Modelling the effect of Bing Chat integrated instruction with analogies on students' computational thinking: A structural approach

Rahma Diani[1*], Hendri Noperi[1], Cindi Almira Putri[1], Mukarramah Mustari[1], Yessy Velina[2]

[1]Physics Education Study Program, Universitas Islam Negeri Raden Intan Lampung, Endro Suratmin Street, Lampung, 35131, Indonesia

[2]Interdisciplinary Environments Department, Kyoto University, Japan

*Corresponding author, email: rahmadiani@radenintan.ac.id

Article History
Received: 28 February 2026
Revised: 20 May 2026
Accepted: 6 July 2026

Keywords
Bing chat
Computational thinking
Teaching with analogy

Abstract
The purpose of this study was to determine the effect of implementing Bing Chat in the Teaching with Analogies (TWA) model on students' computational thinking skills. This study used a quasi-experimental research method with a non-equivalent control group design. This study was conducted at Yasmida Ambarawa High School in August 2024. The research sample consisted of 65 students, consisting of 33 students of class X.1 as the experimental class and 32 students of class X.2 as the control class, using cluster random sampling. The instrument used was a computational thinking ability test consisting of 10 essay questions. Based on the results of hypothesis testing using statistical tests, namely the independent sample t-test with the help of Jamovi software, the results of the independent sample t-test showed a p-value of 0.001, which means a significance value <0.05, so H1 was accepted. This indicates that there is a significant difference between the average Computational Thinking skills of students in the control class and the experimental class after being given different treatments.

Diani, R., Noperi, H., Putri, C. A., Mustari, M., & Velina, Y. (2026). Modelling the effect of Bing Chat integrated instruction with analogies on students' computational thinking: A structural approach. *Momentum: Physics Education Journal*, *10*(1), 162-172. https://doi.org/10.21067/mpej.v10i1.13935

1. Introduction

Physics education at the high school level requires students to possess higher-order thinking skills, particularly Computational Thinking (CT). CT encompasses the processes of decomposition, pattern recognition, abstraction, and algorithms that enable students to systematically solve physics problems (Weintrop et al., 2016; Wing, 2006) . In the context of renewable energy, CT skills are crucial because students must understand abstract concepts, model physical processes, and relate mathematical representations to real-world phenomena in physics, biology, and chemistry (Deissenroth et al., 2017) . Renewable energy topics involve complex and interconnected systems, such as energy conversion processes, power generation, and environmental interactions, which are often difficult for students to directly visualize. In conventional learning, students often experience difficulties in connecting theoretical physics concepts with real-world energy phenomena, especially when interpreting abstract systems and identifying conceptual relationships (Diani & Hartati, 2018) . In the context of learning, the development of computational thinking in renewable energy learning requires the integration of analogy-based pedagogical approaches and AI technology support so that students can build more concrete, systematic, and contextual conceptual representations (Grover & Pea, 2013; Shute et al., 2017) .

One method that has the potential to support students' understanding of abstract physics material is Teaching with Analogies (TWA) (Diani et al., 2019) . This method uses analogies to explain complex concepts in simpler ways, helping students understand abstract concepts through analogies relevant to their everyday experiences, thereby improving their conceptual understanding and scientific reasoning (Duit et al., 2001; Glynn, 1991) . In the analogy process, students compare target concepts in physics with familiar situations from everyday life to identify structural similarities between the two domains. This activity encourages students to recognize patterns, filter relevant

doi: 10.21067/mpej.v10i1.13935
This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2548-9135

information, and build abstract conceptual relationships, which are fundamental processes in computational thinking (Weintrop et al., 2016; Wing, 2006) . Analogy-based learning theoretically supports the core dimensions of computational thinking, particularly abstraction and pattern recognition. Research in Learning and Instruction shows that the effective use of analogies in science learning can improve students' critical thinking skills and enable them to connect newly learned concepts to existing experiences or knowledge (Richland & Simms, 2015; Saregar et al., 2018) . In physics learning, analogies have been shown to improve the quality of students' mental modeling and help them understand the structure of concepts that cannot be directly observed (Harrison & Treagust, 2006). The effectiveness of analogies depends on the teacher's ability to provide appropriate analogies and further explanations to address misunderstandings (Fuadi et al., 2020; Klopfer et al., 2009) . However, the effectiveness of conventional TWA implementation is often limited by the teacher's ability to provide diverse, contextual, and adaptive analogies according to students' learning needs.

In recent years, the integration of technology into learning has become increasingly important, particularly through the use of digital platforms such as Bing Chat. This technology has the potential to be an interactive medium that can increase student engagement and facilitate the use of analogies in a more dynamic and responsive manner (Ali et al., 2024; Susskind, 2018) . Bing Chat, as an AI-based platform, has the ability to respond and explain concepts interactively, which can support the TWA method by offering a variety of additional analogies and helping students overcome difficulties in understanding physics concepts (Bower et al., 2024; Grover & Pea, 2013) .

