ISSN 2502-5872 (Print)
ISSN 2622-3627 (Elektronik)

Volume 9 Number 3, August 2024, 675-690

DEVELOPMENT OF COMPUTATIONAL MATHEMATIC THINKING TEST INSTRUMENTS BASED ON COMPUTERED BASED TEST

Robiah Nur Haniah[1], Mohamad Waluyo[2*]
[1,2]Departement of Mathematics Education, Universitas Muhammadiyah Surakarta, Surakarta, Central Java, Indonesia
*Correspondence: mw192@ums.ac.id

ABSTRACT

This reseacrh is a type of development reseacrh. The aims of this research is to develop an instrument for mathematical computational thinking tests based on Computer Based Testing (CBT) using the Quizizz application, evaluate the validity of the developed test instrument, and analyze the mathematical computational thinking abilities of students. The ADDIE model (Analysis, Design, Development, Implementation, and Evaluation) is adopted as the method in this research. A total of 52 students from classes X and XI of SMA Muhammadiyah 1 Surakarta were involved in this study as research subjects. Data analysis was conducted using the Rasch model with the assistance of Winsteps Rasch software. The results of this research indicate that 20 items are categorized as valid. The overall test reliability is 0.62, classified as sufficient, while the reliability of the test items is 0.88, classified as good. Test analysis of the level of difficulty of test items reveals that there are 3 items categorized as very difficult, 7 items categorized as difficult, 5 items categorized as moderate, and 5 items categorized as easy. In general that the computational thinking skills test does not show a tendency to be biased in any question item, because all probability values for the question items exceed 0.05. Items are considered unbiased if they do not give an unfair advantage to any one individual. Based on the experiment in this study, the average mathematical computational thinking ability of class X students is 75.53, while for class XI students it is 79.37.

Keywords: Computational Thinking, Computer Based Test, Mathematics

How to Cite: Haniah, R. N., & Waluyo, M. (2024). Development of Computational Mathematic Thinking Test Instruments Based On Computered Based Test. *Mathline: Jurnal Matematika dan Pendidikan Matematika*, *9*(3), 675-690. http://doi.org/10.31943/mathline.v9i3.648

PRELIMINARY

Advancements in the field of education are greatly influenced by the rapid development of technology and science (Sudarsana et al., 2019). The Minister of Education and Culture of the Republic of Indonesia, Nadiem Makarim, is committed to formulating a curriculum that includes several key skills that students must possess, including creativity, collaboration, communication, critical thinking, computational thinking, and empathy (Pramesti, 2019). Students' ability in mathematical thinking is crucial to address the challenges of the present era (Insani et al., 2021). Mathematics learning in the 21st century is expected to benefit students in enhancing various skills needed in a rapidly changing global context (Cahdriyana & Richardo, 2020). One important skill in mathematical thinking is computational thinking, which has been recognized by various studies (Barcelos et al., 2018).

Computational Thinking refers to a thinking process aimed at formulating problems and their solutions in a way that allows for efficient implementation by an information processing entity, such as a computer, robot, or human (Bebras, 2019). (Cahdriyana & Richardo, 2020) also describe that Computational Thinking (CT) involves understanding how to solve problems by designing appropriate solutions through algorithm creation. In the context of computational thinking, there are four basic elements that must be considered in efforts to effectively solve problems, namely decomposition, abstraction, pattern recognition, and algorithms (Natali, 2022)

Mathematics is a fundamental subject taught formally to students starting from elementary school level; this subject serves as an indicator in measuring students' understanding of logic. According to (Bebras, 2019) it is a global effort that aims to educate teachers and students from elementary/MI level to the general public about Computational Thinking (Apriani et al., 2021) . The existence of Bebras allows students to determine solutions to CT-related problems in the form of test items with various problems contained within mathematical concepts as an application to daily activities.

Computational Thinking has been integrated into the mathematics test of PISA in 2021, which is an important skill to be assessed (Bebras, 2019).Computational thinking becomes the key to helping students overcome mathematical challenges involving various skills that train students to formulate problems by breaking them down into smaller components to enable easier problem-solving (Lee et al., 2014). Additionally, through this approach, students will be encouraged to test their creativity in facing mathematical problems (Angeli et al., 2020).

Factually, the current learning model narrows students' opportunities to enhance computational thinking skills (Marcelino et al., 2018). Teachers are still accustomed to solving mathematical problems in the form of formulas, which students then memorize and apply to determine solutions to problems during tests (Lee et al., 2014).This causes students to become less proactive in improving computational thinking skills, resulting in low mathematical computational thinking abilities.

