



A framework for assessing computational thinking skills in the physics classroom: study on cognitive test development

Rizki Zakwandi^{1,2}  · Edi Istiyono³ 

Received: 16 June 2022 / Accepted: 5 February 2023 / Published online: 20 February 2023
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2023

Abstract

This study aimed to develop a computational thinking skills (CTs) assessment framework for physics learning. The framework was developed in two stages: theoretical and empirical. Furthermore, the framework was examined by developing questions, a set test instrument, in the form of multiple-choice (3 items), right-wrong answers (2 items) complex multiple-choice (2 items), and essays (15 items) for sound wave topic. There are three stage of framework examination in empirical study involving 108 students to obtain the item characteristic, 108 students for the explanatory factor analysis (EFA), and 113 students for the confirmatory factor analysis (CFA). The sample in this study was senior high school students aged 15–17 years, which were selected randomly. Theoretical study produced seven indicators for assessing CTs consisting of decomposition, redefine problems, modularity, data representation, abstraction, algorithmic design, and strategic decision-making. The empirical study proved that the items were fit to the one parameter logistic (1PL) model. Furthermore, EFA and CFA concluded that the model fits comply the unidimensional characteristics. Hence, the framework can be optimizing the measurement of students CTs in learning physics or science.

Keywords Item response theory · Computational thinking skills · Physics learning · Science learning

✉ Rizki Zakwandi
r.zakwandi@upi.edu

Edi Istiyono
edi_istiyono@uny.ac.id

¹ Department of Physics Education, Faculty of Mathematics and Natural Sciences Education, Universitas Pendidikan Indonesia, Bandung, Indonesia

² Department of Physics Education, Faculty of Mathematics and Natural Sciences, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia

³ Department of Research and Educational Evaluation, Graduate School, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia

Introduction

Physics learning has great potential to develop students thinking skills. Learning content, from concrete to abstract, allows students to organize knowledge from simple to complex. The learning process-oriented to the inquiry process with a scientific approach practice the ability to think critically (critical thinking skills) and creative thinking (creative thinking skills). In addition, the object of physics observations, a problem-oriented in the surrounding environment, indicates that learning physics exercises problem-solving skills in a structured way (algorithmic thinking). Thus, learning physics is an excellent alternative to practice higher-order thinking skills (HOTs).

Practicing of HOTs in physics learning is based on the context and content of learning, the complexity of the material, the learning process, and the evaluation implementation (Bao and Koenig 2019). In experimental-based physics learning (EBPL), all these components become a single unit that students use continuously where context, content, and physics material are integrated in experimental activities. Hence, optimization of EBPL designing can practice the HOTs easier.

However, the development of learning physics looks relatively slow, particularly in the development of experimental activities to practice HOTs. The availability of laboratory equipment and teacher skills are obstacles to developing EBPL. This condition is complicated by the emergence of new competencies along with technological developments. New competencies are shown in changes to the 4C framework (critical thinking skills, creative thinking skills, collaborative skills, and communication skills) (Kivunja 2019) become a 6C framework (Critical Thinking Skills, Creative Thinking Skills, Communication Skill, Collaboration Skills, Compassion, and Computational Thinking) known as A Global Framework of Reference on Digital Literacy for Indicator 4.4.2 (UNESCO 2018).

Innovative physics learning has a unique characteristic when students apply computational logic to solve the problems. Flexibility and accessibility to technology can support students to develop their potential independently (Lye and Koh 2014). Hence, teachers and practitioners have an opportunity to develop CTs through learning physics. However, this is constrained by the agreement on conceptual and operational definitions of CTs (Kong and Abelson 2019; Román-González et al. 2017; Shute et al. 2017). Not only in course learning, the same thing also happens in physics or science learning which emphasizes exploration through experimental activities.

The framework, so far developed, is designed for highly theoretical learning with a numerical orientation, such as in computational physics (Cross et al. 2016; Dwyer et al. 2013; Moreno-León et al. 2020; Moreno León et al. 2015; Nuraisa et al. 2019; Yağcı, 2019). Hence, the process of measuring students CTs with experimental-based learning (EBL) characteristics, such as in EBPL, does not yet have a specific framework. In addition, this problem needs to be solved, bearing in mind that without proper measurement references and indicators, students' CT will not be mapped properly and accurately.

This article discusses the framework for assessing students' CTs especially to EBL and EBPL. There are two research question which are: (1) What is the measured aspect of CTs in physics education? (2) What criterion for cognitive question for assessing CTs?

Computational thinking skills

Computational thinking skills (CTs) are additional competencies that need to be prepared to support digital literacy. CTs were initially introduced to solve cognitive processes using computer programming rules (Lye and Koh 2014). Currently, with the very rapid development of education, the CTs has been adapted to various problems outside the context of programming (Wing 2008), especially in physics and mathematics.