Research conducted by (Weintrop et al., 2016) reported that the use of chatbots can improve the quality of conceptual understanding through interactive dialogue and strengthen students' reasoning skills in solving science problems. Furthermore, AI-based chatbots increase students' motivation, engagement, and the quality of cognitive reflection, primarily due to AI's ability to provide immediate adaptive feedback. The integration of AI-based chatbots can enhance student understanding through active interaction, which in turn supports the development of critical and creative thinking skills (Kumar et al., 2025; Lai & Bower, 2020) .

Previous research has shown that research on CT, TWA, and the use of AI has progressed rapidly but remains fragmented. Some studies have focused on the effectiveness of analogies in physics learning (Harrison & Treagust, 2006; Richland & Simms, 2015) . Meanwhile, other studies have highlighted the role of chatbots in improving conceptual understanding or learning motivation (Rahman et al., 2019) . On the other hand, research on CT in physics education has largely focused on instrument development or conceptual studies without integrating assistive technology (Mulyono et al., 2025) . In renewable energy learning, students are required to understand complex and abstract systems such as energy conversion, electricity generation, and environmental interactions that are difficult to observe directly. These concepts often involve interconnected processes and dynamic relationships, making it difficult for students to abstract and recognize patterns, which are essential components of computational thinking. Although the TWA model can help students connect abstract physics concepts with everyday experiences, conventional analogy-based learning still has limitations because the analogies provided by teachers are often static, limited in number, and unable to represent concepts through various forms of visual, contextual, and interactive representations.

Bing Chat (currently integrated as Microsoft Copilot) offers several advantages that can more effectively support the TWA model. This platform is capable of generating contextual analogies, providing up-to-date references linked to internet sources, and presenting multimodal explanations through text and visual representations (Saregar et al., 2025) . These features enable students to explore renewable energy concepts more interactively and help teachers provide richer and more adaptive analogy explanations. These advantages make Bing Chat more relevant for use in analogy-based learning compared to conventional chatbots that focus solely on textual responses (Diani, Saregar, et al., 2016) . However, to date, no research has specifically investigated the integration of Bing Chat into the TWA model to improve students' computational thinking skills in renewable energy learning. A significant research gap remains regarding how AI-assisted analogy teaching can support abstraction, pattern recognition, and other computational thinking processes in physics education.

Therefore, this study aims to explore the effect of implementing Bing Chat in the TWA model on high school students' computational thinking skills in learning physics material about renewable energy. The novelty of this study lies in the integration of AI as an analogy supporter in the TWA model, as well as its focus on the development of CT, which has not been widely studied in previous research. The results of this study are expected to contribute to the development of adaptive and innovative physics learning strategies that are appropriate to the needs of 21st-century learning.

2. Method

This study used a quantitative approach with a quasi-experimental design. The research design employed was a non-equivalent control group design. In this design, the experimental and control groups were not randomly selected individually, but rather used pre-existing groups (classes). However, both groups were given a pretest to determine their initial abilities and a posttest to measure the impact of the treatment.

The population in this study consisted of all 10th-grade students at Yasmida Ambarawa High School in the 2024/2025 academic year, divided into several classes. The sampling technique used was cluster random sampling. The sampling procedure was carried out by drawing lots from all k experimental classes receiving TWA assisted by Bing Chat. From the results of the draw, two sample classes were selected, namely class X.1 as the experimental class (33 students) who received Teaching by Analogy (TWA) assisted by Bing Chat, and class X.2 as the control class (32 students) who used the conventional learning model. The total sample consisted of 65 students. To minimize potential bias, both classes were taught by the same physics teacher, received the same learning objectives and teaching duration, and were allocated the same teaching time during the study. In addition, both classes studied the same renewable energy material and used equivalent assessment instruments.

In the experimental class, Bing Chat was integrated into several phases of the Teaching with Analogies (TWA) model. During the introduction phase, the teacher used Bing Chat to generate contextual analogies related to renewable energy concepts based on students' everyday experiences. In the analogy exploration phase, students interacted with Bing Chat to ask questions, identify similarities between the target physics concept and everyday situations, and obtain additional explanations or visual representations. Furthermore, in the concept mapping and reflection phase, Bing Chat was used to help students compare limitations and similarities between the analogy source and the target concept, thus supporting the abstraction process and the introduction phase of patterns in computational thinking. Meanwhile, the control class received conventional instruction. without the support of AI-assisted analogies. The study was conducted in four sessions over two weeks.

The data collection instrument consisted of a Computational Thinking (CT) ability test in the form of an essay with 10 questions. This instrument was developed based on CT indicators, which include decomposition, introduction of phase patterns, abstraction, and algorithms. Before being used, the instrument underwent validity and reliability testing using the Rasch Measurement Model analysis. Based on the Rasch analysis, all items were declared valid (appropriate) because they met the Outfit Mean Square (MNSQ) criteria in the range of $0.5 < MNSQ < 1.5$, Outfit ZSTD in the range of $-2.0 < ZSTD < +2.0$, Pt-Measure Corr in the range of $0.4 < PTMEA < 0.8$, and had good reliability, so it is suitable for measuring computational thinking abilities. (Boone et al., 2014; Laliyo et al., 2022; J. Linacre, 2009; JM Linacre, 2012; Sumintono & Widhiarso, 2015) .

