JPPIPA 9(3) (2023)

Jurnal Penelitian Pendidikan IPA

Journal of Research in Science Education

http://jppipa.unram.ac.id/index.php/jppipa/index

Analysis of Computational Thinking Instrument for High School Student Using Rasch Model

Hani Sulsilah[1*], Arif Hidayat[1], Taufik Ramlan Ramalis[1]

[1] Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam, Universitas Pendidikan Indonesia, Kota Bandung, Indonesia.

Received: December 29, 2022
Revised: March 25, 2023
Accepted: March 28, 2023
Published: March 31, 2023

Corresponding Author:
Hani Sulsilah
hanisulsilah@upi.edu

DOI: 10.29303/jppipa.v9i3.2771

Abstract: This study is aimed-to construct and analyze the Computational Thinking Instrument at Physics subject using Rasch Measurement Model. The instrument was developed by the researcher to assess students' computational thinking in Grade Eleventh in high school on heat and transfer topics under STEM Quartet Integrated Learning. The type of the instrument is a multiple choice with multiple-choice reasoning. This test consists of 13 questions to measure five concepts of Computational Thinking which are abstraction, decomposition, algorithm, evaluation, and generalization. The test was tried out on 120 students (87 female and 33 male aged 16-18 years old) in West Java and Banten. The item of the questions on the test was analyzed using Winstep 5.3.2.0. The reliability of the Instrument can be shown at the Cronbach's alpha, Item Reliability, Person Reliability, Item Separation, and Person Separation. The result shows that Cronbach's alpha (KR-20) is 0.75 which means the instrument has high reliability. The value of the item reliability is 0.95 and person reliability is 0.72. This means that although the consistency of student in answering questions is sufficient, the instrument has high reliability. The values of the item and person separation are 1.61 and 4.43 which means that the instrument has great separation. The validity of the instrument can be seen in unidimensionality and item fit order. The unidimensionality shows that the value of Raw Variance Explain by Measure is 32.8%, the Unexplained variance 1-5 contrast is below 15% and the eigenvalues are also below 3%. Overall, 8 of 13 items of the test meet 3 criteria of the item fit order, then 5 items have 2 of 3 criteria of the item fit order. We can conclude that this Instrument is reliable and valid. Then, the instrument can be used to measure students' computational thinking on heat and transfer topics.

Keywords: Computational thinking instrument; Rasch model; STEM quartet

Introduction

The term computational thinking which was popularized again by Wing states that Computational Thinking is a pattern of thinking like a computer scientist, not thinking like a computer (Wing, 2006). However, this computational thinking ability is not limited only to someone in the field of computer science, but rather this ability is important for everyone in all fields to solve problems. From several definitions expressed by experts, Computational Thinking refers to a set of skills that show a series of cognitive processes to be able to solve a complex problem (Grover et al., 2017; Mcclelland et al., 2018; Selby et al., 2013).

Computational thinking focuses on ways that can always be made more effective, accurate, and elegance (Li et al., 2020). This skill is needed to faces challenges or other complex problem in future year. To support this idea, the Next Generation Science Standards (NGSS) authors listed "mathematical and computational thinking" as one of the eight crucial science and engineering practices that K–12 instructors should work to instill in their pupils in 2013. Thus this skills can be included in the classroom (National Resarch Council, 2013). In line with that, the OECD has also included CT in PISA (Programme for International Student Assessment) starting in 2022 (OECD, 2019).

One way to practice CT in class is to give a complex problem to solve. A complex problem can involve various disciplines to be able to produce a solution, thus Computational Thinking can be trained in multidisciplinary learning, namely STEM learning.

How to Cite:
Sulsilah, H., Hidayat, A., & Ramalis, T.R. (2023). Analysis of Computational Thinking Instrument for High School Student Using Rasch Model. *Jurnal Penelitian Pendidikan IPA*, *9*(3), 1445–1450. https://doi.org/10.29303/jppipa.v9i3.2771

STEM in education is a "platform" that integrates Science, Technology, Engineering, and Mathematics in learning (Sokolowski, 2018). Computational Thinking is important in the STEM field where Computational Thinking acts as a set of thought processes that involve formulating problems and converting them into forms that can be carried out effectively and efficiently by computational tools (Tan et al., 2019). {Formatting Citation} have the developed STEM Quartet Model which contains STEM practices in problem-centric a classroom. The fundamental perspective in this framework is to solve real problems that are complex, persistent, and extended by paying attention to the connections between the four STEM disciplines. Complex mean that the problem need to solve using two or more discipline knowledge and skill. Persistent mean that the problem recure often. Extended problem mean that the problem needed a longer discussion to solve. This will support how student carry out a thought process.

