Research Article https://doi.org/10.12973/eu-jer.12.4.1781

European Journal of Educational Research
Volume 12, Issue 4, 1781 - 1794.
ISSN: 2165-8714
http://www.eu-jer.com/

Development of Waves Critical Thinking Test: Physics Essay Test for High School Student

Ani Rusilowati* Universitas Negeri Semarang, INDONESIA
Ridho Adi Negoro Universitas Negeri Semarang, INDONESIA
Mahardika Prasetya Aji Universitas Negeri Semarang, INDONESIA
Bambang Subali Universitas Negeri Semarang, INDONESIA

Received: January 31, 2023 ▪ Revised: April 1, 2023 ▪ Accepted: May 25, 2023

Abstract: This study aims to produce a product to evaluate students' critical thinking skills that departs from physics content where students often have misconceptions. This research is a development research with research stages covering a) research and review literature; (b) planning chapter objectives; (c) developing a preliminary form; (d) field-testing the preliminary form; (e) Revise the preliminary form; (f) conducting a main field-test. The Waves Critical Thinking (WCT) test developed consists of 7 questions with 15 specific domains. Total percentage of content validity test was obtained 87.98% with appropriate criteria and based on the construct validity WCT test, the Goodness of Fit criteria were obtained which were classified as fit. The test instrument being tested consists of 15 objective items. The reliability of WCT test results 0.597 as a Cronbach's alpha score with the medium category and all the components have a good level of composite reliability. The outcome of the study was the WCT test with a valid state for measuring students' CT in a specific domain of physics wave material.

Keywords: *Assessment in physics, essay test, physics essay test, waves critical thinking test.*

To cite this article: Rusilowati, A., Negoro, R. A., Aji, M. P., & Subali, B. (2023). Development of waves critical thinking test: Physics essay test for high school student. *European Journal of Educational Research*, *12*(4), 1781-1794. https://doi.org/10.12973/eu-jer.12.4.1781

Introduction

The technology and science fields have advanced quickly, especially on education world, increasingly emphasizes the need for new skills that are relevant for students (Akbar et al., 2020; Kravchenko et al., 2023; Rahimi, 2020; Rahimi & Park, 2020; Yulianti et al., 2022). This need corresponds to the future of students after taking the educational path, so that they can compete and be able to produce strategic solutions to new problems in life (Chusni et al., 2020; Selman & Jaedun, 2020; Supena et al., 2021). One of the skills that is being discussed in various types of science is critical thinking skills. These skills are considered an important part of an overall problem solving process. Moreover, these skills can be considered as a foundation for both personal and community competency development (Behar-Horenstein & Niu, 2011; Larsson, 2021).

The scope of the problem is more specific, namely in physics learning, the role of critical thinking skills is needed to become a physicist (McMillan et al., 2018; Schmaltz et al., 2017). This is because scientific activities that are closely related to science, especially physics, cannot be separated from critical thinking activities, so that one of the efforts to make physics education and learning effective is to focus on critical thinking skills of students (Osborne, 2014; Sidiq et al., 2021).

The other side shows that learning difficulties that are traced to students' critical thinking skills in studying physical phenomena are often found in the form of misconceptions (Negoro et al., 2020; Rusilowati et al., 2020). The term misconception in general is a mistake about the method or basic knowledge of science which refers to a valid measure of scientific literacy (Kahan, 2017). The causes of misconceptions are often revealed in research through various approaches, both from the material provider, learning media, or from the experience of the students themselves (Gurel et al., 2015; Negoro et al., 2020; Rahmawati et al., 2017).

Several studies have shown the possibility of misconceptions that occur in the conception of physics, especially the wave material, which is proven through a concept mastery test. Apparently, there are still many misconceptions about the interpretation of the concept compared to the valid conception, including the speed of propagation of a mechanical wave

* **Corresponding author:**
Ani Rusilowati, Universitas Negeri Semarang, Jawa Tengah, Indonesia. ✉ rusilowati@yahoo.com

is influenced by the shape of the pulse, frequency, and wavelength regardless of the medium of propagation (Barniol & Zavala, 2016; Negoro et al., 2023; Reyes & Rakkapao, 2018). Then, there is a physical error in understanding that departs from mathematical equations, which are still found, one of which is the difficulty in understanding the meaning of positive and negative signs in the wave equation, the difference in the trigonometric forms used (sin or cos), and so on (Auli et al., 2018; Kennedy & de Bruyn, 2011; Somroob & Wattanakasiwich, 2017; Zaleha et al., 2017).

Starting from students' thinking processes, misconceptions will be closely related to students' critical thinking skills (Dellantonio & Pastore, 2021; Stupple et al., 2017). Several factors that cause misconceptions are physical phenomena from students' daily experiences which are naturally interpreted by students themselves to produce a frame of mind as the meaning of these phenomena (Bozzi et al., 2020; Suprapto, 2020). The framework in question is a pattern which is a simplification of the translation of phenomena using students' thinking instruments naively. The accuracy of this interpretation is strongly influenced by critical thinking skills which in turn with a description of students' critical thinking skills, we will be able to uncover the main problems that cause misconceptions (Negoro et al., 2020). Based on this description, an evaluation instrument is needed which can be called a diagnostic tool that refers to critical thinking skills.

Various experts argue that the critical thinking skills tool used in physics education has another function, namely as a diagnostic tool for physics material misconceptions (Negoro et al., 2020; Tiruneh et al., 2017). Although, there is still little research that reveals this, rationally, it can be understood that revealing students' ways of thinking about a physical phenomenon which is then called a concept will simultaneously reveal their critical thinking skills (Dellantonio & Pastore, 2021; Negoro et al., 2020; Rusilowati et al., 2020). As a result of this, it is clear how critical thinking skills and misconceptions are closely related qualitatively. This relationship can be clarified by describing that critical thinking skills are the main instrument for students to understand the concept of physics (Dellantonio & Pastore, 2021).

Departing from several research results on critical thinking skills, in general, there are still questions about the level of validity of evaluation tools that are still in doubt (Leach et al., 2020; Shaw et al., 2020). One of the confirmations is that there are many debates and proposals that every scientific field should have a different thinking skill domain from other scientific fields. In addition, many aspects of thinking skills are offered by many experts with the support of various arguments that are considered unsuitable for certain scientific fields (Hart et al., 2021; Ngajie et al., 2020). This is the basis for the preparation of a critical thinking skill evaluation tool to date.

Based on the necessity of preparing in critical thinking abilities evaluation tools, then we will be faced with a choice of versions of the evaluation tools so that it can reveal the components of critical thinking skills as a whole. Various types of critical thinking skills evaluation tools and their forms were made by several experts with various basic arguments. In general, essay tests are mostly prepared by several experts because they are very good for measuring high-level mental processes (Dudung & Oktaviani, 2020; Rusilowati, 2014). Aside from, Essay assessments typically have the ability to fully show pupils' ideas with a picture of an authentic thinking sequence, such as when students compose a reasoning or argument that requires a process of analysis, synthesis, and evaluation (Risnita & Bashori, 2020; Rusilowati, 2014).