The essay test was scored using an analytical scoring rubric developed based on computational thinking indicators, namely decomposition, introduction of phase patterns, abstraction, and algorithmic thinking. Each item was scored on a scale of 0–3, with higher scores indicating a more complete and systematic demonstration of computational thinking processes (Wulandari et al., 2025) . Student responses were evaluated by the researcher and a physics education expert to maintain consistency of scoring. Before data were analyzed using Winsteps software, all raw scores were tabulated and checked to ensure accuracy and completeness of the assessment.

Data analysis was conducted in two stages. First, raw scores from the pretest and posttest were converted to interval-scale data (Logit/Measure values) using Winsteps (Rasch Modeling) software (J.M. Linacre, 2012; Sumintono & Widhiarso, 2015) . Conversion to logit was performed to improve

measurement accuracy compared to using only raw scores (Ekstrand et al., 2022; Leung et al., 2022; Planinic et al., 2019; Tesio et al., 2023) . Second, logit data were analyzed using inferential statistics assisted by JAMOVI software. Prerequisite tests included normality and homogeneity tests (Kim & Park, 2019) . After the prerequisites were met, hypothesis testing was conducted using an independent sample t-test. to see significant differences in the improvement of CT abilities between the experimental and control classes (Barber et al., 2022; Calkin et al., 2023; Leung et al., 2022).

3. Results and Discussion

3.1. Computational Thinking Ability Improvement (Individual Measurement)

Based on the analysis using Rasch Modeling, the individual measure values (student ability levels) were obtained in logit units. Summary pretest and posttest statistics for the experimental and control classes are presented in **Table** 1.

Table 1. The mean logit of the pretest and post-test of students in the experimental class and the control class.

Class	Number of Students	Number of Questions	Mean/SE (logit) of Students		
			Pre-test	Post-test	Before and After Changes
Experimental	33	10	0.21/ (0.31)	1.90/ (0.32)	1.69
Control	32	10	-0.04/ (0.33)	1.56/ (0.31)	1.60

Table 1 shows that both classes experienced improved performance. However, the experimental class showed dominance, with a final (post-test) performance score of 1.90 logits, higher than the control class's score of 1.56 logits. The improvement in the experimental class (+1.69 logits) was also greater than that in the control class (+1.60 logits). This indicates that the Bing Chat-assisted TWA model intervention was more effective in improving students' critical thinking skills than conventional learning.

3.2. Analysis of Changes in the Level of Difficulty of Question Items (*Measurement of Question Items*)

To validate this improvement in ability, an analysis of the level of difficulty of the questions (Item Measurement) was conducted. According to the Rasch principle, if a student's ability increases, the same questions will be measured as having a lower level of difficulty (logit) (the questions will feel easier). The data in Table 2 shows an interesting phenomenon.

Table 2. The mean logit of the pretest and posttest scores for the Computational Thinking skill items in the experimental and control classes.

Class	Number of Students	Number of Questions	Mean/SE (logit) Item Computational Thinking		
			Pre-test	Post-test	Before and After Changes
Experimental	33	10	0.78/ (0.17)	-1.13/ (0.23)	-1.91
Control	32	10	1.03/ (0.18)	-0.67/ (0.19)	-1.71

In the experimental class, there was a drastic decrease in the difficulty level of the questions, from 0.78 logits (in the pretest) to -1.13 logits (in the posttest). This extreme change of -1.91 logits indicates that after learning using Bing Chat-assisted analogies, students found the CT questions much easier to solve. A more detailed analysis was conducted to determine which aspects/indicators students found most helpful after the intervention. Next, item difficulty level category can be seen in Table 3.

Table 3. Item difficulty level category

Range Division Reference	Renge Logit Item	Category	Question Item Code
> + 1 SD	> 0.30	Very difficult	-
0.0 logit s/d + 1SD	0.0 to 0.30	Difficult	2PR, 1DC, 4L, 9DC, 10AB, 8AL
-1 SD to 0.0 logit	-0.30 to 0.0	Easy	6PR, 5DC, 7AL
< -1 SD	< -0.30	Very easy	3AB

Based on Table 3, the results of the difficulty level test on 10 Computational Thinking ability test questions show that there are 6 questions in the difficult category, 3 questions in the easy category, and 1 question in the very easy category. Next, Table 4 shown a pre-test and post-test items for computational thinking indicators

Table 4. Changes in pre-test and post-test items for computational thinking indicators

Indicator	Experimental Class			Control Class		
	Pre	Post	Difference	Pre	Post	Difference
Decomposition	0.86	-1.11	-1.97	1.03	-0.66	-1.69
Pattern Recognition	0.94	-1.09	-2.03	1.20	-0.79	-1.99
Abstraction	0.70	-0.99	-1.69	0.79	-0.12	-0.91
Algorithm	0.62	-1.27	-1.89	1.05	-0.61	-1.66
Average	0.78	-1.12	-1.90	1.02	-0.54	-1.56