In physics learning, the ability to think computationally (CT) becomes very important considering the vast role of physics in the development of technology globally (Adeleke and Joshua 2015). In addition, learning physics also practice students to solve theoretical and contextual problems. (Nadapdap et al. 2016). In solving problems, a physicist, including students, must go through systematic stages as characteristic of CT (García-Peñalvo and Mendes 2018).

CTs refer to thinking skills that require students to be able to (1) formulate problems computationally and (2) formulate the best solution through clear procedures (algorithms) or state reasons regarding the factors that cause solutions cannot be found. (Moreno-León et al., 2016). These two main characteristics of CTs have been indirectly taught in physics learning (Shute et al. 2017).

Decomposition

Decomposition is a way of thinking by breaking down complex problems into specific problems (Shute et al. 2017). In learning physics, students can describe the problem more simply. Simplification of the problem can be done by identifying the main problems that can be solved with the available information (Cross et al. 2016). The essential ability to perform decomposition is analyzing and evaluating.

Redefine problems

Identifying problems in learning physics is an early indicator that students can solve problems (Cross et al. 2016; Shute et al. 2017) Students are expected to be able to determine the main problem to be solved by referring to the availability of information presented (Cross et al. 2016). The essential ability to perform redefine problem is analyzing and inference.

Modularity

Modularity is the ability of students to classify information based on the subject matter (Kong and Abelson 2019; Shute et al. 2017). Modularity indicators require students to choose concepts with the information needed to solve problems (Cross et al. 2016). The essential ability to perform modularity is analyzing, evaluating, informing, and explaining.

Data representation

Data representation refers to two forms of ability: understanding data in various representations and presenting data in various representations (Kong and Abelson 2019; Moreno-Leon et al. 2016). The essential ability to perform data representation is interpreting and analyzing.

Abstraction

Students who can think computationally must have the ability of abstraction. Abstraction can generalize solutions from a problem to other problems with similar roots (Moreno León et al. 2015). The essential ability to perform abstraction is analyzing, synthesizing, predicting, and evaluating.

Strategic decision-making

Students who can think computationally must have the ability to make strategic decisions. This indicator requires students to solve complex considerations (Cross et al. 2016). The essential ability to make strategic decisions is analyzing and evaluating.

Computational thinking skill in physics learning

The nature of physics education is reflected in a series of physics learning activities. The implementation of physics learning that needs to be considered is practicing skills according to current global competencies, one of which is CTs. Increasing CTs in physics learning can be practiced by developing supporting skills, i.e. giving questions ranging from simple to complex (Voskoglou and Buckley 2012). The problems presented in the questions are intended to access information related to student performance and guide the course of learning (Indrawati 2010).

CTs in learning physics are seen in learning activities where students are faced with abstract problems. Students cannot solve physics problems without a precise sequence of activities. Complex and abstract problems will usually be broken down into more straightforward problems as a generalization reference. After that, physicists will use a systematic method to find a solution to the problem. Finally, after the solution has been obtained, the physicist will make a generalization of the results to

be used as a general reference for solving similar problems. Hence, the activities in learning physics follow the characteristics and indicators of other CTs.

In the EBL, for example, decomposition activities are carried out by students when identifying the objectives of the experimental activities to be carried out. Furthermore, in the problem-solving laboratory (PSL) and higher-order thinking laboratory (HOT Lab) experimental model, students are asked to identify problem to direct experimental activities in formulating solutions (Malik 2015; Malik et al. 2018). Simultaneously, preliminary activities also indirectly practice the redefine problem indicators. Hence, the emphasis on analytical skills is essential in learning physics to practice decomposition and redefine problems indicators.

EBL is presenting many variables related to data analysis. However, not all variable should be considered to solve the problems. CTs emphasize this aspect where students are required to only focus on relevant variables and not be distracted by other variables. The ability to select and focus variables in EBL is called modularity (Kong and Abelson 2019; Shute et al. 2017).

The EBL provide so many information in various representation i.e. numbers, tables, or graphs. Thus, it needs the ability to understand all these representations. Data processing on a computer is capable of extracting data from various representations. On that basis, indicators data representation are essential to be mastered by students studying physics (Gebre 2018; Kohl et al. 2007; Moore et al. 2020). In addition to understanding data from various representations, physics students are also required to make generalizations from their findings. Generalization aims to make general conclusions and will be the initial assumption for future researchers. Generalization or abstraction in physics learning also aims to communicate the findings during experimental activities so that readers will more easily understand the findings.

The last indicator is algorithmic design and strategic decision-making. These two indicators are characteristics emphasized in the CT framework. This statement refers to the pattern of a computer system that is systematic, effective, and efficient. The characteristics of the indicators of algorithmic design and strategic decision-making in physics learning determine the most effective solution to the problems. Physics or science learning, in general, produces many possible solutions to a problem, but using a structured pattern such as in algorithmic design, students will be able to make effective and efficient strategic decisions.

