

Enhancing Students' Critical Thinking in Physics through Discovery Learning-Based Worksheets: Integrating Facione's Framework in a Quasi-Experimental Study

Ananda Mulyana¹⁾, Susilawati^{1),2),*}, Susanna¹⁾

¹⁾Department of Physics Education, Teacher Training and education Faculty, Universitas Syiah Kuala,

²⁾Science, Technology, Engineering, and Mathematics (STEM) Research Center, Universitas Syiah Kuala

*Corresponding Author: susila@usk.ac.id

ABSTRAK

This study aims to investigate the effect of discovery learning-based student worksheets on improving students' critical thinking skills in physics, particularly on the topic of straight motion. A quasi-experimental design with a non-equivalent control group was employed, involving 66 eleventh-grade students divided into an experimental group ($n = 34$) and a control group ($n = 32$), selected through purposive sampling. Data were collected using a set of 20 validated multiple-choice items designed to measure critical thinking skills. The results reveal that the experimental group achieved a higher mean posttest score ($M = 75.64$) compared to the control group ($M = 71.91$), with a moderate N-gain value (0.57). Statistical analysis using the independent samples t-test indicates a significant difference between the two groups ($p = 0.01 < 0.05$), confirming the effectiveness of the intervention. Notably, the highest improvement was observed in the inference indicator (91.18%), suggesting that the Discovery Learning approach particularly enhances students' ability to draw logical conclusions. This study contributes an empirically-tested worksheet model that explicitly maps each discovery learning stage to a specific Facione critical thinking indicator. The differentiated N-Gain outcomes observed per indicator (ranging from 0.27 for self-regulation to 0.84 for inference) serve as empirical evidence of the mapping's differential instructional effect, providing measurable diagnostic value beyond a holistic improvement score. It should be noted that the causal contribution of each individual worksheet stage to each specific indicator has not been directly tested; the observed differentiation represents a theoretically-grounded design inference supported by indicator-level outcome patterns. It is recommended that educators adopt discovery learning-based instructional materials that are explicitly aligned with critical thinking indicators to optimize students' higher-order cognitive skills. Future research should examine the long-term impact of such interventions and explore their applicability across different scientific domains and educational contexts.

Keywords: Discovery Learning; Critical Thinking Skills; Physics Education; Student Worksheets; Facione Framework

This is an open access article under the CC-BY license.



INTRODUCTION

Physics learning plays a strategic role in developing students' scientific competencies in the era of Industry 4.0 and Society 5.0, which are characterized by rapid technological advancement. The primary goal of physics learning encompasses not only mastery of scientific concepts but also the cultivation of critical thinking skills that equip students to face real-world challenges (Azizah et al., 2021). Critical thinking skills are directed and reflective cognitive processes used to evaluate evidence, analyze arguments, and formulate logical conclusions (Facione, 2015). The 2022 Programme for International Student Assessment (PISA) study reveals that Indonesian students' science competencies particularly scientific reasoning and problem-solving still require substantial improvement (OECD, 2023). This situation reflects a systemic issue in classroom practice: the continued dominance of teacher-centered, one-way information transfer that prevents students from actively constructing their own knowledge (Laili et al., 2022).

Initial observations conducted at an eleventh-grade high school in Banda Aceh revealed that physics learning on the topic of linear motion relies heavily on conventional lecture methods and procedural worksheets. Students are not encouraged to independently identify problems, collect and process data, or draw empirically grounded conclusions. Empirical evidence from school administrative records further supports this concern: the average daily quiz score for eleventh-grade physics was 58.3, which falls below the minimum

competency criterion (KKM) of 70, indicating a systemic gap in students' higher-order thinking skills prior to any intervention. As a result, critical thinking indicators such as interpretation, analysis, inference, and self-regulation have not developed optimally (Novayani, 2021; Kusprayanto et al., 2022).