Referring to Table 4, all CT indicators in the experimental class experienced a significant decrease in logit difficulty (mean decrease of -1.90 logit), which was better than the control class (mean decrease of -1.56 logit). The largest decrease in difficulty in the experimental class occurred in the Pattern Recognition (decrease of 2.03 logit) and Decomposition (decrease of 1.97 logit). These findings indicate that Bing Chat assistance is very effective in helping students identify problem patterns and decompose complex problems into simpler components. Cognitively, Bing Chat supports pattern recognition by providing a variety of contextual analogies and diverse representations of renewable energy concepts. When students compare the various sources of analogies generated by Bing Chat with the target physics concept, they engage in a similarity mapping process that helps them recognize recurring structures, relationships, and cause-and-effect patterns across situations. This process reduces students' cognitive load in identifying abstract relationships because the AI-generated analogies provide immediate support and adaptive explanations. As a result, students become better able to systematically recognize conceptual patterns, which strengthens the pattern recognition dimension of computational thinking.

To visualize the differences in treatment impact more comprehensively between the two classes, a comparison of the magnitude of the logit reduction for each indicator is presented in Figure 1.

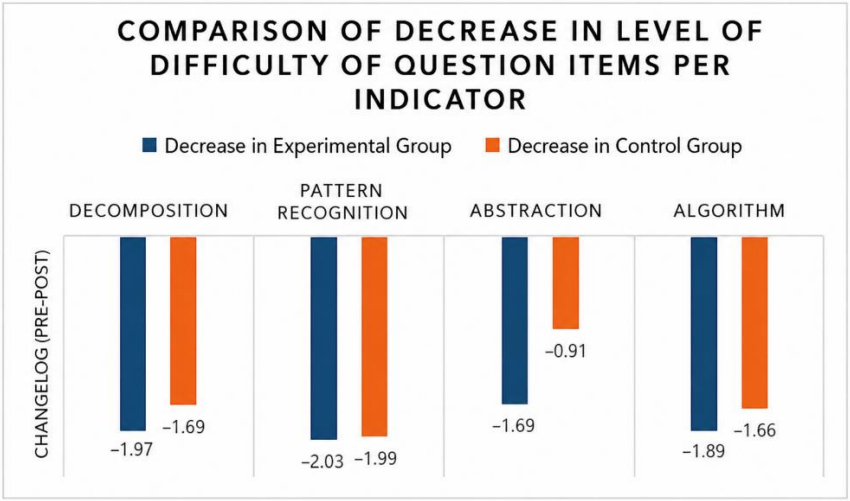


Figure 1. Comparison of the Level of Decrease in Question Item Difficulty (Question Item Size) per Indicator between the Experimental Class and the Control Class.

Based on the visualization in Figure 1, it is clear that the experimental class (blue bars) consistently outperformed the control class (orange bars) in terms of the magnitude of the decline across all indicators. In addition to the aforementioned decomposition and pattern recognition aspects, the graph also highlights a significant effectiveness gap in the Abstraction indicator. In this indicator, the experimental class was able to record a decrease in difficulty of -1.69 logits, far exceeding the control class which only achieved -0.91 logits. This confirms that the integration of Bing Chat in the TWA model is superior in helping students build an understanding of abstract

concepts, which is usually the weakest point in conventional learning. Next, wright map analysis can be seen in Figure 2.

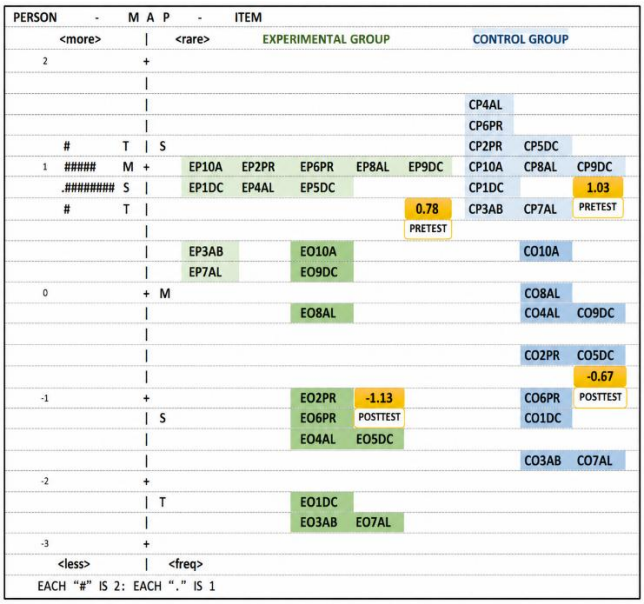


Figure 2. Wright map analysis

The Wright Map above shows the distribution of students in the experimental and control classes based on Rasch analysis on the same logit scale. The left side of the map displays the distribution of student abilities, while the right side displays the distribution of grades. The Wright Map results indicate that after the treatment (posttest), student abilities in the experimental class improved significantly compared to those in the control class.

3.3. Prerequisite test analysis

Before testing the hypothesis about the effectiveness of the learning model, a prerequisite test was carried out to ensure that the data met the assumptions of parametric statistics, namely that it was normally distributed and had homogeneous variance.