Method

This research is cognitive instrument development to examine the framework research using a psychometric analysis approach. The research stage consists of three main stages. At the initial stage, the researcher developed questions following the defined indicators of CTs. The second stage is instrument testing. The instruments that have been compiled were tested to obtain information about the item and test characteristics. The last stage is the implementation stage as the final data collection. The items in the implementation test are questions that have passed the

selection in the previous stage. The data on the implementation test were analyzed using Explanatory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA).

The sample in this study was 354 senior high school students aged 15–17 years, which were selected randomly. It is certain that each student has studied the physics material being tested, in this case, the sound wave material, and is involved in independent EBPL during the pandemic. 108 students participated in the initial instrument test to obtain the characteristics of the instrument items. In the EFA test, 108 students were used as research subjects, and in the CFA test, there were 113 students.

Results and discussion

This article discusses the framework for assessing students computational thinking skills. The framework developed is intended for EBPL. According to the characteristics of physics learning, which depends not only on the description of the material but also the ability to collect data, analyze data, and describe the findings.

The CTs framework developed is different from the previously developed aspects. Yağcı (2019) is developing a framework for assessing CTs that generally apply essential components in the form of problem-solving skills, algorithmic thinking skills, critical thinking, collaborative learning, and creative thinking. Five aspects that serve as a reference for the development of Yağcı (2019) then analyzed in-depth by integrating the opinions of Moreno-Leon et al., (2016) which states that there are four taxonomies of CTs in science learning, namely data practice, modeling, and simulation practices, computational problem-solving practice, and systems thinking practice (for more information see Cross et al. 2016; Moreno-Leon et al. 2016; Wing 2011; and Yağcı, 2019). By considering the learning characteristics, seven indicators representing four aspects of computational thinking skills in physics learning are obtained: decomposition, redefine problems, modularity, data representation, abstraction, algorithmic design, and strategic decision-making.

Item question

The items were developed by referring to the seven assessment indicators developed. 22 item questions of sound wave was developed are a mixture of multiple-choice, complex, and essays. Sample questions for each indicator are listed in Table 1.

According to the theoretical analysis, this study declare that the characteristics of the items for CTs cognitive test in science or physics on consisted of several rules that are (1) the stimulus has a complex problems; (2) the stimulus is in the form of multiple representations; (3) the problems presented are contextual; (4) emphasizes higher-order thinking skills; (5) the formulation of the solution must be general and applicable, and (6) problem solving which is expected to trigger the ability of analysis, evaluation, synthesis, and creation.

Table 1 Sample item question

Indicator	Question																				
Decomposition	<p>Q9. A clothing company received complaints from employees because of the noise of a vast sewing machine. Based on recommendations from the occupational health and safety (K3) section, they must reduce 50% of the noise level so that it does not negatively impact employees. The company has 10,000 sewing machines, each producing noise of 75 dB and operating simultaneously</p> <table border="1"> <thead> <tr> <th>Merk</th> <th>TI (dB)</th> <th>P (W)</th> <th>Price</th> </tr> </thead> <tbody> <tr> <td>P</td> <td>30</td> <td>210</td> <td>Rp 5.000.000</td> </tr> <tr> <td>Q</td> <td>40</td> <td>220</td> <td>Rp 4.000.000</td> </tr> <tr> <td>R</td> <td>50</td> <td>200</td> <td>Rp 3.000.000</td> </tr> <tr> <td>S</td> <td>60</td> <td>250</td> <td>Rp 2.000.000</td> </tr> </tbody> </table> <p>The solution that can be done is to replace the sewing machine with another sewing machine or reduce the sewing machine, which impacts decreasing production. The most appropriate solution to do so as not to reduce the company's productivity is...</p>	Merk	TI (dB)	P (W)	Price	P	30	210	Rp 5.000.000	Q	40	220	Rp 4.000.000	R	50	200	Rp 3.000.000	S	60	250	Rp 2.000.000
Merk	TI (dB)	P (W)	Price																		
P	30	210	Rp 5.000.000																		
Q	40	220	Rp 4.000.000																		
R	50	200	Rp 3.000.000																		
S	60	250	Rp 2.000.000																		
Redefine problems	<p>Q10. Statement: The sound intensity level of 100 sewing machines is worth 2 times that of one sewing machine because the relationship between the intensity level and the number of sound sources is logarithmic, and $\log 100 = 2$.</p> <p>Based on the statements and reasons above, determine the truth (true/false) and explain each statement!</p>																				
Modularity	<p>Q7. In an expedition into the cave, you and your team will enter the cave vertically using webbing. Because the cave you are exploring is very dark and deep, the flashlight does not reach the bottom of the cave, which causes you and your team to be unable to predict the depth of the cave, while the available webbing is limited because it has been installed on several cliffs before. While preparing the equipment, you remember that one way to determine the depth of a well is to drop a stone. Finally, you discuss this with your team and get ready to take measurements with makeshift equipment. Some of the data that you managed to collect include 1) rock mass, 2) stone dimensions, 3) air humidity, 4) temperature, 5) fall time, 6) speed of sound in air, 7) acceleration of gravity, and 8) air pressure. Based on the data that has been collected, the variables needed to estimate the cave are....</p>																				
Data representation	<p>Q11. Sound intensity expresses the amount of wave power that is transferred through a plane with a specific cross-sectional area perpendicularly. Furthermore, the sound intensity level is expressed as a comparison between the sound intensity and the threshold intensity. Draw a graph showing the relationship between Intensity and Intensity Level!</p>																				
Abstraction	<p>Q5. Arda wants to expand the textile company he owns by adding more workers. If Arda has 10 employees who operate 10 sewing machines with an average noise level of 74.04 dB, calculate the maximum number of employees that Arda can add to violate the established rules when using a machine with the exact specifications!</p>																				
Algorithmic design	<p>Q13. When doing school renovations, the homeroom teacher of class XI IPA C provided input to replace the loudspeakers used. This is because the loudspeakers commonly used have started to break down, so the announcements delivered through the loudspeakers do not reach class XI IPA C clearly, which is 50 m away from the sound source. Therefore, the vice principal began to look for information about the type of loudspeaker that would be used as a substitute. The following is data on the types of speakers available:</p> <table border="1"> <thead> <tr> <th>Merk</th> <th>I (Watt/m²)</th> <th>P (W)</th> <th>R_{max} (m)</th> </tr> </thead> <tbody> <tr> <td>P</td> <td>30</td> <td>210</td> <td>100</td> </tr> <tr> <td>Q</td> <td>40</td> <td>220</td> <td>50</td> </tr> <tr> <td>R</td> <td>50</td> <td>200</td> <td>120</td> </tr> <tr> <td>S</td> <td>60</td> <td>250</td> <td>40</td> </tr> </tbody> </table> <p>Based on the available data, the type of loudspeaker that is suitable to be used to replace a damaged loudspeaker is.... (It is known that the noise threshold limit in the school environment is 55 dB and $\log 0.5 = -0.3$)</p>	Merk	I (Watt/m ²)	P (W)	R _{max} (m)	P	30	210	100	Q	40	220	50	R	50	200	120	S	60	250	40
Merk	I (Watt/m ²)	P (W)	R _{max} (m)																		
P	30	210	100																		
Q	40	220	50																		
R	50	200	120																		
S	60	250	40																		

