Universida de Vigo

Revista de Investigación en Educación, 2026, 24(2), 219-250
DOI: https://doi.org/10.35869/reined.v24i2.6938
https://revistas.webs.uvigo.es/index.php/reined
ISSN 1697-5200 | e-ISSN 2172-3427

Virtual laboratories with interactive simulations and computational thinking to enhance higher-order thinking skills

Laboratorios virtuales con simulaciones interactivas y pensamiento computacional para mejorar las habilidades de pensamiento de orden superior

Didik Setyawarno[1], Mohamad Termizi Borhan[2], Maryati[3], Dita Puji Rahayu[4]

[1] Universitas Negeri Yogyakarta *didiksetyawarno@uny.ac.id*
[2] Universiti Pendidikan Sultan Idris *termizi@fsmt.upsi.edu.my*
[3] Universitas Negeri Yogyakarta *maryati@uny.ac.id*
[4] Universitas Negeri Yogyakarta *ditapuji@uny.ac.id*

Received: 2/10/2025
Accepted: 22/4/2026

Contact address:
Didik Setyawarno
Department of Science Education,
Faculty of Mathematics and Natural Sciences
Universitas Negeri Yogyakarta.Colombo Street No. 1 Yogyakarta, Indonesia.

Abstract

Virtual laboratories with interactive simulations offer an innovative solution for science learning in the digital era by supporting students' understanding of complex, abstract concepts. Integrating computational thinking (CT) skills, decomposition, pattern recognition, abstraction, and algorithms, within virtual laboratories further strengthens higher-order thinking skills (HOTs). This study investigated the effect of Virtual Laboratories with Interactive Simulations and Computational Thinking (VL-ISCT) on HOTs. A quasi-experimental design with the matching pretest-posttest control group was used, involving 90 pre-service science teachers from Universitas Negeri Yogyakarta. Participants were divided into three classes: Experiment Class 1 (four CT skills), Experiment Class 2 (three CT skills), and Control Class (two CT skills). Instruments included validation and observation sheets, and pretest–posttest questions, which were analyzed using the Content Validity Ratio, Fleiss's Kappa, and the Rasch Model. The results showed that VL-ISCT significantly improved HOTs. Analysis of Variance ($F = 56{,}5$, $p < 0{,}001$, $\eta^2 = {,}565$) confirmed significant group differences, with post hoc tests indicating the strongest effect in Experiment Class 1. These findings highlight VL-ISCT as an innovative approach in science education, offering a meaningful model for strengthening HOTs and preparing future teachers for more effective, contemporary classroom practice.

Key Words

Virtual Laboratories, Interactive Simulations, Computational Thinking, Science Learning, Higher Order Thinking Skills

Resumen

Los laboratorios virtuales con simulaciones interactivas y pensamiento computacional (LV-SIPC) ofrecen una solución innovadora para el aprendizaje de las ciencias en la era digital, al apoyar la comprensión de conceptos complejos y abstractos por parte de los estudiantes. La integración de habilidades de pensamiento computacional (PC), descomposición, reconocimiento de patrones, abstracción y algoritmos, dentro de LV-SIPC fortalece aún más las habilidades de pensamiento de orden superior (HPOS). Este estudio investigó el efecto de LV-SIPC en las HPOS. Se aplicó un diseño cuasiexperimental con un grupo de control con pretest y postest, en el que participaron 90 futuros docentes de ciencias de la Universidad Estatal de Yogyakarta. Los participantes se dividieron en tres clases: Clase Experimental 1 (cuatro habilidades de PC), Clase Experimental 2 (tres habilidades de PC) y Clase de Control (dos habilidades de PC). Los instrumentos incluyeron hojas de validación, hojas de observación y preguntas de pretest y postest, analizadas mediante el Índice de Validez de Contenido, el coeficiente Kappa de Fleiss y el modelo de Rasch. Los resultados mostraron que LV-SIPC mejoró significativamente las HPOS. El análisis de varianza ($F = 56{,}5$, $p < {,}001$, $\eta^2 = {,}565$) confirmó diferencias significativas entre los grupos, y las pruebas post hoc indicaron el efecto más fuerte en la clase experimental 1. Estos hallazgos resaltan LV-SIPC como un enfoque innovador en la enseñanza de las ciencias, que ofrece un modelo significativo para fortalecer las HPOS y preparar a los futuros docentes para una práctica en el aula más eficaz y contemporánea.

Palabras clave

Laboratorios Virtuales, Simulaciones Interactivas, Pensamiento Computacional, Aprendizaje de Ciencias, Habilidades de Pensamiento de Orden Superior

1. INTRODUCTION

Conventional, non-digital science learning does not meet current educational demands, especially for abstract concepts. Traditional lectures, labs, and textbooks cannot fully address students' varied abilities and learning styles (Abdulrahaman et al., 2020; Shaidullina et al., 2023). Science learning in the 21st century demands higher-order thinking, yet students' skills remain suboptimal in both schools and higher education (Kareem, 2022; Yanti & Thohir, 2024). One approach to improving higher-order thinking skills is to integrate computational thinking (CT) into science teaching, supported by virtual laboratories with interactive simulations, especially in abstract science materials (Angeli & Giannakos, 2020; Chevalier et al., 2020). CT develops logical, analytical, and systematic thinking to tackle complex scientific problems.

In science learning, students engage in observing, experimenting (physically or virtually), and analyzing data as part of the scientific process (Carin & Sund, 1989; Chiappetta & Koballa, 2010). Scientific work aligns with CT because it involves decomposition, pattern recognition, abstraction, and algorithmic problem-solving (Haines et al., 2019; Sneider et al., 2014). CT aligns with how computers solve problems in computing and data science, which are very important in science learning, but many science learners today have not integrated CT into the scientific process (Irvani et al., 2024; Misir, 2023). Although CT has been recognized as an essential skill in 21st-century education, its implementation in science learning, both in schools and in college, remains suboptimal (Abidin et al., 2023; Hidayat et al., 2024; Threekunprapa & Yasri, 2020).

Several factors limit CT integration in science learning, including teachers' limited understanding and resources, and the utilization of digital technology (Angeli & Giannakos, 2020; Shute et al., 2017). CT can comprehensively support science thinking, investigation, knowledge, and technology interaction, emphasizing immersive digital labs as essential in modern education, as shown in Figure 1. This is a strong reason why science learning in this era can integrate immersive technologies, including virtual laboratory activities with interactive simulations (Chiappetta & Koballa, 2010).

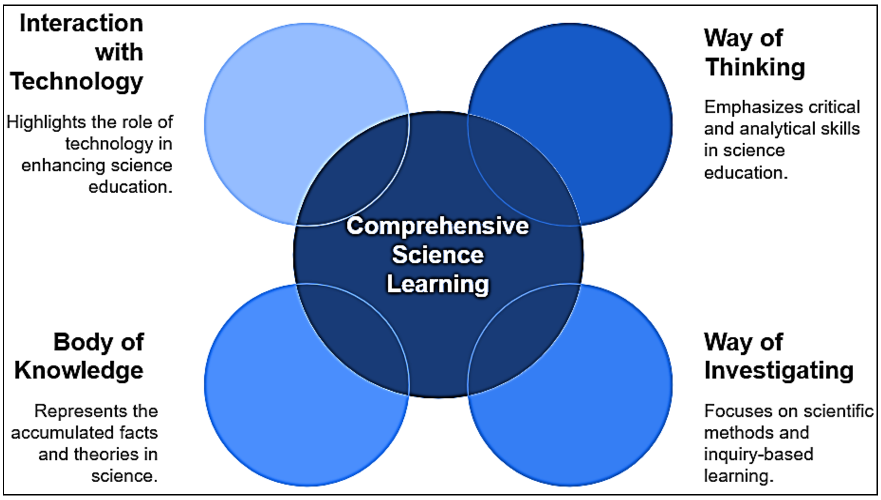


Figure 1. Comprehensive science domain

Digital-era science learning can be enhanced with immersive technology that supports CT habituation and higher-order thinking by aligning with scientific processes and domains. Immersive technology in the context of this research is through virtual laboratory activities and interactive simulation and computational thinking (VL-ISCT) which is able to facilitate scientific investigation in a more structured, safe, and able to develop a critical and creative way of scientific thinking (Angeli & Giannakos, 2020; Carin & Sund, 1989; Contant et al., 2018). The limited use of virtual labs and CT in science learning highlights the need for innovation, such as integrating VL-ISCT into discovery-based models aligned with the Technological Pedagogical Content Knowledge (TPACK) framework.

Discovery learning is one of the innovative learning models that encourages students to be active in scientific work, both hands-on and mind-on, in the exploration and discovery of concepts (Artiningsih & Nurohman, 2020; Carin & Sund, 1989; Chiappetta & Koballa, 2010). This model has been shown to positively influence learners, as evidenced by increased critical thinking and problem-solving skills, and a deeper understanding of the material being studied (Alfieri et al., 2011; Mayer, 2004). One problem in science learning today is the limited integration of discovery learning with VL-ISCT, especially for abstract concepts such as electrical materials (Abidin et al., 2023). Various studies show that integrating CT into learning can help students develop stronger analytical, logical, and innovative skills (Barr & Stephenson, 2011; Lye & Koh, 2014; Man et al., 2020). However, discovery learning implementations are often not designed to explicitly integrate CT into the learning process (Chevalier et al., 2020; Kirschner et al., 2006). Another challenge was also found for science educators, namely, how to teach abstract science concepts in virtual laboratory activities.

The use of virtual laboratories has been proven to provide various advantages in science learning (Deshmukh et al., 2020; Veza et al., 2022). Some of them are increased conceptual understanding, time and cost efficiency, greater implementation flexibility, and increased student motivation to learn. In addition, virtual laboratories are ideal for presenting abstract materials that are difficult to visualize directly in a real laboratory, such as atomic processes, astronomical phenomena, or energy changes in chemical reactions (Mayer, 2004; Zacharia & Olympiou, 2011). This study aims to examine the impact of VL-ISCT on pre-service science teachers' HOTs contributing to adaptive teacher education and promoting TPACK-based science learning in the digital era. The questions from this study are as follows.

- What is the role of a virtual laboratory with interactive simulation (VL-IS) in science learning?
- How is CT integrated into science learning?
- What are the results of the implementation of VL-ISCT in science learning for HOTs in pre-service science teachers?

This study addresses a critical gap in science education by integrating virtual laboratories, interactive simulations, and computational thinking into an instructional approach to enhance HOTs. While previous studies have examined these components separately, limited research has explored their combined impact, particularly in the context of pre-service teacher education. Therefore, this study contributes to both research and educational innovation by proposing and empirically testing a novel framework that supports more interactive, inquiry-based, and technology-enhanced learning environments. The findings are expected to provide practical insights for designing innovative science instruction and to strengthen the preparation of future teachers in developing advanced thinking skills required in contemporary classrooms.

2. LITERATURE REVIEW

2.1. Virtual laboratory with interactive simulation

A virtual laboratory (VL) is defined as a computer-based environment, whether connected to an internet network or not, that allows learners to conduct simulative experiments in the form of animation and digital interaction to observe and understand complex and abstract scientific concepts (Carnevali & ButtazzQ, 2003; Ma & Nickerson, 2006). Interactive simulations in virtual labs allow learners to change variables, run experiments, and observe results in real time, making learning more contextual and engaging (Montoya et al., 2023; Smetana & Bell, 2012; Zacharia & Olympiou, 2011). This highlights immersive technology and virtual laboratory with interactive simulations as a science-learning innovation that supports scientific thinking and investigation through digital experimentation, enabling students to build a body of knowledge and relate to their interactions with technology, as shown in Figure 2.

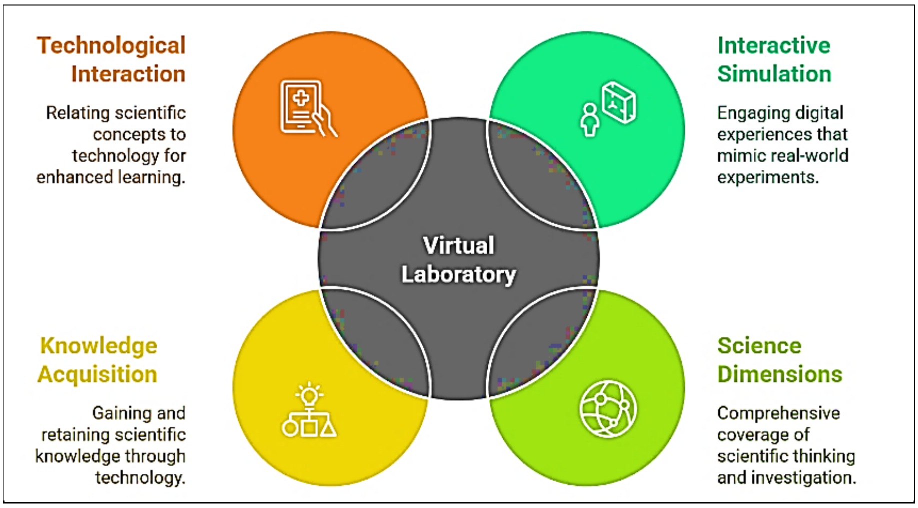


Figure 2. Revolutionizing science education with virtual laboratory

Essential components of a virtual lab include an interactive interface, dynamic visual simulations, and systematic experimental instructions. These components are designed to facilitate students' conceptual understanding through discovery-based learning (Herga et al., 2014; Rutten et al., 2012). The intuitive, interactive interface helps learners navigate the experiment with ease, while simulations allow visualization of abstract concepts such as chemical reactions, gravitational forces, and electrical phenomena that cannot be observed directly (Makransky et al., 2019; Parong & Mayer, 2018). Effective virtual lab implementation requires a systematic approach in which students actively simulate experiments, record data, analyze results, and make conclusions.