Various things behind the need for critical thinking skills evaluation tools in accordance with developing a critical thinking skill assessment as a result of the prior description by researchers that comprehensively raised the theme of waves that often occured misconceptions. The test tool was developed in an effort to reveal the critical thinking capabilities or skills of physics students. The content of test tool raised as a problem for each item focuses on content where students experience misconceptions. Abstraction of a physics concept is an important part in terms of uncovering students' thinking skills.

Methodology

Research Design

This study falls under the category of research on educational development, or R&D (Research and Development) for short. This kind of research and development (R&D) is a procedure for creating and evaluating a product (Sugiyono, 2013). Learning tools are products developed here in after referred to as Waves Critical Thinking (WCT) Test for intermediate level students. The Waves Critical Thinking (WCT) Test's research and development process is divided into six stages, which are adapted from Borg and Gall (2007) R&D model.

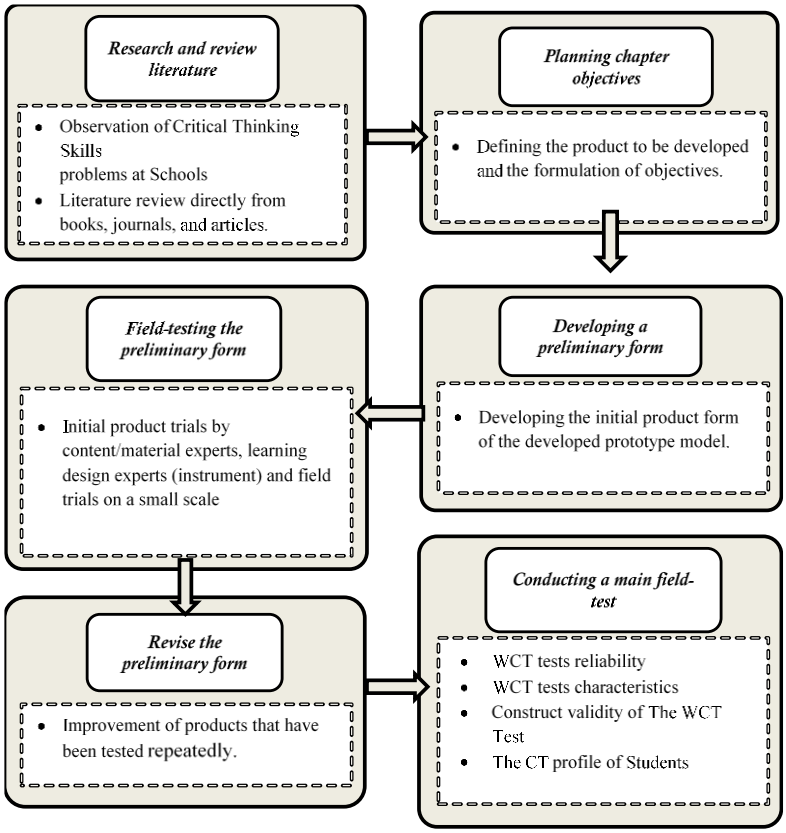


Figure 1. Step of The WCT test Production

Sample and Data Collection

Field trials were conducted in this investigation, specifically at the primary field-test stage. The trial was conducted at Teuku Umar High School Semarang and Kesatrian 2 High School Semarang with 106 students as samples. The method of sampling employed on this study is purposive sampling where this technique determines the sample with certain considerations. The considerations are (a) Students are part of a group that has studied physics material, namely waves; (b) Students are part of a group that has studied physics material, namely motion kinematics, particle dynamics, circular motion, and simple harmonic motion.

Analyzing of Data

The initial stage starts from the analysis of potential and problems where research is carried out to obtain various main elements to unravel the research problem. The design of the wave material's critical thinking test was then put together using the information gathered in the first phase. In steps 4 and 6, the product trial was conducted twice, in which the first trial could be categorized as a small-scale trial with a limited sample and the trial in step 6 was a large-scale trial by taking into account the large number of samples required for the fulfillment of statistical analysis. The Waves Critical Thinking (WCT) Test, a product of this study, is a description test with three themes and a variety of critical thinking ability descriptions that will be used to analyze and rate the critical thinking abilities of middle-level pupils. The wave material raised consists of several topics that fall into the category of mechanical waves, namely traveling waves, stationary waves, and wave characteristics in a medium.

Prior to the field trial, the test instrument was tested for feasibility by making a feasibility analysis using professional judgment. The purpose of this feasibility analysis is to examine the relevance of the content in terms of (a) the material component specifically the compatibility of the content's presentation with the relevant theory; (b) The aspect of construction, which includes the content's completeness, the test questions' applicability to the many aspects, the instrument's ability to disclose critical thinking abilities, and (c) the aspect of language. Specifically, the employment of strong phrases that adhere to the principles of physics.

The formula was used to examine the feasibility test findings using expert judgment (Anas, 2008)

$$P = \frac{f}{N} x\ 100\%,$$

where, P : assessment percentage; f: score achieved; N : total score.

The interpretation of the results obtained are listed in Table 1 as follows the criteria are as follows:

Table 1. Eligibility Criteria

Percentage (P)	Description
1 % < P ≤ 50 %	Not feasible
50% < P ≤ 70 %	Quite feasible
70% < P ≤ 85 %	Feasible
85% < P ≤ 100%	Very feasible

Furthermore, The Waves Critical Thinking Test was put through testing to determine its characteristics as well as its reliability and construct validity. Using the construct validity with the goodness of fit value, structural equation modeling (SEM) analysis offers a model with a fit level. Additionally, a second order confirmatory factor analysis (CFA) is used to demonstrate the construct validity of the WCT by determining the indicator value for the latent variable. In this case, CFA is used because the latent variables already have clear factors that adapt from the critical thinking domain according to Halpern Critical Thinking Theory (Tiruneh et al., 2016). Trials were also done to gather test findings as an analytical tool, including the level of difficulty, power of difference regarding each distinct domain of critical thinking skills, construct validity, and profiles of students' critical thinking skills. In addition, to test the reliability of the instrument, the Alpha Cronbach formula was used. The essay test reveals students' answers with a score that refers to the rubric that had been described in the indicators for each specific CT domain. The essay test reveals students' answers with a score that refers to the rubric that has been described in the indicators for each specific CT domain. The intended rubric sample can be seen in the Table 2.