Table 1 (continued)

Indicator	Question
Strategic decision-making	Q14. In a debate competition related to public health issues, Yosef stated that the safe distance between settlements from industrial areas is 0.5 km. This is based on the research results, which states that the average noise level produced by industrial areas is 85 dB during operating hours and 40 dB when not operating. Yosef added that the maximum noise level when resting at night is 33 dB to rest optimally. Based on the text above, give your opinion about Yosef’s argument regarding the distance between residential and industrial areas! Include your argument with the results of mathematical calculations regarding the noise level received by the public during operating hours and when not operating!

Instrument characteristics

The instrument items that have been developed are then tested to obtain information about the characteristics of the items. This study used the Item Response Theory (IRT) 1PL approach to estimate item characteristics. Thus, only one parameter is estimated, namely the difficulty level of item (*b*). The 1 PL estimation is based on the Chi-Square score in Parscale (PSL) and Quest software. The Chi-Square value for 1PL estimation is more significant than 2PL, while for 3PL, it cannot be estimated. The results of 1PL estimation using PSL software can be seen in Fig. 1 and Table 2.

Figure 1 shows the suitability of the items in the estimation of 1 PL which requires the INFIT MNSQ score range to be 0.77 to 1.33. Based on Fig. 1, all items match the 1PL model so that each item is stated to be empirically valid (Boone et al. 2014).

Furthermore, based on Table 2, the student’s ability as measured using the questions in the developed framework is 0.02 logit. Conversion of logit scores to a scale of 100 using the equation $Score = 50 + 10z$ resulted in the average score obtained by students being 50.34. With a standard curve, this value indicates that the items have good quality because they can produce estimates of students’ abilities in the medium range.

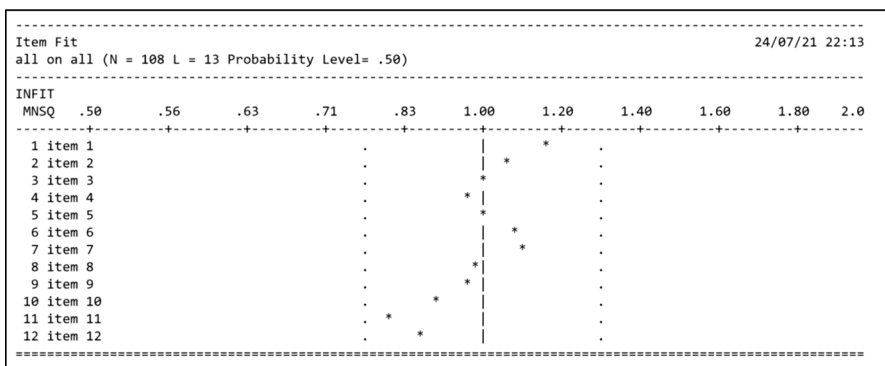


Fig. 1 Item fit MNSQ

The estimated reliability values in Table 2 indicate that the developed test can measure students' abilities well. This score is indicated by a reliability score of 0.71 with the "Fair" interpretation (Fisher Jr 2007). The result of the subsequent analysis is the parameter estimation for the developed test. These are listed in Table 3 for tests and Fig. 2 for items.