2.2. Computational thinking skills

Computational thinking (CT) is a concept discussed in the context of computer science since the 1960s (Denning, 2009). CT stems from a programmer's ability to solve problems in the computer world. Further, Wing (2006) defines CT as problem-solving skills, system design, and the understanding of human behavior using computer science concepts. This aligns with the statement by Kong et al. (2018) that it is essential for learners to acquire CT skills to cultivate a generation capable of solving problems through creativity and technology. The explanation emphasizes that the main dimension of CT is problem-solving using the logic of the computer world. CT in computer science has four basic foundations for problem-solving: decomposition, pattern recognition, abstraction, and algorithms, as shown in Figure 3 (Aziz et al., 2023; Denning, 2009; Wing, 2006).

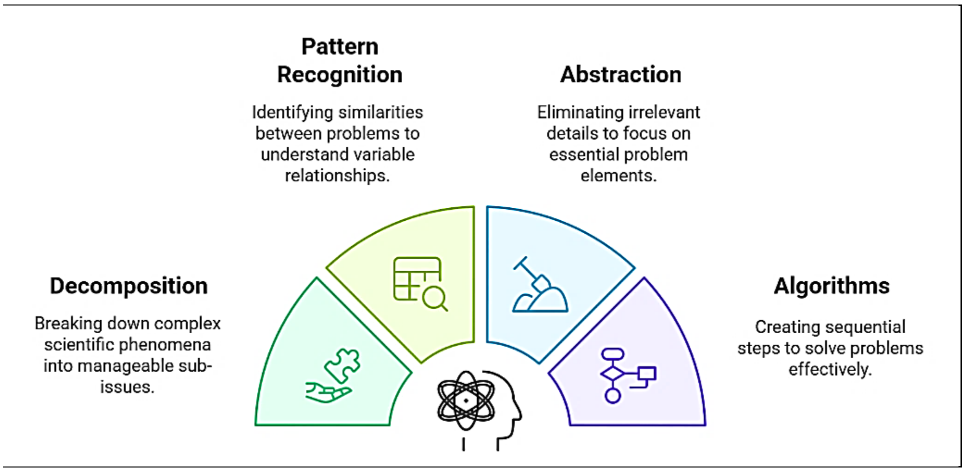


Figure 3. Foundations of computational thinking (CT) skills

CT foundations, decomposition, pattern recognition, abstraction, and algorithms help students grasp scientific phenomena from facts to laws. Decomposition is carried out by describing scientific phenomena into sub-problems (problem statements based on facts) to be studied (Grippo & Sciandrone, 2023; Rich et al., 2019). Pattern recognition is the observation or analysis of the similarities between problems, enabling the identification of relationships between variables (Jain et al., 2000; Maratos et al., 2023).

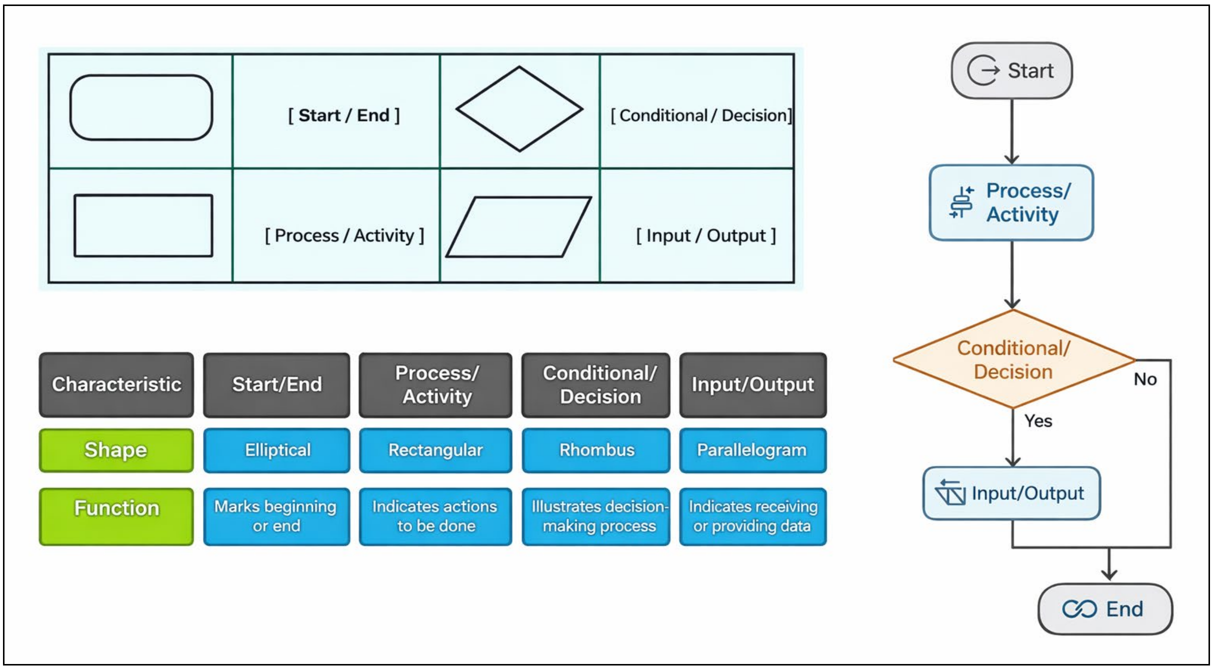


Figure 4. Basic symbols in the algorithm

These variables are consistently observed across experiments, yielding a general pattern that can be formulated as concepts or principles. Abstraction is the process of eliminating irrelevant parts of a problem (Mirolo et al., 2021; Yanti & Thohir, 2024). Abstraction creates problem-solving blueprints, revealing pure mathematical relationships between variables as theories or laws. Finally, algorithmic skills are the ordered steps for solving a problem that allow individuals to provide clear, followable instructions for a particular task (Cansu & Cansu, 2019; Sadykova & Il'bahtin, 2020).

The elements and symbols of the algorithm, as shown in Figure 4, are essentially computer programming that has been adapted to solve problems beyond the realm of computers (Ensmenger, 2016; Obilikwu et al., 2019). This study focuses on using CT with immersive virtual labs and simulations to help students systematically understand scientific phenomena, theories, laws, and their applications.

2.3. Science learning

Science is a collection of knowledge obtained through a systematic process of observation, experimentation, and logical reasoning to understand natural phenomena (Carin & Sund, 1989; Merchant, 2018). Science includes not only products (facts, concepts, principles, laws, and theories), but also scientific processes and attitudes that shape scientific thinking. Learning is the process of interaction among students, educators, and learning resources in a learning environment that aims to build knowledge, skills, and attitudes (Joyce et al., 2015; Kilbane & Natalie, 2014). In the context of science education, learning not only transfers knowledge but also emphasizes how students acquire, process, and test knowledge through scientific activities as in discovery learning.

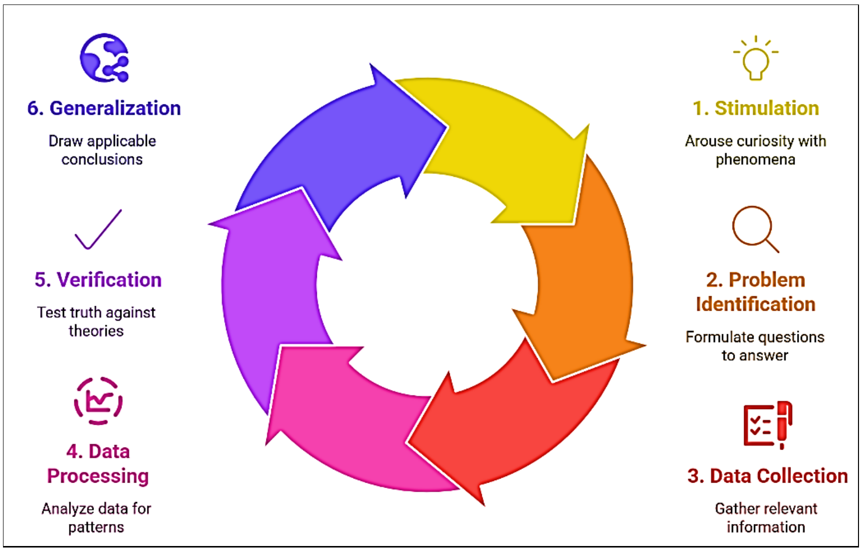


Figure 5. Stages of discovery learning

Science learning is an instructional process designed to develop an understanding of natural phenomena through a scientific approach, foster curiosity, and foster critical thinking skills and science process skills (Bybee, 2013). Science learning should engage students as active learners who observe, question, experiment, and conclude. Discovery learning, rooted in constructivism, supports this by guiding students to actively uncover concepts, principles, and laws through scientific processes. Jerome Bruner, one of the pioneers in this approach, stated that learning will be more effective if learners are directly involved in the process of discovery (Bruner, 1961). In this study, discovery learning aligns well with CT and promotes HOTs. Combined with immersive tech like virtual laboratory, it deepens students' understanding of complex science through computational logic and real visualization.

2.4. Higher-order thinking skills

A number of cognitive taxonomies have been developed in education, but the most widely used are Bloom's and Marzano's taxonomy (Anderson & Krathwohl, 2001; Marzano & Kendall, 2007). The cognitive domain encompasses intellectual skills, with Bloom's taxonomy dividing them into six levels: knowledge, comprehension, application, analysis, evaluation, and creation. The first three represent lower-order thinking, focused on recall and basic understanding. In contrast, the other three levels of Bloom's taxonomy require learners to use higher-order thinking skills, thus driving their learning performance (Brookhart, 2010; Forehand, 2010). Marzano's Taxonomy, developed by Marzano and Kendall, offers a more comprehensive framework than Bloom's, grounded in cognitive psychology and brain research to better represent thinking and learning processes. In contrast to Bloom, who compiled a cognitive hierarchy from low to high, Marzano organized his taxonomy based on the mental systems involved in learning, including emotional systems, self-awareness, cognitive control, and the application of knowledge, as shown in Figure 6 (Dubas & Toledo, 2016; Marzano & Kendall, 2007).

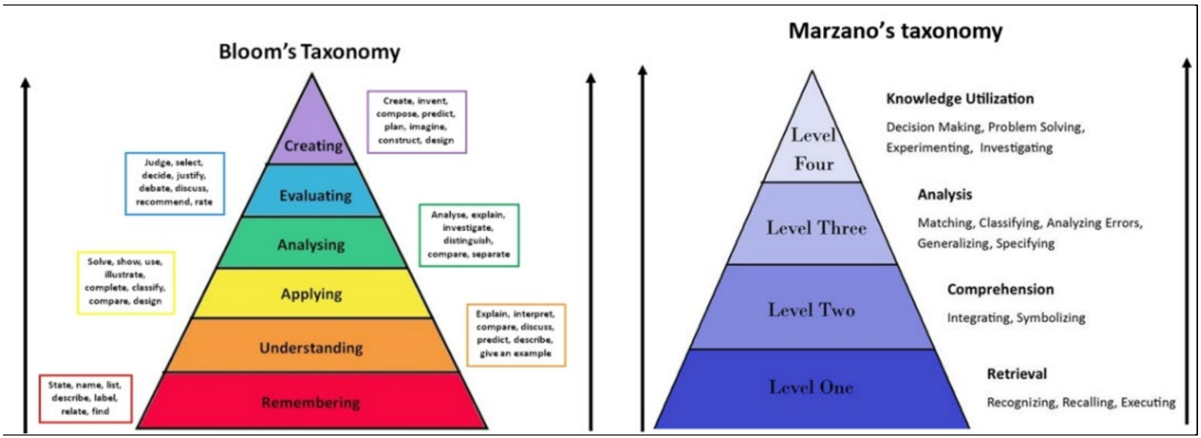


Figure 6. Bloom's taxonomy Vs. Marzano's taxonomy

In addition to Bloom and Marzano's taxonomy, the Brookhart approach defines higher-order thinking skills (HOTs) as the ability to transfer knowledge to new situations, think critically, and think creatively (Brookhart, 2010). Brookhart divides HOTS into three components: transfer (applying knowledge in new contexts), critical thinking (evaluating information and making evidence-based decisions), and creative thinking (generating innovative and original ideas). The dimensions of HOTs, as measured in this study, include the ability to analyze, evaluate, and create, as shown in Figure 7.

