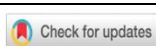


Assessing Higher-Order Thinking Skills in Physics: A Comparative Study of Paper-Based and Computer-Based Method

Rahman Eka Cahya^{1*}, Budi Jatmiko¹, Mita Anggaryani¹, Dwikoranto¹, Salma Hasna Hamiydah²

¹State University of Surabaya, Surabaya, Indonesia

²Al-Azhar University, Cairo, Egypt



DOI : <https://doi.org/10.63230/jdpe.1.2.41235>

Sections Info

Article history:

Submitted: June 1, 2025

Final Revised: June 21, 2025

Accepted: June 21, 2025

Published: July 12, 2025

Keywords:

21st Century Skills;

Computer-Based Test (CBT);

Higher-Order Thinking Skills

(HOTs);

Paper-Based Test (PBT);

Physics Assessment.

ABSTRACT

Objective: This research is a quasi-experimental investigation with a non-equivalent control group design, aiming to analyze the differences in the use of CBT and PBT methods for HOTs' question assessments. **Method:** This research employed a quantitative approach, conducting statistical tests using SPSS 26 on each of the 35 students by comparing the N-Gain results in the control group using PBT and the experimental group using CBT. Furthermore, normality tests, homogeneity tests, and independent sample-T tests were conducted to determine the significance of the differences between CBT and PBT results. **Results:** The N-Gain of the experimental group is 0.7652 higher than the N-Gain of the control group is 0.2272. Therefore, the CBT method is more effective than the PBT method. The data were distributed normally and homogeneously in both the control group and the experimental group, resulting in a significant difference. (2-tailed) in the independent sample-T test is $0.000 < 0.05$. It is concluded that there is a significant difference between the use of the PBT and CBT methods for assessing HOT questions. **Novelty:** This research compares the use of digital and conventional technology in physics learning, particularly to enhance students' cognitive domain in developing high-order thinking skills (HOTs), which is a crucial aspect of 21st-century education.

INTRODUCTION

21st-century education emphasizes the development of higher-order thinking skills (HOTs), which are the primary foundation for producing a critical, creative, and solution-oriented generation (Muliastirini, 2020). The main context of The 21st Century Skills focuses on creative thinking skills, critical thinking, problem-solving, communication, and collaboration, which students are not only required to memorize and understand the material conceptually but also to be able to analyze, evaluate, and create new ideas that are relevant to real situations (Nisa et al., 2022). Therefore, the development of HOTs is an integral part of learning and assessment at every level of education, especially at the middle and high school levels, which are crucial transition periods in the formation of scientific thinking (Martatiyani et al., 2023).

Physics is one of the subjects studied at the middle and high school levels, which is part of the Natural Sciences. It has characteristics that are ideal for training and developing students' HOT abilities (Choiroh et al., 2022). Physics material is not always focused on logical thinking to solve complex and abstract problems; furthermore, it aims to develop critical, creative, and innovative thinking, as well as communication skills, collaboration skills, and self-confidence after learning HOTs (Fransiska et al., 2021). Through physics learning, students aim to develop the ability to connect theory with actual phenomena in their surroundings. In addition, physics teaches not only scientific concepts but also instills critical and systematic thinking patterns, which are essential in 21st-century skills (Fananni, 2018).

The essential factors required to measure HOT's ability are determined by the assessment method used by teachers to succeed in the assessment processes (Istiyono,

2020). The assessment method is accustomed to having evaluation tools that are purposed to measure high-level cognitive aspects accurately and in-depth (Nurseha et al., 2021). Questions designed to measure HOTS must be open, contextual, and challenging and include stimuli that encourage students to analyze problems, generate alternative solutions, and design appropriate solutions (Fitriana & Sitompul, 2024). However, the effectiveness of these questions depends on the assessment instruments used. The assessment instruments connect both the questions and the students, which practically influences the way students understand, interpret, and answer the questions given (Miladanta et al., 2024).

Generally, there are two assessment methods commonly used in education: Paper-Based Tests (PBT) and Computer-Based Tests (CBT) (Ningsih & Kalamudin, 2022). PBT is a conventional form of assessment that utilizes paper-based tests and has been widely applied in the formal education system in Indonesia (Murni, 2024). PBT has advantages in terms of ease of use, does not require special devices, and is well known to most students and teachers (Hidayah, 2021). PBT does not rely too much on the readiness of the digital learning system, making it easier to access, especially in schools with limited technological facilities (Wardani, 2021).