Several studies have developed many computational thinking instruments through portfolios, tests, surveys, and interviews (Tang et al., 2020). In the field of physics, Computational Thinking instruments are still rarely developed. Several studies using computational thinking instruments generally refer to kinematics (Handayani et al., 2022; Hutchins et al., 2019; Orban et al., 2020; Ridlo et al., 2022), dynamics such as Force (Aksit et al., 2020) and momentum (Dwyer et al., 2013); sound (Zakwandi et al., 2023) and electricity (Yin et al., 2020). Orban et al. (2020) said despite the adoption of K-12 computer science standards in many nations, there is still no universal consensus on the specific definition or application of CT, and efforts to measure CT are still in their infancy. In this research, we developed a computational thinking instrument in physics subjects on the topic of heat and transfer under STEM Quartet Learning. The characteristics of the problems in the STEM Quartet are seen as suitable as a basis for developing issues/problems to bring up computational thinking.

This instrument is made with the concept of presenting problems and then solving them by using computational thinking. STEM Quartet which has a complex, persistent and extended problem is used as a baseline for how computational thinking instrument are made. Connections between STEM subjects on instruments were made to map how the STEM Quartet was used to measure students' computational thinking skills.

A good instrument must at least meet two criteria, there are valid and reliable. The instrument is said to be valid, meaning that it can measure what it wants to measure. The instrument is said to be reliable if the instrument is used repeatedly to produce the same value (steady) (Fraenkel et al., 2012). The analysis instrument that will be used in this study uses Rasch analysis. Rasch analysis was chosen based on the advantages it has, including it can identify error responses, predictable missing data scores, the ability does not only to depend on the number of correct answers, and can identify guesses. The response pattern given shows the accuracy of the response from each respondent to each item. This advantage causes Rasch modeling to allow us to determine the conceptual validity of each item. The superiority of Rasch modeling which can predict missing data makes the results of statistical analysis more accurate in the research conducted (Sumintono et al., 2014).s

Method

This research purposed to develop the instrument then analyze using Rasch Model. Thus, the method of this research is using Research and Development. The stages used are following Plomp's Model (Gustiani, 2019) with simplification. The stages are: Investigation, Designing, Realization/construction, Testing, Evaluation and Revision; These stages are shown in Figure 1.

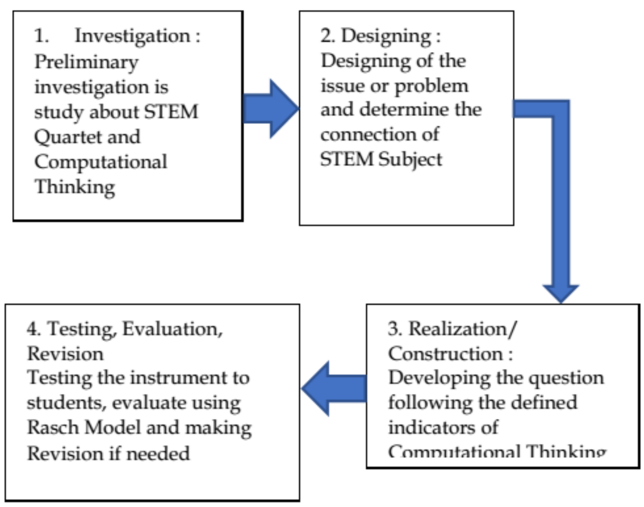

Figure 1. The stages of the research

The Instrument was developed by the researcher to measure students' computational thinking ability. This test is multiple-choice. Preventing students from guessing the answer, the test was made with reason and the type of reasoning are also multiple choice. The Computational Thinking framework used to develop the test is based on Dagiene (Dagiene et al., 2017) where the CT aspects measured consist of abstraction, decomposition, algorithmic thinking, evaluation, and generalization. Five aspects of CT in that reference are measured by 14 indicators. We developed at least two questions for each Indicator. But in this study, we will present several indicators but still measure five aspects of Computational Thinking Skills. The test is composed of 3 reading texts and 13 questions. Topics/problems raised in this issue are Pan and Cooking Steak, Liquid Cooling System for PC, and Making Heat Transfer Simulation.