This problem is not solely attributed to students' cognitive limitations, but rather to learning approaches that do not encourage active discovery. The dominance of teacher-centered patterns with lecture methods and the use of conventional, algorithm-based worksheets that only require students to follow procedures without conceptual understanding are major obstacles to the development of critical thinking skills (Wahyu et al., 2021; Fuadi et al., 2022). The eleventh-grade physics teacher confirmed that learning was still teacher-centered, students tended to be passive, and they had difficulty analyzing physics problems based on everyday phenomena.

One promising solution to address these challenges is the implementation of discovery learning-based worksheets. Discovery learning is a cognitive learning model that encourages students to actively construct scientific concepts through structured investigative activities (Hosnan, 2022). This model consists of six systematic stages: stimulation, problem statement, data collection, data processing, verification, and generalization, each of which has the potential to simultaneously train multiple dimensions of critical thinking skills. The integration of structured worksheets in the discovery learning process is considered a strategic approach for directing students' thinking systematically. Worksheets designed based on discovery learning principles not only serve as learning activity guides but also function as instruments that explicitly train specific critical thinking competencies at each stage of investigation (Ardiansyah et al., 2021). When each stage of the worksheet is explicitly linked to a particular cognitive indicator, students receive more targeted and measurable scaffolding compared to conventional worksheets that merely present sequential procedures without fostering conceptual understanding (Amaniyah & Sugiarto, 2021).

Facione (2015) identifies six core critical thinking indicators: interpretation, analysis, evaluation, inference, explanation, and self-regulation. These indicators are hierarchically interconnected and collectively represent the full spectrum of higher-order cognitive competencies. In the context of physics learning, these indicators are particularly relevant because physics problems require students not only to recall facts, but also to interpret phenomena, construct logical arguments, evaluate data validity, and draw empirically grounded conclusions (Ekawati & Ilyas, 2022). Despite the theoretical alignment between discovery learning stages and Facione's indicators, empirical studies that explicitly operationalize this mapping through structured worksheets in high school physics remain limited, which forms the central gap this study addresses. Crucially, prior studies that do incorporate both discovery learning and critical thinking assessment tend to measure only global thinking scores without specifying which worksheet stage drives improvement in which indicator – a diagnostic granularity that is absent from the existing literature (Sapitri et al., 2021; Nurrohmi et al., 2022).

Previous studies have demonstrated the effectiveness of discovery learning in improving critical thinking skills across various subjects (Sapitri et al., 2021; Aditya, 2021; Fahaluddin et al., 2021; Nurrohmi et al., 2022; Ristiasari et al., 2023; Zulyusri et al., 2023). However, studies that explicitly map all six stages of discovery learning to the six indicators of Facione's critical thinking framework through structured worksheets in the context of high school physics remain scarce. A systematic review by Zulyusri et al. (2023) covering 47 discovery learning studies found that none operationalized all six Facione indicators within a single worksheet instrument. Similarly, Fadillah et al. (2025) and Rahmawati et al. (2026) identified this explicit one-to-one stage-indicator mapping as a critical gap. Unlike Sapitri et al. (2021) who applied discovery learning without explicit indicator mapping, and Nurrohmi et al. (2022) who addressed only three of the six indicators, this study operationalizes all six DL stages in direct alignment with all six Facione indicators in a single integrated worksheet. The novelty of this study lies in three aspects: (1) the explicit one-to-one alignment of each discovery learning stage with a specific Facione critical thinking indicator, validated empirically through differentiated per-indicator N-Gain outcomes; (2) the use of multiple-choice questions with reasoning that are more sensitive in measuring critical thinking; and (3) the implementation context of linear motion, a topic typically taught procedurally, demonstrating the model's applicability to procedural physics topics.

Based on the above background, this study aims to: (1) describe the profile of students' critical thinking skills before and after treatment; (2) examine differences in critical thinking improvement between classes using discovery learning-based worksheets and those using the direct instruction model; and (3) analyze the

contribution of each critical thinking indicator in both groups. This study is expected to provide both theoretical and practical contributions to the design of 21st-century physics learning.