In the digital era of teaching, the development of information technology has influenced the education system, transforming it into digital learning, a modern and efficient assessment alternative known as computer-based testing (CBT) (Mustari, 2023). CBT offers several advantages that increase its popularity, including faster and more accurate implementation and assessment processes, flexibility in presenting questions (such as interactive, visual, or simulation-based questions), and efficiency in resource utilization (Saidah, 2025). CBT provides the integration between technology and learning. Furthermore, CBT increases students' motivation through more interesting and dynamic learning outcomes (Amalia & Hadi, 2020).

However, the application of CBT faces some obstacles, such as the fact that not all students possess adequate digital literacy, which causes them to struggle with navigating the system, understanding digital instructions, or completing questions optimally (Ummah, 2019). Another obstacle is the readiness of hardware, network, and interface design that affects students' comfort and performance in digital learning (Ariani et al., 2023). The conventional method, which applies PBT, is easier to apply due to the efficiency of the processing results. However, some students argue that PBT is too monotonous and less supportive of the optimal visualization or exploration of questions (Sandra et al., 2022).

In the digital learning era, it is crucial to consider the appeal of CBT as a learning method. Several previous studies have indeed compared the technical effectiveness of CBT and PBT. However, few have specifically examined the effects on students' HOT achievement in physics subjects that require complex thinking (Aldalia, 2023). In addition, consideration factors such as students' work speed, stress levels, comfort while studying, and technological readiness should be compared as variables that affect the different results between the two methods (Annisak & Pathoni, 2017).

According to the background, it is necessary to research digital learning methods, comparing CBT and PBT integrated with HOTS' assessment in physics materials. The purpose of this research was to determine the extent to which differences in learning methods affect students' achievement outcomes and to identify the factors that may contribute to these differences. This study is expected to provide empirical data that can

be widely cited as a reference in making educational policies, especially in selecting the most appropriate and effective learning methods to measure HOTS as essentials in 21st-century skills.

RESEARCH METHOD

This research employed a quantitative approach in a quasi-experimental study using a non-equivalent control group design to compare the results of the HOTS learning outcomes between the control and experimental groups, conducted at Dr. Soetomo High School, Surabaya, in the even semester of the 2024/2025 academic year. The design flow of this research was adapted and modified from the study by Purwaningsih et al. (2020), as outlined in Table 1.

Table 1. Non-equivalent control group design

<i>Group</i>	<i>Pre-Test</i>	<i>Treatment</i>	<i>Post-Test</i>
X-1	O ₁	X ₁	O ₂
X-2	O ₃	X ₂	O ₄

Information:

- X-1 = control group
- X-2 = experimental group
- O₁ = control group pre-test
- O₂ = control group post-test
- O₃ = experimental group pre-test
- O₄ = experimental group post-test

This research aimed at all students in the first grade of senior high school. Two groups were selected using a purposive sampling method, namely X-1 and X-2, each consisting of 35 students. The X-1 group was treated with the PBT method, while the X-2 group was treated with the CBT method. The research samples were selected based on the same ability as referred to by the summative test outcomes. The outcomes of this research were obtained by analyzing the comparison between the learning outcomes of the conventional learning method using PBT and the digital learning method using CBT, which can be described in the form of a comparative flow diagram adapted from the research of Latifah and Suprihatiningrum (2024), as shown in Table 2.

Table 2. Comparative flow diagram.

Pre-Test (O ₁)	→	Treatment (X ₁)	→	Post-Test (O ₂)	} Comparison → Outcome
Pre-Test (O ₃)	→	Treatment (X ₂)	→	Post-Test (O ₄)	

The objective results were obtained by comparing several statistical tests of learning outcomes for both the CBT and PBT methods. The N-Gain test, normality test, homogeneity test, and independent sample t-test were conducted to assess changes in learning outcomes using both methods. The n-gain test was determined by the equation adapted from the research of Choiroh et al. (2020).

$$N - Gain = \frac{(score_{post-test} - score_{pre-test})}{(score_{ideal} - score_{post-test})} \quad (1)$$

According to the n-gain test equation, the next step is to determine the n-gain result range criteria, adjusted from Hake's research (1999).

Table 3. Criteria of the n-gain score range

n-Gain Score Range	Criteria
N-Gain Score > 0,7	High
0,3 < N-Gain Score < 0,7	Moderate
0,0 < N-Gain Score < 0,3	Low
N-Gain Score < 0,0	Failed

The n-gain effectiveness percentage criteria need to be determined as a benchmark for the learning outcomes of both the CBT method and the PBT method, according to Table 3 from Saraswati and Hatibe's (2023) research adaptation.

Table 3. Criteria of the percentage n-gain score range

Percentage of n-Gain Score Range	Criteria
n-Gain Score > 70%	High
30% < n-Gain Score < 70%	Moderate
0% < n-Gain Score < 30%	Low
n-Gain Score < 0%	Failed

The determination of the percentage n-gain score range used as references to explain the criteria for the effectiveness of the both learning methods. Referring to the research of Himmah et al. (2023), the n-gain score effectiveness criteria can be described as Table 4.