The learning process has the potential to change students' behavior through repeated learning processes. Through evaluation processes, understanding can be gained about changes in students' behavior post-learning (Thobroni, 2016). Therefore, teachers require assessment instruments as tools for evaluating student (Desilva et al., 2020). Assessment is the process of collecting data through measurement, then interpreting, describing, and analyzing the data based on information gathered from measurement results (Pramesti, 2019). The evaluation process becomes an important stage that must be carried out to determine the effectiveness of learning.

Initially, the assessment process was conducted through conventional methods using print media such as paper, which was considered less efficient because it would take longer time for students to complete and for teachers to correct, as well as requiring funds to provide questions (Lestari et al., 2022). Online technology in the form of computer-based tools can be used to initiate the development of assessment instruments. Conventional assessment instruments and techniques can provide an overview of the use of technology-based evaluation tools (Hamidah et al., 2021). Computer Based Test (CBT) is a type of test that is conducted through computer devices, minimizing the use of paper, pens, or pencils to answer questions. Mathematics problems are presented in digital format on the computer, and answers can also be inputted through the computer system, allowing exam participants to choose the correct answers or provide written responses if there are essay questions. Computer-based assessments like this are widely used in various fields, including education, usually to assess participants' understanding of specific subjects or courses (Lidya et al., 2018). This step is taken to overcome the shortcomings of traditional evaluation both in process and results (Farman, et al., 2021)

Quizizz is one of the platforms in multimedia-based interactive learning that can be implemented in the classroom as a learning evaluation tool, as well as being useful for learning evaluation processes equipped with various interesting features including music, themes, and the ability to embed images in the background of questions (Jahring et al., 2022). Quizizz has advantages such as time limits on questions that can train students to think quickly and accurately in the process of solving problems. Another advantage is that students' responses to questions will be displayed through images and can be monitored on the teacher's computer as the operator (Citra et al., 2020). Quizizz has flexible and narrative characteristics, meaning that besides being used as a tool for delivering material, it can also function as a fun and engaging evaluation tool in the learning process (Salsabila et al., 2020).

Since 2014, the UK government has introduced computational thinking to students in primary and secondary schools to improve their decision-making and problem-solving abilities (Malik et al., 2018). Research conducted by (Ardania et al., 2022) the test instrument created consists of 20 multiple choice questions and is accompanied by a guide to using the Quizizz application. The data obtained from this development was analyzed quantitatively descriptively. After validation, the results showed that the multiple choice test instrument assisted by the Quizizz application was considered very good, with a success rate reaching 99.3%. The research concluded that the multiple choice test instrument using Quizizz could be applied well to students. Based on the background stated above, the researcher intends to conduct research with the following objectives (1) developing a computational thinking test instrument, (2) determining validity and reliability, (3) determining the level of difficulty.

METHODS

This research implements the Research and Development (R&D) type. Reseach and Development (R&D) involves research followed by development activities (Zahra et al., 2022). The ADDIE model designed by Dick and Carry (1996) is applied as the development design in this study. According to Sutama, R&D is an effort to improve available products or create new ones (Waluyo et al., 2023). The ADDIE model generally consists of five stages, namely, Analysis, which involves observing through interviews with teachers at SMA Muhammadiyah 1 Surakarta; Design, which involves designing the question grid, then creating 20 items with their answer keys; Development, which is the process of validating an instrument by experts; Implementation, which is the actual step to implement the created test instrument; and Evaluation, which is a process to review the results of students' mathematical computational thinking abilities through the created instrument.

This research was conducted at SMA Muhammadiyah 1 Surakarta, involving students from classes X and XI, each consisting of 26 students. The test instrument used was mathematical computational multiple-choice items totaling 20 items, designed using the Quizizz application. Data collection for this research was carried out through online tests using the Quizizz application. The instrument used was a test instrument containing items that students had to answer correctly.

The analysis used for the obtained data includes instrument validity analysis, instrument reliability, and instrument difficulty level analysis. In instrument validity testing, two tests were conducted: content validity and item validity. The content validity analysis of the instrument used the Aiken index.

Tabel 1. Index Content Validity According to Aiken (1985)

Index	Level of Validity
≤ 0.4	Less validity
0.4 – 0.8	Moderate Validity
≥ 0.8	High Validity

If the value approaches an Aiken index of 1, it indicates better validity and greater relevance to the indicators (Retnawati, 2016). Then, Rasch analysis assisted by the Winsteps application was conducted to test item validity, reliability, and difficulty level.