Table 3 shows that the overall level of item difficulty (odd ratio) of the CTs test questions with the developed framework is 0.626. This score has a problematic tendency because it has a score of more than 0.00. However, in the item difficulty range, this score is still quite acceptable with the aim of mapping students' abilities, not for selecting subjects (Chan et al. 2020).

Figure 2 shows the level of item difficulty, discriminating power, and guessing factor for each item. As seen in Fig. 2, the most difficult items are items number 9, 13, and 14 (in figure number 10 to 12), which represent algorithmic design and strategic decision-making indicators. At the same time, the items that show the lowest level of difficulty are numbered items that represent modularity.

Figure 3 shows the correlation between the average score and students' abilities. While this is not the only problem inherent in interpreting test scores, we focus on regression to the mean because it is widely ignored and misunderstood. (Smith and Smith 2005). Overall, the questions developed in the computational thinking ability test could estimate students' abilities in a suitable category.

Figure 4 is the total test information collected from the information function on each item. Based on Fig. 4, the test items developed could estimate students' ability with an ability range of -2.00 to 2.75 . This value indicates that the questions and the framework developed can estimate students' abilities over a wide range (Boone et al. 2014). A wide range of measurements in the framework of computational thinking skills makes it possible to estimate students' abilities reasonably, considering that computational thinking skills tend to think at higher levels by solving complex problems. (Adams and Wieman 2015; Lye and Koh 2014; Nuraisa et al. 2019; Shute et al. 2017).

Table 2 Summary statistics of test estimates

Item Estimate	Score
Mean	0.02
SD	0.70
SD (Adjusted)	0.59
Reliability	0.71

Table 3 Summary statistics of parameter estimates

Parameter	Mean	St. Dev	N
Slope	0.626	0.000	12
Log (Slope)	-0.468	0.000	12
Threshold	0.423	1.377	12
Guessing	0.000	0.000	0

ITEM	BLOCK	SLOPE	S.E.	LOCATION	S.E.	GUESSING	S.E.
0001	1	0.626	0.000	-0.133	0.292	0.000	0.000
0002	2	0.626	0.000	-0.559	0.287	0.000	0.000
0003	3	0.626	0.000	-1.209	0.602	0.000	0.000
0004	4	0.626	0.000	-0.271	0.309	0.000	0.000
0005	5	0.626	0.000	0.128	0.272	0.000	0.000
0006	6	0.626	0.000	0.851	0.284	0.000	0.000
0007	7	0.626	0.000	0.666	0.304	0.000	0.000
0008	8	0.626	0.000	0.978	0.274	0.000	0.000
0009	9	0.626	0.000	-1.848	0.633	0.000	0.000
0010	10	0.626	0.000	1.226	0.278	0.000	0.000
0011	11	0.626	0.000	2.235	2.781	0.000	0.000
0012	12	0.626	0.000	3.016	2.528	0.000	0.000

Fig. 2 Item parameter information

Explanatory factor analysis (EFA)

Table 4 shows that the KMO value is more significant than 0.05. This score indicates that the number of samples used is sufficient to analyze the data. According to Hair (2009), the number of samples is related to determining the loading factor.