Table 4. Criteria of the percentage of n-gain score range

N-Gain Score Percentage	Criteria
0% - 20%	Not Effective
21% - 40%	Less Effective
41% - 60%	Quite Effective
61% - 80%	Effective
81% - 100%	Very Effective

Physics education experts have validated the HOT's physics research instruments. The instruments consisted of questions that had the same indicators, difficulty level, and material scope yet were presented through different methods. The data obtained were analyzed statistically to determine the distribution of learning outcomes between two student groups using an independent sample t-test. Prerequisite tests were conducted in the form of normality tests and homogeneity tests with the help of a statistical program (SPSS version 26) to test statistically the research hypothesis:

- H_0 : There is no significance difference of learning outcomes between using CBT method and PBT method.
- H_a : There is significance difference of learning outcomes between using CBT method and PBT method.

RESULTS AND DISCUSSION

Results

The learning outcomes from both using the CBT method and PBT methods for each control class and experimental class can be displayed in the graph Figure 1 and Figure 2.

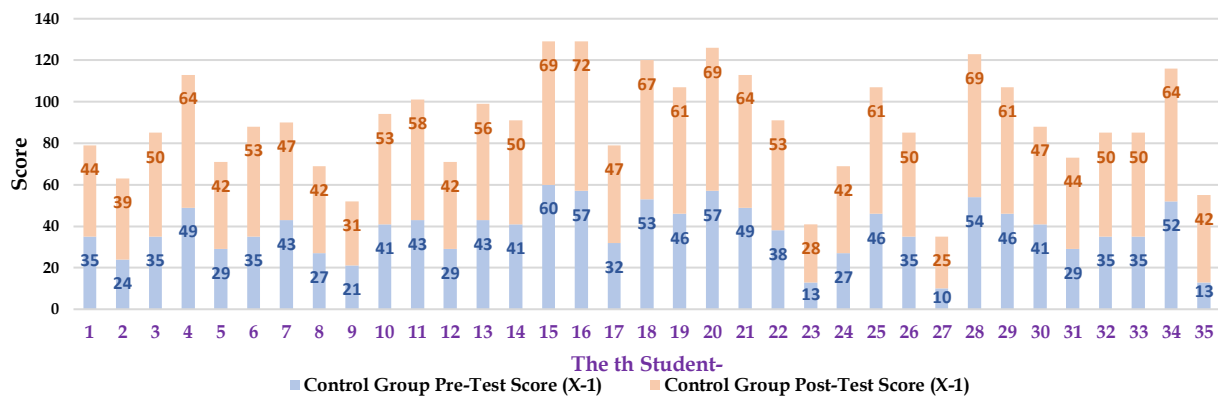


Figure 1. Pre-test and post-test results of control group (X-1)

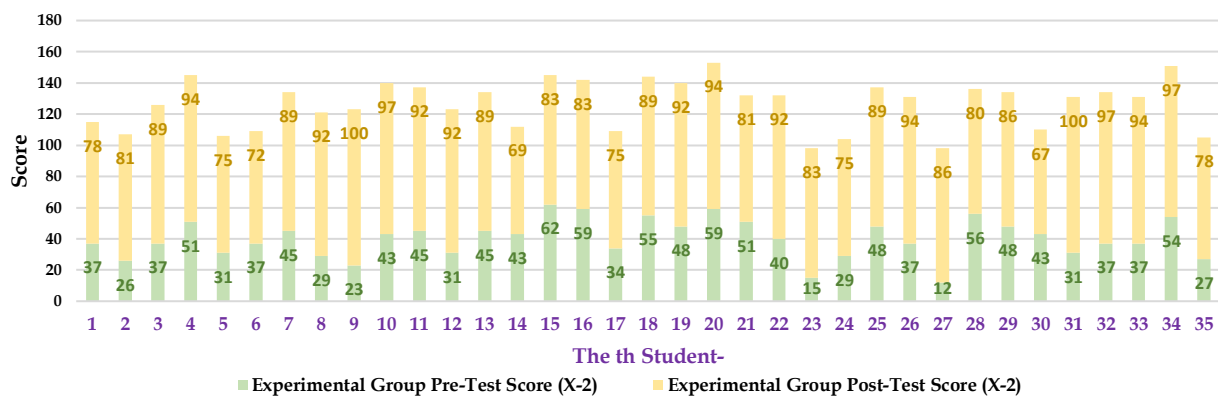


Figure 2. Pre-test and post-test results of experimental group (X-2)

According to the learning outcomes shown in Figures 1 and 2 above, it is necessary to determine the changes that occur with both methods by displaying the following Table 5.