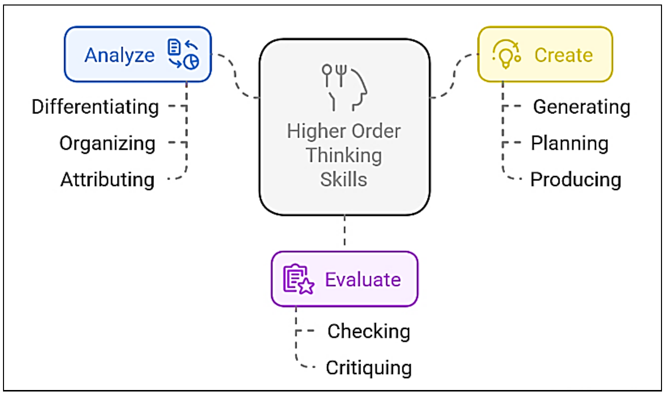


Figure 7. Dimension of higher-order thinking skills (HOTs)

2.5. Research objectives

This article examines the role of virtual laboratory with interactive simulations (VL-IS) integrated with CT in enhancing pre-service science teachers' HOTs. It also highlights the application of the Technological Pedagogical Content Knowledge (TPACK) framework using immersive VL-IS technology to simplify the learning of abstract science concepts.

3. METHOD

3.1. Study approach and design

The approach used in this study is quantitative with the matching-only pretest-posttest control group design (Bhattacherjee, 2012; Fraenkel et al., 2023; Jackson, 2009). This design can be used to assess the integration of computational thinking (CT) skills into discovery learning assisted by immersive technology on higher-order thinking skills (HOTs). The research method uses a quasi-experimental design with three classes: Experiment Class 1, Experiment Class 2, and Control Class, as shown in Figure 8.

Groups	Pretest	Treatment	Posttest
Experiment 1	O_1	X1	O2
Experiment 2	O_1	X2	O2
Control	O_1	X3	O2

Figure 8. Research design

Experiment Class 1 integrates four CT skills: decomposition, pattern recognition, abstraction, and algorithms; Experiment Class 2 integrates three CT skills: decomposition, pattern recognition, and abstraction; and Control Class integrates only two CT skills: decomposition and pattern recognition. The treatment was carried out in three meetings: meeting 1 with material on force and electric field; meeting 2 with material on the motion of charged particles in an electric field and electrical potential; and meeting 3 with capacitor material and capacitor series.

3.2. Study group

The research study group was determined using a multistage sampling technique, including cluster random sampling and purposive sampling (Bhattacherjee, 2012; Creswell, 2017; Fraenkel et al., 2023). In the first stage, cluster random sampling was applied to select classes at Universitas Negeri Yogyakarta that were accessible to the researcher and equipped with immersive technology. In the second stage, purposive sampling was conducted to select pre-service science teachers studying static electricity, proficient in computer-based learning, and who met the normal distribution of HOTs. Initially, 134 pre-service science teachers were included, but the final sample was reduced to 90 who met the normal distribution criteria, as presented in Table 1.

Class	Initial Number	After Reduction		Total
		Man	Woman	
Experiment Class 1 (Class A)	46	8	22	30
Experiment Class 2 (Class C)	46	5	25	30
Control Class (Class D)	42	7	23	30
	134			All participants 90

Table 1. Study group

The samples shown in Table 1 have met the statistical test, both the difference test (inferential) (Linacre, 1994; P. Louangrath, 2014; P. I. Louangrath, 2017). An adequate number of samples is an important requirement for inferential tests to provide reliable and valid results.

3.3. Research instruments

The research instruments used to collect data consisted of learning tool validation sheets, learning observation sheets, and question sheets (Canals, 2017). Learning tool validation sheets are used to ensure that teaching tools, including lesson plans, learning media, and student worksheets, meet the validity of the concept, content, and language (Froment et al., 2025; Setyawarno et al., 2025; Waugh & Gronlund, 2013). Observation sheets ensure proper implementation of science learning using the six-stage discovery learning model: stimulation, problem statement, data collection, data processing, verification, and generalization (Figure 5). Question sheets are expert-validated for construction, content, and language to accurately assess HOTs (Anderson & Krathwohl, 2001; Brookhart, 2010). The aspects of HOTs are measured as in Table 2.

The quality of all research instruments used was evaluated by five science education experts and analyzed for validity and reliability. Content Validity Ratio (CVR) is used to ensure that all instruments used meet the validity of constructs, content, and language (Gilbert & Prion, 2016; Lawshe, 1975).

Dimension	Sub dimension	Question Form	No. Item	Total
Analyze	Differentiating	Multiple Choice	B1, D1, G1	7
	Organizing	Multiple Choice	C1, F1	
	Attributing	Multiple Choice	A1, E1	
Evaluate	Checking	Description	A2, G2	3
	Critiquing	Description	E2,	
Create	Generating	Description	F2,	4
	Planning	Description	B2,	
	Producing	Description	C2, D2,	
			Total	14 Item

Table 2. Dimension of higher-order thinking skills (HOTs)

Content Validity Ratio (CVR) according to Lawshe (1975) using the following mathematical equations.

$$\mathrm{CVR} = \frac{\left(n_e - \frac{N}{2}\right)}{\frac{N}{2}}$$

Description:

ne: the number of validators who declare it valid by giving a score of 1

N : number of validator members or expert team members

After calculating the CVR for each aspect, the CVI (Content Validity Index) or average CVR is calculated. The instrument of the question set used is revealed valid if the CVI value is 0,99 (Lawshe, 1975). In addition, Fleiss's Kappa is used to assess the reliability of instruments based on expert ratings (inter-rater reliability), providing agreement information for categorical variables. Fleiss's Kappa provides information on the level of agreement among evaluators (McHugh, 2012; Orts-Cortés et al., 2013). Mathematical equations for calculating Fleiss's Kappa coefficients.

$$K = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)}$$

Cohen suggested interpreting the Kappa result as follows: values ≤ 0 indicate no agreement; 0,01–0,20 indicate none to slight; 0,21–0,40 indicate fair; 0,41–0,60 indicate moderate; 0,61–0,80 indicate substantial; and 0,81–1,00 indicate almost perfect agreement (McHugh, 2012). The question instruments used to measure HOTs were also empirically tested, including the construct validity and the fit of the Rasch model. This test was carried out on other classes that had studied static electricity, so its validity and empirical reliability were known. Principal Component Analysis (PCA) is used to empirically test the validity of constructs by analyzing the data's internal structure. Empirical tests with PCA to ensure that the question instrument measures higher-order thinking skills from three main dimensions: analyze, evaluate, and create, each of which consists of several subdimensions and question items. Some important requirements for the PCA test, as shown in Table 3 (Dimitrov, 2012; Greenacre et al., 2022).

Aspects	Conditions
Raw variance explained by measures	At least 20%
Unexplained variance (contrast)	Does not exceed 15%

Table 3. Interpretation of principal component analysis (PCA)

Determination of items fit the Rasch model using several criteria, as shown in Table 4 (Dimitrov, 2012; Linacre, 2002; Meyer, 2014).

Model Fit Criteria	Value Range
Outfit Mean Square (MNSQ)	0,5 < Outfit MNSQ < 1,5
Outfit Z-Standard (ZSTD)	-2,0 < ZSTD < 2,0
Point Measure Correlation (Pt Mean Corr)	0,4 < Pt Mean Corr < 0,85

Table 4. Item fit criteria in the Rasch model

3.4. Data analysis

The data obtained in the study are quantitative, including scores/ percentages for the implementation of the science learning process and HOTs. The details of data analysis to meet the research objectives are as follows.

- Descriptive statistical analysis to get an overview of the research results without generalizing (Fraenkel et al., 2023; Jackson, 2009). This analysis was conducted to provide an overview of the results of implementing VL-ISCT in science learning on pre-service science teachers' HOTs.
- Conversion of quantitative to qualitative scale from the implementation score of the science learning process and the average score of the posttest of higher-order thinking skills with the norm curve referenced with the provisions as shown in Figure 9 and Table 5 (Mangal & Mangal, 2019; Waugh & Gronlund, 2013).

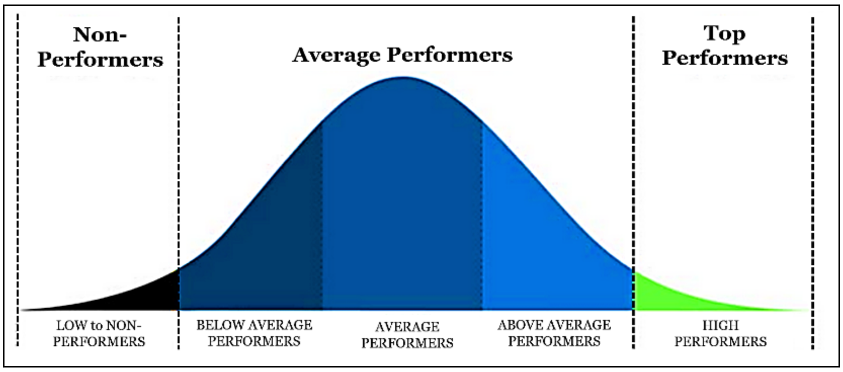


Figure 9. Conversion criteria in the normal curve

Score Interval	Learning Process	Criteria
$x > \bar{x}_i + 1{,}8SD_i$	$x > 80$	High Performers
$\bar{x}_i + 0{,}6SD_i < x \leq \bar{x}_i + 1{,}8SD_i$	$60 < x \leq 80$	Above Average Performers
$\bar{x}_i - 0{,}6SD_i < x \leq \bar{x}_i + 0{,}6SD_i$	$40 < x \leq 60$	Medium Performers
$\bar{x}_i - 1{,}8SD_i < x \leq \bar{x}_i - 0{,}6SD_i$	$20 < x \leq 40$	Below Average Performers
$x \leq \bar{x}_i - 1{,}8SD_i$	$x \leq 20$	Low Performers

Table 5. Scale conversion in the normal curves

- Prerequisite tests in ANOVA include: normality and homogeneity tests (Jackson, 2009; Navarro & Foxcroft, 2019).
- Inferential statistical analysis in this case uses analysis of variance (ANOVA) to determine whether there is a difference in pretest and posttest scores for pre-service science teachers' HOTs (Jackson, 2009; Navarro & Foxcroft, 2019).
- Effect size analysis to determine the effect size of implementing VL-ISCT in science learning on pre-service science teachers' HOTs (Brydges, 2019; Cohen, 1988; Navarro & Foxcroft, 2019).

4. RESULT

This section presents analysis results in three parts: instrument quality, learning implementation, and pre-service science teachers' higher-order thinking skills (HOTs).

4.1. Quality of research instruments

All research instruments were validated by experts and empirically tested to ensure accuracy and reliability, enabling precise conclusions about the impact of VL-ISCT on pre-service science teachers' HOTs. The first instrument used was the VL-ISCT learning tool validation sheet, which was analyzed for content validity using a Content Validity Ratio (CVR) involving five lecturers in science education; the results are shown in Table 6.

Learning Tools	Aspects	Validator (Experts)					CVR
		I	II	III	IV	V	
Lesson Plan	Construct	1	1	1	1	1	1,00
	Konten	1	1	1	1	1	1,00
	Language	1	1	1	1	1	1,00
Student Worksheets	Construct	1	1	1	1	1	1,00
	Konten	1	1	1	1	1	1,00
	Language	1	1	1	1	1	1,00
Learning Media	Construct	1	1	1	1	1	1,00
	Konten	1	1	1	1	1	1,00
	Language	1	1	1	1	1	1,00
Test Items	Construct	1	1	1	1	1	1,00
	Konten	1	1	1	1	1	1,00
	Language	1	1	1	1	1	1,00
	CVI			1,00			

Table 6. Analysis results of content validity ratio

Table 6 shows that the CVI is 1,00, indicating that all aspects of the learning tools used in the VL-ISCT activities, including lesson plans, student worksheets, learning media, and test questions, are valid (Gilbert & Prion, 2016; Lawshe, 1975). Furthermore, the assessment results from these experts were also analyzed for inter-rater reliability using Fleiss's Kappa, as shown in Table 7. Table 7 shows that the Fleiss's Kappa coefficient is 1,00. The coefficient is in the range of ,81-1,00, indicating perfect agreement and suitable for use in research (McHugh, 2012; Orts-Cortés et al., 2013). The coefficient values confirm expert agreement on the device.