3.3.1. Normality Test

Normality testing was performed using the Kolmogorov-Smirnov technique with the aid of JAMOVI software. The testing criteria state that the data is normally distributed if the significance value (p) is greater than 0.05 ($p > 0.05$). The results of the normality test for the pretest and posttest data are presented in Table 5.

Table 5. Normality test

Kolmogorov Smirnov	Pre-test	Post-test
P	0.215	0.428
A	0.05	0.05
Decision	Normally Distributed	Normally Distributed

Based on Table 5, the significance value (p) for the pretest data is 0.215 and for the posttest data is 0.428. Since both p-values are greater than the 0.05 significance level, it can be concluded that the Computational Thinking ability data in both classes are normally distributed.

3.3.2. Homogeneity Test

After ensuring that the data were normally distributed, a homogeneity test was performed using Levene's test to determine the equality of variance between groups. The test results are presented in Table 6.

Table 6. Homogeneity Results *of Pretest* and *Posttest*

Levene's	Pre-test	Post-test

P	0.187	0.805
A	0.05	0.05
Decision	Homogeneous	Homogeneous

Based on the results in Table 6 above, it can be seen that both data sets are homogeneous. This can be concluded from the Levene values for both data sets. For the pretest results, the Levene value with a p value of 0.187, which means >0.05, indicating that the pretest results are homogeneous. For the posttest results, the Levene value with a p value of 0.805, which means >0.05, indicating that the posttest results are also homogeneous.

3.4. Hypothesis Testing

After ensuring that the data met the assumptions of normality and homogeneity, the next step was to test the hypotheses to answer the research questions. The statistical test used was the Independent Samples T-Test using JAMOVI software. This test focused on post-test data to determine whether there were significant differences in Computational Thinking skills between students who used the Bing Chat-assisted TWA model (Experimental Class) and students who used the conventional model (Control Class) after the treatment was completed (Love et al., 2022; Rohman et al., 2021). A summary of the hypothesis test results is presented in Table 7.

Table 7. Hypothesis testing post-test

	Post-test (Independent T-test)
p (Significance)	0.001
A	0.05
Decision	H o is rejected H 1 is accepted

Based on Table 7, the results of the posttest data analysis show a significance value (p) of 0.001. Because the p value <0.05, the test decision is H_0 is rejected and H_1 is accepted. This statistically proves that there is a significant difference in the average ability of Computational Thinking of students between the experimental class and the control class. In other words, the application of Bing Chat in the Teaching by Analogy (TWA) model has a real and significant influence on students' Computational Thinking abilities on renewable energy material compared to the conventional learning model.

In Rasch measurement, higher individual logit scores indicate greater student proficiency in completing computational thinking tasks. Posttest results showed that students in the experimental class scored above the average item difficulty level, indicating that most students were able to successfully complete the assigned computational thinking tasks after the intervention. Furthermore, the relatively small standard error values indicate that the individual measurements are quite stable and consistent. These findings suggest that the Bing Chat-assisted TWA model not only improved students' average scores but also improved the accuracy and consistency of students' computational thinking performance.

3.5. Discussion

This study aims to determine the effect of Bing Chat implementation in the Teaching by Analogy model on student abilities. This study was conducted at Yasmida Ambarawa High School with 65 students, focusing on physics and renewable energy subjects. The Teaching by Analogy (TWA) model was applied to the experimental class, class X.1, while the conventional learning model was applied to the control class, class X.2.

The Teaching by Analogy (TWA) model, by implementing Bing Chat, can increase student learning motivation. This is evident from students' positive responses to the use of Bing Chat in learning. Students feel more challenged and motivated to complete the assigned tasks. This finding is supported by research (Diani et al., 2018; Hoerunnisa & Fauziah, 2023) , which shows that the use of technology in learning can be developed in various forms and plays a role in improving student learning outcomes. This shows a positive influence, with significant results on Computational Thinking skills. This is evident from the average logit results of students' pretest and posttest in the experimental and control classes. The average logit score for the pretest was 0.21 and the average logit score for the posttest was 1.90 in the experimental class, with an average logit change of 1.69. Meanwhile, in the control class, the average logit score for the pretest was -0.04 and the average logit score for the posttest was 1.56, indicating an average logit change of 1.60. Based on the data, it

appears that the experimental class experienced a higher change in scores than the control class. It can be concluded that the use of the Teaching With Analogies (TWA) model, which implements the use of Bing Chat, shows a systematic increase in computational thinking skills. This increase is in line with research findings (Azmi, 2017; Diani, Yuberti, et al., 2016; Rudianto et al., 2022) , which state that analogy is a learning strategy that can systematically improve mathematical and problem-solving skills.

This indicates that the experimental class was able to analyze each indicator in the Computational Thinking ability test more easily after being given a learning treatment using the Teaching With Analogies (TWA) model by implementing Bing Chat in it. This is compared to the control class that used conventional learning. This finding is also supported by research by Lutfi Azizatul Laila and Rini Verary Shanthi, which revealed that the application of Computational Thinking skills in learning can improve students' analytical and logical abilities (Maskur et al., 2021) . This is because analogies are used in learning so that students will more easily visualize how two different things can have similarities (Diani et al., 2019; Mawarni, 2019; Nursafitri et al., 2021) . Therefore, an interactive and responsive learning environment is created with the Teaching With Analogies (TWA) model by implementing the use of Bing Chat in it.

