes students can transform a little information but do it accurately, or transform a lot of information and do it inaccurately. These two descriptors allow different scores to be assigned.

The next aspect of critical thinking is supposition and integration. Determining scores in this aspect refers to students' ability to consider and find reasons based on existing theory or knowledge. As a result of discussions with the raters, information was obtained that the descriptors provided could be used to provide a good assessment. Learners get high scores when they are able to find reasons for the questions and integrate theories or facts. Figure 3 below shows students getting a score of 3 (giving reasons correctly but the supporting theory is not correct) and giving a score of 1 because the reasons given are not correct. One of the critical thinking skills is seen in students' ability to interpret information using their previous knowledge to explain the meaning of something, make conclusions, match data with predictions, and extract patterns from data.

	<p>Translate: This occurs because the muscles in the legs contract due to standing too much and too much. And when he woke up in the morning, it was no longer there because the muscles were already at rest.</p>
---	---

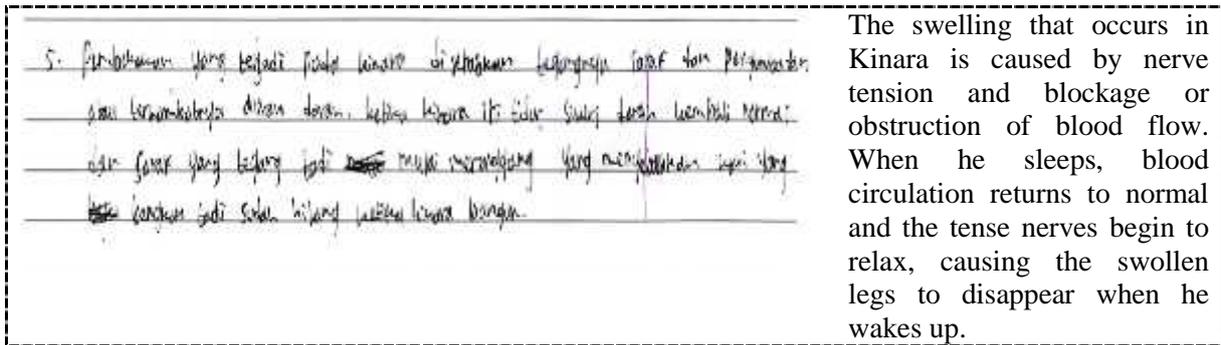


Figure 3. Examples of Student answers on the supposition and integration aspect

The final aspect of the critical thinking rubric developed is strategies and tactics. This aspect obtains input from the rater to provide details regarding the reasons for providing a solution. Based on this input, descriptors for the strategies and tactics aspects are added by providing reasons by considering appropriate criteria. This reason needs to be accompanied by appropriate arguments. Arguments are also a key aspect of critical thinking so that a critical thinker must be able to form well-structured and valid arguments (Lai, 2011). Determining strategy and tactics requires the ability to connect several pieces of information. Of course, students need to identify relationships between different pieces of information or concepts. This process can be used as a basis for determining strategy and tactics (Wale & Bishaw, 2020). For example, in this rubric, students need to gather information about kidney function and the working mechanisms of the kidneys in order to determine the right strategy when someone only has one kidney.

This critical thinking assessment rubric also looked at the interrater agreement of the raters using Cohen Kappa analysis. The results of the analysis can be seen in Table 3.

Table 3. Results of Cohen Kappa Analysis on the Critical Thinking Rubric

No	Aspects of Critical Thinking	Value of Agreement Kappa	Interpretation
1	Basic clarification	0.333	Fair
2	Bases for a decision	0.571	Moderate
3	Inference	0.455	Moderate
4	Advanced clarification	0.250	Fair
5	Supposition and integration	0.250	Fair
6	Strategies and tactic	0.182	Poor

A discussion of descriptors for each aspect of critical thinking has been explained previously. To determine the reliability of this rubric, the scores given by the assessors have been reviewed using Cohen's Kappa. Table 3 shows that every aspect of the critical thinking rubric can be used to assess critical thinking abilities. In the aspects of bases for decision and inference, the aggregate kappa values are 0.571 and 0.455 respectively, which means they are in the moderate category. Three other aspects are in the fair category, namely basic clarification, advance clarification, supposition and integration. These numbers show that the rubric is reliable for assessing critical thinking. A rubric must be reliable so that different assessors can use the rubric consistently (Reynders et al., 2020). Interrater agreement is the most common indicator of good research when two or more raters are involved to provide an assessment of an instrument so that the instrument's reliability can be assessed (Button et al., 2020; Tong et al., 2020).

The strategic and tactical aspect refers to the kappa value which is still in the poor category. Although the inter-rater agreement scores were much lower than desired, the adjacent agreement scores met the reliability threshold. We discussed with the raters the



differences in scoring in this aspect. The rater suggested that the descriptor in this aspect be made more detailed. To fulfill this suggestion, we added giving reasons when students answered questions. Apart from that, differences in understanding of the content can also cause raters to give exactly the same scores to the same student's answers (Reynders et al., 2020).