			95% CI	
Ratings	**Fleiss's kappa**	**SE**	**Lower**	**Upper**
Overall	1,000	0,148	0,987	1,000
3	0,982	0,148	0,988	0,962
4	0,983	0,148	0,987	0,963

Table 7. Analysis results of Fleiss's Kappa

HOTs question quality was further empirically validated using Principal Component Analysis (PCA) for construct validity and the Rasch model for empirical fit. The results of the empirical test of construct validity are shown in Table 8, which uses Principal Component Analysis (PCA).

Component	Empirical (Eigenvalue)	Empirical (%)	Modeled (%)
Total of raw variance in observations	34,5	100,0%	100,0%
Variance explained by measures	14,5	44,1%	43,8%
• Explained by persons	8,5	24,5%	22,3%
• Explained by items	6,1	17,6%	17,4%
Unexplained variance (total)	20,0	57,9%	58,2%
• Unexplained variance in 1st contrast	2,4	7,0%	12,1%
• Unexplained variance in 2nd contrast	1,8	5,3%	9,2%
• Unexplained variance in 3rd contrast	1,6	4,6%	7,9%
• Unexplained variance in 4th contrast	1,4	4,1%	7,1%

Table 8. Analysis Results of principal component analysis (PCA)

The results of the construct validity test using Principal Component Analysis (PCA), shown in Table 8, indicate that the main construct accounts for 44,1% of the variance. Based on the PCA, the value exceeds the minimum threshold of 20% and is in the good category because it is above 40%. Meanwhile, the residual contrast values were 7.0%, 5,3%, 4,6%, and 4,1%, respectively, all below 10% and none above 15%, indicating that no other dominant latent dimensions were present. In addition, no more than two contrasts exceed 10% according to the unidimensionality criteria (Dimitrov, 2012; Greenacre et al., 2022). Thus, these results support that the instrument on HOTs has met the conditions of construct validity. The results of the Rasch model validity test are summarized in Table 9.

Statistics	Measure	Infit		Outfit	
		MNSQ	ZSTD	MNSQ	ZSTD
Mean	,00	1,00	-,1	1,01	,0
Standard Deviation	,74	,17	1,3	,19	1,3
Maximum	1,26	1,35	2,4	1,46	2,5
Minimum	-1,81	,75	-2,1	,75	-2,1

Table 9. Summary of the results of the Rasch model

Component	Value
Separation (Real)	3,80
Item Reliability (Real)	,94
Separation (Model)	3,94
Item Reliability (Model)	,94
S.E. of Item Mean	,17

Table 10. Reliability and item separation

A summary of the Rasch model analysis of 14 HOTs items indicates that all items fit the model. The average Outfit Mean Square (MNSQ) value of 1,01 is within the ideal range of ,5 < MNSQ < 1,5, and the average Outfit Z-Standard (ZSTD) value of ,0 is within the limit of -2,0 < ZSTD < 2,0 according to the interpretation criteria of the Rasch model (Dimitrov, 2012; Linacre, 2002; Meyer, 2014). This indicates that none of the items deviates significantly from the model's expectations, so the student's response to the item can be considered logical and consistent. The item reliability is ,94, and the item separation index is 3,94, indicating that the instrument has an excellent ability to distinguish item difficulty levels (Crocker & Algina, 2008; Meyer, 2014). The range of measure values from -1,81 to 1,26 logit also indicates that the item has sufficient difficulty level variation to measure different levels of higher-order thinking skills. Thus, this instrument has proven valid and reliable for use in HOTs assessments. Furthermore, the results of the Rasch model fit analysis of each item are shown in detail in Table 11.

Measure	Outfit MNSQ	Outfit ZSTD	Pt Measure Corr	Item	Result
-,54	,62	-1,74	,73	A1	Fit
-,42	,51	-1,92	,64	A2	Fit
,56	,98	-,05	,54	B1	Fit
,42	1,11	,54	,80	B2	Fit
,51	1,46	1,95	,67	C1	Fit
-1,81	,95	-,08	,72	C2	Fit
-1,19	1,43	1,89	,59	D1	Fit
1,03	1,16	,95	,63	D2	Fit
,34	,57	-1,92	,71	E1	Fit
1,26	,73	-1,51	,63	E2	Fit
,40	,59	-1,91	,42	F1	Fit
,62	1,01	-,13	,52	F2	Fit
-,12	,78	-,41	,72	G1	Fit
-,24	,76	-1,07	,77	G2	Fit
,00	,93	-,45	–	–	
,66	,33	1,52	–	–	

Table 11. Analysis results of each fit item

The Rasch model analysis of 14 items showed that all items fit the model. The value of the Mean Square Outfit (MNSQ) ranges from ,51 to 1,46, and the ZSTD Outfit is in the range of -1,92 to 1,95, all of which are still within the tolerance limit (,5 < MNSQ < 1,5 and -2,0 < ZSTD < 2,0) (Dimitrov, 2012; Linacre, 2002; Meyer, 2014). This indicates that none of the items are too deviant from the model. The Point Measure Correlation (Pt Mean Corr) value ranged from ,42 to ,80, all above the minimum threshold of 0,4 indicating that all items contributed positively to the construct measurement of higher-order thinking skills (Linacre, 2002). The average Observed Match value of 79,7% was also close to the expected value (78,6%), indicating consistency between the actual response and the model prediction. Thus, all items can be revealed valid and feasible for objectively and accurately measuring HOTs.

4.2. Implementation of VL-ISCT in science learning

Virtual laboratories with interactive simulations and computational thinking (VL-ISCT) is implemented using the discovery learning in all classes. Student worksheets from learning implementation are used as a guide in the discovery learning stages, as shown in Figure 5. Learning on static electrical materials during three meetings is observed by two observers at each meeting. Data on the implementation of learning from three classes during three meetings is shown in Table 12.

Table 12 shows that VL-ISCT in learning is well implemented (100%). Experiment Class 1 uses VL-ISCT with four CT skills: decomposition, pattern recognition, abstraction, and algorithm; Experiment Class 2 uses three CT skills: decomposition, pattern recognition, and abstraction; and Control Class uses two: decomposition and pattern recognition.

Class	Syntax	Meeting 1		Meeting 2		Meeting 3	
		I	II	I	II	I	II
Eksperimen 1	Opening	√	√	√	√	√	√
	Stimulation	√	√	√	√	√	√
	Problem Statement	√	√	√	√	√	√
	Data Collection	√	√	√	√	√	√
	Data Processing	√	√	√	√	√	√
	Verification	√	√	√	√	√	√
	Generalisation	√	√	√	√	√	√
	Closing	√	√	√	√	√	√
		100% Accomplished (High Performers)					
Eksperimen 2	Opening	√	√	√	√	√	√
	Stimulation	√	√	√	√	√	√
	Problem Statement	√	√	√	√	√	√
	Data Collection	√	√	√	√	√	√
	Data Processing	√	√	√	√	√	√
	Verification	√	√	√	√	√	√
	Generalisation	√	√	√	√	√	√
	Closing	√	√	√	√	√	√
		100% Accomplished (High Performers)					
Control	Opening	√	√	√	√	√	√
	Stimulation	√	√	√	√	√	√
	Problem Statement	√	√	√	√	√	√
	Data Collection	√	√	√	√	√	√
	Data Processing	√	√	√	√	√	√
	Verification	√	√	√	√	√	√
	Generalisation	√	√	√	√	√	√
	Closing	√	√	√	√	√	√
		100% Accomplished (High Performers)					

Table 12. Results of the implementation of virtual laboratories with interactive simulations and computational thinking (VL-ISCT)

4.3. Higher-order thinking skills

The pre-service science teachers' higher-order thinking skills (HOTs) were measured using fourteen pretest and posttest questions that had met the quality of both validity and reliability. The results of descriptive statistics from data of HOTs from three learning classes as shown in Table 13. Table 13 presents the results of descriptive statistical analysis of the higher-order thinking skills of pre-service science teachers, including mean (average), median, mode, standard deviation, kurtosis, skewness, range, minimum, maximum, and count, which describe the distribution of the data.

Statistics	Exp. Class 1		Exp. Class 2		Control Class	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
Mean	32,67	83,93	29,80	75,27	28,90	67,03
Median	33,00	84,00	30,00	75,00	28,50	67,00
Mode	22,00	73,00	41,00	75,00	25,00	64,00
Std. Deviation	7,11	6,95	7,25	5,23	7,83	6,17
Kurtosis	-1,04	-,94	- ,59	-1,04	-,66	-,34
Skewness	,06	-,19	,12	-,14	,22	-,09
Range	22,00	22,00	24,00	16,00	28,00	25,00
Minimum	22,00	73,00	18,00	67,00	16,00	53,00
Maximum	44,00	95,00	42,00	83,00	44,00	78,00
Count (N)	30,00	30,00	30,00	30,00	30,00	30,00

Table 13. Analysis results of descriptive statistics

The average value for each dimension describing HOTs is complete, as shown in Table 14.

Domains of Hots		Exp. Class 1		Exp. Class 2		Control Class	
		Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
Analysis	Differentiating	35,73	88,59	34,73	74,62	31,28	71,89
	Organizing	39,95	81,74	26,96	75,26	32,54	74,91
	Attributing	26,77	82,83	34,2	71,81	24,11	61,52
Evaluation	Checking	28,96	79,45	20,9	78,65	31,98	72,89
	Critiquing	30,12	83,72	31,93	82,13	25,5	69,42
Creative	Generating	32,85	78,83	27,58	70,42	28,34	65,37
	Planning	32,85	82,96	29,95	75,46	26,16	61,02
	Producing	33,84	89,74	32,49	81,92	28,63	62,71
Total, Average Score		32,67	83,93	29,80	75,27	28,90	67,03
Category		High Performers		Above Average Performers		Above Average Performers	

Table 14. Average of higher-order thinking skills (HOTs)

Table 13 shows that pre-service science teachers' HOTs improved after the learning process, especially in Experimental Class 1. The highest average posttest score was achieved by Experimental Class 1 (83,93; high performers), followed by Experimental Class 2 (75,27; above-average performers), and Control Class (67,03; above-average performers) (Mangal & Mangal, 2019; Waugh & Gronlund, 2013). Significant HOTs improvements, differentiating, organizing, attributing, and producing, occurred, especially in Experimental Class 1, which showed a sharp pretest-to-posttest gain. Control Class showed minimal improvement, indicating the treatment's greater effectiveness in Experimental Class 1.

Inferential statistical tests were performed to ensure that Experimental Class 1 was more effective than Experimental Class 2 and Control Class. The data analyzed included pretest and posttest scores that met the analysis prerequisite, namely, normal and homogeneous distributions. The results of the prerequisite test analysis, focusing on the normality of the pretest data, are shown in Table 15. Since this value is greater than ,05, the pretest data is normally distributed. In addition, the results of the variance homogeneity test using Levene's test showed an F value of ,334 with a significance of ,717.

Normality Test (Shapiro-Wilk)		Homogeneity of Variances Test (Levene's)			
Statistic	p	F	df1	df2	p
,967	,051	,334	2	87	,717

Table 15. Analysis result of test of normality and homogeneity from pretest

Significance values exceeding ,05 indicate that the variance between groups is homogeneous or equivalent. In addition, these results are reinforced by the Q–Q Plot graph in Figure 10.

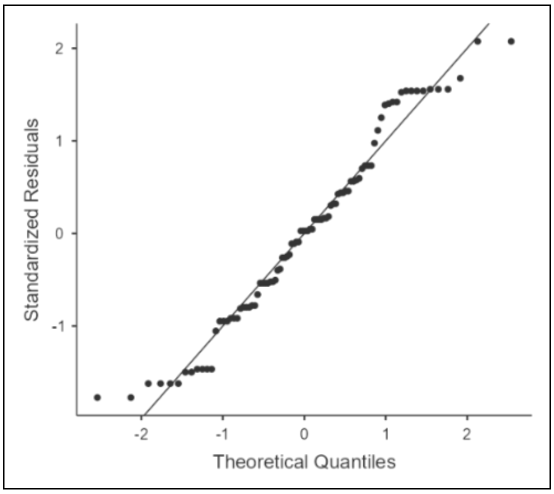


Figure 10. Q-Q plot from pretest

The graph in Figure 10 shows that the residuals are nearly normally distributed, as most points lie on a straight line that approximates the theoretical normal distribution (Jackson, 2009; Navarro & Foxcroft, 2019). Minor deviations in the tail are normal and do not significantly interfere with the assumption of normality. Therefore, the residual normality assumption is met, and the regression model or parametric test used remains valid. Thus, pretest data meet the assumptions of normality and homogeneity, allowing to be analyzed using analysis of variance (ANOVA). The results of ANOVA from pretest are shown in Table 16.