Table 5. Description of learning outcomes on both group

Learning Outcomes	N	Min.	Max.	Mean
Pre-Test Control Group	35	10	60	37,80
Post-Test Control Group	35	25	72	51,60
Pre-Test Experimental Group	35	12	62	40,14
Post-Test Experimental Group	35	67	100	86,40

The learning outcomes in Table 5 above indicate the results of the experimental group, which obtained minimum and maximum scores of 12 and 62 in the pre-test, with an average score of 40.14. In contrast, the minimum and maximum scores in the post-test were 67 and 100, with an average score of 86.40. Different results were obtained in the control group, with a minimum and maximum score of 10 and 60 in the pre-test and an average score of 37.80. In contrast, the minimum and maximum scores in the post-test were 25 and 72, with an average score of 51.60.

An additional display of comparison for learning outcomes of both methods is also required through the N-Gain Table 6 and Table 7.

Table 6. n-gain score of control group

	N	Min.	Max.	Mean	Std. Deviation
Control Group N-Gain Score	35	0,0702	0,3488	0,2272	0,06461
Effectivity Percentage of Control Group N-Gain Score	35	7	35	22,72	6,461

Table 7. n-gain score of experimental group

	N	Min.	Max.	Mean	Std. Deviation
Experimental n-Gain Score	35	0,4211	1,000	0,7652	0,15446
Effectivity Percentage of Experimental Group n-Gain Score	35	42	100	76,52	15,446

According to Table 6 and Table 7, it can be inferred that the practical learning outcomes of the CBT method reach 76.52%, which is greater than those of the PBT method. Furthermore, according to the data in Table 6 and Table 7, the frequency of each N-Gain score range is described in Table 8 and Table 9.

Table 8. Frequency of n-gain score effectivity of control group

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid: 0% - 20% (Not Effective)	10	7,1	28,6	28,6
21% - 40% (Less Effective)	25	17,9	71,4	100,0

Table 9. Frequency of n-gain score effectivity of experimental group

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid: 41% - 60% (Quite Effective)	6	4,3	17,1	17,1
61% - 80% (Effective)	13	9,39	37,1	54,3
81% - 100% (Very Effective)	16	11,4	45,7	100,0

The data shown in Table 8 and Table 9 are then visualized in the form of a frequency category graph of the effectiveness of the n-Gain score for both the control group and the experimental group, as shown in Figures 3 and 4.

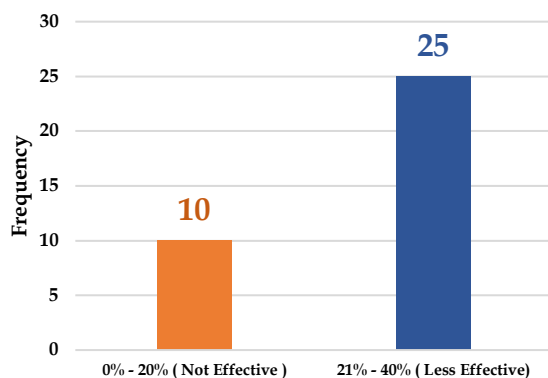


Figure 3. Frequency of n-gain score effectivity of control group

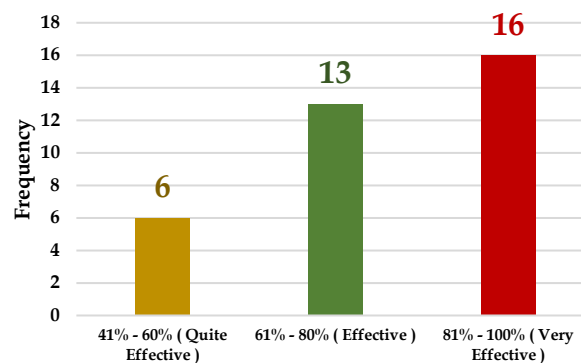


Figure 4. Frequency of n-gain score effectivity of experimental group

The data from Figures 3 and 4 above indicate that 29 students in the experimental group achieved a significant increase in learning outcomes, as supported by the results, with n-gain scores categorized as “effective” and “very effective.” In contrast, in the control group, there was no significant increase in learning outcomes for all students, proved by the n-gain scores categorized as “not effective” and “less effective” (Adapted from Himmah et al., 2023).

Discussion

The learning outcomes, as indicated in Table 5, show that the CBT method and the PBT method have had different impacts on the results of students' HOT assessments. Other findings indicate that for 35 students in each group, the learning outcomes using the CBT method were higher than those using the PBT method. According to Tables 6 and 7, other findings emerged that the n-gain score of the control group was 0.2272, which falls within the low category, while the n-gain score of the experimental group was 0.7652, which falls within the high category (Hake, 1999).