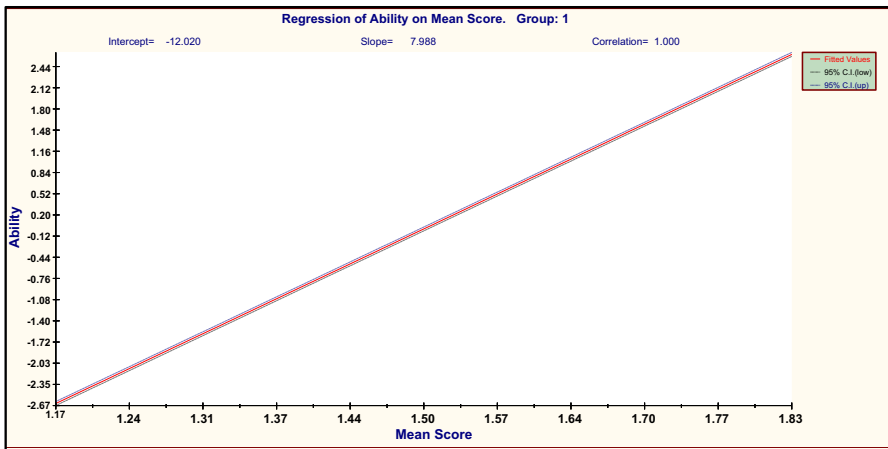


Fig. 3 Bivariate plot

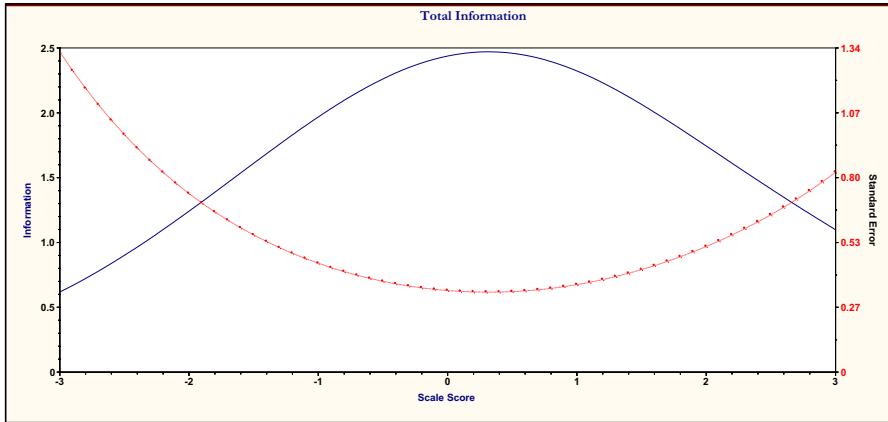


Fig. 4 Test total information

Table 4 Summary EFA analysis

Variable	Score
KMO Measure of Sampling Adequacy	0.543
Bartlett’s Test of Sphericity	0.001
Total Variance	7.000
% Cumulative	73.636

With a sample size of 100–120 students in the EFA, the minimum loading factor required is 0.50.

Based on the criteria set by Hair (2009), The resulting factor loading is declared valid with a significant correlation > 0.5 for a sample of 100–120 people (Table 5).

Confirmatory factor analysis (CFA)

In order to validate the interpretation of scores by evaluating a stand-alone fit model, it is necessary to use a fit index for which good cutoff criteria, or at least some rules of thumb, are available. Other characteristics of the fit index need to be taken into account, particularly sensitivity to sample size and model specification errors (Sun 2005).

CFA testing is carried out by distributing questions that match the 1PL model and meet the criteria in the EFA. The model’s fit based on the CFA is shown in Table 6.

Based on Table 6, all of the indicators resulting from the SEM analysis have the goodness of statistics better than the recommended value. This indicates that the current model is statistically adequate. Then, the results of the SEM analysis and the significance of all factor loading (Sun 2005).

The p -value = 0.25, $RMSA$ = 0.037, CFI = 0.94, IFI = 0.96, NFI = 0.74, GFI = 0.95 and $AGFI$ = 0.87 also indicate that the index of conformity from the CFA analysis

Table 5 Rotated component matrix

Question	Component	Correlation
Q1	Decomposition	0.950
Q2	Modularity	0.647
Q3	Abstraction	0.888
Q4	Algorithmic Design	0.399
Q5	Abstraction	0.565
Q7	Modularity	0.582
Q8	Data Representation	0.753
Q9	Decomposition	0.776
Q10	Redefine Problems	0.815
Q11	Data Representation	0.901
Q13	Algorithmic Design	0.785
Q14	Strategic Decision-Making	0.668

Table 6 Characteristic of CFA goodness of fit

No	Statistic	Value	Criteria	Category
1	p-value	0.25	>0.05	Fit
2	RMSEA	0.037	<0.05	Poor Fit
3	CFI	0.94	>0.90	Fit
4	IFI	0.96	>0.90	Fit
5	NFI	0.74	>0.70	Good Fit
6	PNFI	0.37	-	Good Fit
7	GFI	0.95	>0,9	Fit
8	AGFI	0.87	<0,9	Good Fit

is at an acceptable level. (Marsh et al. 2020). These findings indicate that the four-factor model obtained as a result of the CFA has a sufficient level of fitness.

As can be understood from the relevant literature, the results of the EFA and CFA analyzes are acceptable. The findings related to the validity and reliability of the scale indicate that the scale can be used to measure the computational thinking ability of students at the secondary school level, especially in learning physics or science. It is believed that this study will significantly contribute to the literature, considering that there are only a handful of valid and reliable measuring tools to measure computational thinking ability. (Korkmaz et al. 2017; Yağcı 2019).

Conclusion

The framework for assessing computational thinking skills in physics learning with characteristics in EBL consists of decomposition, redefining problems, modularity, data representation, abstraction, AQ2 algorithmic design, and strategic decision-making. The CTs question has some rules that are (1) the stimulus presented in the form of complex problems; (2) the stimulus AQ3 presented is in the form of multiple representations; (3) the problems presented are contextual; (4) emphasizes AQ4 higherorder thinking skills; (5) the formulation of the solution must be general and applicable, and (6) problem solving which is expected to trigger the ability of analysis, evaluation, synthesis, and creation.