	Sum of Squares	df	Mean Square	F	p
Learning Model	232	2	116,1	2,12	,126
Residuals	4768	87	54,8		

Table 16. Result of analysis of variance (ANOVA) from Pretest

The result of ANOVA shows ($p = ,126$) and post hoc tests (all $p > ,05$), indicating no significant difference in pretest of HOTs among the three classes, confirming equal initial abilities before treatment (Jackson, 2009; Navarro & Foxcroft, 2019).

In addition, the results of the prerequisite normality and homogeneity tests for the posttest data are shown in Table 17.

Normality Test (Shapiro-Wilk)		Homogeneity of Variances Test (Levene's)			
Statistic	p	F	df1	df2	p
,976	,090	1,08	2	87	,344

Table 17. Analysis result of test of normality and homogeneity from posttest

The posttest data from HOTs met parametric test assumptions, with Shapiro-Wilk ($p = ,090$) and Levene's Test ($p = 1,08$) both $> ,05$, indicating normality and homogeneity of variance for valid group comparisons (Jackson, 2009; Navarro & Foxcroft, 2019). In addition, the results were further strengthened by the Q–Q Plot shown in Figure 11.

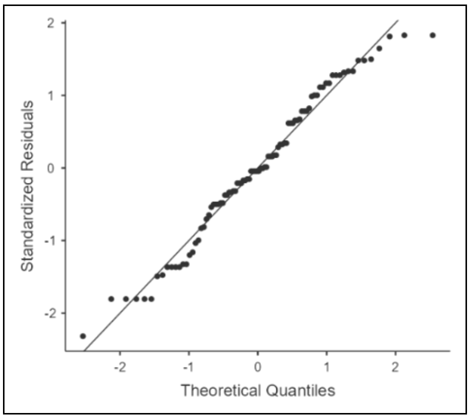


Figure 11. Q-Q plot from posttest

Q–Q residual plots on the higher-order thinking skills posttest data show that most of the data points are along a diagonal line, which indicates that the residuals are normally distributed. Minor distribution deviations remain within acceptable limits, supporting the Shapiro-Wilk result ($p = ,090 > ,05$). Thus, the post-test data meet the normality assumption and are suitable for ANOVA (Jackson, 2009; Navarro & Foxcroft, 2019). The results of ANOVA and post hoc analysis from posttest data are shown in Tables 18 and 19.

	Sum of Squares	df	Mean Square	F	p	h^2
Learning Model	4285	2	2142,5	56,5	< ,001	,565
Residuals	3297	87	37,9			

Table 18. Result of analysis of variance (ANOVA) from posttest

Learning Model		Mean Difference	df	t	Phonferroni	Cohen's d
1	2	8,67	87,0	5,45	< ,001	1,41
1	3	16,90	87,0	10,63	< ,001	2,75
2	3	8,23	87,0	5,18	< ,001	1,34

Table 19. Result of post hoc from posttest

Bonferroni correction showed that all group pairs differed significantly ($p < ,001$), with high effectiveness values based on Cohen's d = 1,41 (Experiment Class 1 vs. Experiment Class 2), d = 2,75 (Experiment Class 1 vs. Control Class), and d = 1,34 (Experiment Class 2 vs. Control Class) (Cohen, 1988; Faul et al., 2007). Thus, it can be concluded that the learning model applied to the Experiment Class was significantly more effective in improving HOTs than Control Class, and the difference was also practically significant.

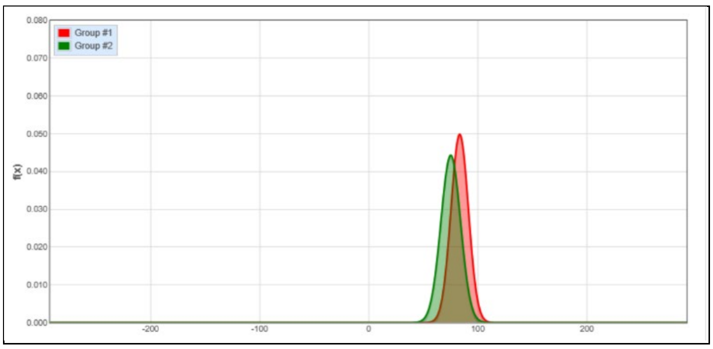


Figure 12. Effect size (Cohen's d) experiment class 1 vs experiment class 2

According to Cohen's criteria, Cohen's d = 1,41 represents a very large effect size, as it exceeds the threshold of ,80 (Cohen, 1988; Faul et al., 2007). This shows that the difference in HOTs between Experiment Class 1 and Experiment Class 2 is practically meaningful, not just statistically significant.

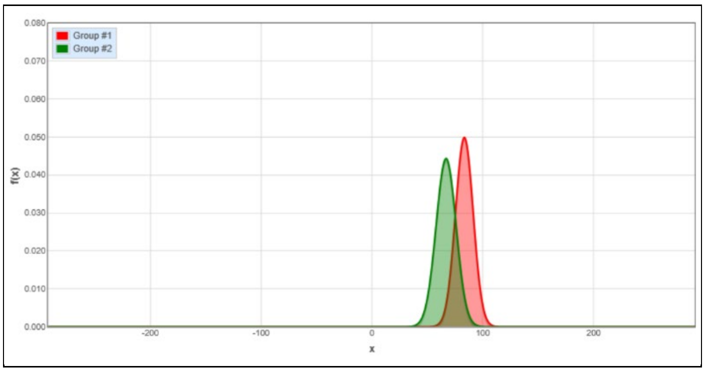


Figure 13. Effect size (Cohen's d) experimental class i vs control class

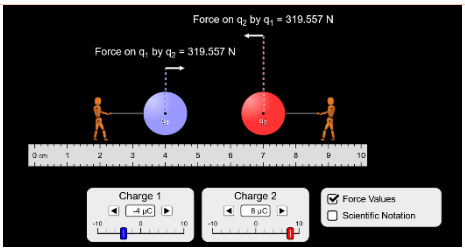
According to Cohen's conventional criteria, Cohen's d = 2,75 represents a very large effect size, as it exceeds the threshold of ,80 (Cohen, 1988; Faul et al., 2007). This shows that the learning model used in Experiment Class 1 had a stronger influence on HOTs than the one used in the Control Class. This distribution visualization supports these statistical data: the overlap between the curves is very small, confirming that the learning outcomes of the two groups are very different.

5. DISCUSSION

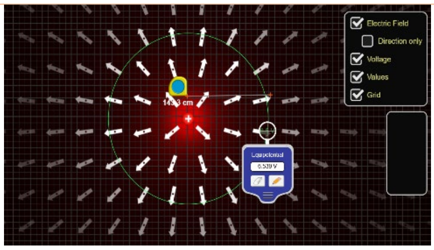
This section presents three sections of discussion to answer the purpose of writing an article. This section includes the role of the virtual laboratory with interactive simulation (VL-IS) in science learning, computational thinking (CT)-integrated science learning, and the results of implementing VL-ISCT in science learning on pre-service science teachers' higher-order thinking skills (HOTs).

5.1. Virtual laboratory with interactive simulation

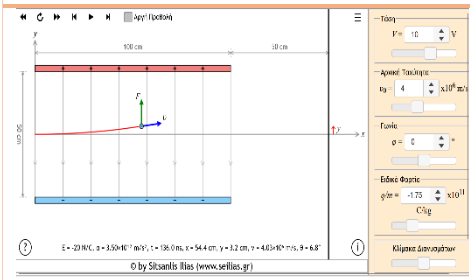
VL-IS is an innovative digital-era approach that bridges conceptual understanding and higher-order thinking in abstract science topics. This study supports learning about electric force and field, motion of charged particles, and capacitors (Figure 14). Each meeting is designed to provide an exploration-based learning experience with digitally interactive simulations based on science phenomena (Rutten et al., 2012; Walters et al., 2017; Zacharia & Olympiou, 2011). With this learning technology, students not only observe phenomena but also organize variables, collect data, and draw conclusions through virtual experiments (De Jong et al., 2013; Smetana & Bell, 2012).



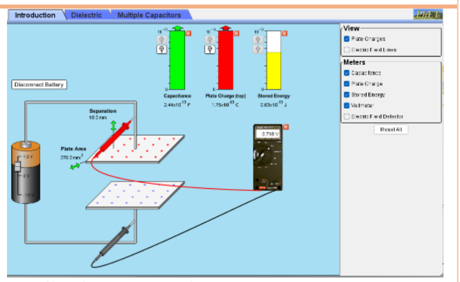
Application 1. Electric Force



Application 2. Electric Field and Electrical Potential



Application 3. Particle Motion in an Electric Field



Application 4. Capacitor

Figure 14. Application for virtual laboratory with interactive simulations (VL-IS)

This approach is intended not only to provide an understanding of science concepts but also to encourage students to develop computational thinking skills, which are an important foundation for data-driven problem-solving (Fei et al., 2025; Wing, 2006). For instance, Application 3 helps explore particle motion in an electric field, as seen in lightning, which can strike between clouds, within a cloud, or from cloud to ground.

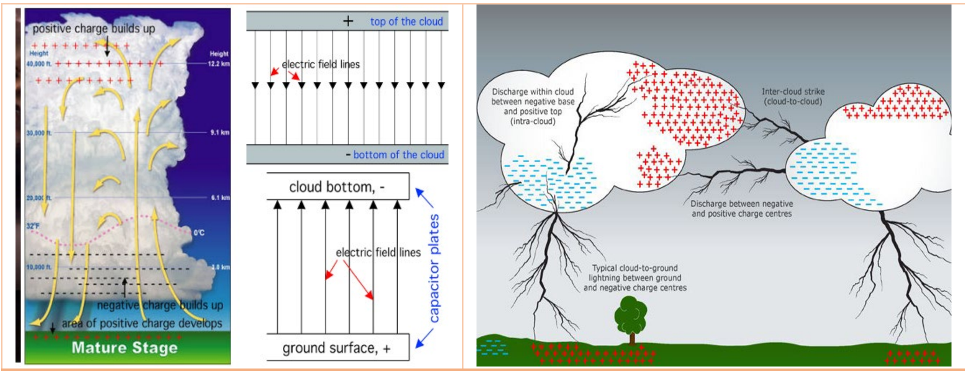


Figure 15. Scientific phenomena of lightning and science concepts

Electric Field		Time* t (ns)	Position (cm)		Velocity v_t (m/s)	Particle Direction
$\vec{\mathbf{E}}$ (V/m)	Direction		x	y		
-40	Down	100,4	40,2	-2,5	4,03	Down
		150	60	-5,6	4,07	Down
		201	80,4	-10,1	4,12	Down
		249,2	99,7	-15,5	4,19	Down
		316,2	126,5	-25	4,3	Down
40	Up	34,8	13,9	0,3	4,02	Up
		76,6	30,6	1,5	4,02	Up
		88,6	35,4	2	4,02	Up
		150,6	60,2	5,7	4,07	Up
		235,4	94,2	13,9	4,17	Up

Table 20. Example of electrostatic force simulation results

VL-IS enables students to explore phenomena in science that are unsafe or impossible to replicate in a physical laboratory (Makransky et al., 2019; Parong & Mayer, 2018). As part of scientific investigation, it provides accurate data through virtual experiments, such as lightning (Figure 15), which involves massive electrical energy discharge and cannot be safely studied directly. Through virtual simulations, students can observe the distribution of positive and negative charges in clouds and model various atmospheric conditions to analyze the formation of lightning strikes. Furthermore, with interactive simulations, students can modify variables such as humidity, temperature, and load distribution to obtain accurate data and dynamic visualizations. This activity reflects the dimension of science as a way of investigating, as it involves observation, simulation-based experiments, data collection, and the drawing of scientific conclusions (Makransky et al., 2016; Rutten et al., 2012). VL-IS not only presents theory visually but also fosters scientific thinking skills and problem-solving, and increases students' active involvement in science learning.

5.2. CT integrated science learning

The integration of CT into the science learning process is increasingly attracting attention as a strategic approach to learning science, spanning four science domains. This approach aligns closely with the discovery learning stage, which prioritizes students' active exploration through a series of systematic stages: stimulation, problem statement, data collection, data processing, verification, and generalization (Bruner, 1961). Students, pre-service science teachers, are invited to elaborate complex phenomena into smaller, specific, and scientifically traceable sub-problems (Brennan & Resnick, 2012; Wing, 2006). For example, in observing the motion of particles in an electric field (lightning phenomenon), as shown in Figure 15, students can break it down into the following problems:

- How does the electric field affect the trajectory of charged particles?
- How does this trajectory explain the beginning of lightning strikes?
- What is the effect of the initial speed, angle of fire, and magnitude of the electric field on the path of motion of the charge?

This process not only helps students focus on important aspects of scientific phenomena but also lays the foundation for scientific reasoning in crafting subsequent investigations (Angeli et al., 2016; Shute et al., 2017). After defining the problem, students collect and analyze data, building pattern-recognition skills by exploring relationships between variables such as electric potential and field, entry angle and trajectory, or field direction and motion. Based on the data from the virtual laboratory work with interactive simulation, as shown in Table 20, the explanation of pattern recognition is provided in Table 21.