Table 2. Sample of Rubric WCT Test

Theme 1 General Domain: Argument Analysis Items (Question Number 2)			
Specific Domain	**Indicators**		**Score**
Argument Part: Identify the important parts of an argument	Students answer with show variable the waves that are referencing arguments precisely according to the wave concept	Relationship of equation $v = \lambda f$	1
		Melde equation of waves	1
		Other variables from the environment (time-based disturbance instability, homogeneity of the rope material, the direction of the thrust relative to the acceleration due to gravity)	1
Inference of Data: Criticize validity deep generalization test	Students answered with show deficiencies or adequacy of analysis/data of the argument according to the concept of waves	Dependent and control variables of the equation $v = \lambda f$	1
		Dependent and control variables of waves	1
		Other variables from the environment that can influence (time-based disturbance instability, homogeneity of the rope material, the direction of the thrust relative to the acceleration due to gravity)	1
Generalization Validity: Summing up statements correct from the data set which are given	Students answer with choose the right argument along with the analysis exactly according to the wave concept	Choose based on relationship of equation $v = \lambda f$	1
		Choose based on relationship of Melde equation of waves	1
		Choose based on relationship of other variables from the environment with the phenomenon (time-based disturbance instability, homogeneity of the rope material, the direction of the thrust relative to the acceleration due to gravity)	1
Flawed Argument: Identifying information missing relevance an argument	Students answer with show that analysis hiatus of the argument corresponds to the concept of waves	Shows the part that is less based on equation $v = \lambda f$	1
		Shows the part that is less based on Melde equation of waves	1

Findings/Results

The Waves Critical Thinking (WCT) Test instrument is the final product of this development research. One of the strategic uses of this tool is to evaluate middle-level students' critical thinking skills with wave content. Many physics misconceptions are found, one of which is wave material in students, which is the main reference point for how critical thinking skills are closely related to the problem. If pupils possess critical thinking abilities, the issue of wave material misconceptions with a high level of abstraction is thought to be solved. Additionally, the Waves Critical Thinking (WCT) Test, which aims to expose critical thinking abilities, is anticipated to enable the results to be utilized as a benchmark for students' wave-related analytical skills. The concepts or conceptual connections that pupils believe more logically can be demonstrated by this analytical skill.

Research & Review Literature

At this stage, a study of information gathering is carried out by means of a literature study and field observations as a basis for determining the formulation of the problem and determining the orientation and strategy of product development to be produced. The results of the search at this stage are broadly (a) there is a connection between the issue of physics misconceptions, particularly the wave material, and the critical thinking abilities of middle-level students; (b) It is very challenging to find an evaluation tool for critical thinking skills that refers to the issue of physics misconceptions. In addition, this research will result in a product that addresses these issues in the form of an assessment instrument called the Waves Critical Thinking Test (WCT).

Planning Chapter Objectives

The planning stage includes the activities of defining the product to be developed, formulating objectives, estimating funding requirements, estimating labor and time, and working procedures required during the research. Figure 2 illustrates the material covered by the Waves Critical Thinking Test (WCT), which was developed based on the need analysis conducted during the Research and literature review stage.

Based on Figure 2, the WCT test content developed covers five domains of CT which aims to reveal critical thinking skills based on wave mechanics content which is segmented into three question themes. The raised mechanical wave content is a phenomenon that is closely related to everyday life by providing opportunities for involvement of variables that may not be controlled as in the laboratory. This becomes material for examining the details of students' critical thinking skills.

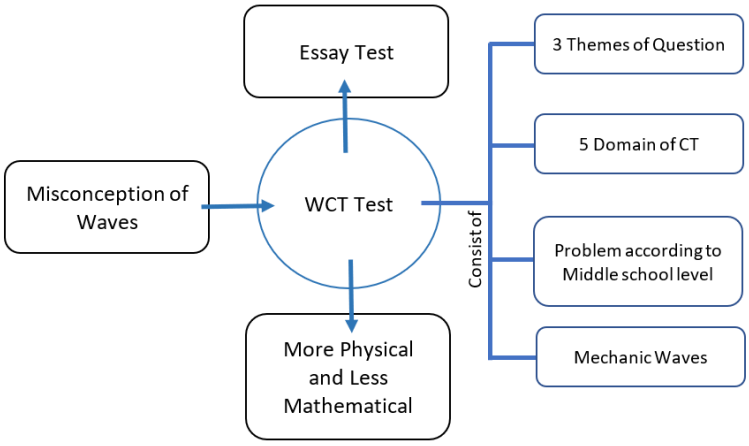


Figure 2. WCT Test Contents Break-down

Developing a Preliminary Form

The stage of developing the preliminary form of product, includes the development of the initial product form of the prototype model being developed, including preparing assessment instruments, and others needed for product testing. The initial product was developed by focusing on the specification of critical thinking domains that are more directed to wave content. The five core areas of critical thinking—hypothesis testing, argument analysis, reasoning, likelihood and uncertainty analysis, and problem-solving and decision-making—are divided into specific domains for the Waves Critical Thinking Test (WCT) (Tiruneh et al., 2016). The foundational elements of pupils' critical thinking abilities will consist of these five elements. Additionally, the construct is detailed in a second-order supporting component that we will refer to as the Waves Critical Thinking Test for Domain-Specific Problems (WCT). Domain-specific disaggregation produces 18 main indicators which can be seen in Table 3.

The WCT test is structured in an essay test consisting of 3 themes referring to the content of mechanical waves. Theme 1 contains traveling wave content. Theme 2 contains experimental content to determine the main factors determining the

speed of wave propagation in a medium. Theme 3 contains experimental content to validate Melde's equation regarding the speed of wave propagation in a medium. Each theme is strived to be very contextual with real life. Experimental content discussed in theme 3 is related and closely related to students' experiences in everyday life.

Table 3. Specific Domain Thinking Skills of WCT test

No	Aspect of CT	WCT Test Spesific Domain
1	Hypothesis testing	Variable relation Completeness of Information Causality Principle Generalization bias
2	Argument Analysis	Argument part Inference of Data Generalization validity Flawed Argument
3	Reasoning	Evaluating data Measuring error Interpreting result Ambiguity result
4	Likelihood and uncertainty analysis	Predicting by probability Making decision Addition Information
5	Problem-solving and decision-making	Alternative decision Work procedure Evaluate Solution

Wave content that is too mathematical and has concepts that are difficult to teach is lifted and packaged in a WCT test to make it easier to convey the constraints of each item so that the analysis bias of students' critical thinking between physics and mathematics orientation is reduced. This is in line with the issue, which is that many students still have misconceptions regarding wave phenomena, which are directly tied to daily life (Chen et al., 2020; Karim et al., 2016; Xu et al., 2020). The WCT test is based on this issue and aims to demonstrate the critical power in students' thinking, as well as how they comprehend mathematical formulations and assess the phenomena of wave in life.

Field-testing The Preliminary Form

This stage is then continued with preliminary field testing, namely initial product testing activities by content/material experts, learning design experts (instrument) and field trials on a small scale. Material experts examine the criteria for the test instrument, including (a) the items' fit for the test's objectives and the test population; (b) the quality of the data given in the items; and (c) the clarity of the words, phrases, and representations of each item. The overall proportion with extremely feasible criteria according to the feasibility test questionnaire was 87.98%.