Acknowledgements The author is grateful to the Indonesia Endowment Fund for Education (LPDP) for the scholarship given to study in the Master of Physics Education program and support this research's publication.

Author contributions Conceptualization, RZ. and EI; methodology, RZ.; software, RZ.; validation, EI.; formal analysis, RZ.; data curation, RZ.; writing—original draft preparation, RZ.; writing—review and editing, RZ.; visualization, RZ.; supervision, EI.. All authors have read and agreed to the published version of the manuscript.

Funding This study was supported by Lembaga Pengelola Dana Pendidikan

Data availability The data that support the findings of this study are available from the corresponding author, RZ, upon reasonable request.

Declarations

Conflict of interest The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Ethical approval Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Consent to participate The purpose of this study is to develop a framework for assessing Computational thinking skills in physics classroom following the test instrument approach. Participation is voluntary. There is no penalty if you decide not to participate or withdraw from the study, and your relationship with RZ, the Physics Education Department and Universitas Pendidikan Indonesia will not be affected by this decision. The estimated time of participation is three weeks. You will be expected to learn physics using self-laboratory due to COVID 19. After that, student the computational thinking skill was examined. Potential benefits for participating include framework and guidelines for developing Computational Thinking Skill test specially for physics classroom. The criteria of student participated in this study was student in 11th grade of senior high school which learn sound wave topic. In order to maintain confidentiality, participant name will not be connected to any publication or presentation that uses the information and data collected about you or with the research findings from this study. The researcher used anonymizing method, such as number or pseudonym to identify participants. Participant identifiable information will only be shared if required by law or in written permission.

References

- Adams WK, Wieman CE (2015) Analyzing the many skills involved in solving complex physics problems. *Am J Phys* 83(5):459–467. <https://doi.org/10.1119/1.4913923>
- Adeleke AA, Joshua EO (2015) Development and validation of scientific literacy achievement test to assess senior secondary school students' literacy acquisition in physics. *J Educ Pract* 6(7):28–43
- Bao L, Koenig K (2019) Physics education research for 21st century learning. *Discip Interdiscip Sci Educ Res* 1(1):1–12. <https://doi.org/10.1186/s43031-019-0007-8>
- Boone WJ, Yale MS, Staver JR (2014) Rasch analysis in the human sciences. *Rasch Anal Hum Sci.* <https://doi.org/10.1007/978-94-007-6857-4>
- Chan S, Looi C, Sumintono B (2020) Title author (s) Source published by assessing computational thinking abilities among Singapore secondary students : a rasch model measurement analysis journal of computers in education, (2020) Copyright © 2020 Springer this is a post-peer-review, p. J Comput Educ 8(2):213–236
- Cross J, Hamner E, Zito L, Nourbakhsh I (2016) Engineering and computational thinking talent in middle school students: a framework for defining and recognizing student affinities. *Proc—Front Educ Conf.* <https://doi.org/10.1109/FIE.2016.7757720>
- Dwyer, H. A., Boe, B., Hill, C., Franklin, D., & Harlow, D. (2013). Computational thinking for physics: Programming models of physics phenomenon in elementary school. *Engelhardt, Churukian, & Jones (Eds.) 2013 PERC Proceedings*, (pp. 133–136). Portland. <https://doi.org/10.1119/perc.2013.pr.021>
- Fisher WP Jr (2007) Living capital metrics. *Rasch Meas Trans* 21(1):1092–1093
- García-Peñalvo FJ, Mendes AJ (2018) Exploring the computational thinking effects in pre-university education. Elsevier
- Gebre E (2018) Learning with multiple representations: Infographics as cognitive tools for authentic learning in science literacy. *Can J Learn Technol* 44(1):1–24. <https://doi.org/10.21432/cjlt27572>
- Hair JF (2009) *Multivariate data analysis*, 7th edn. Prentice Hall, Hoboken
- Indrawati. (2010). *Evaluasi Pembelajaran Fisika*. Universitas Jember. https://repository.unej.ac.id/bitstream/handle/123456789/10230/EVALUASI_PEMBELAJARAN_FISIKA_IBARU.pdf?sequence=1
- Kivunja, C. (2019). A Pedagogy to Embed into Curricula the Super 4C Skill Sets Essential for Success in Sub-Saharan Africa of the 21st Century. In *Global trends in Africa's development*. Centre for Democracy, Research and Development (CEDRED). <https://hdl.handle.net/1959.11/27328>. 295–302
- Kohl PB, Rosengrant D, Finkelstein ND (2007) Strongly and weakly directed approaches to teaching multiple representation use in physics. *Phys Rev Spec Top Phys Educ Res* 3(1):1–10. <https://doi.org/10.1103/PhysRevSTPER.3.010108>
- Kong S, Abelson H (2019) Computational Thinking Education. In: Kong S, Abelson H (eds) *Computational Thinking Education*. Springer, Singapore
- Korkmaz Ö, Çakir R, Özden MY (2017) A validity and reliability study of the computational thinking scales (CTS). *Comput Hum Behav* 72:558–569. <https://doi.org/10.1016/j.chb.2017.01.005>
- Lye SY, Koh JHL (2014) Review on teaching and learning of computational thinking through programming: what is next for K-12? *Comput Hum Behav* 41:51–61. <https://doi.org/10.1016/j.chb.2014.09.012>
- Malik A, Setiawan A, Suhandi A, Permanasari A, Sulasman S (2018) HOT Lab-Based Practicum Guide for Pre-Service Physics Teachers. *IOP Conf Ser: Mater Sci Eng* 288(1):012027. <https://doi.org/10.1088/1757-899X/288/1/012027>
- Malik, Adam. (2015). Model Problem Solving Laboratory to Improve Comprehension the Concept of Students. In *Proc. Inter. Sem. Math. Sci. Comp. Sci. Educ.* (pp. 43–48). Universitas Pendidikan Indonesia
- Marsh HW, Guo J, Dicke T, Parker PD, Craven RG (2020) Confirmatory factor analysis (CFA), exploratory structural equation modeling (ESEM), and set-ESEM: optimal balance between goodness of fit and parsimony. *Multivar Behav Res* 55(1):102–119
- Moore TJ, Brophy SP, Tank KM, Lopez RD, Johnston AC, Hynes MM, Gajdzik E (2020) Multiple representations in computational thinking tasks: a clinical study of second-grade students. *J Sci Educ Technol* 29(1):19–34. <https://doi.org/10.1007/s10956-020-09812-0>
- Moreno León J, Robles G, Román González M (2015) Dr. scratch: automatic analysis of scratch projects to assess and foster computational thinking. *RED Rev De Educ a Distancia* 46:1–23