This success can be explained by TWA's function in helping students understand abstract concepts through relevant analogies. Analogies act as cognitive bridges that facilitate the processes of decomposition, pattern recognition, and algorithm construction. Bing Chat engagement enriches students' analogies, provides additional explanations, and allows students to ask questions independently when encountering difficulties. This interaction encourages students to explore physics concepts in greater depth, thus contributing directly to improved logit scores across all CT indicators. When compared with previous research, these findings are consistent with studies (Harrison & Treagust, 2006; Richland & Simms, 2015; Yunita et al., 2024) which state that the use of analogies strengthens conceptual understanding and scientific reasoning. Similarly, the results of this study align with studies (Hoerunnisa & Fauziah, 2023) which found that the use of educational technology can improve motivation and learning outcomes. Studies by (Setiawan et al., 2023) and (Maskur et al., 2021) also show that learning approaches that encourage independent exploration can enhance aspects of CT such as algorithmic thinking and pattern recognition. Thus, this research strengthens empirical evidence that combining analogy strategies and interactive technology can support the development of students' computational skills.

Although the results of this study show significant improvements, there are several limitations that need to be considered. First, the difference in initial abilities between the experimental and control classes was quite significant, as reflected in the pretest results. This condition may influence the interpretation of the magnitude of the treatment effect, even though the statistical analysis still shows a significant difference. Second, the use of Bing Chat is highly dependent on students' internet connection and digital skills, so learning experiences may vary across students. Third, this study used a limited sample size and was conducted in only one school, so the findings should be generalized with caution. Furthermore, the relatively short duration of the intervention and its application to only one physics topic do not fully guarantee the long-term retention of students' computational thinking skills after intensive Bing Chat use is discontinued. (Diani, Saregar, et al., 2016) . The improvements observed in this study may still reflect short-term cognitive support facilitated by AI-assisted interactions, rather than the stable and sustained development of computational thinking. Therefore, longitudinal studies are needed to examine whether students are able to maintain abstraction, pattern recognition, and algorithmic thinking skills independently over a longer period of time and across different physics topics.

This study provides three main findings. First, it demonstrates that combining analogies and AI-based technology is effective in enhancing critical thinking (CT), particularly in renewable energy materials that require an understanding of complex systems. Second, it presents a hybrid learning model that integrates Technology-Based Analysis (TWA) and chatbots, an approach that has not been widely studied in the physics education literature. Third, it opens up opportunities for further research to explore how variations in analogies, interaction patterns with chatbots, and students' cognitive profiles may influence the effectiveness of CT-based learning. For future research, several strategic directions need to be considered. First, future studies should employ longitudinal designs or delayed posttests to examine the long-term retention of computational thinking skills after AI

assistance is reduced or removed. Second, further research is needed to investigate the extent to which students become cognitively dependent on AI-generated explanations during analogy-based learning. Third, future studies could analyze the quality and characteristics of AI-generated analogies to determine which types of analogy representations are most effective in supporting abstraction and pattern recognition processes. Furthermore, a broader sample involving different schools, educational levels, and physics topics is needed to increase the generalizability of the findings. A qualitative approach focusing on student-AI interaction patterns can also provide deeper insights into how AI-assisted dialogue facilitates the development of computational thinking in physics learning.

4. Conclusion

Based on the research results, it can be concluded that the Teaching With Analogies (TWA) model, which implements the use of Bing Chat, is effective in improving the computational thinking skills of 10th-grade students at Yasmida Ambarawa High School in physics lessons on the theme of renewable energy. This result can be proven by the results of the hypothesis test which states that H0 is rejected and H1 is accepted, meaning there is a difference in students' computational thinking skills before and after the treatment.

Furthermore, from the pretest and posttest results of students in the experimental class, the average pretest score was 0.21 logit and the average posttest score was 1.90 logit, with a change of 1.69 logit. Meanwhile, in the control class, the average pretest score was -0.04 logit and the average posttest score was 1.56 logit, indicating a change of 1.60 logit. In this case, students in both classes showed positive changes in their abilities. However, the mastery of questions in the experimental class was far superior to that in the control class. In addition, it can also be seen that the average change in the experimental class was much greater than that in the control class. This indicates that students in the experimental class put in greater effort than students in the control class. Thus, it can be concluded that the use of the Teaching With Analogies (TWA) model by implementing the use of Bing Chat has an effect on students' computational thinking abilities.

Author Contributions

All authors have equal contributions to the paper. All the authors have read and approved the final manuscript.

Funding

No funding support was received.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Ali, A., Maniboey, L. C., Megawati, R., Djarwo, C. F., & Listiani, H. (2024). *Interactive learning media: Comprehensive theory and development of interactive learning media in elementary schools.* PT Sonpedia Publishing Indonesia.

Azmi, M. P. (2017). Developing mathematical analogy skills. *Cendekia: Jurnal Pendidikan Matematika, 1*(1), 100–111.