Although the achievement of eligibility is high, there are suggestions for input from material experts to improve the preparation of the WCT test. The suggestion in question is the need for some improvements in the form of adjusting the choice of words that are more easily understood by middle-level students.

In addition, instrument testing was conducted on several middle-level students who had been taught the Wave material. The qualitative approach was carried out by digging up data through several interviews with students who carried out the WCT test. According to the findings of the interviews, students generally thought that the context and content of the WCT test were simple to understand. Also, each item's context is thought to be strongly tied to students' everyday lives so that they can respond truthfully based on their observations.

Revise The Preliminary Form

Revise the preliminary form stage, is the stage of refinement or improvement of products that have been tested. The implementation is carried out repeatedly so that a better product is obtained, which is called the main product that is ready.

The focus of improvement is to improve sentence structure based on word selection that is more appropriate for middle-level students. This repair is carried out repeatedly with the assistance of a material expert.

Conducting a Main Field-test

In the conducting a main field-testing stage, the main product testing activities are carried out on a wide scale. The trial was conducted on 106 students from 2 schools, namely SMA Teuku Umar Semarang and SMA Kesatrian 2 Semarang. The input obtained from this trial phase is used as input for revising the product.

Broadly speaking, there are 2 main activities in the Conducting a main field-test stage. These activities are operation field testing and final review of products. Operational field testing is a type of test that is also referred to as an empirical test. This action is done to evaluate the reliability of the product. The results are evaluated and compared to see the strengths and weaknesses and to assess whether the developed WCT test is feasible. The final product review stage is the final revision stage of the resulting product, in order to obtain a product that is ready to be disseminated and implemented.

Characteristics of WCT Test

Hypothesis testing, Argument Analysis, Reasoning, Likelihood and Uncertainty Analysis, and Problem-Solving and Decision-Making are the critical thinking domains identified by Halpern that are sought to be examined in the WCT test (Tiruneh et al., 2016).

Table 4. Critical Thinking Skills Achievement

Item	Difficulty Level	Power of Difference	Interpretation
1	0.66[b)]	0.61	Item used
2	0.82 [a)]	0.57	Item used
3	0.70 [b)]	0.43	Item used
4	0.29 [c)]	0.69	Item used
5	0.68 [b)]	0.55	Item used
6	0.54 [b)]	0.52	Item used
7*	0.26 [c)]	0.14	Item not used
8	0.71 [a)]	0.44	Item used
9	0.89 [a)]	0.59	Item used
10	0.86 [a)]	0.47	Item used
11	0.71 [a)]	0.63	Item used
12	0.63 [b)]	0.44	Item used
13	0.46 [b)]	0.45	Item used
14	0.54 [b)]	0.43	Item used
15	0.70 [b)]	0.42	Item used
16*	0.89 [a)]	0.13	Item not used
17	0.51 [b)]	0.50	Item used
18*	0.80 [a)]	0.16	Item not used

[a)] Easy [b)]Medium [c)]Hard

The description of the questions' degree of difficulty and differentiating power in Table 3 provides quantitative parameters of the WCT test. The WCT test consists of 7 questions with a total of 18 objective items, and it has been determined that there are 7 easy, 9 middle, and 2 difficult items. Then, the discriminatory power of each item shows that 3 items do not meet the minimum value of distinguishing power so that they must be removed from the WCT test item objectives.

The test results were also examined in order to determine the WCT test's reliability value. The exam results of 106 middle school pupils demonstrate the test's dependability. There were 15 objective elements on the WCT exam instrument. According to the reliability findings, the medium category's Cronbach's alpha score is 0.597.

Construct Validity of WCT Test

The construct validity test was used to evaluate how well the WCT exam results used to gauge students' critical thinking abilities described the construct of the variable being assessed. The construct referred to here is a specific critical thinking skill domain, which consists of 15 constructs as well as a reduction of 18 constructs based on the value of the distinguishing power obtained.

The construct validity of the WCT test uses confirmatory factor analysis (CFA) second order, namely by calculating the estimated value of the indicator on the latent variable. The validity of an observed variable can be seen from the factor loading (factor loading = λ) of the variable on the latent variable. If the t-count result is higher than the critical value with a significance level of 0.05, which is 1.96, then the variables are considered to have strong validity versus constructs or other variables (Rusilowati, 2014). Construct validity was analyzed by second order confirmatory factors to produce a Goodness of Fit construct model. The Structural Equation Modeling (SEM) from the Lisrel 8.80 software suite was utilized for data analysis in this validity test. According to Ferdinand, SEM involves the following steps (a) creation of a model based on theory; (b) creation of a path diagram; (c) assessment of the Goodness-of-Fit standards; (d) in this step, the model's adequacy is assessed through an examination of several Goodness-of-Fit criteria. Table 5 and Figure 3 show the outcomes of the Goodnes-of-fit criteria achieved and the structural modeling together with t-values.

Table 5. Goodness of Fit Criteria for Critical Thinking Test Instrument

No	Goodness of fit Index	Cut off Value	Result	Category
1	X^2 Chi square	≤ α.df (df: 105)	1.285	FIT
2	Significance probability	≥ 0,05	all items are significant	FIT
3	GFI	≥ 0,90	0,91	FIT
4	AGFI	≥ 0,90	0,96	FIT
5	CFI	≥ 0,95	0,93	FIT
6	TLI / NNFI	≥ 0,95	0,98	FIT
7	NFI	≥ 0,95	0,96	FIT
8	RMSEA	≤ 0,08	0,01	FIT
9	SRMR	≤ 0,08	0.04	FIT

Table 5 shows that the fit category met all nine of the Goodness of Fit Standards for Critical Thinking Test Instruments. As there is no discernible difference between the covariance matrices and the chi-square value is less than the cut-off value of 5.25, the model is deemed to be fit. In addition, the values of GFI, AGFI, CFI, TLI, NFI, and RMSEA have values that satisfy the cutoff value thresholds, at which point the index value is said to be model-fit. It is evident from the index's findings that the instrument has a high level of validity. In addition, this achievement is in accordance with content validation which is carried out through adjusting content with the variables to be measured.