METHOD

This study employed a quantitative approach with a quasi-experimental design (non-equivalent control group type). This design was chosen because full randomization was not feasible, as class groupings within the school were determined institutionally (Anjarsari & Prasetyo, 2022). The study was conducted at a high school in Banda Aceh, Aceh Province.

The independent variable was the discovery learning-based worksheet, while the dependent variable was students' critical thinking skills, measured using reasoned multiple-choice questions. The experimental class received discovery learning-based worksheets, while the control class used a direct instruction approach with conventional worksheets.

The study population comprised all eleventh-grade students at the school. The sample was selected using purposive sampling based on equivalence of initial abilities determined from prior semester grades. Both classes demonstrated comparable academic profiles: similar mean semester scores (experimental class $M=71.2$, control class $M=70.8$), comparable gender composition (experimental: 18 female/16 male; control: 17 female/15 male), and similar socioeconomic backgrounds as verified through school administrative data. These characteristics minimize potential selection bias inherent in purposive sampling. To further strengthen internal validity, the pretest was used as an additional check of group equivalence prior to treatment; the non-significant pretest t -test result ($p = 0.079 > 0.05$) confirmed that the two groups were statistically equivalent at baseline. The sample consisted of an experimental class ($n = 34$) and a control class ($n = 32$), totaling 66 students.

Data were collected using a critical thinking skills test comprising 20 multiple-choice items with reasoning. The distribution of items per indicator was as follows: Interpretation (items 1–3, $n=3$), Analysis (items 4–7, $n=4$), Evaluation (items 8–11, $n=4$), Inference (items 12–15, $n=4$), Explanation (items 16–18, $n=3$), and Self-Regulation (items 19–20, $n=2$). These items covered six indicators according to Facione (2015): interpretation, analysis, evaluation, inference, explanation, and self-regulation. Before use, the instrument was piloted in a class outside the sample. Validity was assessed using point-biserial correlation (all items valid, $r > 0.30$), and reliability was tested using the Kuder-Richardson formula ($KR-20 = 0.78$, categorized as high).

The discovery learning-based worksheet was implemented across two meetings (4×45 minutes). This duration was determined by the school's official lesson schedule for the linear motion topic and represents a real-world classroom constraint; it is acknowledged that two sessions may be insufficient for fully developing higher-order indicators such as self-regulation, which requires sustained metacognitive practice over time. Each meeting followed six stages in sequence: (1) Stimulation the teacher presented straight motion phenomena through videos and demonstrations; (2) Problem Statement students formulated questions and hypotheses; (3) Data Collection students conducted measurements and recorded observations; (4) Data Processing students analyzed data in groups; (5) Verification students presented and verified their findings; (6) Generalization students drew conclusions and generalized concepts.

Data analysis was conducted in three stages. First, descriptive analysis of pretest, posttest, and N-Gain scores for both classes. N-Gain was calculated using the Hake (1999) formula: $g = (S_{\text{post}} - S_{\text{pre}}) / (S_{\text{max}} - S_{\text{pre}})$, with categories: high ($g > 0.70$), medium ($0.30 \leq g \leq 0.70$), and low ($g < 0.30$). Second, prerequisite tests included a normality test using the Shapiro-Wilk test and a homogeneity test (Fisher's exact test). The Shapiro-Wilk test was selected because it is more statistically powerful than the Chi-Square goodness-of-fit test for small samples ($n < 50$), making it the most appropriate choice for the present sample sizes of $n = 32$ and $n = 34$ (Field, 2018). Third, hypothesis testing was conducted using an independent samples t -test at $\alpha = 0.05$.

RESULTS AND DISCUSSION

Pretest and Posttest Results of Critical Thinking Skills

Table 1 presents the descriptive statistics of pretest and posttest scores for both classes. The pretest mean scores for the control class (39.97) and the experimental class (42.56) were relatively comparable, indicating

homogeneous initial abilities prior to treatment. Following the intervention, the experimental class demonstrated a higher posttest mean (75.64) compared to the control class (71.91), with a difference of 3.73 points. This finding indicates that discovery learning-based worksheets produced a greater positive impact on the development of critical thinking skills than conventional learning.