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Springer.

Bower, M., Torrington, J., Lai, J. W. M., Petocz, P., & Alfano, M. (2024). How should we change teaching and assessment in response to increasingly powerful generative artificial intelligence? Outcomes of the ChatGPT teacher survey. *Education and Information Technologies, 29*(12), 15403–15439. https://doi.org/10.1007/s10639-024-12551-1

Deissenroth, M., Klein, M., Nienhaus, K., & Reeg, M. (2017). Assessing the plurality of actors and policy interactions: Agent-based modeling of renewable energy market integration. *Complexity,* 2017, Article 7494313. https://doi.org/10.1155/2017/7494313

Diani, R., Asyhari, A., & Julia, O. N. (2018). The influence of the RMS (reading, mind mapping and sharing) model on students' higher-order thinking skills on the topic of impulse and momentum. *Edutama Education Journal, 5*(1), 31–44.

Diani, R., & Hartati, N. S. (2018). Islamic literacy-based flipbook: Development of physics learning media with professional 3D pageflip. *Jurnal Inovasi Pendidikan IPA, 4*(2), 234–244.

Diani, R., Herliantari, H., Irwandani, I., Saregar, A., & Umam, R. (2019). Search, solve, create, and share (SSCS) learning model: The impact on students' creative problem-solving ability on the concept of substance pressure. *Jurnal Penelitian & Pengembangan Pendidikan Fisika, 9*(1), 65–77. https://doi.org/10.21009/1.09107

Diani, R., Saregar, A., & Ifana, A. (2016). Comparison of problem-based learning and guided inquiry models on students' critical thinking skills. *Jurnal Penelitian Pembelajaran Fisika, 7*(2), 125–135.

Diani, R., Yuberti, Y., & Syafitri, S. (2016). Testing the effect size of the scramble learning model with video media on the physics learning outcomes of class X students at MAN 1 Pesisir Barat. *Jurnal Ilmiah Pendidikan Fisika Al-Biruni, 5*(2), 265–275. https://doi.org/10.24042/jpifalbiruni.v5i2.126

Duit, R., Roth, W.-M., Komorek, M., & Wilbers, J. (2001). Fostering conceptual change by analogies—Between Scylla and Charybdis. *Learning and Instruction, 11*(4–5), 283–303. https://doi.org/10.1016/S0959-4752(00)00034-7

Ekstrand, J., Westergren, A., Årestedt, K., Hellström, A., & Hagell, P. (2022). Transformation of Rasch model logits for enhanced interpretability. *BMC Medical Research Methodology, 22*, Article 145. https://doi.org/10.1186/s12874-022-01816-1

Fuadi, H., Robbia, A. Z., Jamaluddin, J., & Jufri, A. W. (2020). Analysis of factors causing low scientific literacy skills of students. *Jurnal Ilmiah Profesi Pendidikan, 5*(2), 108–116. https://doi.org/10.29303/jipp.v5i2.122

Glynn, P. W. (1991). Coral reef bleaching in the 1980s and possible connections with global warming. *Trends in Ecology & Evolution, 6*(6), 175–179. https://doi.org/10.1016/0169-5347(91)90056-J

Grover, S., & Pea, R. (2013). Computational thinking in K–12: A review of the state of the field. *Educational Researcher, 42*(1), 38–43. https://doi.org/10.3102/0013189X12463051

Harrison, A. G., & Treagust, D. F. (2006). Teaching and learning with analogies: Friend or foe? In P. J. Aubusson, A. G. Harrison, & S. M. Ritchie (Eds.), *Metaphor and analogy in science education* (pp. 11–24). Springer. https://doi.org/10.1007/1-4020-3830-5_2

Hoerunnisa, M., & Fauziah, S. R. (2023). Utilization of technology in science learning as an effort to improve learning outcomes. *Journal of Science Education Studies, 3*(2), 272–278.

Kim, T. K., & Park, J. (2019). More about the basic assumptions of t-test: Normality and sample size. *Korean Journal of Anesthesiology, 72*(4), 331–335. https://doi.org/10.4097/kja.d.18.00292

Klopfer, E., Osterweil, S., & Salen, K. (2009). *Moving learning games forward: Obstacles, opportunities & openness*. The Education Arcade, Massachusetts Institute of Technology.

Kumar, S., Kumar, V., Chaudhuri, R., Chatterjee, S., & Vrontis, D. (2025). AI capability and environmental sustainability performance: Moderating role of green knowledge management. *Technological Forecasting and Social Change, 81*, Article 102870. https://doi.org/10.1016/j.techsoc.2025.102870

Lai, J. W. M., & Bower, M. (2020). Evaluation of technology use in education: Findings from a critical analysis of systematic literature reviews. *Journal of Computer Assisted Learning, 36*(3), 241–259. https://doi.org/10.1111/jcal.12391

Laliyo, L. A. R., Sumintono, B., & Panigoro, C. (2022). Measuring changes in hydrolysis concept of students taught by inquiry model: Stacking and racking analysis techniques in Rasch model. *Heliyon, 8*(3), Article e09001. https://doi.org/10.1016/j.heliyon.2022.e09001

Leung, Y.-Y., Thumboo, J., Yeo, S., Wylde, V., & Tannant, A. (2022). Validation and interval scale transformation of the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) in patients undergoing knee arthroplasty, using the Rasch model. *Osteoarthritis and Cartilage Open, 4*, Article 100322. https://doi.org/10.1016/j.ocarto.2022.100322

Linacre, J. (2009). Unidimensional models in a multidimensional world. *Rasch Measurement Transactions, 23*(3), 1224–1227.