Table 1. The Aspect of Computational Thinking that is Measured

Computational Thinking Aspect	Indicator	Item	The Connection Between STEM Subject
Abstraction	1.a. Removing unnecessary skills	SO2	S-T
	1.b. Spotting key element	SO13	T-E
	1.c. Choosing a representation of a system	SO11	
Algorithmic thinking	2.c. Creating an algorithm	SO3, SO7, SO12	S-E S-E S-E
Decomposition	3.a. Breaking down tasks	SO1, SO9	S-E S-T
Evaluation	4.a. Finding the best solution	SO5	S-E
	4.b. Making decisions about good use of resources	SO10	S-E
Generalization	5.a. Identifying patterns as well as similarities and connections	SO4, SO8	S-E S-E
	5.c. Utilizing the general solutions	SO6	S-M

Table 1 shows the distribution of 13 questions that measure the five aspects of computational thinking and the connection between STEM subjects developed. In solving a problem, the relationship between STEM subjects is described as having at least a dominant relationship between 2 subjects.

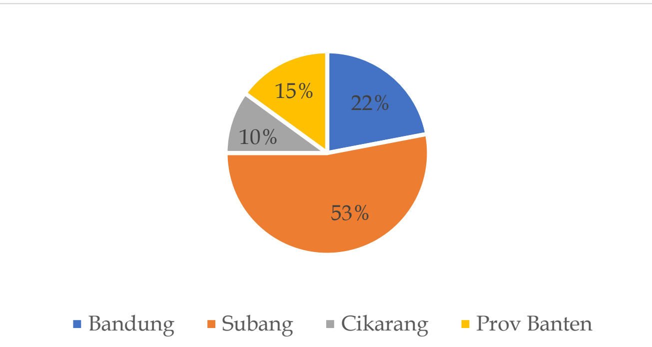

Figure 2. Distribution of sample-based on region in West Java

The Instrument was tried out on 120 high school students in West Java who have learned about heat and transfer in Physics Subject. The distribution of the sample based on Region can be seen in Figure 2. The student was composed of 87 (72%) females and 33 (28.5%) males aged 16-18 years old.

We analyzed the items by using Winstep 5.3.2.0 then reliability was determined based on three criteria including Cronbach's alpha, item and person reliability, and item and person separation. These criteria shows in Table 2.

Table 2. Reliability in Rasch Analysis (Krishnan et al., 2014)

Statistics	Fit Indices	Interpretation
Cronbach's alpha (KR-20)	<0.5	Low
	0.5 – 0.6	Moderate
	0.6 – 0.7	Good
	0.7 – 0.8	High
	>0.8	Very High
Item and Person Reliability	<0.67	Low
	0.67 – 0.8	Sufficient
	0.81 – 0.90	Good
	0.91 – 0.94	Very Good
	>0.94	Excellent
Item and Person Separation		A high separation value indicates that the instrument has good quality since it can identify the group and respondent

The validity of the instrument can be analyzed from unidimensionality which shows the value of raw variance explained by measure and raw unexplained variance in the first until fifth contrast. The criteria raw variance explained by measure can be seen in Table 3. The raw unexplained variance in the first to fifth contrast must be below 15%.

Table 3. The Value of Raw Variance Explained by Measures

Statistics	Fit Indices	Interpretation
Raw variance explained by measure	>20%	Acceptable
	>40%	Good
	>60%	Excellent

We also need information about Item Fit Order to know if the item fits or not. Outfit means outlier-sensitive fit, which means measuring the sensitivity of the response pattern to items with a certain level of difficulty in person or vice versa. An easy example is a wrong response from a person such as not being able to do easy questions even though they have high abilities (careless) or being able to do difficult questions with low abilities (lucky guess). Mean square fit statistic, showing a measure of randomness, namely the amount of distortion in the measurement system. Statistically, the mean-square is the chi-square statistical value divided by the degrees of freedom and the value is always positive. The standardized fit statistic (ZSTD) is a t-test for the hypothesis, "do the data fit the model?" The result is a z-value that is unit deviation. This explains the probability of the data i.e. its significance if the data does fit the model.

Table 4. Fit Indices for MNSQ, ZSTD, and PTMEA-CORR

Statistics	Fit Indices
Outfit mean square values (MNSQ)	0.50 – 1.50
Outfit z-standardized values (ZSTD)	-2.00 – 2.00
Point Measure Correlation (PTMEA-CORR)	0.40 – 0.85

We can also determine the level of difficulties using the comparation beetween the measure and the Standard deviation that we have.