RESULTS AND DISCUSSION

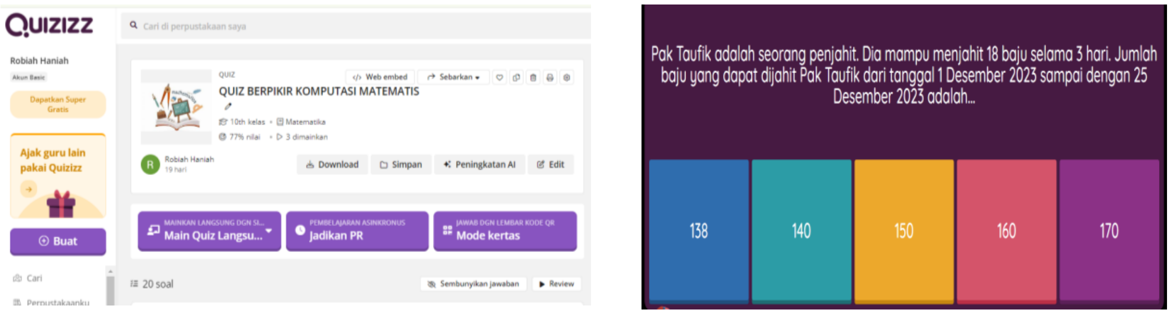
The online-based test instrument is the outcome of the conducted research. This result will be implemented through the Quizizz application. This research produces a product in the form of a quizizz-based test instrument which can be accessed online on that page https://quizizz.com/admin/quiz/65a51d83ec7f171a417441b2?source=quiz_share.In this development research, the ADDIE model was applied. The following are the processes for each sequence in the model.

1. Analysis

The analysis process begins with interviews and observations of mathematics teachers at SMA Muhammadiyah 1 Surakarta. Based on the interview results, it was found that teachers have not yet implemented mathematical computational thinking test instruments. This stage is carried out to determine that the development of the test instrument meets the needs and can be useful in solving mathematical problems in learning evaluation. Based on the pre-research analysis, it was concluded that the development of mathematical computational thinking test instruments is needed to support the development of 21st-century skills. SMA Muhammadiyah 1 Surakarta already has various adequate facilities such as computer labs, LCD projectors, Android devices, laptops, and wifi networks, which can be supporting factors in the creation and implementation of online test instruments based on Quizizz. After interviewing the teachers, the following results were obtained: 1) There is no evaluation instrument to test students' mathematical computational thinking skills, 2) Quizizz application has not yet been implemented in the learning process at SMA Muhammadiyah 1 Surakarta.

2. Design

In the design process, it starts with designing the question grid, question items, and answer keys. Twenty multiple-choice question items related to mathematical computational thinking are created. These question items are then inputted into the Quizizz platform, which includes creating a logo, selecting the appropriate subject, adding quiz names, creating a logo, and inputting question items with answer keys on the Quizizz platform, as well as setting the tempo for when students respond to questions. The instrument's appearance is shown in figure 1.



English Version

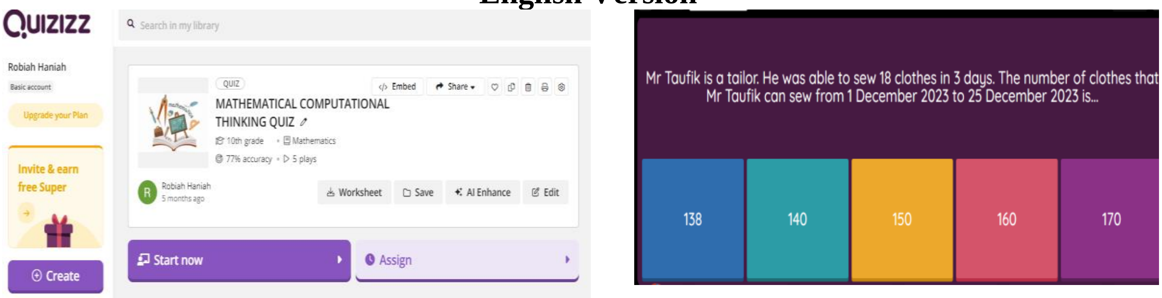


Figure 1. Display of the Design of the Mathematical Computational Thinking Test Instrument on Quizizz

3. Development

In the development stage, the instrument that has been designed is then validated by experts according to construction, content, and language aspects. Validation in the development of this instrument is carried out by 2 expert parties, namely 1 Mathematics Education Lecturer and 1 Mathematics Education Teacher from Senior High School. Referring to the validity test with the Aiken index, it was found that 10 items of the instrument were declared valid with a high validity category. Expert validation analysis of the mathematical computational thinking test instrument based on the Quizizz online application indicates that it falls into the valid or highly valid category with a score of 0.83333. Experts also provide several suggestions for improvement such as using correct language