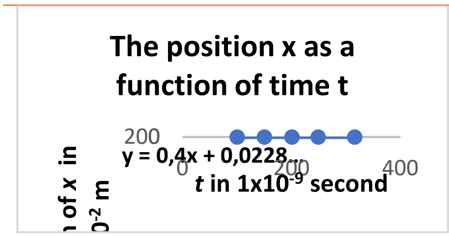
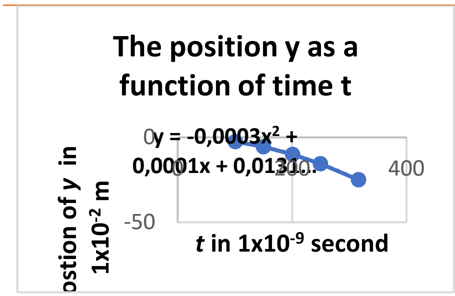
Data Analysis	Pattern Recognition
	• The type of motion on the x-axis is uniform linear motion • The equation of linear motion $y = mx$, where the position x (in the order of $1x10^{-2}$ m) is indicated by y, and the time t (in the order $1x10^{-9}$) is indicated by x, so the equation is obtained: $$x = vt = 0{,}4\ t$$
	• The type of motion on the y-axis is non-uniform linear motion • The parabolic motion equation $y = ax^2 + bx + c$, where the position x (in the order $1x10^{-2}$ m) is indicated by y, and the time t (in the order $1x10^{-9}$) is indicated by x, so the equation is obtained: $$y = \pm\frac{1}{2}at^2 + v_0t + y_0$$ $$y = -0{,}0003x^2 + 0{,}0001x + 0{,}0131$$
Findings from Pattern Recognition: • The direction of motion of charged particles is affected by the type of charge and the direction of the electric field. • Positively charged particles will move towards the negatively charged conductor plate, and conversely the negatively charged particles will travel towards the positively charged conductive plate • The trajectory traveled by the particles forms parabolic motion	

Table 21. Example of applying pattern recognition

Graph analysis shows charged particles move linearly along the x-axis and parabolically along the y-axis, demonstrating pattern recognition that supports conceptual understanding and problem-solving. Pre-service science teachers apply abstract thinking to obtain concepts, laws, or theories (as in Table 22).

Thinking Abstraction I	Thinking Abstraction II
Along the x-axis, the electron moving in uniform linear motion because it is not affected by the downward electric field, so that the equation is obtained: $$\mathbf{x = v_o . t}$$ $$\mathbf{t = \frac{x}{v_o}}$$ Along the y-axis, the electron moves in non-uniform linear motion because it is affected by an electric field pointing downwards, so the equation is obtained: $$\mathbf{y = d = -\frac{1}{2}at^2}$$	If the electrostatic force $\overrightarrow{F_e} = e\vec{E}$ is connected to Newton's Law II, then the equation is obtained: $$\mathbf{F = ma}$$ $$\mathbf{a = \frac{F}{m} = \frac{eE}{m}}$$ If the equation of abstract thinking I is combined with abstract thinking II, then the equation is obtained: $$\mathbf{d = -\frac{1}{2}\frac{(-e)E}{m}\left(\frac{x}{v_o}\right)^2}$$ $$\mathbf{d = \frac{eE}{2mv_o^2}x^2}$$

Table 22. Examples of the application of abstract thinking

Through abstraction, students simplify physical events into mathematical models, identifying uniform motion on the x-axis and uniformly accelerated motion on the y-axis. By engaging in abstraction, students learn to simplify problem representations and to form conceptual models that are useful for explaining or predicting phenomena (Grover & Pea, 2013; Voogt et al., 2013). By using electric force and Newton's law, these equations support algorithmic thinking by logically linking parameters to acceleration, time, and position. Thus, abstract thinking serves as a conceptual foundation, while algorithmic thinking provides a procedural guide for solving data- and logic-based science problems, as shown in Table 23.

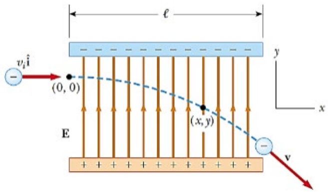
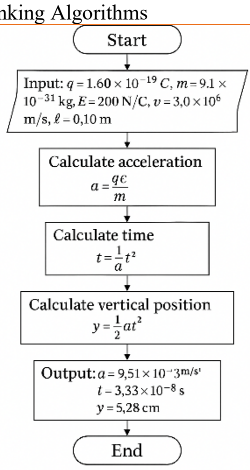
Problem Discussion	Thinking Algorithms
An electron with a charge of $q = -1{,}60 \times 10^{-19}$ C and a mass of $m = 9{,}1 \times 10^{-31}$ kg enters a uniform electric field as shown in Figure 3.  Figure 3. Motion of an Electron in an Electric Field If the initial velocity of the electron is $v_x = 3{,}0 \times 10^6$ m/s and the electric field strength is E = 200 N/C, calculate the acceleration of the electron in the field, the time required to travel a distance ℓ = 0,10 m, and the vertical position of the particle when it leaves the field, respectively ... . A. $a = 9{,}51 \times 10^{13}$ m/s²; $t = 3{,}33 \times 10^{-8}$ s; and y = 5,28 cm B. $a = 6{,}51 \times 10^{13}$ m/s²; $t = 5{,}33 \times 10^{-8}$ s; and y = 1,95 cm C. $a = 3{,}51 \times 10^{13}$ m/s²; $t = 3{,}33 \times 10^{-8}$ s; and y = 1,95 cm	

Table 23. Examples of the application of algorithmic thinking

In the final stage, pre-service science teachers apply algorithmic thinking by designing systematic steps for data analysis and scientific argumentation. Integrating CT into discovery learning from problem statement to data processing creates a robust framework for active and problem-solving science learning (as in Figure 16).

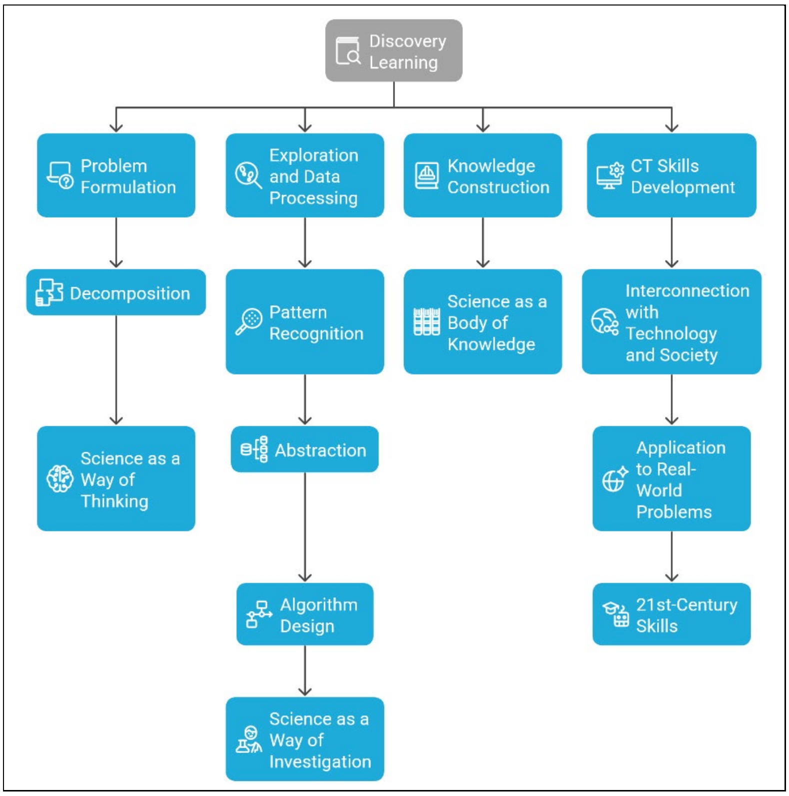


Figure 16. Integration of computational thinking (CT) in discovery learning

Based on this description, integrating CT into stage of discovery learning not only enhances the effectiveness of science learning but also strengthens the dimension of HOTs. This emphasizes the importance of designing science learning that not only focuses on concept transfer but also equips students with a way of thinking and strategic skills relevant to future challenges.

5.3. Implementation of VL-ISCT in science learning on HOTs

HOTs in this research are part of the cognitive learning outcomes. Every research, including this research, urgently needs a data collection tool, a research instrument, that meets feasibility and validity criteria (Canals, 2017; Sudaryono et al., 2019). The instruments used in the study have met the criteria for validity and reliability, as well as expert ratings (Lawshe, minimum CVI > ,99) and Fleiss's Kappa. In addition, the question instruments for pretest and posttest have been analyzed by testing the Rasch model fit items with the criteria of MNSQ Outfit values with a range of ,5 < MNSQ Outfit < 1,5, Z-Standard Outfit values with a range of -2,0 < ZSTD < 2,0 and Point Measure Correlation values with a range of ,4 < Pt Mean Corr < ,85 (Boone et al., 2014; Dimitrov, 2012; Linacre, 2002). The results of this analysis indicate that the quality standards of the instruments used in this study have been met, so the conclusions obtained are accurate.

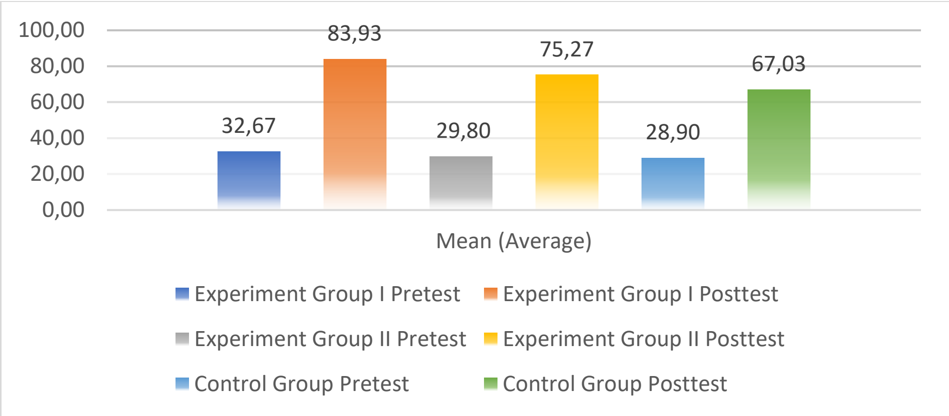


Figure 17. Average pretest and posttest

This study shows that implementing science learning that integrates virtual laboratory and CT significantly improves the higher-order thinking skills of pre-service science teachers. Based on the descriptive data in Table 13, there was a substantial increase in posttest scores for digital literacy skills in Experiment Class 1 (M = 83,93) and Experiment Class 2 (M = 75,27), compared to the control class (M = 67,03). Experimental Class 1, which integrated all four components of CT, showed the highest results across mean, data distribution, and minimum–maximum value distribution. This confirms that the more complete the integration of CT skills, the greater their influence on the development of HOTs (Brennan & Resnick, 2012; Ching et al., 2018; Sneider et al., 2014).

Furthermore, the results of ANOVA on the posttest data confirmed that the differences between groups were statistically significant ($F = 56,5$, $p < ,001$) with a high effect size ($\eta^2 = ,565$), which shows that integrating virtual lab and CT in the discovery learning model plays a major role in student achievement (Jackson, 2009; Navarro & Foxcroft, 2019). Bonferroni's post hoc follow-up test showed that Experiment Class 1 was significantly superior to Experiment Class 2 (MD = 8,67; p < ,001) and to Control Class (MD = 16,90; p < ,001), with Cohen's d exceeding 1,3 which falls into the large effects category. This strengthens the evidence that implementing CT within discovery learning, supported by interactive simulations, drives meaningful, in-depth learning achievement.

A more detailed analysis is shown in Figures 17 and 18, which outline the results for each dimension of HOTs, from analysis, evaluation, and creativity, in the average posttest scores.