Linacre, J. M. (2012). *Winsteps® Rasch measurement computer program user's guide*. Winsteps.com.

Maskur, M. M., Anwar, M. K., & Trianah, T. T. (2021). Implementation of blended learning in MI. *MAGISTRA: Media Pengembangan Ilmu Pendidikan Dasar dan Keislaman, 12*(2), 120–133. https://doi.org/10.31942/mgs.v12i2.5506

Mawarni, T. A. T. (2019). Mathematical analogy reasoning ability through hands-on mathematics learning model with concrete object media reviewed from visualizer and verbalizer cognitive styles. *Jurnal Penelitian Pendidikan dan Pengajaran Matematika, 5*(1), 1–10.

Mulyono, M., Hayat, M. S., & Roshayanti, F. (2025). Analysis of junior high school students' computational thinking and science reasoning abilities on motion and force. *Didaktika: Jurnal Kependidikan, 14*(3), 5827–5838.

Nursafitri, M., Santoso, A., & Sumari, S. (2021). *The effect of the 5E learning cycle learning model with analogy on students' conceptual understanding of reaction rate material* [Master's thesis, Universitas Negeri Malang].

Planinic, M., Boone, W. J., Sušac, A., & Ivanjek, L. (2019). Rasch analysis in physics education research: Why measurement matters. *Physical Review Physics Education Research, 15*(2), Article 020111. https://doi.org/10.1103/PhysRevPhysEducRes.15.020111

Rahman, N. F., Yasiron, A. M., Roziqi, M. A., & Ilmiani, A. M. (2019). Computer assisted language learning. *Computer Assisted Language Learning, 32*(8), 827–877.

Richland, L. E., & Simms, N. (2015). Analogy, higher order thinking, and education. *Wiley Interdisciplinary Reviews: Cognitive Science, 6*(2), 177–192. https://doi.org/10.1002/wcs.1336

Rudianto, R., Diani, R., Subandi, S., & Widiawati, N. (2022). Development of assessment instruments 4C skills (critical thinking, collaboration, communication, and creativity) on parabolic motion materials. *Journal of Advanced Sciences and Mathematics Education, 2*(2), 65–79.

Saregar, A., Irwandani, I., Abdurrahman, A., Parmin, P., Septiana, S., Diani, R., & Sagala, R. (2018). Temperature and heat learning through SSCS model with scaffolding: Impact on students' critical thinking ability. *Journal for the Education of Gifted Young Scientists, 6*(3), 39–54. https://doi.org/10.17478/JEGYS.2018.80

Saregar, A., Putra, F. G., Diani, R., Anugrah, A., Misbah, M., & Umam, R. (2025). Innovative integrated disaster education in physics learning: An effort to enhance students' disaster literacy skills. *Indonesian Journal of Science Education, 14*(2).

Setiawan, A., Widyasari, M., & Aprinastuti, C. (2023). Application of computational thinking with simple experiments on temperature and heat material for grade V. *PEDADIDAKTIKA: Jurnal Ilmiah Pendidikan Guru Sekolah Dasar, 10*(3), 392–403.

Shute, V. J., Sun, C., & Asbell-Clarke, J. (2017). Demystifying computational thinking. *Educational Research Review, 22*, 142–158. https://doi.org/10.1016/j.edurev.2017.09.003

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan Rasch pada assessment pendidikan*. Trim Komunikata.

Susskind, J. (2018). *Future politics: Living together in a world transformed by technology*. Oxford University Press.

Tesio, L., Caronni, A., Kumbhare, D., & Scarano, S. (2023). Interpreting results from Rasch analysis 1. The "most likely" measures coming from the model. *Disability and Rehabilitation, 46*, 591–603. https://doi.org/10.1080/09638288.2023.2169771

Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., & Wilensky, U. (2016). Defining computational thinking for mathematics and science classrooms. *Journal of Science Education and Technology, 25*(1), 127–147. https://doi.org/10.1007/s10956-015-9581-5

Wing, J. M. (2006). Computational thinking. *Communications of the ACM, 49*(3), 33–35. https://doi.org/10.1145/1118178.1118215

Wulandari, D. T., Diani, R., & Noperi, H. (2025). Development of diagnostic instruments for critical thinking skills of students in physics learning on socio-scientific issues (SSI). *Journal Corner of Education, Linguistics, and Literature, 5*(1), 94–112.

Yunita, Y., Diani, R., Noperi, H., & Velina, Y. (2024). Development of interactive e-modules with STEM approach using Canva application to improve students' critical thinking ability in physics learning. *JINoP (Jurnal Inovasi Pembelajaran), 10*(2), 192–210.