Table 5. Level of Difficulties of the Items

StatisticsC	Fit Indices
Easy	Measure < -1SD
Medium	-1SD< Measure < +1SD
Difficult	Measure > +1SD

Result and Discussion

Table 6 shows several values that indicate whether an instrument is reliable or not. We have a Cronbach's alpha of 0.75 which means the instrument has good reliability. Meanwhile, if the Cronbach's alpha that we get is 0.71-0.90 then the instrument is acceptable as it is at the best level (Bond et al., 2007). The values of item reliability and person reliability are 0.95 and 0.72 which means although consistency of person answering the questions is sufficient, the item has a high reliability (Sumintono et al., 2014).

Table 6. Result value for Reliability

Statistics	Result Value
Cronbach's alpha (KR-20)	0.75
Item Reliability	0.95
Person Reliability	0.72
Item Separation	4.43
Person Separation	1.61

The grouping of people and items can be seen from the separation value. The greater value of separation, the better the quality of the instrument because it can identify groups of respondents and groups of items (Sumintono et al., 2014). We have item separation values of 4.43. (Krishnan et al., 2014) stated that if the value of item separation is higher than 1.00, the items have sufficient spread. The greater of the separation, the greater to categorized the level of item difficulties. The person separation that we have is 1.61, then we can calculate the value of H is [(1.61 x 4) +1]/3 = 2.5, then we up into three so, we have three different levels for student ability, there are low, medium, and high. The person separation reliability must be more than 1.00 to warrant that the students are measured across the continuum (Krishnan et al., 2014).

Table 7. Standardized Residual Variance

Raw Variance Explained by Measures	Unexplned Variance in 1st Contrast
Observed : 32.8%	Observered: 9.3% Eigenvalue: 1.7986

In terms of the validity of the instrument, firstly, we need the information about raw variance explained by measures and unexplained variance in 1st contrast. The result can be seen in Table 7. The value of raw variance explained by measures is 32.8 % which means that the instrument is acceptable. Then, the value of unexplained variance in 1st contrast is 9.3% and that eigenvalue is 1.79. Because the value is under 15% and the eigenvalue is also under 3%, Sumintono et al. (2014) states that the instrument is valid.

Second, we need the information about Outfit MNSQ, Outfit ZSTD, and PTMEASURE-CORR to identify whether the items are fit or not.

Table 8. Item Fit Order

NAME	OUTFIT		PTMA	MEASURE
	MSQ	ZSTD		
SO1	1.2827	2.0513	0.4007	-0.53
SO2	0.8449	-0.3492	0.5173	1.31
SO3	1.0716	0.5211	0.5321	-0.13
SO4	1.2486	1.7112	0.3848	-0.28
SO5	1.2256	1.4312	0.3227	-0.1
SO6	1.1134	0.8911	0.485	-0.62
SO7	1.0236	0.201	0.4953	-0.1
SO8	0.7001	-1.1493	0.6397	0.79
SO9	0.9395	-0.2791	0.4905	0.18
SO10	0.6043	-3.1894	0.6438	-0.21
SO11	0.7296	-2.0093	0.5956	-0.18
SO12	0.7219	-2.2093	0.5081	-0.3
SO13	1.0143	0.141	0.5494	0.19

Firstly, we need evaluate in Pt Measure Corr (PTMA). Table 8 shows that there are no negative value in Pt Measure Corr and we have the logit value between 0.32-0.64. Second we evaluate the value of outfit MNSQ. Based on Table 8, the value of Outfit MNSQ for all items are between 0.6-1.28. This means that all of the items have a good condition for measurement. The Outfit ZSTD value for questions number 1-9 and 11 is between -2 and 2 although question number 1 slightly above 2 and number 8 slightly under -2, these means that questions have a logical prediction. However, the ZSTD value in question number 12 and 13 show a number smaller than -2. (Boone et al., 2014) suggest that an item is said to be fit if it meets the three criteria for the value of outfit MSNQ (0.50-1.50), Outfit ZSTD (-2.00–2.00), and PTMEASURE-CORR (0.40-0.85). Meanwhile, (Sumintono et al., 2014) stated that an item is said to be suitable if it meets at least two of the three predetermined criteria. Therefore, all things can still be maintained and can be said to be fit. We have 0.51 logit SD, based on Table 5 we have 2 easy items, there are SO1 (-0.53<-1 SD) and SO6 (-0.62<-1SD). The items in medium difficulty have a logit -0.51 until +0.51, so we have 9 items (SO3, SO4, SO5, SO7, SO9, SO10, SO11, SO12, SO13). The high difficulty of the items are SO 2 and SO 8 which have the logit >+0.51.