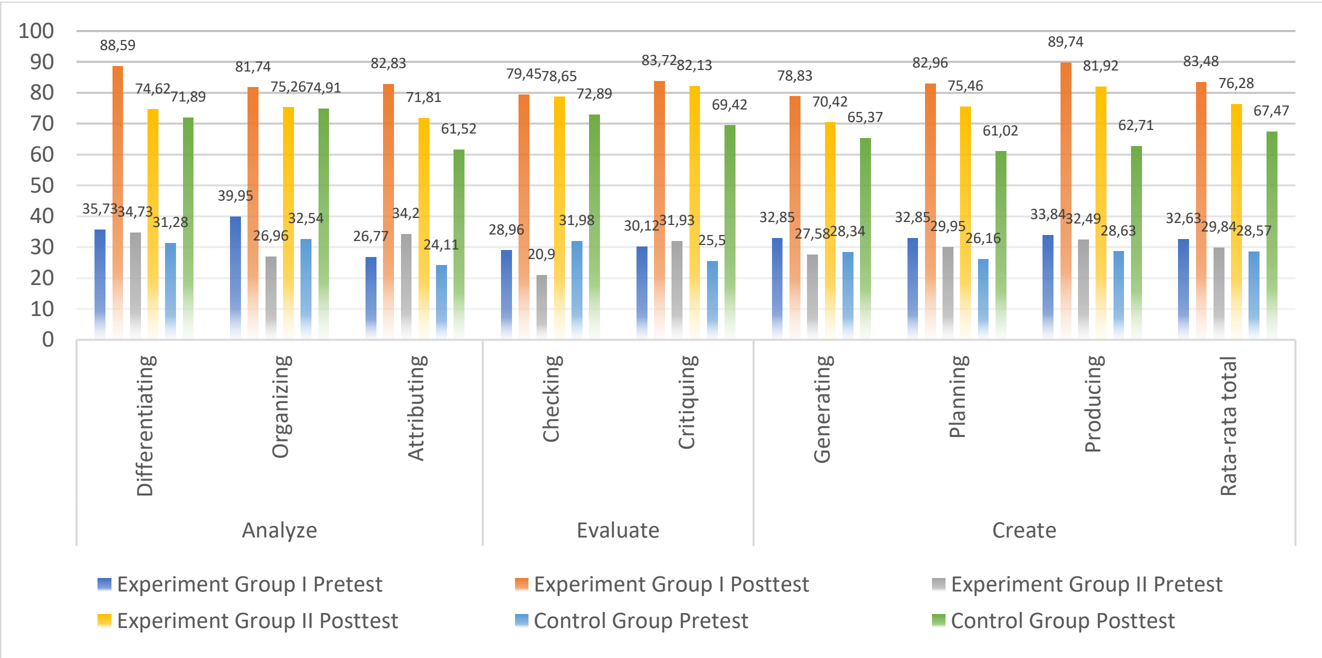


Figure 18. Average comparison of all dimensions of higher-order thinking skills (HOTs)

Pre-service science teachers in Experiment Class 1 showed very high scores, especially in the differentiating (88,59), organizing (81,74), and attributing (82,83) indicators. This success is attributed to decomposition and pattern recognition skills, which help students break down scientific phenomena into smaller pieces and recognize relationships between variables from the simulated data (Aziz et al., 2023; Denning, 2009; Wing, 2006). Experiment Class 2 also produced good results, although not as high as Experiment Class 1, underscoring the importance of algorithmic components for reinforcing logical thinking. The evaluation dimension also showed that Experiment Class 1 achieved a high score, with a checking score of 79,45 and a critiquing score of 83,72. This dimension is closely related to abstraction and algorithmic skills, as pre-service science teachers are required to filter relevant information and systematically compile a sequence of evaluative steps. Likewise, in the creative dimension, Experiment Class 1 excelled in the indicators of generating (78,83), planning (82,96), and producing (89,74), which shows that the combination of CT and virtual laboratory is able to generate ideas, design solutions, and create products based on scientific processes as core competencies in 21st-century science learning (Cansu & Cansu, 2019; Sadykova & Il'bahtin, 2020). These results demonstrate that integrating interactive virtual laboratories with computational thinking not only improves conceptual mastery and higher-order thinking skills but also makes an innovative contribution to science education research and practice. By combining digital experimentation with structured cognitive processes, this approach extends existing inquiry-based and constructivist models into a more technology-enriched learning framework. Its relevance is particularly significant in science teacher education, where it provides a practical and research-based model for preparing future teachers to design meaningful, student-centered, and innovation-oriented learning experiences in 21st-century classrooms.

6. CONCLUSION AND FURTHER RESEARCH

This study shows that integrating virtual labs with interactive simulations and CT covering decomposition, pattern recognition, abstraction, and algorithms significantly enhances prospective science teachers' logical and systematic thinking compared to conventional methods. The results of the experiment showed that students who participated in learning with the integration of CT in its entirety and virtual lab obtained significantly higher HOTs scores compared to other classes. These findings have important implications for the development of science curriculum and learning strategies, particularly in preparing prospective teachers to think critically, analytically, and adaptively to the challenges of 21st-century learning. Integrating interactive simulations with CT offers an innovative approach that enhances the quality of science education and prepares graduates to apply higher-order thinking in real-world and teaching contexts. However, this study has some limitations. The research sample came from only one institution, consisted of a limited number of participants, and did not evaluate the long-term impact of the learning interventions provided. Therefore, follow-up research is recommended with a wider scope and a longitudinal design to evaluate the consistency of results over the long term. In addition, aspects such as learning motivation, emotional engagement, and context-based problem-solving abilities need to be explored to gain a more thorough understanding of this approach's effectiveness.

ACKNOWLEDGMENTS

We are grateful to Faculty of Mathematics and Natural Sciences, Universitas Negeri Yogyakarta for providing the grant (T/395/UN34.9/PT.01.03/2025) that enables this study to be carried out.

REFERENCES

Abdulrahaman, M.D., Faruk, N., Oloyede, A.A., Surajudeen-Bakinde, N.T., Olawoyin, L.A., Mejabi, O.V., Imam-Fulani, Y.O., Fahm, A.O. & Azeez, A.L. (2020). Multimedia tools in the teaching and learning processes: A systematic review. *Heliyon*, *6*(11), e05312. *https://doi.org/10.1016/j.heliyon.2020.e05312*

Abidin, Z., Herman, T. & Wahyudin, W. (2023). The Computational Thinking in Elementary School in the Indonesia New Curriculum: A Teacher's Perspective. *Sekolah Dasar: Kajian Teori Dan Praktik Pendidikan*, *32*(2), 178. *https://doi.org/10.17977/um009v32i22023p178-185*

Alfieri, L., Brooks, P.J., Aldrich, N.J. & Tenenbaum, H.R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, *103*(1), 1-18. *https://doi.org/10.1037/a0021017*

Anderson, L.W. & Krathwohl, D.R. (2001). *A Taxonomy for Learning Teaching and Assessing: A Revision of Bloom`s Taxonomy of Educational Objetives.* Longman.

Angeli, C. & Giannakos, M. (2020). Computational thinking education: Issues and challenges. *Computers in Human Behavior*, *105*, 106185. *https://doi.org/10.1016/j.chb.2019.106185*

Angeli, C., Voogt, J., Fluck, A., Webb, M., Cox, M., Malyn-Smith, J. & Zagami, J. (2016). A K-6 Computational Thinking Curriculum Framework: Implications for Teacher Knowledge. *Journal of Educational Technology and Society*, *19*(3), 47-57. *https://pure.uva.nl/ws/files/8964271/A_K_6_Computational_Thinking_Curriculum_Framework.pdf*

Artiningsih, A. & Nurohman, S. (2020). Analysis of investigative skills based on the use of tracker video analysis for 21th century skill. *Journal of Science Education Research, 3*(2). *https://doi.org/10.21831/jser.v3i2.30621*

Aziz, S.A., Fauzan, A. & Helma, H. (2023). Computational Thinking Analysis of 21st Century Professional Teacher Candidates in Solving Math Problems Viewed from The Four Foundations of CT. *JURNAL EKSAKTA PENDIDIKAN (JEP)*, *7*(2), 211-224. *https://doi.org/10.24036/jep/vol7-iss2/779*

Barr, V. & Stephenson, C. (2011). Bringing computational thinking to K-12: What is Involved and what is the role of the computer science education community? *ACM Inroads*, *2*(1), 48-54. *https://doi.org/10.1145/1929887.1929905*

Bhattacherjee, A. (2012). *Social Science Research: Principles, Methods, and Practices*. University of South Florida.

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

Brennan, K. & Resnick, M. (2012). New Frameworks for Studying and Assessing the Development of Computational Thinking. *Proceedings of the 2012 Annual Meeting of the American Educational Research Association*. *http://scratched.gse.harvard.edu/ct/files/AERA2012.pdf*

Brookhart, S.M. (2010). *How to Assess Higher-Order Thinking Skills in Your Classroom*. ASCD.

Bruner, J.S. (1961). The Act of Discovery. *Harvard Educational Review*, *31*, 21-32. *https://digitalauthorship.org/wp-content/uploads/2015/01/the-act-of-discovery-bruner.pdf*

Brydges, C.R. (2019). Effect Size Guidelines, Sample Size Calculations, and Statistical Power in Gerontology. *Innovation in Aging*, *3*(4), 1-8. *https://doi.org/10.1093/geroni/igz036*

Bybee, R.W. (2013). *The Case for STEM Education: Challenges and Opportunities*. National Science Teachers Association - NSTA Press.

Canals, L. (2017). Instruments for Gathering Data. In E. Moore & M. Dooly (Eds.). *Qualitative approaches to research on plurilingual education / Enfocaments qualitatius per a la recerca en educació plurilingüe / Enfoques cualitativos para la investigación en educación plurilingüe* (pp. 390-401). Research-publishing.net. *https://doi.org/10.14705/rpnet.2017.emmd2016.637*

Cansu, F.K. & Cansu, S.K. (2019). An Overview of Computational Thinking. *International Journal of Computer Science Education in Schools*, *3*(1), 17-30. *https://doi.org/10.21585/ijcses.v3i1.53*

Carin, A.A. & Sund, R.B. (1989). *Teaching Science Through Discovery*. Merrill Publishing Company.

Carnevali, G. & ButtazzQ, G. (2003). A virtual laboratory environment for real-time experiments. *IFAC Proceedings Volumes*, *36*(12), 31-36. *https://doi.org/10.1016/S1474-6670(17)32509-0*

Chevalier, M., Giang, C., Piatti, A. & Mondada, F. (2020). Fostering computational thinking through educational robotics: A model for creative computational problem solving. *International Journal of STEM Education*, *7*(1), 39. *https://doi.org/10.1186/s40594-020-00238-z*

Chiappetta, E.L. & Koballa, T.R. (2010). *Science Instruction in The Middle and Secondinary Schools Developing Fundamental Knowledge and Skills* (7th ed.). Pearson Education Inc.

Ching, Y.-H., Hsu, Y.-C. & Baldwin, S. (2018). Developing Computational Thinking with Educational Technologies for Young Learners. *TechTrends*, *62*(6), 563-573. *https://doi.org/10.1007/s11528-018-0292-7*

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates.

Contant, T.L., Bass, J.L., Tweed, A.A. & Carin, A.A. (2018). *Teaching Science Through Inquiry-Based Instruction* (13th ed.). Pearson Education.

Creswell, J.W. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE Publications.

Crocker, L. & Algina, J. (2008). *Introduction to Classical and Modern Test Theory*. Holt, Reinhart, and Winston, Inc.

De Jong, T., Linn, M.C. & Zacharia, Z.C. (2013). Physical and Virtual Laboratories in Science and Engineering Education. *Science*, *340*(6130), 305-308. *https://doi.org/10.1126/science.1230579*

Denning, P.J. (2009). The profession of ITBeyond computational thinking. *Communications of the ACM*, *52*(6), 28-30. *https://doi.org/10.1145/1516046.1516054*

Deshmukh, S., Deshmukh, W. & Mulla, A. (2020). Survey on Comparative Study of Virtual Lab and Traditional Lab. *International Journal of Advanced Research in Science, Communication and Technology*, *12*(2). *https://doi.org/10.48175/568*

Dimitrov, D.M. (2012). *Statistical Methods for Validation of Assessment Scale Data in Counseling and Related Fields.* Wiley.

Dubas, J.M. & Toledo, S.A. (2016). Taking Higher Order Thinking Seriously: Using Marzano's Taxonomy in The Economics Classroom. *International Review of Economics Education*, *21*, 12-20. *https://doi.org/10.1016/j.iree.2015.10.005*

Ensmenger, N. (2016). The Multiple Meanings of a Flowchart. *Information & Culture: A Journal of History*, *51*(3), 321-351. *https://doi.org/10.1353/lac.2016.0013*

Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*Power 3: A Flexible Statistical Power Analysis Program for The Social, Behavioral, and Biomedical Sciences. *Behavior Research Methods*, *39*(2), 175-191. *https://doi.org/10.3758/BF03193146*

Fei, J., Dong, S. & Fei, D. (2025). Computational thinking in complex problem solving based on data analysis: Exploring the role of problem representation using the Tower of Hanoi. *Advances in Continuous and Discrete Models*, *2025*(1), 6. *https://doi.org/10.1186/s13662-025-03866-3*

Forehand, M. (2010). *Bloom's taxonomy. In M. Orey (Ed.), Emerging perspectives on learning, teaching, and technology.* University of Georgia.

Fraenkel, J.R., Wallen, N.E. & Hyun, H.H. (2023). *How to design and evaluate research in education (11th ed.).* (11th ed.). McGraw Hill.