This research is in line with other research which explains that it is better for critical thinking assessment rubrics to use descriptions so that assessors can provide feedback (Reynders et al., 2020). Apart from that, assessments that refer to certain criteria mean that students can assess their level of self-confidence and performance so that they can improve learning (Pui et al., 2020).

Conclusion

The research results show that the critical thinking rubric developed can be used to assess critical thinking abilities. The critical thinking aspects assessed are basic clarification, bases for a decision, inference, advanced clarification, supposition and integration, strategies and tactics. Each aspect in this rubric has been equipped with a descriptor. This rubric is quite reliable, according to the Cohen Kappa analysis.

Recommendation

Critical thinking rubrics cannot cover all types of critical thinking questions tested. This rubric is intended for essay tests only. For future researchers, to accommodate the wider need for critical thinking rubrics, it is also necessary to develop rubrics for question types other than essays. Apart from that, this rubric has not yet reached the very good category when tested statistically because it is still in the moderate category. Therefore, there is still an opportunity to improve the existing rubric. We also have not confirmed that the use of this rubric can be applied to scientific disciplines other than biology, so there are still opportunities to conduct further research. This research was supported by Universitas Islam Riau, Indonesia.

References

- Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2014). Strategies for Teaching Students to Think Critically: A Meta-Analysis. *Review of Educational Research*, 1–40. <https://doi.org/10.3102/0034654314551063>
- Anderson, T., & Soden, R. (2001). Peer interaction and the learning of critical thinking skills. *Psychology Learning & Teaching*, 1(1), 37–40.
- Bezanilla, M. J., Fernández-Nogueira, D., Poblete, M., & Galindo-Domínguez, H. (2019). Methodologies for teaching-learning critical thinking in higher education: The teacher's view. *Thinking Skills and Creativity*, 33(100584), 1–10. <https://doi.org/10.1016/j.tsc.2019.100584>
- Biggs, J. (2014). Constructive alignment in university teaching. *HERDSA Review of Higher Education*, 1(1), 5–22.
- Brookhart, S. M., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, 67(3), 343–368. <https://doi.org/10.1080/00131911.2014.929565>
- Butler, H. A. (2012). Halpern Critical Thinking Assessment Predicts Real-World Outcomes of Critical Thinking. *Applied Cognitive Psychology*, 26(5), 721–729. <https://doi.org/10.1002/acp.2851>
- Button, C. M., Snook, B., & Grant, M. J. (2020). Inter-Rater Agreement, Data Reliability, and The Crisis of Confidence in Psychological Research. *Quant Methods Psychol*,



- 16(5), 467–471.
- Cooke, N. A. (2017). Posttruth, Truthiness, and Alternative Facts: Information Behavior and Critical Information Consumption for a New Age. *The Library Quarterly*, 87(3), 211–221. <https://doi.org/10.1086/692298>
- Cooper, K. M., & Brownell, S. (2020). *Student Anxiety and Fear of Negative Evaluation in Active Learning Science Classrooms Chapter 56 Student anxiety and fear of negative evaluation in active learning science classrooms. February.* <https://doi.org/10.1007/978-3-030-33600-4>
- Din, M. (2020). Evaluating university students' critical thinking ability as reflected in their critical reading skill: A study at bachelor level in Pakistan. *Thinking Skills and Creativity*, 35(100627), 1–11. <https://doi.org/10.1016/j.tsc.2020.100627>
- Ennis, R. H. (2011). *The Nature of Critical Thinking: An Outline of Critical Thinking Dispositions.* 1–8. http://faculty.education.illinois.edurhennis/documents.TheNatureofCriticalThinking_51711_000.pdf
- Facione, P. A. (2000). The Disposition Toward Critical Thinking: Its Character, Measurement, and Relationship to Critical Thinking Skill. *Informal Logic*, 20(1), 61–84.
- Golter, P., Brown, G., & Wie, B. Van. (2016). Preliminary Validation of a Critical Thinking Rubric for Fluid Mechanics and Heat Transfer. *Journal of STEM Education*, 17(4), 19–28.
- Gray, K., & Koncz, A. (2018). *The key attributes employers seek on students' resumes.* <http://www.naceweb.org/about-us/press/2017/the-key-attributes-employers-seek-on-students-resumes/>
- Halpern, D. F. (2010). *Halpern Critical Thinking Assessment.* SCHUHFRIED. <http://www.schuhfried.com/vienna-test-system-vts/all-tests-from-a-z/test/hcta-halpern-critical-thinking-assessment-1/>
- Hart, C., Da, C., Souza, D. D., Kimpton, A., & Ljbusic, J. (2021). Exploring higher education students' critical thinking skills through content analysis. *Thinking Skills and Creativity*, 41(100877), 1–9. <https://doi.org/10.1016/j.tsc.2021.100877>
- Hattie, J., & Gan, M. (2011). *Instruction based on feedback. Handbook of research on learning and instruction.*
- Kholid, M. N., Hamida, P. S., Pradana, L. N., & Maharani, S. (2020). Students' Critical Thinking Depends On Their Cognitive Style. *International Journal of Scientific & Technology Research*, 9(01), 1045–1049.
- Khonamri, F., & Karimabadi, M. (2015). Collaborative Strategic Reading and Critical Reading Ability of Intermediate Iranian Learners. *Theory and Practice in Language Studies*, 5(7), 1375–1382. <http://dx.doi.org/10.17507/tpls.0507.09>
- Kong, L.-N., Qin, B., Zhou, Y., Mou, S., & Gao, H.-M. (2014). The effectiveness of problem-based learning on development of nursing students' critical thinking: A systematic review and meta-analysis. *International Journal of Nursing Studies*, 51, 458–469. <https://doi.org/10.1016/j.ijnurstu.2013.06.009>
- Ku, K. Y. L., Kong, Q., Song, Y., Deng, L., Kang, Y., & Hu, A. (2019). What predicts adolescents' critical thinking about real-life news? The roles of social media news consumption and news media literacy. *Thinking Skills and Creativity*, 33(100570), 1–12. <https://doi.org/10.1016/j.tsc.2019.05.004>
- Kuhn, D. (2019). Critical Thinking as Discourse. *Human Development*, 62(3), 146–164. <https://doi.org/10.1159/000500171>
- Lai, E. R. (2011). Critical Thinking: A Literature Review. *Pearson's Research Reports*, 6(1),