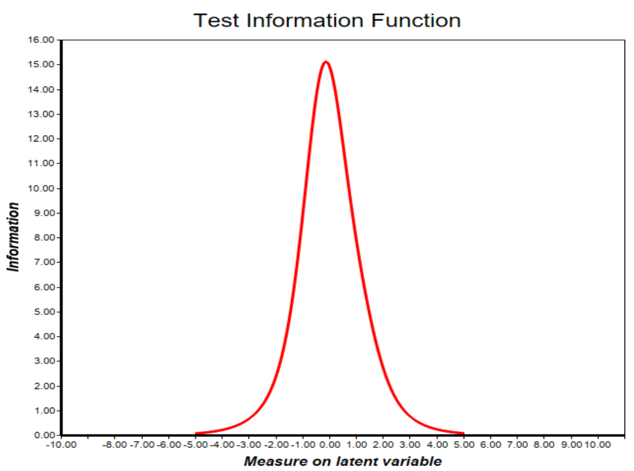

Figure 3. Test information function

In general, Test Information Function in Figure 3. Test information is based on the competence and abilities of students. Test information can be calculated by adding up all competency information. The test results of the instrument show that the amount of information has a maximum value at an ability level of about 0 at an information value of 15. In other words, the competency model is most informative when the student's ability is equal to the level of difficulty of the competency and is less informative when the student's ability moves away from the difficulty level of the competency (i.e. when the competency is too easy or too difficult for students). Thus, this instrument can give good information for the student in a middle ability.

Conclusion

According to this research, the Computational Thinking Instrument has a good Cronbach's Alpha and sufficient item and person reliability. This indicates that the instrument has good reliability to measure high school students' computational thinking on heat and transfer topics. Meanwhile, the item and person separation (4.43 and 1.61) that we got indicate that the person can be distinguished into three levels of ability (low, medium, high) and the item can be categorized into three levels (low, medium, and high) ability. In terms of the validity of the instrument, we have the value of raw variance explained by measures at 29.7% which means that the instrument is acceptable. The value of unexplained variance in the 1st – 5th contrast is also below 15% and the eigenvalue is also below 3.00 which means that the instrument is valid.

Acknowledgements

We wish to thank all the parties who help us in this research.

References

Aksit, O., & Wiebe, E. N. (2020). Exploring Force and Motion Concepts in Middle Grades Using Computational Modeling: a Classroom Intervention Study. *Journal of Science Education and Technology*, *29*(1), 65–82. https://doi.org/10.1007/s10956-019-09800-z

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch Model : Fundamental Measurement in the Human Sciences Second Edition University of Toledo*.

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch Analysis in the Human Sciences*.

Dagiene, V., Sentance, S., & Stupuriene, G. (2017). Developing a Two-Dimensional Categorization System for Educational Tasks in Informatics. *Informatica (Netherlands)*, *28*(1), 23–44. https://doi.org/10.15388/Informatica.2017.119

Dwyer, H. A., Boe, B., Hill, C., Franklin, D., & Harlow, D. (2013). Computational Thinking for Physics: Programming Models of Physics Phenomenon in Elementary School. In Engelhardt, Churukian, & Jones (Eds.), *PERC Proceeding* (pp. 133–136). American Association of Physics Teacher. https://doi.org/10.1119/perc.2013.pr.021

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How To Design and Evaluate Research in Education* (S. Keifer (ed.); Eigth Edit). McGraw-Hill.

Grover, S., & Pea, R. (2017). Computational thinking: A competency whose time has come. *Computer Science Education: Perspectives on Teaching and Learning in School*, *December*, 20–35.