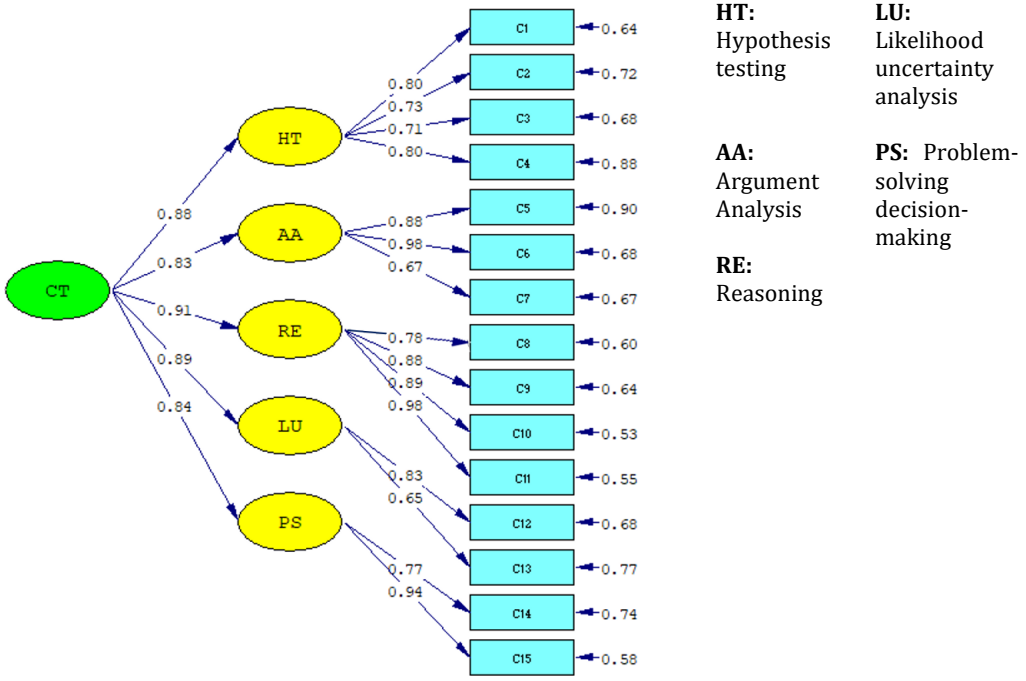


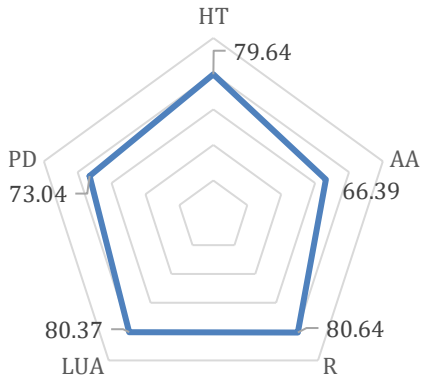
Figure 3. Basic Model for Estimation

The value of the construct variable's loading factor to the latent variable is shown in Figure 3. Based on the results, it is clear that the overall loading factor value is greater than 0.50, which suggests that the construct variable has strong validity as an observable indicator to represent the circumstance or description of the hidden variable. Because sufficient factor loading was achieved, the re-estimate as a model enhancement was not performed. all the components have a good level of composite reliability (more than 0.7) including (a) Hypothesis testing=0.855; (b) Argument analysis=0.813; (c) Reasoning=0.842; (d) Likelihood uncertainty analysis=0.721; (e) Problem-solving decision-making=0.788. Based on the trial of the WCT test instrument, a description of the reliability of the test was found in the medium category which refers to the Cronbach alpha value obtained at 0.597.

High School Students' Capacity for Critical Thinking

The trial of the WCT test on several middle grade students provides several descriptions regarding students' critical thinking skills while at the same time providing a more detailed picture of how their critical thinking is written in the answers to each WCT test problem. Each item is attempted to be able to reveal each specific domain more deeply through essay tests.

According to the results in Figure 4, the domain with the highest achievement is likelihood and uncertainty analysis, and the domain with the lowest achievement relative to the others is argument analysis. Several factors cause the incomplete achievement of indicators for each construct from argument analysis from the core of the problem, namely misconceptions. The construct of the argument part in general can still be achieved well even though students experience misconceptions. This is because, the argument part analyzes the whole view of the argument by presenting several important parts without further interpretation from students. Problems arise in the Inference of Data construct, namely students are required to categorize the truth of the conclusions from the argument part that has been presented. The wave misconception in students will greatly affect the completeness of Inference of Data analysis in an argument. Because in general, wave misconceptions are still found, quantitatively, the domain relationship with the background of good concept mastery will have a low percentage of achievement.



HT: Hypothesis testing
AA: Argument Analysis
R: Reasoning
LUA: Likelihood uncertainty analysis
PD: Problem-solving decision-making

Figure 4. CT Skills Achievement

Scoring in the WCT test considers the completeness of the analysis based on the wave matter in detail. This is done with the intention of maintaining the applicability of the critical thinking abilities provided by mastery of the concept of wave matter as a construct to the domain. As an illustration, for each theme there are many physical quantities and their relationships as materials that are used as problems to be analyzed by students so that the results of the analysis require that the concept of waves be presented. On the other hand, utilizing this type of scoring approach will make it very simple to remove students' description bias and identify the key components of critical thinking skill.

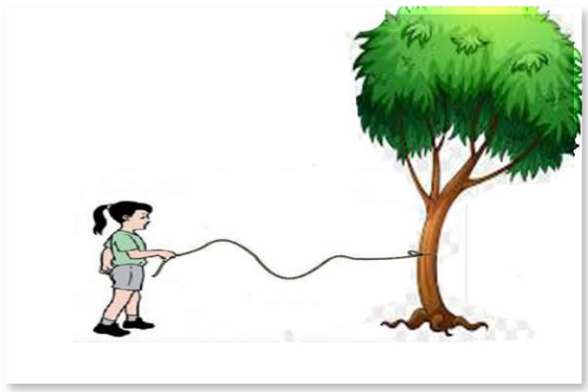
Mastery of the wave concept becomes the starting point for evaluating the results of the WCT test. This has implications for students, namely students must go through wave learning before carrying out the test. This kind of consideration is important so that the WCT test function is in the right position, namely analyzing critical thinking skills specifically on wave material so that the expected domain depicted from the construct appears in its entirety without any errors due to student answers that are too broad due to the fact that the test is an essay answer.

The consistency of the answers from students tested by the WCT test was not between themes, but each theme instructed several interrelated questions. This is intended to make the correction of student answers in the form of an unlimited essay (essay) easier. Then, tracing the consistency of student answers on the WCT test will be simple because the scoring system focuses on the physical conception of wave matter only so that descriptions that are not related to the conception will be easy to be eliminated.

Qualitatively, many things can be studied from the descriptions made by students. So that the results of the WCT test are not limited to students' critical thinking skills, but many other areas that may not be expected to appear but have benefits for further learning improvement materials. This is an advantage of the essay test where new and unexpected things can be important findings.

Based on Table 6, it can be observed how the theme presents a phenomenon that is closely related to students' daily lives. This approach is called a contextual approach which is intended to make students more motivated to analyze the phenomenon.