4. Implementation

The validated test instrument is then implemented in the classroom. This implementation stage involves students from classes X and XI of SMA Muhammadiyah 1 Surakarta, each consisting of 26 students per class. The workflow of this implementation stage is that each student completes a set of 20 multiple-choice questions online through the Quizizz platform. Then, the scores obtained by the students are used for calculating the validity of the items. Validity testing is conducted by linking the scores obtained for each question item with the total scores of each individual. Data processing in this validity test uses the Rasch model assisted by the Winsteps program. Item validity testing is conducted for two classes with a total of 52 respondents. Meanwhile, decision-making is based on the values of Outfit Z-Standard (ZSTD), Outfit Mean Square (MNSQ), and Point Measure Correlation (Pt Mean Corr). Validity testing using the Rasch Model, referring to (Sumintono & Widhiarso, 2015) is conducted to review whether the question items are considered valid according to the following criteria: Accepted outfit ZSTD value: -2.0 < ZSTD < +2.0; Accepted outfit MNSQ value: 0.5 < MNSQ < 1.5; Accepted Pt Measure Corr value: 0.4 < Pt Measure Corr < 0.85. Based on calculations from 20 questionnaire items that have been tested on respondents, a total of 20 valid questionnaire items were obtained as shown in Figure 2.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	PTMEASUR-AL EXP.	EXACT MATCH OBS%	EXACT MATCH EXP%	Item
17	20	52	2.32	.32	1.12	.99	1.38	1.80	A .31	.43	64.0	68.6	S17
13	25	52	1.84	.31	1.12	1.22	1.23	1.44	B .31	.42	60.0	64.7	S13
2	38	52	.54	.34	1.16	.97	1.09	.45	C .26	.37	70.0	76.0	S2
16	44	52	-.28	.41	.93	-.19	1.16	.51	D .33	.32	86.0	84.4	S16
18	42	52	.03	.38	1.09	.47	1.12	.46	E .26	.34	80.0	81.4	S18
20	41	52	.17	.37	1.11	.58	1.04	.23	F .28	.35	78.0	79.9	S20
19	41	52	.17	.37	1.07	.42	.98	.04	G .31	.35	78.0	79.9	S19
3	44	52	-.28	.41	1.05	.26	.81	-.34	H .32	.32	82.0	84.4	S3
5	42	52	.03	.38	1.03	.21	.91	-.17	I .33	.34	80.0	81.4	S5
8	48	52	-1.16	.54	.96	.03	1.00	.22	J .26	.25	92.0	91.9	S8
7	31	52	1.27	.31	.92	-.68	.95	-.25	j .45	.40	78.0	67.4	S7
11	50	52	-1.94	.74	.95	.13	.55	-.19	i .25	.18	96.0	96.0	S11
10	46	52	-.66	.46	.93	-.11	.81	-.24	h .34	.29	88.0	87.9	S10
6	35	52	.87	.32	.92	-.56	.83	-.82	g .46	.39	74.0	72.1	S6
9	49	52	-1.49	.61	.92	-.01	.75	-.08	f .29	.22	94.0	93.9	S9
15	49	52	-1.49	.61	.92	-.01	.75	-.08	e .29	.22	94.0	93.9	S15
12	49	52	-1.49	.61	.91	-.03	.51	-.48	d .32	.22	94.0	93.9	S12
14	46	52	-.66	.46	.90	-.23	.78	-.29	c .37	.29	88.0	87.9	S14
4	47	52	-.89	.49	.88	-.25	.67	-.45	b .37	.27	90.0	89.9	S4
1	13	52	3.10	.36	.87	-.62	.76	-.70	a .54	.44	82.0	79.2	S1
MEAN	40.0	52.0	.00	.44	.99	.13	.90	.05			82.4	82.7	
P.SD	10.1	.0	1.32	.12	.09	.51	.22	.63			9.9	9.2	

Figure 2. Calculation Results of Item Validity Test using Rasch Model

Reliability analysis of the online-based mathematical computational thinking skills test instrument through the Rasch model is presented in Figure 3.

CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .62 SEM = 1.61

SUMMARY OF 20 MEASURED (NON-EXTREME) Item

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	40.0	52.0	.00	.44	.99	.13	.90	.05
SEM	2.3	.0	.30	.03	.02	.12	.05	.15
P.SD	10.1	.0	1.32	.12	.09	.51	.22	.63
S.SD	10.3	.0	1.35	.12	.09	.53	.23	.65
MAX.	50.0	52.0	3.10	.74	1.16	1.22	1.38	1.80
MIN.	13.0	52.0	-1.94	.31	.87	-.68	.51	-.82

REAL RMSE .46 TRUE SD 1.23 SEPARATION 2.67 Item RELIABILITY .88
MODEL RMSE .46 TRUE SD 1.24 SEPARATION 2.70 Item RELIABILITY .88
S.E. OF Item MEAN = .30

Item RAW SCORE-TO-MEASURE CORRELATION = -.98

Figure 3. Calculation Results of Reliability Test Using Rasch Model

In the figure 3, the Cronbach's alpha value obtained for the reliability of mathematical computational thinking abilities is 0.62, which falls into the category of "fair". In the table above, the reliability value of the items is 0.88, categorized as "good" as stated by (Sumintono & Widhiarso, 2015).