Gustiani, S. (2019). Research and Development (R&D) Method As A Model Design in Educational Research and Its Alternatives. *Holistic Journal*, *11*(2), 12–22. https://jurnal.polsri.ac.id/index.php/holistic/article/view/1849

Handayani, D., Handono, S., Prastowo, B., Prihandono, T., & Nuraini, L. (2022). Computational Thinking : Students ' Abstraction on the Concepts of Kinematics. *Jurnal Penelitian Pendidikan IPA*, *8*(1), 114–118. https://doi.org/10.29303/jppipa.v8i1.1188

Hutchins, N. M., Biswas, G., Maróti, M., Lédeczi, Á., Grover, S., Wolf, R., Blair, K. P., Chin, D., Conlin, L., Basu, S., & McElhaney, K. (2019). C2STEM: a System for Synergistic Learning of Physics and Computational Thinking. *Journal of Science Education and Technology*, *29*(1), 83–100. https://doi.org/10.1007/s10956-019-09804-9

Krishnan, S., & Idris, N. (2014). Investigating Reliability and Validity for the Construct of Inferential Statistics. *International Journal of Learning, Teaching and Educational Research*, *4*(1), 51–60. https://ijlter.net/index.php/ijlter/article/download/1198/1204

Li, Y., Schoenfeld, A. H., diSessa, A. A., Graesser, A. C., Benson, L. C., English, L. D., & Duschl, R. A. (2020). Computational Thinking Is More about Thinking than Computing. *Journal for STEM Education Research*, *3*(1), 1–18. https://doi.org/10.1007/s41979-020-00030-2

Mcclelland, K., & Grata, L. (2018). A review of the importance of computational thinking in K-12. *The Tenth International Conference on Mobile, Hybrid, and On-Line Learning*, *c*, 32–34. https://rb.gy/tk5z

National Resarch Council. (2013). *Next Generation Science Standards: For States, By States*. National Academies Press. https://doi.org/10.17226/18290

OECD. (2019). PISA 2018 Assessment and Analytical Framework, PISA. In *OECD Publishing*. PISA,OECD Publishing. https://doi.org/10.1787/b25efab8-en

Orban, C. M., & Teeling-Smith, R. M. (2020). Computational Thinking in Introductory Physics. *The Physics Teacher*, *58*(4), 247–251. https://doi.org/10.1119/1.5145470

Ridlo, Z. R., Wahyuni, S., Mahardika, I. K., Wicaksono, I., & Marsya, E. (2022). The Analysis of Implementation Project-Based Learning Model of Teaching Integrated with Computer Programming in Improving Computational Thinking Skills in a Classical Mechanics Course. *Jurnal Penelitian Pendidikan IPA*, *8*(4), 1734–1742. https://doi.org/10.29303/jppipa.v8i4.1789

Selby, C., & Woolard, J. (2013). Computational Thinking : The Developing Definition. *ITiCSE Conference 2013*, 5–8. https://eprints.soton.ac.uk/356481/

Sokolowski, A. (2018). Scientific inquiry in mathematics - Theory and practice: A STEM perspective. In *Scientific Inquiry in Mathematics - Theory and Practice: A STEM Perspective* (Issue May). Springer International Publishing AG. https://doi.org/10.1007/978-3-319-89524-6

Sumintono, B., & Widhiarso, W. (2014). *Aplikasi Model Rasch Untuk Penelitian Ilmu-Ilmu Sosial* (B. Trim (ed.)). Trim Komunikata Publishing House.

Tan, A.-L., Teo, T. W., Choy, B. H., & Ong, Y. S. (2019). The S-T-E-M Quartet. *Innovation and Education*, *1*(1), 1–14. https://doi.org/10.1186/s42862-019-0005-x

Tang, X., Yin, Y., Lin, Q., Hadad, R., & Zhai, X. (2020). Assessing computational thinking: A systematic review of empirical studies. *Computers and Education*, *148*(January), 103798. https://doi.org/10.1016/j.compedu.2019.103798

Wing, J. M. (2006). Computational thinking. In *Communication of The ACM* (Vol. 49, Issue 3, pp. 33–35). https://doi.org/10.1201/b16812-43

Yin, Y., Hadad, R., Tang, X., & Lin, Q. (2020). Improving and Assessing Computational Thinking in Maker Activities: the Integration with Physics and Engineering Learning. *Journal of Science Education and Technology*, *29*(2), 215. https://doi.org/10.1007/s10956-020-09822-y

Zakwandi, R., & Istiyono, E. (2023). A framework for assessing computational thinking skills in the physics classroom : study on cognitive test development. *SN Social Sciences*. https://doi.org/10.1007/s43545-023-00633-7