Table 6. Sample of Argument Analysis Item

Theme 1	
Phenomenon	Hana tied a rope to a tree and stretched it as shown in the picture. Then Hana moved the rope up and down periodically so that it formed a wave.
Question	Three of Hana's friends, namely Anto, Dilan, and Desi noticed the rope being moved by Hana. Anto considers that if Hana's up and down rope move faster, the waves produced will reach the tree faster because the frequency has an effect on the speed of wave propagation. Dilan has another opinion, namely the fast and slow movement of the rope up and down has no effect on the speed of the rope wave reaching the tree because the frequency changes accompanied by a change in wavelength so that the speed of propagation remains constant. Meanwhile, Desi argues that the bigger the waves created by Hana, the faster the waves will reach the tree, this is because the energy possessed by the waves will be greater, thus providing more speed for the waves to propagate. What do you think caused them to disagree? Then who is right? Explain according to the context of physics.
Argument Analysis item: Demonstrate the variable that refers to two identical arguments.	
Student 1	Based on the arguments of Hana's three friends, only Anto has the right arguments. Anto explained the phenomenon based on the frequency value which is proportional to the wave speed. The frequency increases if the movement up and down the rope is getting faster then it will cause faster propagation. The wavelength, in this case, will always be the same * Score = 1 (relationship of equation v = λf)
Student 2	Hana's three friends argued by considering the frequency-influenced propagation rate, according to the relationship between wavelength and a period. Based on the most appropriate relationship, Dilan's analysis is most suitable for the situation, where the speed of propagation will adjust to the type of propagation medium. This concept will explain the relationship of frequency values that are not proportional to the speed of propagation so that when the frequency increases, the speed of wave propagation will remain the same, with the consequence that the wavelength is getting smaller. * Score = 2 (relationship equation v = λf), Melde Equation of waves)

Discussion

This study highlights the development of critical thinking skills tests with wave content in the physics domain. The development of this test departs from the basis that each scientific domain has its own characteristics so that if a skill is to be measured through the content of a particular scientific domain, it is necessary to specify the domain of a skill. This test was developed by taking into account wave content which is often a concern for misconceptions. The tendency that arises is the misconception of wave content caused by low critical thinking skills. Other studies confirm that these two things are a unity of problems in physics learning that are interrelated (Negoro & Karina, 2019; Prince et al., 2012).

The description instructed in the WCT test does not require excessive mathematical calculations. Physical quantities and their associations become the focus of the problem. This is in accordance with the basis for the preparation of the WCT test, namely elevating the evaluation of critical thinking skills from a general domain that is specified into a construct with physics content, namely waves so that the product of students' thinking in the form of essays emphasizes understanding physical concepts compared to mathematical concepts. However, mathematical concepts are not completely eliminated, because physical concepts, especially waves, still require mathematical relationships (Aghvinian et al., 2021; Getto, 2020).

It is clear from data analysis that the WCT test is feasible as an instrument to test and become the basis for measuring critical thinking skills specifically on wave content. The WCT test is able to explain the 5 main domains of critical thinking skills with new constructs to further detail the independence between domains so that critical thinking skills are measured more thoroughly and without bias.

As a continuation with the critical thinking skills test outcomes using the WCT test, the test results can be studied more deeply to qualitatively find out the conceptions of students along with students' ways of thinking and perceptions of a phenomenon. In line with this, it is believed that the WCT test, used to assess critical thinking abilities, can help students demonstrate understanding of physics concepts. According to research, this WCT test ability supports the advancement of students' conceptual mastery by providing detailed and concrete identification connected to critical thinking features (Tiruneh et al., 2017).

On the other hand, Numerous critical thinking assessments, such as Halpern's Critical Thinking Assessment and California's Critical Thinking Skills, use a multiple-choice style. This seeks to control participants' judgements by reducing the bias of their responses. (Bassett, 2016; Shavelson et al., 2019). The controlled nature of the answers prevents the critical thinking participants' originality from being revealed (Bassett, 2016; Kleemola et al., 2022; Shavelson et al., 2019). By taking into account response biases through a domain-specific design and extensive indicators, the WCT test in essay format demonstrates the originality of participants' critical thinking.

The WCT test has been expanded to include components from the reasoning, likelihood and uncertainty, argument analysis, hypothesis testing, and problem-solving and decision-making domains. This is intended so that measurements are more specific in physics content, because each scientific family has different characteristics. By taking into account the analysis of thinking carefully about a phenomenon offered in each theme of the WCT test, hypothesis testing, argument analysis, and reasoning are prepared. After that, the topics of likelihood and uncertainty analysis, problem-solving and decision-making are put together by taking into account how well students can break down difficulties using a thorough and logical analysis based on the concept of opportunity.

The CT domain in the WCT test was made independent of the other CT domains. Independent referred to here is independent in terms of content raised in the form of different wave concepts, mental conditions that arise from demands for ways of thinking to unravel phenomena. This is intended so that there is a clear distinction that the critical thinking domain has independent dimensions but can be generalized as a critical way of thinking.

In addition, the presentation of phenomena in the WCT test is made in the form of story telling and modeling of children's activities so that it will be more memorable when students read it. This is an adjustment to the generalization of several studies related to contextual approaches (Aghvinian et al., 2021; Getto, 2020).

Conclusion

It may be concluded that the generated product, namely the WCT test, is valid as a Critical Thinking Instrument based on the findings of the research and discussion. The results of the content validity test demonstrate this; the overall percentage is 87.98% with extremely practicable standards. The exam's reliability is demonstrated by the test results of 106 students, which included 15 objective items, and which showed a Cronbach's alpha score of 0.597 with a medium category. The construct validity test findings demonstrate that all items satisfy the Goodness of Fit requirements, demonstrating the validity of all WCT test items.

Recommendations

The WCT test's potential utility as a tool for evaluating critical thinking for research purposes is our next goal. The WCT test is designed to be used as a foundation for responding to additional crucial research questions. We hope that this effort has a good approach and can be used as a reference for developing and validating assessments of critical thinking skills in other science content, despite the fact that the validation technique presented in this work still has many flaws. Some things that can be taken into consideration for future research include how to make an essay test that accommodates various representations of critical thinking models because the WCT test is limited to solving narrative problems. The representation in question might lead to the theory of learning modalities and multiple intelligences.

Limitations

Validation of test instruments, measurement of latent components of critical thinking, and evaluation of students' thinking profiles can be considered as limitations for our study. The specific critical thinking domain that was developed refers to the field of physics with the involvement of senior high school students which is considered a limitation.