Based on the Rasch Model test results, an analysis of the difficulty level of the questions is presented with a series of items ranging from difficult to easy. Figure 4 contains the results of the analysis of the difficulty level of mathematical computational thinking abilities questions.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE
1	13	52	3.10
17	20	52	2.32
13	25	52	1.84
7	31	52	1.27
6	35	52	.87
2	38	52	.54
19	41	52	.17
20	41	52	.17
5	42	52	.03
18	42	52	.03
3	44	52	-.28
16	44	52	-.28
10	46	52	-.66
14	46	52	-.66
4	47	52	-.89
8	48	52	-1.16
9	49	52	-1.49
12	49	52	-1.49
15	49	52	-1.49
11	50	52	-1.94
MEAN	40.0	52.0	.00
P.SD	10.1	.0	1.32

Figure 4. Difficulty Level Testing of Item Questions Using Rasch Model

Based on the figure 4 above, the measure values indicate the sequence of item questions from the most difficult to the easiest. Item question number 1 is the most difficult, while item question number 11 falls into the easy category for students. The categorization of the difficulty levels of questions can be reviewed based on the mean value of 0.00 and a standard deviation of 1.32. The data obtained from the analysis of difficulty levels are grouped into four categories: items with the most difficult category are indicated by items with measures greater than +1.32 SD, difficult items fall within the interval of 0.00-1.32, moderate difficulty items fall within the interval of -1.32-0.00, and easy items have measures below -1.32. The difficulty level test can also be seen in the Wright person item map. The following Figure 5 displays the results of the Wright person item map.

```
MEASURE    Person - MAP - Item
               <more>|<rare>
    4            XX T+
                     |
                 XX  |
                     |
                     |
                     |
                     |
                     |
                     |
                     |  S1
    3                +
               XXXX S|
                     |
                     |
                     |T
                     |
                     |
          XXXXXXXXX  |  S17
                     |
                     |
    2                +
                     |
      XXXXXXXXXXXXX M|  S13
                     |
                     |
                     |
         XXXXXXXXXX  |
                     |S S7
                     |
               XXXX  |
    1                +
                     |  S6
                  X  |
                     |
                    S|
                  X  |  S2
                     |
                     |
                 XX  |  S19    S20
                     |
    0                +M S18    S5
                  X  |
                     |
                 XX  |  S16    S3
                     |
                    T|
                  X  |
                     |  S10    S14
                     |
                     |  S4
```

```
 -1              +
                 |
                 |  S8
                 |S
                 |
                 |  S12    S15    S9
                 |
                 |
                 |
                 |  S11
 -2              +
           <less>|<freq>
```

Figure 5. Result of Wright Person Item Map

Based on figure 4 and figure 5 above, there are 5 questions categorized as easy, 5 questions categorized as moderate, 7 questions categorized as difficult, and 3 questions categorized as very difficult. Questions categorized as very difficult can only be correctly answered by a few students with very high computational thinking abilities, namely questions number 1, 13, and 17. Questions categorized as very difficult can only be solved by students with high computational thinking abilities, namely questions number 2, 6, 19, 20, 5, 18, and 7. Questions categorized as moderate are those that can be easily solved by students with high mathematical computational thinking abilities, but some students with low mathematical computational thinking abilities may find them somewhat difficult. The items categorized as moderate are numbers 3, 4, 10, 14, and 16. Meanwhile, questions categorized as low difficulty level are questions number 8, 9, 11, 12, and 15. Questions categorized as low difficulty level mean that they can be answered by all students. The differences that differentiate each test item are seen based on the DIF value or characteristics of each item. Evaluation of the characteristics of each test item can be found in the following figure 6.

Person CLASSES	SUMMARY DIF CHI-SQUARED	D.F.	PROB.	BETWEEN-CLASS/GROUP UNWTD MNSQ	ZSTD	Item Number	Name
6	2.0517	5	.8419	.4425	-.92	1	Q1
6	6.3995	5	.2689	1.9510	1.39	2	Q2
6	7.5180	5	.1846	3.0772	2.37	3	Q3
6	.5285	5	.9910	.1491	-2.02	4	Q4
6	4.1997	5	.5208	1.4887	.88	5	Q5
6	4.4745	5	.4830	1.4315	.81	6	Q6
6	7.0009	5	.2203	2.2732	1.70	7	Q7
6	2.3513	5	.7986	1.0127	.23	8	Q8
6	1.4840	5	.9149	.4925	-.79	9	Q9
6	3.3878	5	.6402	1.1019	.37	10	Q10
6	1.6359	5	.8968	.4616	-.87	11	Q11
6	2.2302	5	.8164	.7800	-.17	12	Q12
6	9.8310	5	.0800	2.9541	2.27	13	Q13
6	.9961	5	.9629	.4546	-.89	14	Q14
6	1.2307	5	.9419	.4295	-.95	15	Q15
6	1.3269	5	.9321	.5175	-.72	16	Q16
6	5.8073	5	.3251	1.8901	1.33	17	Q17
6	7.0461	5	.2169	2.4179	1.83	18	Q18
6	8.7824	5	.1178	3.2115	2.47	19	Q19
6	4.9698	5	.4193	1.5494	.96	20	Q20