The minimum n-gain score in the control group was 0.0702, while in the experimental group, it was 0.4211. The maximum n-gain score in the control group was 0.3488, whereas in the experimental group, it was 1.0000. The n-gain percentage in the control group was 22.72%, which falls within the low category. In contrast, the experimental group's rate was 76.52%, which falls within the high category (Saraswati & Hatibe, 2023).

According to the data on n-gain scores obtained through the CBT method and the PBT method, it is necessary to compare the level of significance of the differences between the two methods using the independent sample t-test, as well as to determine whether the final hypothesis (H_0) is accepted or rejected. The first requirement is to conduct a data normality test, as presented in Table 10.

Table 10. Normality test on learning outcomes on CBT method and PBT method

	Group	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Results:	Control Group Pre-Test	0,099	35	0,200	0,972	35	0,492
	Control Group Post-Test	0,099	35	0,200	0,965	35	0,314
	Experimental Group Pre-Test	0,087	35	0,200	0,980	35	0,747
	Experimental Group Post-Test	0,158	35	0,027	0,953	35	0,139

According to Table 10, it was found that all significant differences were observed. $> 0,05$. Furthermore, the pre-test significance is 0.747 for the PBT method, while the post-test significance is 0.139. In contrast, the pre-test significance is 0.492, and the post-test significance is 0.314 for the CBT method. Those findings indicate that the learning outcomes for both methods are typically distributed across the population. According to prior research by Pora et al. (2023), if sig. Value $> 0,05$, including normally distributed. Therefore, it can be continued to analyze the data homogeneity test as presented in the following Table 11.

Table 11. Homogeneity test based on normality test results

		Levene Statistic	df ₁	df ₂	Sig.
Results:	Based on Mean	1,259	3	136	0,291

According to Table 11, the sig. The value based on the arithmetic mean is 0.291, which is greater than 0.05. These findings indicate that both the pre-test and post-test results for both the CBT method and the PBT method are homogeneous. Therefore, the results of sig. $> 0,05$ and corroborated by similar prior research by Mardiyanti & Jatmiko (2022), which stated that a sig. Value > 0.05 , including homogeneous data, allows for an independent sample t-test to be carried out, as presented in the following Table 12.

Table 12. Independent sample-t test based on homogeneity test results

	Levene's Test for Equality of Variances:		T-Test for Equality of Means:						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Results: Equal variances assumed	0,116	0,734	-4,676	68	0,000	-13,800	2,951	-19,689	-7,911
Equal variances not assumed			-4,676	67,728	0,000	-13,800	2,951	-19,690	-7,910

The sig. (2-tailed) According to Table 12, the value is 0.000, which is smaller than 0.05, statistically supporting the rejection of H_0 and the acceptance of H_a . This result is corroborated by prior research by Safitri et al. (2022), which stated that if sig. A value of $p < 0.05$ indicates a significant influence between variables, and H_a is accepted. Therefore, there is a significant difference in learning outcomes between using the CBT method and the PBT method. This result is reinforced by similar research conducted by Kurniawan et al. (2024), which states that the application of the CBT method for assessment can significantly improve student learning outcomes compared to the conventional method (PBT), as indicated by the results of the pre-test and post-test in both the control group and the experimental group.

In addition, the application of the CBT method in digital learning can have a positive impact on students, making it easier for them to work on and understand questions. The results of this research are reinforced by Tanjung et al. (2021), who stated that the CBT method used for learning makes it easier for students to work on questions because the appearance of the questions is more attractive and can be developed well.

CONCLUSION

Fundamental Finding: This research provides strong empirical evidence that the Computer-Based Test (CBT) method results in significantly higher outcomes in assessing higher-order thinking skills (HOTs) in physics learning compared to the Paper-Based Test (PBT) method. HOTs, as defined by Kusumaningtyas et al. (2024), encompass advanced cognitive processes, including analysis, evaluation, and problem-solving. The CBT approach effectively supports these demands by offering interactive tasks that promote deeper cognitive engagement. These findings reinforce previous studies, including Sulistyarini (2022), which emphasize the advantages of digital learning platforms in fostering higher-order cognitive development. **Implication:** The CBT method enables more authentic and interactive assessment designs, encouraging students to visualize, simulate, analyze, and evaluate complex problems. Its integration into physics education can enhance students' engagement and better align with 21st-century learning goals. **Limitation:** This study used a quasi-experimental design without random assignment, which may affect internal validity. Moreover, differences in students' familiarity and competence with digital tools may have introduced confounding variables that affected the results. **Future Research:** Further studies should involve randomized controlled trials with larger sample sizes and explore the impact of CBT-based assessments across different subjects and educational levels. Research could also investigate how digital literacy training influences the effectiveness of CBT in promoting HOTs.