- Moreno-León J, Robles G, & Román-González M (2016) Comparing computational thinking development assessment scores with software complexity metrics. *IEEE Glob Eng Educ Conf EDUCON* 10(13):1040–1045. <https://doi.org/10.1109/EDUCON.2016.7474681>
- Moreno-León J, Robles G, Román-González M (2020) Towards data-driven learning paths to develop computational thinking with scratch. *IEEE Trans Emerg Top Comput* 8(1):193–205. <https://doi.org/10.1109/TETC.2017.2734818>
- Moreno-Leon, J., Robles, G., & Roman-Gonzalez, M. (2016). Comparing computational thinking development assessment scores with software complexity metrics. *IEEE Global Engineering Education Conference, EDUCON, 10–13-April*(April), 1040–1045. <https://doi.org/10.1109/EDUCON.2016.7474681>
- Nadapdap, A. T. Y., Lede, Y., & Istiyono, E. (2016). Authentic assessment of problem solving and critical thinking skill for improvement in learning physics. In *Proc. Inter. Sem. Sci. Edu. (ISSE)* (pp. 37–42). Universitas Negeri Yogyakarta
- Nuraisa D, Azizah AN, Nopitasari D, Maharani S (2019) Exploring Students Computational Thinking based on Self-Regulated Learning in the Solution of Linear Program Problem. *JIPM (Jurnal Ilmiah Pendidikan Matematika)* 8(1):30. <https://doi.org/10.25273/jipm.v8i1.4871>
- Román-González M, Pérez-González JC, Jiménez-Fernández C (2017) Which cognitive abilities underlie computational thinking? criterion validity of the computational thinking test. *Comput Hum Behav* 72:678–691. <https://doi.org/10.1016/j.chb.2016.08.047>
- Shute VJ, Sun C, Asbell-Clarke J (2017) Demystifying computational thinking. *Educ Res Rev* 22:142–158. <https://doi.org/10.1016/j.edurev.2017.09.003>
- Smith G, Smith J (2005) Regression to the mean in average test scores. *Educ Assess* 10(4):377–399. https://doi.org/10.1207/s15326977ea1004_4
- Sun J (2005) Assessing goodness of fit in confirmatory factor analysis. *Meas Eval Couns Dev* 37(4):240–256. <https://doi.org/10.1080/07481756.2005.11909764>
- UNESCO (2018) A Global framework of reference on digital literacy for indicator 442. *Inf Pap* 51(51):1–146
- Voskoglou MG, Buckley S (2012) Problem solving and computational thinking in a learning environment. *arXiv preprint arXiv* 36(4):28–46
- Wing JM (2008) Computational thinking and thinking about computing. *Philos Trans R Soc a: Math, Phys Eng Sci* 366(1881):3717–3725
- Wing, J. (2011). Research notebook: Computational thinking—What and why? *The Link Magazine*. <http://www.cs.cmu.edu/link/research-notebook-computational-thinking-what-and-why>. Accessed 23 June 2015
- Yağcı M (2019) A valid and reliable tool for examining computational thinking skills. *Educ Inf Technol* 24(1):929–951. <https://doi.org/10.1007/s10639-018-9801-8>

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.