Figure 6. Different Power Test Results Test Questions

Based this figure 6, it can be concluded in general that the computational thinking skills test does not show a tendency to be biased in any question item, because all probability values for the question items exceed 0.05. Items are considered unbiased if they do not give an unfair advantage to any one individual.

5. Evaluation

In this stage, the mathematical computational thinking abilities of students can be observed after they have completed the questions. The assessment results of students' mathematical computational thinking abilities can be seen in the following bar chart :

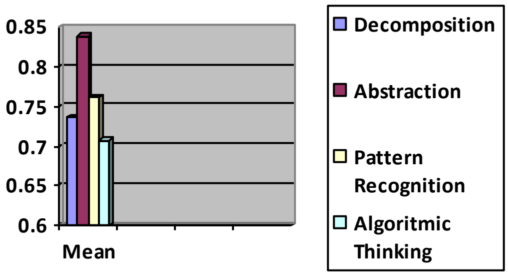


Figure 7. Bar Chart Average Computational Thinking Ability for each Aspect of Class X

Based on figure 7 results of the average scores for each aspect of mathematical computational thinking abilities in class X indicate that students' abilities in the abstraction aspect have improved more compared to other aspects. This implies that students in class XI have the ability to identify important objects in creating models/representations to solve a problem. The student test results yield an average score of 75.5357, with the highest score being 90 and the lowest score being 50. The detailed scores of the class X students' tests are presented in the table below.

Table 2. Details of Mathematical Computational Thinking Ability Test Results for Class X

Description	Score
The highest score	90
Lowest value	50
Mean	75.5357

The following is the average computational thinking ability of class XI students.

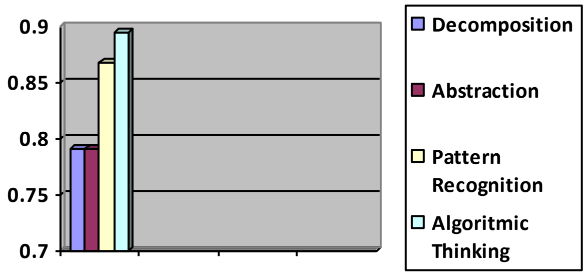


Figure 8. Bar Chart Average Computational Thinking Ability for each Aspect of Class XI

The mean scores obtained from the bar chart for each aspect of mathematical computational thinking abilities in class XI indicate that students' abilities in the algorithmic thinking aspect have improved more compared to other aspects. This means that class XI students have the ability to design a series of steps accurately to obtain a solution to a problem. In contrast to the average scores in class X, the mathematical computational thinking ability in the abstraction aspect has the lowest average compared to other aspects. The test results for class XI students show an average of 79.375, with the lowest score being 40 and the highest score being 100. A detailed description of the test results for class XI students is provided in Table 3.

Table 3. Details of Evaluation Results of Mathematical Computational Thinking Abilities for Class XI

Description	Score
The highest score	100
Lowest velue	40
Mean	79.375

Based on the description table 3 of the test results for class X and XI students, it can be concluded that the computational thinking abilities of class XI students show a better improvement compared to the computational thinking abilities of class X.

Generally based on the validity analysis, reliability test, and difficulty level of the questions, it can be concluded that the mathematical computational thinking test instrument is valid and reliable. The online test instrument through the Quizizz application is suitable for implementation and utilization in the classroom as mathematics learning evaluation tool (Jahring et al., 2022). Mathematical computational thinking abilities can design the learning process with the aim of understanding computational thinking approaches to solve a problem and develop solutions to address problems effectively when needed (Kalelioğlu et al., 2016). Developing online-based test instruments using Quizizz is more flexible, and the utilization of information and communication technology in learning activities can support the quality of education as part of implementing digital learning programs. This is proven in research (Ardania et al., 2022) that the results of developing a multiple choice question instrument using the Quizizz application show a very high level of adequacy, reaching a percentage of 99.3%. Based on evaluations from experts, the multiple choice questions instrument supported by the Quizizz application can be used effectively for students.