AUTHOR CONTRIBUTIONS

Rahman Eka Cahya: Conceptualization, Methodology, Investigation, Formal Analysis, Writing – Original Draft, Visualization; **Budi Jatmiko:** Supervision, Validation, Writing – Review & Editing, Project Administration; **Mita Anggaryani:** Methodology, Resources, **Dwikoranto:** Validation, Writing – Review & Editing; **Salma Hasna Hamiyda:** Investigation, Writing – Original Draft. All authors have read and approved the final version of this manuscript.

DECLARATION OF COMPETING INTEREST

The authors declare no known financial conflicts of interest or personal relationships that could have influenced the work reported in this manuscript.

DECLARATION OF ETHICS

The authors declare that the research and writing of this manuscript adhere to ethical standards of research and publication, in accordance with scientific principles, and are free from plagiarism.

DECLARATION OF ASSISTIVE TECHNOLOGIES IN THE WRITING PROCESS

The authors declare that generative artificial intelligence (Gen AI) and other AI-assisted tools were used prudently, not excessively, during the research and preparation of this manuscript. Specifically, ChatGPT was used for drafting analytical explanations, refining paragraph structure, and checking conceptual clarity; Grammarly for grammar and style correction; and ChatPDF for summarizing reference materials and extracting key findings from cited papers. All AI-generated material was reviewed and edited for accuracy, completeness, and compliance with ethical and scholarly standards. The authors accept full responsibility for the final content of the manuscript.

REFERENCES

- Aldilla, D. I. (2023). Gambaran hasil evaluasi proses belajar mahasiswa menggunakan metode paper based test dan computer based test. *Jurnal Medika Utama*, 4(4), 3546-3555. <https://doi.org/10.53039/jmh.v4i04Juli.661>
- Amalia, D., & Hadi, W. (2020). Analisis kesalahan siswa dalam menyelesaikan soal HOTS berdasarkan kemampuan penalaran matematis. *Transformasi: Jurnal Pendidikan Matematika Dan Matematika*, 4(1), 219-236. <https://doi.org/10.36526/tr.v4i1.904>
- Annisak, W., & Pathoni, H. (2017). Desain pengemasan test diagnostik miskonsepsi berbasis CBT (Computer Based Test). *Edufisika: Jurnal Pendidikan Fisika*, 2(1), 1-12. <https://doi.org/10.22437/edufisika.v2i01.3939>
- Ariani, M., Zulhawati, Z., Haryani, H., Zani, B. N., Husnita, L., Firmansyah, M. B., ... & Hamsiah, A. (2023). *Penerapan media pembelajaran era digital*. PT. Sonpedia Publishing Indonesia.
- Choiroh, S. S., Prastowo, S. H. B., & Nuraini, L. (2022). Pengaruh penggunaan e-LKPD interaktif berbantuan live worksheets terhadap kemampuan berpikir kognitif HOTS fisika siswa SMA. *Jurnal Ilmiah Pendidikan Fisika*, 6(3), 694-705. <https://doi.org/10.20527/jipf.v6i3.6795>
- Fanani, M. Z. (2018). Strategi pengembangan soal HOTS pada kurikulum 2013. *Edudeena: Journal of Islamic Religious Education*, 2(1), 57-76. <https://doi.org/10.30762/ed.v2i1.582>
- Fitriana, F., & Sitompul, S. S. (2024). Analisis kemampuan kognitif peserta didik dalam menyelesaikan soal HOTS fisika materi getaran harmonis. *Jurnal Dunia Pendidikan*, 5(2), 553-565. <https://doi.org/10.55081/jurdip.v5i2.2893>
- Fransiska, A., Prasetyo, E., & Jufriansah, A. (2021). Desain LKPD fisika terintegrasi HOTS untuk meningkatkan kemampuan berpikir kritis peserta didik. *Jurnal Pendidikan Fisika Dan Teknologi*, 7(2), 153-158. <http://dx.doi.org/10.29303/jpft.v7i2.3098>
- Hake, R. R. (1999). *Analyzing change/gain scores*. Dept. of Physics, Indiana University.
- Hidayah, A. S. (2021). *Efektivitas penggunaan Google Form pada penilaian harian aspek kognitif mata pelajaran Al Qur'an Hadis di MTs Ma'arif Nu 1 Cilongok Kabupaten Banyumas* [Master's thesis, Institut Agama Islam Negeri Purwokerto (Indonesia)].
- Himmah, F., Rofi'i, R. I., & Wiyarno, Y. (2023). Pengembangan aplikasi asesmen diagnostik berbasis Computer Based Test (CBT) menggunakan moodle. *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, 8(3), 1022-1032. <https://doi.org/10.29100/jupi.v8i3.4380>
- Istiyono, E., Dwandaru, W. S. B., Erfianti, L., & Astuti, W. (2020). Applying CBT in physics learning to measure students' higher order thinking skills. *Journal of Physics:*