Authorship Contribution Statement

Rusilowati: Concept and design, data acquisition, data analysis / interpretation, drafting manuscript, critical revision of manuscript, technical or material support, supervision, final approval. Negoro: Concept and design, data acquisition, data analysis / interpretation, drafting manuscript, critical revision of manuscript. Aji: Critical revision of manuscript, supervision, final approval. Subali: Critical revision of manuscript

References

Aghvinian, M., Santoro, A. F., Gouse, H., Joska, J. A., Linda, T., Thomas, K. G. F., & Robbins, R. N. (2021). Taking the test: A qualitative analysis of cultural and contextual factors impacting neuropsychological assessment of Xhosa-Speaking South Africans. *Archives of Clinical Neuropsychology*, *36*(6), 976–980. https://doi.org/10.1093/arclin/acaa115

Akbar, A., Disman, & Kusnendi. (2020). The role of problem solving, problem-based learning, and critical thinking in the era of globalization. In T. H. Dahlan, H. Yulindrasari, Y. Rachmawati, V. Adriany, T, Aryanti, K. Banda, F. D. Hutagalung & B. Sumintono (Eds.), *Proceedings of the International Conference on Educational Psychology and Pedagogy - "Diversity in Education" (ICEPP 2019)* (pp. 15–17). Atlantis Press. https://doi.org/10.2991/assehr.k.200130.070

Anas, S. (2008). *Pengantar statistik pendidikan* [Introduction to educational statistics]. Raja Grafindo Persada.

Auli, S., Diana, N., & Yuberti. (2018). Analisis miskonsepsi siswa SMP pada materi fisika [Analysis of misconception of junior high school]. *Indonesian Journal of Science and Mathematics Education*, *1*(2), 155–161. https://bit.ly/44vT0Qv

Barniol, P., & Zavala, G. (2016). Mechanical waves conceptual survey: Its modification and conversion to a standard multiple-choice test. *Physical Review Physics Education Research*, *12*, Article 010107. https://doi.org/10.1103/PhysRevPhysEducRes.12.010107

Bassett, M. H. (2016). Teaching critical thinking without (Much) writing: Multiple-choice and metacognition. *Teaching Theology and Religion*, *19*(1), 20–40. https://doi.org/10.1111/teth.12318

Behar-Horenstein, L. S., & Niu, L. (2011). Teaching critical thinking skills in higher education: A review of the literature. *Journal of College Teaching & Learning*, *8*(2), 25–42. https://doi.org/10.19030/tlc.v8i2.3554

Borg, W. R., & Gall, M. D. (2007). *Educational research: An introduction*. Longman.

Bozzi, M., Ghislandi, P., & Zani, M. (2020). Misconceptions in physics at Politecnico di Milano: Preliminary results. In C. Andrà, D. Brunetto, & F. Martignone (Eds.), *Theorizing and measuring affect in mathematics teaching and learning* (pp. 211–216). Springer. https://doi.org/10.1007/978-3-030-50526-4_21

Chen, C., Sonnert, G., Sadler, P. M., Sasselov, D., & Fredericks, C. (2020). The impact of student misconceptions on student persistence in a MOOC. *Journal of Research in Science Teaching*, *57*(6), 879–910. https://doi.org/10.1002/tea.21616

Chusni, M. M., Saputro, S., Suranto, & Rahardjo, S. B. (2020). Review of critical thinking skill in Indonesia: Preparation of the 21st century learner. *Journal of Critical Reviews*, *7*(9), 1230–1235. https://doi.org/10.31838/jcr.07.09.223

Dellantonio, S., & Pastore, L. (2021). Ignorance, misconceptions and critical thinking. *Synthese*, *198*, 7473–7501. https://doi.org/10.1007/s11229-019-02529-7

Dudung, A., & Oktaviani, M. (2020). Mathematical connection ability: An analysis based on test forms. *International Journal of Advanced Science and Technology*, *29*(6), 4694–4701. https://bit.ly/3XRwfny

Getto, G. (2020). The story/test/story method: A combined approach to usability testing and contextual inquiry. *Computers and Composition*, *55*, Article 102548. https://doi.org/10.1016/j.compcom.2020.102548

Gurel, D. K., Eryilmaz, A., & McDermott, L. C. (2015). A review and comparison of diagnostic instruments to identify students' misconceptions in science. *Eurasia Journal of Mathematics, Science and Technology Education*, *11*(5), 989–1008. https://doi.org/10.12973/eurasia.2015.1369a

Hart, C., Da Costa, C., D'Souza, D., Kimpton, A., & Ljbusic, J. (2021). Exploring higher education students' critical thinking skills through content analysis. *Thinking Skills and Creativity*, *41*, Article 100877. https://doi.org/10.1016/j.tsc.2021.100877

Kahan, D. M. (2017, May 24). *Misconceptions, misinformation, and the logic of identity-protective cognition*. SSRN. https://doi.org/10.2139/ssrn.2973067

Karim, S., Saepuzaman, D., & Sriyansyah, S. P. (2016). The learning reconstruction of particle system and linear momentum conservation in introductory physics course. *Journal of Physics: Conference Series*, *739*, Article 012111. https://doi.org/10.1088/1742-6596/739/1/012111

Kennedy, E. M., & de Bruyn, J. R. (2011). Understanding of mechanical waves among second-year physics majors. *Canadian Journal of Physics*, *89*(11), 1155–1161. https://doi.org/10.1139/P11-113

Kleemola, K., Hyytinen, H., & Toom, A. (2022). Exploring internal structure of a performance-based critical thinking assessment for new students in higher education. *Assessment and Evaluation in Higher Education*, *47*(4), 556–569. https://doi.org/10.1080/02602938.2021.1946482

Kravchenko, O., Dokuchaieva, V., Valentieva, T., Sbitnieva, L., & Chornobryva, N. (2023). The use of technology-based model of critical thinking development to reshape students' self-study process. *European Journal of Educational Research*, *12*(1), 281–296. https://doi.org/10.12973/eu-jer.12.1.281

Larsson, K. (2021). Using essay responses as a basis for teaching critical thinking: A variation theory approach. *Scandinavian Journal of Educational Research*, *65*(1), 21–35. https://doi.org/10.1080/00313831.2019.1650824

Leach, S. M., Immekus, J. C., French, B. F., & Hand, B. (2020). The factorial validity of the cornell critical thinking tests: A multi-analytic approach. *Thinking Skills and Creativity*, *37*, Article 100676. https://doi.org/10.1016/j.tsc.2020.100676

McMillan, C., Loads, D., & McQueen, H. A. (2018). From students to scientists: The impact of interactive engagement in lectures. *New Directions in the Teaching of Physical Sciences*, *13*(1), 1–12. https://doi.org/10.29311/ndtps.v0i13.2425

Negoro, R. A., & Karina, V. (2019). Development of a Four-Tier Diagnostic Test for Misconception of Oscillation and Waves. *Jurnal Penelitian & Pengembangan Pendidikan Fisika*, *5*(2), 69–76. https://doi.org/10.21009/1.05201

Negoro, R. A., Rusilowati, A., & Aji, M. P. (2023). Scratch-assisted waves teaching materials : ICT literacy and students ' critical thinking skills. *Journal of Turkish Science Education*, *20*(1), 189–210. https://doi.org/10.36681/tused.2023.011

Negoro, R. A., Rusilowati, A., Aji, M. P., & Jaafar, R. (2020). Critical thinking in physics: Momentum critical thinking test for pre-service teacher. *Jurnal Ilmiah Pendidikan Fisika Al-Biruni*, *9*(1), 73–86. https://doi.org/10.24042/jipfalbiruni.v9i1.4834