- 1–49. <https://doi.org/10.1046/j.1537-2995.1995.35395184278.x>
- Lin, S. (2013). Science and non-science undergraduate students' critical thinking and argumentation performance in reading a science news report. *International Journal of Science and Mathematics Education*, 12(5), 1023–1046. <https://doi.org/10.1007/s10763-013-9451-7>
- Liu, O. L., Frankel, L., & Roohr, K. C. (2014). Assessing Critical Thinking in Higher Education: Current State and Directions for Next-Generation Assessment. *ETS Research Report Series*, 1, 1–23.
- Pearl, A. O., Rayner, G. M., Larson, I., & Orlando, L. (2018). Thinking about critical thinking: An industry perspective. *Industry and Higher Education*, 1–11. <https://doi.org/10.1177/0950422218796099>
- Pui, P., Yuen, B., & Goh, H. (2020). Using a Criterion-referenced Rubric to Enhance Student Learning: a case study in a critical thinking and writing module. *Higher Education Research & Development*. <https://doi.org/10.1080/07294360.2020.1795811>
- Reynders, G., Lantz, J., Ruder, S. M., Stanford, C. L., & Cole, R. S. (2020). Rubrics to assess critical thinking and information processing in undergraduate STEM courses. *International Journal of STEM Education* <https://doi.org/10.1186/s40594-020-00208-5>
- Saxton, E., Belanger, S., & Becker, W. (2012). The Critical Thinking Analytic Rubric (CTAR): Investigating intra-rater and inter-rater reliability of a scoring mechanism for critical thinking performance assessments. *Assessing Writing*, 17, 251–270.
- Shaw, A., Liu, O. L., Gu, L., Kardonova, E., Chirikov, I., Li, G., Hu, S., Yu, N., Ma, L., Guo, F., Su, Q., Shi, J., & Loyalka, P. (2019). Thinking critically about critical thinking: validating the Russian HEIghten ® critical thinking assessment. *Studies in Higher Education*, 2–16. <https://doi.org/10.1080/03075079.2019.1672640>
- Smit, R., & Birri, T. (2014). Assuring the quality of standards-oriented classroom assessment with rubrics for complex competencies. *Studies in Educational Evaluation*, 1–9. <https://doi.org/10.1016/j.stueduc.2014.02.002>
- Stein, B., & Haynes, A. (2011). Engaging Faculty in the Assessment and Improvement of Students' Critical Thinking Using the Critical Thinking Assessment Test By Barry Stein and Ada Haynes. *Change: The Magazine of Higher Learning*, 43(2), 44–49.
- Tong, F., Tang, S., Irby, B. J., Lara-alecio, R., & Guerrero, C. (2020). The determination of appropriate coefficient indices for inter-rater reliability: Using classroom observation instruments as fidelity measures in large-scale randomized research. *International Journal of Educational Research*, 99(101514), 1–10. <https://doi.org/10.1016/j.ijer.2019.101514>
- Undergraduate Professional Education in Chemistry ACS Guidelines and Evaluation Procedures for Bachelor's Degree Programs*. (2015). American Chemical Society Committee on Professional Training.
- VALUE Rubric Development Project*. (2019). Association of American Colleges and Universities. <https://www.aacu.org/value/rubrics>
- Vision for science and mathematics education*. (2014). The Royal Society Science Policy Centre.
- Wale, B. D., & Bishaw, K. S. (2020). Effects of using inquiry-based learning on EFL students' critical thinking skills. *Asian-Pacific Journal of Second and Foreign Language Education*, 5(9), 1–14.