Froment, F., López Medialdea, A., De-Besa Gutiérrez, M., & Gil Flores, J. (2025). Adaptación y validación del Inventario Breve de Claridad Docente en población universitaria española. *Revista de Investigación en Educación*, *23*(1), 26-41. *https://doi.org/10.35869/reined.v23i1.6107*

Gilbert, G. E. & Prion, S. (2016). Making Sense of Methods and Measurement: Lawshe's Content Validity Index. *Clinical Simulation in Nursing*, *12*(12), 530-531. *https://doi.org/10.1016/j.ecns.2016.08.002*

Greenacre, M., Groenen, P.J.F., Hastie, T., D'Enza, A.I., Markos, A. & Tuzhilina, E. (2022). Principal component analysis. *Nature Reviews Methods Primers*, *2*(1), 100. *https://doi.org/10.1038/s43586-022-00184-w*

Grippo, L. & Sciandrone, M. (2023). Decomposition Methods. In L. Grippo & M. Sciandrone, *Introduction to Methods for Nonlinear Optimization* (Vol. 152, pp. 589-638). Springer International Publishing. *https://doi.org/10.1007/978-3-031-26790-1_26*

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*

Haines, S., Krach, M., Pustaka, A., Li, Q. & Richman, L. (2019). The Effects of Computational Thinking Professional Development on STEM Teachers' Perceptions and Pedagogical Practices. *Athens Journal of Sciences*, *6*(2), 97-122. *https://doi.org/10.30958/ajs.6-2-2*

Herga, N R., Grmek, M.I. & Dinevski, D. (2014). Virtual laboratory as an element of visualization when teaching chemical contents in science class. *The Turkish Online Journal of Educational Technology*, *13*(4).

Hidayat, H., Zulhendra, Z., Efrizon, E., Delianti, V.I., Dewi, F.K., Mohd Isa, M.R., Harmanto, D., & Tanucan, J.C.M. (2024). Computational Thinking Skills in Engineering Education: Enhancing Academic Achievement Through Innovations, Challenges, and Opportunities. *TEM Journal*, 3.454-3.467. *https://doi.org/10.18421/TEM134-78*

Irvani, A.I., Rochintaniawati, D., Riandi, R. & Sinaga, P. (2024). Analyzing the Integration of Computational Thinking in Science and Physics Education within the Indonesian Curriculum. *Kasuari: Physics Education Journal (KPEJ)*, *7*(1), 182-194. *https://doi.org/10.37891/kpej.v7i1.620*

Jackson, S.L. (2009). *Research Methods and Statistics: A Critical Thinking Approach* (3rd ed.). Cengage Learning.

Jain, A.K., Duin, P.W. & Mao, J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(1), 4-37. *https://doi.org/10.1109/34.824819*

Joyce, B., Weil, M. & Calhoun, E. (2015). *Models of Teaching* (9th ed.). Pearson Education, Inc.

Kareem, A. (2022). Higher-Order Thinking Skills and Scientific Attitudes Components as Predictors of Scientific Creativity Among Preservice Biology Teachers. *International Journal of Progressive Education*, *18*(4), 21-30. *https://doi.org/10.29329/ijpe.2022.459.2*

Kilbane, C.R. & Natalie, B.M. (2014). *Teaching Models Designing Instruction for 21st Century Learners*. Pearson Education.

Kirschner, P.A., Sweller, J. & Clark, R.E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist*, *41*(2), 75-86. *https://doi.org/10.1207/s15326985ep4102_1*

Kong, S.-C., Chiu, M.M. & Lai, M. (2018). A study of primary school students' interest, collaboration attitude, and programming empowerment in computational thinking education. *Computers & Education*, *127*, 178-189. *https://doi.org/10.1016/j.compedu.2018.08.026*

Lawshe, C.H. (1975). A Quantitative Approach to Content Validity. *Personnel Psychology*, *28*, 563-575. *https://doi.org/http://dx.doi.org/10.1111/j.1744-6570.1975.tb01393.x*

Linacre, J.M. (1994). Sample Size and Item Calibration Stability. *Rasch Measurement Transactions*, *7*(4), 328.

Linacre, J.M. (2002). Optimizing Rating Scale Category Effectiveness. *Journal of Applied Measurement*, *3*(1), 85-106. *https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=5778427ee30e063c3af1108be44a0146814a503a*

Louangrath, P. (2014). Sample Size Determination for Non-Finite Population. *International Conference on Discrete Mathematics and Applied Sciences*. International Conference on Discrete Mathematics and Applied Sciences, Bangkok, Thailand. *https://doi.org/10.13140/2.1.4785.7286*

Louangrath, P.I. (2017). Minimum Sample Size Method Based on Survey Scales. *International Journal of Research & Methodology in Social Science*, *3*(3), 44-54. *https://doi.org/10.5281/ZENODO.1322593*

Lye, S.Y. & Koh, J.H.L. (2014). Review on teaching and learning of computational thinking through programming: What is next for K-12? *Computers in Human Behavior*, *41*, 51-61. *https://doi.org/10.1016/j.chb.2014.09.012*

Ma, J. & Nickerson, J.V. (2006). Hands-on, simulated, and remote laboratories: A comparative literature review. *ACM Computing Surveys*, *38*(3), 7. *https://doi.org/10.1145/1132960.1132961*

Makransky, G., Terkildsen, T.S. & Mayer, R.E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction*, *60*, 225-236. *https://doi.org/10.1016/j.learninstruc.2017.12.007*

Makransky, G., Thisgaard, M.W. & Gadegaard, H. (2016). Virtual Simulations as Preparation for Lab Exercises: Assessing Learning of Key Laboratory Skills in Microbiology and Improvement of Essential Non-Cognitive Skills. *PLOS ONE*, *11*(6), e0155895. *https://doi.org/10.1371/journal.pone.0155895*

Man, X., Zhu, X. & Sun, C. (2020). The Positive Effect of Workplace Accommodation on Creative Performance of Employees with and Without Disabilities. *Frontiers in Psychology*, *11*, 1.217. *https://doi.org/10.3389/fpsyg.2020.01217*

Mangal, S.K. & Mangal, S. (2019). *Assessment for Learning*. PHI Learning Private Limited.

Maratos, F.A., Chu, K., Lipka, S., Stupple, E.J.N. & Parente, F. (2023). Exploring pattern recognition: What is the relationship between the recognition of words, faces and other objects? *Cognitive Processing*, *24*(1), 59-70. *https://doi.org/10.1007/s10339-022-01111-3*

Marzano, R.J. & Kendall, J.S. (2007). *The New Taxonomy of Educational Objectives* (2nd ed.). Corwin PressAll.

Mayer, R.E. (2004). Should There Be a Three-Strikes Rule Against Pure Discovery Learning? *American Psychologist*, *59*(1), 14-19. *https://doi.org/10.1037/0003-066X.59.1.14*

McHugh, M.L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276-282. https://doi.org/10.11613/BM.2012.031

Merchant, C. (2018). *Science and Nature. Past, Present, and Future*. Routledge (Taylor & Francis Group).

Meyer, J.P. (2014). *Applied Measurement with jMetrik*. Routledge.

Mirolo, C., Izu, C., Lonati, V. & Scapin, E. (2021). Abstraction in Computer Science Education: An Overview. *Informatics in Education*, *20*(4), 615-639. *https://doi.org/10.15388/infedu.2021.27*

Misir, S.K. (2023). A Review of Computational Thinking: Focus on Science Education. *International Journal of Education Technology and Scientific Researches*, *24*. *https://doi.org/10.35826/ijetsar.701*

Montoya, M., Pablo-Lerchundi, I., Revuelta, F. & Almendros, P. (2023). Prácticas de laboratorio de Ampliación de Química en modalidad a distancia. *Revista de Investigación en Educación*, *21*(3), 500-515. *https://doi.org/10.35869/reined.v21i3.4984*

Navarro, D. & Foxcroft, D. (2019). *Learning Statistics with Jamovi: A Tutorial for Psychology Students and Other Beginners (Version 0.70)*. University of New South Wales.

Obilikwu, P., Kwaghtyo, K.D. & Udo, E.N. (2019). The Practicality of Engineering Principles in Software Engineering. *International Journal of Advanced Research*, *7*(12), 923-934. *https://doi.org/10.21474/IJAR01/10234*

Orts-Cortés, M.I., Moreno-Casbas, T., Squires, A., Fuentelsaz-Gallego, C., Maciá-Soler, L. & González-María, E. (2013). Content validity of the Spanish version of the Practice Environment Scale of the Nursing Work Index. *Applied Nursing Research*, *26*(4), e5–e9. *https://doi.org/10.1016/j.apnr.2013.08.006*

Parong, J. & Mayer, R.E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology*, *110*(6), 785-797. *https://doi.org/10.1037/edu0000241*

Rich, P.J., Egan, G. & Ellsworth, J. (2019). A Framework for Decomposition in Computational Thinking. *Proceedings of the 2019 ACM Conference on Innovation and Technology in Computer Science Education*, 416-421. *https://doi.org/10.1145/3304221.3319793*

Rutten, N., Van Joolingen, W.R. & Van Der Veen, J.T. (2012). The learning effects of computer simulations in science education. *Computers & Education*, *58*(1), 136-153. *https://doi.org/10.1016/j.compedu.2011.07.017*

Sadykova, O.V. & Il'bahtin, G.G. (2020). The Definition of Algorithmic Thinking. *Proceedings of the International Session on Factors of Regional Extensive Development (FRED 2019)*. International Session on Factors of Regional Extensive Development (FRED 2019), Irkutsk, Russia. *https://doi.org/10.2991/fred-19.2020.85*

Setyawarno, D., Rosana, D. & Kuswanto, H. (2025). 'Assessment as learning' in science instruction: A case study of Indonesian science teachers and notion for optimising in teaching and learning process. *Cogent Education*, *12*(1), *2538337*. *https://doi.org/10.1080/2331186X.2025.2538337*

Shaidullina, A.R., Orekhovskaya, N.A., Panov, E.G., Svintsova, M.N., Petyukova, O.N., Zhuykova, N.S. & Grigoryeva, E.V. (2023). Learning styles in science education at university level: A systematic review. *Eurasia Journal of Mathematics, Science and Technology Education*, *19*(7), em2293. *https://doi.org/10.29333/ejmste/13304*

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*

Smetana, L.K. & Bell, R.L. (2012). Computer Simulations to Support Science Instruction and Learning: A critical review of the literature. *International Journal of Science Education*, *34*(9), 1.337-1.370. *https://doi.org/10.1080/09500693.2011.605182*

Sneider, C., Stephenson, C., Schafer, B. & Flick, L. (2014). Computational Thinking in High School Science Classrooms: Exploring the Framework and NGSS. *The Science Teacher*, *81*(5), 53-59. *https://doi.org/10.2505/4/tst14_081_05_53*

Sudaryono, Rahardja, U., Aini, Q., Isma Graha, Y. & Lutfiani, N. (2019). Validity of Test Instruments. *Journal of Physics: Conference Series*, *1.364*(1), 012050. *https://doi.org/10.1088/1742-6596/1364/1/012050*

Threekunprapa, A. & Yasri, P. (2020). Unplugged Coding Using Flowblocks for Promoting Computational Thinking and Programming among Secondary School Students. *International Journal of Instruction*, *13*(3), 207-222. *https://doi.org/10.29333/iji.2020.13314a*

Veza, I., Sule, A., Putra, N.R., Idris, M., Ghazali, I., Irianto, Pendit, U.C., Mosliano, G. & Arasmatusy (2022). Virtual Laboratory for Engineering Education: Review of Virtual Laboratory for Students Learning. *Engineering Science Letter*, *1*(02), 41-46. *https://doi.org/10.56741/esl.v1i02.138*

Voogt, J., Fisser, P., Pareja Roblin, N., Tondeur, J. & Van Braak, J. (2013). Technological pedagogical content knowledge – a review of the literature. *Journal of Computer Assisted Learning*, *29*(2), 109-121. *https://doi.org/10.1111/j.1365-2729.2012.00487.x*

Walters, B., Potetz, J. & Fedesco, H.N. (2017). Simulations in the Classroom: An Innovative Active Learning Experience. *Clinical Simulation in Nursing*, *13*(12), 609-615. *https://doi.org/10.1016/j.ecns.2017.07.009*

Waugh, C.K. & Gronlund, N.E. (2013). *Assessment of Student Achievement*. Pearson Education.

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

Yanti, F.A. & Thohir, M.A. (2024). Higher order thinking skills in science learning: A systematic review from 2014-2023. *International Journal of Evaluation and Research in Education (IJERE)*, *13*(4), 2.419. *https://doi.org/10.11591/ijere.v13i4.28082*

Zacharia, Z.C. & Olympiou, G. (2011). Physical versus virtual manipulative experimentation in physics learning. *Learning and Instruction*, *21*(3), 317-331. *https://doi.org/10.1016/j.learninstruc.2010.03.001*