- Conference Series*, 1440(1), 012061. IOP Publishing. <https://doi.org/10.1088/1742-6596/1440/1/012061>
- Kurniawan, E. S., Mundilarto, E. I., & Istiyono, E. (2024). Improving student higher order thinking skills using Synectic-HOTS-oriented learning model. *International Journal of Evaluation and Research in Education*, 13(2), 1133-1140. <https://doi.org/10.11591/ijere.v13i2.25002>
- Kusumaningtyas, D. A., Manyunu, M., Kurniasari, E., Awal, A. N., Rahmaniati, R., & Febriyanti, A. (2024). Enhancing learning outcomes: A study on the development of higher order thinking skills based evaluation instruments for work and energy in high school physics. *Indonesian Journal on Learning and Advanced Education (IJOLAE)*, 6(1), 14-31. <https://doi.org/10.23917/ijolae.v6i1.23125>
- Latifah, U. N., & Suprihatiningrum, J. (2024). The effect of inquiry-based learning on students' critical thinking ability and activeness in reaction rate material. *Lectura: Jurnal Pendidikan*, 15(1), 95-106. <https://doi.org/10.31849/lectura.v15i1.17368>
- Mardiyanti, N. E. A., & Jatmiko, B. (2022). Keefektifan pembelajaran fisika dengan model inkuiri terbimbing berbantuan PhET Interactive Simulations untuk meningkatkan kemampuan berfikir kritis siswa SMA. *Jurnal Ilmiah Pendidikan Fisika*, 6(2), 328. <https://doi.org/10.20527/jipf.v6i2.5281>
- Martatiyana, D. R., Derlis, A., Aviarizki, H. W., Jurdil, R. R., Andayani, T., & Hidayat, O. S. (2023). Analisis komparasi implementasi kurikulum merdeka dan kurikulum 2013. *Muallimuna: Jurnal Madrasah Ibtidaiyah*, 9(1), 96-109. <https://doi.org/10.31602/muallimuna.v9i1.11600>
- Miladanta, A. N., Nuryantini, A. Y., Farida, I., & Cahyanto, T. (2024). Pengembangan Instrumen Higher Order Thinking Skills (HOTS) materi alat optik melalui validitas, reliabilitas, daya pembeda dan tingkat kesukaran. *Jurnal Penelitian Sains dan Pendidikan (JPSP)*, 4(2), 148-158. <https://doi.org/10.23971/jpsp.v4i2.7828>
- Muliastri, N. K. E. (2020). New literacy sebagai upaya peningkatan mutu pendidikan sekolah dasar di abad 21. *Jurnal Pendidikan Dasar Indonesia (PENDASI)*, 1(4), 115-125. <https://doi.org/10.23887/jpdi.v4i1.3114>
- Murni, S. (2024). *Manajemen penilaian hasil belajar berbasis komputer dalam menguatkan kompetensi profesional guru di MIN 2 Kota Bengkulu* [Doctoral dissertation, UIN Fatmawati Sukarno Bengkulu].
- Mustari, M. (2023). *Teknologi informasi dan komunikasi dalam manajemen pendidikan*. Gunung Djati Publishing Bandung.
- Ningsih, I. S., & Kamaludin, A. (2022). Development of textbooks with Higher Order Thinking Skills (HOTS) on colligative properties of solutions. *Jurnal Pendidikan dan Pembelajaran Kimia*, 11(3), 38-47. <https://doi.org/10.23960/jppk.v11.i3.2022.05>
- Nisa, H., Junus, M., & Komariyah, L. (2022). Penerapan model problem based learning berbantuan simulasi PhET berbasis instrumen HOTS terhadap hasil belajar siswa. *Jurnal Ilmiah Pendidikan Fisika*, 6(3), 560-567. <https://doi.org/10.20527/jipf.v6i3.5514>
- Nurseha, M. L. E., Budiono, A. N., & Wahyuni, W. (2021). Pengaruh ujian ulangan harian menggunakan computer based test terhadap motivasi belajar siswa. *Pedagogika*, 21-31. <https://doi.org/10.37411/pedagogika.v12i1.623>
- Pora, L. L., Habibi, & Sukroyanti, B. A. (2023). Pengaruh penerapan model pembelajaran discovery learning untuk meningkatkan hasil belajar fisika siswa kelas X. *Journal Transformation of Mandalika*, 4(4), 105-115. <https://doi.org/10.36312/jtm.v4i4.1268>