CONCLUSION

The developed mathematical computational thinking test instrument consists of 20 multiple-choice questions. Based on the findings of this research, all 20 questions were found to be valid. The overall test reliability score is 0.62, categorized as sufficient, while the reliability test for individual questions yielded a score of 0.88, categorized as highly reliable. The difficulty level test revealed that there are 5 questions classified as easy, 5 questions classified as moderate, 7 questions classified as difficult, and 3 questions classified as very difficult. The computational thinking skills test did not show any bias in any question item because all probability values for the question items exceeded 0.05. Based on the average computational thinking ability per aspect, there is still a need to improve mathematical computational thinking abilities in aspects such as decomposition, abstraction, pattern recognition, and algorithmic thinking. The description table of test results for students in both classes X and XI indicates that the mathematical computational thinking ability of class XI students has improved compared to that of class X students.

Referring to the findings of this research, further steps are needed in the implementation of the test to assess mathematical computational thinking abilities to meet the challenges of the 21st century. Additionally, improvements are necessary to produce better instruments that can be tested on a wider scale.

REFERENCES

Aiken, L. R. (1985). Three Coefficients for Analyzing the Reliability and Validity of Ratings. *Educational and Psychological Measurement*, *45*(1), 131–142. https://doi.org/https://doi.org/10.1177/00131644854

Angeli, C., & Giannakos, M. (2020). Computational thinking education: Issues and challenges. *Computers in Human Behavior*, *105*, pp.1-16. https://doi.org/10.1016/j.chb.2019.106185

Apriani, A., Ismarmiaty, I., Susilowati, D., Kartarina, K., & Suktiningsih, W. (2021). Penerapan Computational Thinking pada Pelajaran Matematika di Madratsah

Ibtidaiyah Nurul Islam Sekarbela Mataram. *ADMA : Jurnal Pengabdian Dan Pemberdayaan Masyarakat*, *1*(2), 47–56. https://doi.org/10.30812/adma.v1i2.1017

Ardania, R., Fitriyah, L, A., & Kuswanti, N. (2022). Pengembangan Instrumen Soal Pilihan Ganda Berbantu Aplikasi Quizizz Materi Sistem Pernapasan Pada Manusia Untuk Peserta Didik Kelas Viii Smp. *Discovery : Jurnal Ilmu Pengetahuan*, *7*(1), 12–18. https://doi.org/10.33752/discovery.v7i1.2414

Barcelos, T. S., Munoz, R., Villarroel, R., Merino, E., & Silveira, I. F. (2018). Mathematics Learning through Computational Thinking Activities: A Systematic Literature Review. *Journal of Universal Computer Science*, *24*(7), 815–845. https://d1wqtxts1xzle7.cloudfront.net/96594322/jucs_24_07_0815_0845_barcelos-libre.pdf?1672437306=&response-content-disposition=inline%3B+filename%3DMathematics_Learning_throug

Bebras, I. (2019). *Tantangan Bebras Indonesia-2018 Tingkat SD (Siaga)*. Bebras International. http:/bebras.or.id

Cahdriyana, R. A., & Richardo, R. (2020). Berpikir Komputasi Dalam Pembelajaran Matematika. *LITERASI*, *XI*, 50–56. https://doi.org/10.1007/978-3-319-52691-1_13

Citra, C. A., & Rosy, B. (2020). Keefektifan Penggunaan Media Pembelajaran Berbasis Game Edukasi Quizizz Terhadap Hasil Belajar Teknologi Perkantoran Siswa Kelas X SMK Ketintang Surabaya. *Jurnal Pendidikan Administrasi Perkantoran (JPAP)*, *8*, 261–272. https://doi.org/http://dx.doi.org/10.26740/jpap.v8n2.p261-272

Desilva, D., Sakti, I., & Medriati, R. (2020). Pengembangan Instrumen Penilaian Hasil Belajar Fisika Berorientasi Hots (Higher Order Thinking Skills) Pada Materi Elastisitas Dan Hukum Hooke. *Jurnal Kumparan Fisika*, *3*(1), 41–50. https://doi.org/10.33369/jkf.3.1.41-50

Farman, Hali, F., & Rawal, M. (2021). Development of E-LKPD Using Live Worksheets for Online Mathematics Learning during Covid-19. *JME (Journal of Mathematics Education)*, *6*(1), 36–42. https://doi.org/10.31327/jme.v6i1.1626

Hamidah, M. H., & Wulandari, S. S. (2021). Pengembangan Instrumen Penilaian Berbasis Hots Menggunakan Aplikasi "Quizizz" This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. *Kajian Ilmu Administrasi*, *18*(1), 105–124. https://doi.org/10.21831/efisiensi.v17i1.34895

Insani, K., Hobri, Prihandoko, A. C., Sa'id, I. A., & Safik, M. (2021). Developing of learning tools based on science, technology, engineering, and mathematics (STEM) based on learning community to improve critical thinking ability in class X student's arithmetic sequences and arithmetic materials. *Journal of Physics: Conference Series*, *1839*(1), 1–9. https://doi.org/10.1088/1742-6596/1839/1/012020