Ngajie, B. N., Li, Y., Tiruneh, D. T., & Cheng, M. (2020). Investigating the effects of a systematic and model-based design of computer-supported argument visualization on critical thinking. *Thinking Skills and Creativity*, *38*, Article 100742. https://doi.org/10.1016/j.tsc.2020.100742

Osborne, J. (2014). Teaching critical thinking ? New directions in science education. *School Science Review*, *95*(352), 53–62. https://bit.ly/3JX6Enz

Prince, M., Vigeant, M., & Nottis, K. (2012). Development of the heat and energy concept inventory: Preliminary results on the prevalence and persistence of engineering students' misconceptions. *Journal of Engineering Education*, *101*(3), 412–438. https://doi.org/10.1002/j.2168-9830.2012.tb00056.x

Rahimi, R. A. (2020, November). Understanding the Key Factors That Influence Employee Loyalty in Public Organizations. In *2020 International Conference on Decision Aid Sciences and Application (DASA)* (pp. 376-381). IEEE. https://doi.org/10.1109/DASA51403.2020.9317256

Rahimi, R. A., & Park, K.-H. (2020). The changing role of teachers in preparing future generations: The implications of globalization & technology on the education system. *Asian Journal of Canadian Studies*, *26*(1), 82–94.

Rahmawati, I., Sutopo, S., & Zulaikah, S. (2017). Analysis of students' difficulties about rotational dynamics based on resource theory. *Jurnal Pendidikan IPA Indonesia*, *6*(1), 95–102. https://doi.org/10.15294/jpii.v6i1.9514

Reyes, M. G., & Rakkapao, S. (2018). Impact of context on students ' conceptual understanding about mechanical wave speed. *Journal of Physics: Conference Series*, *1144*, Article 012046. https://doi.org/10.1088/1742-6596/1144/1/012046

Risnita, R., & Bashori, B. (2020). The effects of essay tests and learning methods on students' chemistry learning outcomes. *Journal of Turkish Science Education*, *17*(3), 332–341. https://bit.ly/3rtlcVy

Rusilowati, A. (2014). Pengembangan instrumen penilaian [Assessment instrument development]. Unnes Press.

Rusilowati, A., Subali, B., Aji, M. P., & Negoro, R. A. (2020). Development of teaching materials for momentum assisted by scratch: Building the pre-service teacher's skills for 21st century and industry revolution. *Journal of Physics: Conference Series*, *1567*, Article 022010. https://doi.org/10.1088/1742-6596/1567/2/022010

Schmaltz, R. M., Jansen, E., & Wenckowski, N. (2017). Redefining critical thinking: Teaching students to think like scientists. *Frontiers in Psychology*, *8*, Article 459. https://doi.org/10.3389/fpsyg.2017.00459

Selman, Y. F., & Jaedun, A. (2020). Evaluation of the implementation of 4C skills in Indonesian subject at senior high schools. *Jurnal Pendidikan Indonesia*, *9*(2), 244–257. https://bit.ly/3pX88Hv

Shavelson, R. J., Zlatkin-Troitschanskaia, O., Beck, K., Schmidt, S., & Marino, J. P. (2019). Assessment of university students' critical thinking: Next generation performance assessment. *International Journal of Testing*, *19*(4), 337–362. https://doi.org/10.1080/15305058.2018.1543309

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., Shi, H., & Loyalka, P. (2020). Thinking critically about critical thinking: Validating the Russian HEIghten® critical thinking assessment. *Studies in Higher Education*, *45*(9), 1933–1948. https://doi.org/10.1080/03075079.2019.1672640

Sidiq, Y., Ishartono, N., Desstya, A., Prayitno, H. J., Anif, S., & Hidayat, M. L. (2021). Improving elementary school students'

critical thinking skill in science through hots-based science questions: A quasi-experimental study. *Jurnal Pendidikan IPA Indonesia*, *10*(3), 378–386. https://doi.org/10.15294/JPII.V10I3.30891

Somroob, S., & Wattanakasiwich, P. (2017). Investigating student understanding of simple harmonic motion. *Journal of Physics: Conference Series*, *901*, Article 012123. https://doi.org/10.1088/1742-6596/901/1/012123

Stupple, E. J. N., Maratos, F. A., Elander, J., Hunt, T. E., Cheung, K. Y. F., & Aubeeluck, A. V. (2017). Development of the Critical Thinking Toolkit (CriTT): A measure of student attitudes and beliefs about critical thinking. *Thinking Skills and Creativity*, *23*, 91–100. https://doi.org/10.1016/j.tsc.2016.11.007

Sugiyono. (2013). Metode penelitian pendidikan pendekatan kuantitatif, kualitatif dan R&D [Educational research methods quantitative, qualitative and R&D approaches]. Alfabeta.

Supena, I., Darmuki, A., & Hariyadi, A. (2021). The influence of 4C (constructive, critical, creativity, collaborative) learning model on students' learning outcomes. *International Journal of Instruction*, *14*(3), 873–892. https://doi.org/10.29333/iji.2021.14351a

Suprapto, N. (2020). Do we experience misconceptions?: An ontological review of misconceptions in science. *Studies in Philosophy of Science and Education*, *1*(2), 50–55. https://doi.org/10.46627/sipose.v1i2.24

Tiruneh, D. T., De Cock, M., Weldeslassie, A. G., Elen, J., & 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*, 663–682. https://doi.org/10.1007/s10763-016-9723-0

Tiruneh, D. T., Weldeslassie, A. G., Kassa, A., Tefera, Z., De Cock, M., & Elen, J. (2016). specific and domain-general critical thinking skills. *Educational Technology Research and Development*, *64*, 481–505. https://doi.org/10.1007/s11423-015-9417-2

Xu, W., Liu, Q., Koenig, K., Fritchman, J., Han, J., Pan, S., & Bao, L. (2020). Assessment of knowledge integration in student learning of momentum. *Physical Review Physics Education Research*, *16*, Article 010130. https://doi.org/10.1103/PhysRevPhysEducRes.16.010130

Yulianti, D., Sugianto, S., & Ngafidin, K. (2022). Scratch assisted physics learning with a STEM approach in the pandemic era to develop 21st century learning skills. *Jurnal Pendidikan IPA Indonesia*, *11*(1), 102–105. https://doi.org/10.15294/jpii.v11i1.32607

Zaleha, Z., Samsudin, A., & Nugraha, M. G. (2017). Pengembangan instrumen tes diagnostik VCCI bentuk four-tier test pada konsep getaran [Development of the VCCI diagnostic test instrument in the form of a four-tier test on the vibration concept]. *Jurnal Pendidikan Fisika dan Keilmuan*, *3*(1), 36-42. https://doi.org/10.25273/jpfk.v3i1.980