- Purwaningsih, E., Sari, S. P., Sari, A. M., & Suryadi, A. (2020). The effect of STEM-PJBL and discovery learning on improving students' problem-solving skills of impulse and momentum topic. *Jurnal Pendidikan IPA Indonesia*, 9(4), 465-476. <https://doi.org/10.15294/jpii.v9i4.26432>
- Safitri, H., Agatha, F. L., Syiarah, H., & Simamora, N. N. (2022). Analisis pengaruh minat pembelajaran fisika terhadap hasil belajar fisika kelas X di SMAN 4 Kota Jambi. *Integrated Science Education Journal*, 3(2), 55-61. <https://doi.org/10.37251/isej.v3i2.258>
- Saidah, I. S. (2025). *Implementasi pembelajaran pendidikan agama islam dan budi pekerti berbasis media digital dalam meningkatkan kualitas pembelajaran di SD Islam Khalifah Palu* [Doctoral dissertation, Universitas Islam Negeri Datokarama Palu].
- Sandra, R. O., Maison, M., & Kurniawan, D. A. (2022). Pengembangan instrument miskonsepsi menggunakan dreamweaver berbasis web pada materi tekanan berformat five-tier. *Jurnal Fisika: Fisika Sains Dan Aplikasinya*, 7(1), 22-28. <https://doi.org/10.35508/fisa.v7i1.6575>
- Saraswati, S., & Hatibe, A. (2023). Pembelajaran modul fisika berbasis inkuiri di masa pandemi di kelas XI SMA Negeri 1 Tinombo. *JPFT (Jurnal Pendidikan Fisika Tadulako Online)*, 11(3), 136-141. <https://doi.org/10.22487/jpft.v11i3.2110>
- Sulistiyarini, W. (2022). Analysis of cognitive aspects of test techniques in islamic education learning. *EDUKASI: Jurnal Pendidikan Islam*, 10(2), 166-190. <https://doi.org/10.54956/edukasi.v10i2.278>
- Tanjung, Y. I., Wulandari, D., Bakar, A., & Ramadhani, I. (2021). The development of online physics test system at SMA CT Foundation Medan. *Journal of Physics: Conference Series*, 1819(1), 012054. IOP Publishing. <https://doi.org/10.15294/jpfi.v19i1.41745>
- Ummah, K. K. (2019). *Pengaruh ujian system CBT (Computer Based Test) mata pelajaran PAI terhadap motivasi belajar siswa di SMAN 10 Surabaya* [Undergraduate thesis, UIN Sunan Ampel Surabaya].
- Wardani, S. U. K. (2021). Efektivitas penggunaan sistem computer based test dan paper based test dalam pelaksanaan ujian tengah semester Bahasa Indonesia di SMPN 6 Singaraja. *Jurnal Pendidikan Bahasa Dan Sastra Indonesia Undiksha*, 11(4), 491-500. <https://doi.org/10.23887/jjpbs.v11i4.39676>

***Rahman Eka Cahya (Corresponding Author)**

Master of Physics Education, Faculty of Mathematic and Natural Sciences,
State University of Surabaya
Jl. Ketintang, Surabaya 60321, Indonesia, Kampus Ketintang Unesa, Gedung C3
Lantai 1
Email: 24031635005@mhs.unesa.ac.id

Budi Jatmiko

Department of Physics, Faculty of Mathematic and Natural Sciences,
State University of Surabaya
Jl. Ketintang, Surabaya 60321, Indonesia, Kampus Ketintang Unesa, Gedung C3
Lantai 1
Email: budijatmiko@unesa.ac.id

Mita Anggaryani

Department of Physics, Faculty of Mathematic and Natural Sciences,
State University of Surabaya
Jl. Ketintang, Surabaya 60321, Indonesia, Kampus Ketintang Unesa, Gedung C3
Lantai 1
Email: mitaanggaryani@unesa.ac.id

Dwikoranto

Department of Physics, Faculty of Mathematic and Natural Sciences,
State University of Surabaya
Jl. Ketintang, Surabaya 60321, Indonesia, Kampus Ketintang Unesa, Gedung C3
Lantai 1
Email: dwikoranto@unesa.ac.id

Salma Hasna Hamiydah

Al-Azhar University
Youssef Abbas, Gameat Al Azhar, Qesm Thani, Nasr City, Cairo, Egypt
Email: salma.isna2145@gmail.com