Jahring, J., Herlina, H., Nasruddin, N., & Astrinasari, A. (2022). Pengembangan Instrumen Evaluasi Pembelajaran Matematika Berbasis Online Menggunakan Aplikasi Quizizz. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *11*(2), 872–881. https://doi.org/10.24127/ajpm.v11i2.4932

Kalelioğlu, F., Gülbahar, Y., & Kukul, V. (2016). A Framework for Computational Thinking Based on a Systematic Research Review. *Baltic J. Modern Computing*, *4*(3), 583–596. https://doi.org/26908185

Lee, T. Y., Mauriello, M. L., Ahn, J., & Bederson, B. B. (2014). CTArcade: Computational thinking with games in school age children. *International Journal of Child-Computer Interaction*, *2*(1), 26–33. https://doi.org/10.1016/j.ijcci.2014.06.003

Lestari, N. T., & Ulfa, S. W. (2022). Pengembangan Instrumen Penilaian Menggunakan Aplikasi Quizizz Pada Mata Pelajaran Biologi Materi Sistem Pencernaan. *Jurnal Paedagogy: Jurnal Penelitian Dan Pengembangan Pendidikan*, *9*(3), 602–608. https://doi.org/10.33394/jp.v9i3.5374

Malik, S., Prabawa, H. W., & Rusnayati, H. (2018). Peningkatan Kemampuan Berpikir Komputasi Siswa melalui Multimedia Interaktif Berbasis Model Quantum Teaching and Learning. *International Journal of Computer Science Education in Schools, 8*(November), pp.1-41. https://doi.org/10.13140/RG.2.2.34438.83526

Marcelino, M. J., Pessoa, T., Vieira, C., Salvador, T., & Mendes, A. J. (2018). Learning Computational Thinking and scratch at distance. *Computers in Human Behavior, 80,* 470–477. https://doi.org/10.1016/j.chb.2017.09.025

Natali, V. (2022). *Computational Thinking Mata Kuliah Pilihan.* Direktorat Kementrian Pendidikan Profesi Guru Direktorat Jenderal Guru Dan Tenaga Kependidikan Kementerian Pendidikan, Kebudayaan, Riset Dan Teknologi.

Pramesti, A. N. (2019). *Mendikbud Paparkan Kepemimpinan 2. Direktorat Pembinaan Tenaga Kependidikan, Direktorat Jendral Guru Dan Tenaga Kependidikan*. Kementrian Pendidikan Dan Kebudayaan. https://tendik.kemdikbud.go.id/berita/detail/134-mendikbud-paparkankepemimpinan-20-pada-sistem-pendidikan-di-simposium-internasional-kepala-sekolah-danpengawas-sekolah-tahun-2019

Retnawati, H. (2016). *Analisis Kuantitatif Instrumen Penelitian (Panduan Peneliti, Mahasiswa, dan Psikometrian)*.

Salsabila, U. H., Habiba, S., Amanah, I. L., Istiqomah, N. A., & Difany, S. (2020). Pemanfaatan Aplikasi Quizizz Sebagai Media Pembelajaran Ditengah Pandemi Pada Siswa SMA. *Jurnal Ilmiah Ilmu Terapan Universitas Jambi*, *4*(2), 163–172. https://doi.org/http://dx.doi.org/10.22437/jiituj.v4i2.11605

Sudarsana, I. K., Nakayanti, A. R., Sapta, A., Haimah, Satria, E., Saddhono, K., Achmad Daengs, G. S., Putut, E., Helda, T., & Mursalin, M. (2019). Technology Application in Education and Learning Process. *Journal of Physics: Conference Series, 1363*(1), 1–6. https://doi.org/10.1088/1742-6596/1363/1/012061

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelah Rasch Pada Assessment Pendidikan*. Cimahi : Tim Komunikata Publishing House.

Thobroni. (2016). *Belajar dan Pembelajaran Teori dan Praktik*. Yogyakarta : Ar-Ruzz Media.

Waluyo, M., & Bima, W. P. (2023). Pengembangan Lkpd Inquiry Based Learning Untuk Mendukung Kemampuan Penalaran Dan Pembuktian Matematis. *JMPM: Jurnal Matematika Dan Pendidikan Matematika*, *8*(2), 127–140. https://doi.org/10.26594/jmpm.v8i2.3536

Zahra, N. A., & Cahdriyana, R. A. (2022). Development Of Website-Based Learning Media Combined With Geogebra On Pythagoras Materials For VIII-Class Students Of Smp Negeri 1 Pasawahan, Kuningan District. *Mathline : Jurnal Matematika Dan Pendidikan Matematika*, *8*(1), 279–294. https://doi.org/10.31943/mathline.v8i1.299