Table 1. Descriptive Statistics of Pretest and Posttest Scores of Critical Thinking Skills

Statistical Data	Pretest		Posttest	
	Control	Experiment	Control	Experiment
Highest Score	49.35	56.70	81.90	87.15
Lowest Score	22.05	30.45	58.80	61.95
Mean	39.97	42.56	71.91	75.64
Median	42.19	42.71	72.92	77.08
Mode	35.42	42.71	72.92	76.04
Standard Deviation	7.24	7.46	4.39	5.67
Number of Students	32	34	32	34

A visual comparison of the pretest and posttest scores of both classes is presented in Figure 1. The graph clearly illustrates that both groups showed improvement from pretest to posttest; however, the experimental class consistently outperformed the control class on posttest scores, demonstrating the added value of discovery learning-based worksheet instruction.

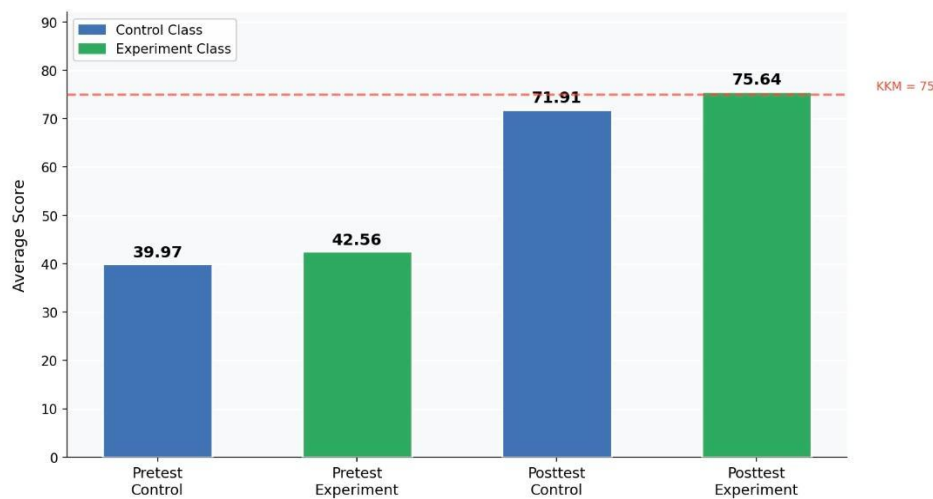


Figure 1. Comparison of Pretest and Posttest Scores of the Control and Experimental Classes

Achievement of Each Critical Thinking Skills Indicator

Indicator-level analysis was conducted to obtain a detailed picture of the development of each dimension of critical thinking skills. Table 2 and Figure 2 present the percentage achievement results for each indicator in both classes across the pretest and posttest phases.

Table 2. Percentage Achievement of Each Critical Thinking Skills Indicator

No	Indicator	Control Class (%)			Experimental Class (%)		
		Pretest	Posttest	N-Gain	Pretest	Posttest	N-Gain
1	Interpretation	46.68	70.31	0.44	49.82	79.60	0.59
2	Analysis	43.36	76.56	0.59	47.61	77.02	0.56
3	Evaluation	43.75	78.13	0.61	43.01	79.23	0.64
4	Inference	42.19	83.98	0.72	45.04	91.18	0.84
5	Explanation	36.13	76.76	0.64	40.44	78.13	0.63
6	Self-Regulation	27.73	45.70	0.25	29.41	48.71	0.27
	Average	39.97	71.91	0.54	42.56	75.65	0.59

Figure 2 illustrates the achievement pattern of each critical thinking indicator in both classes. As clearly visible, the inference indicator showed the most pronounced improvement in the experimental class, while self-regulation remained the lowest across both groups.

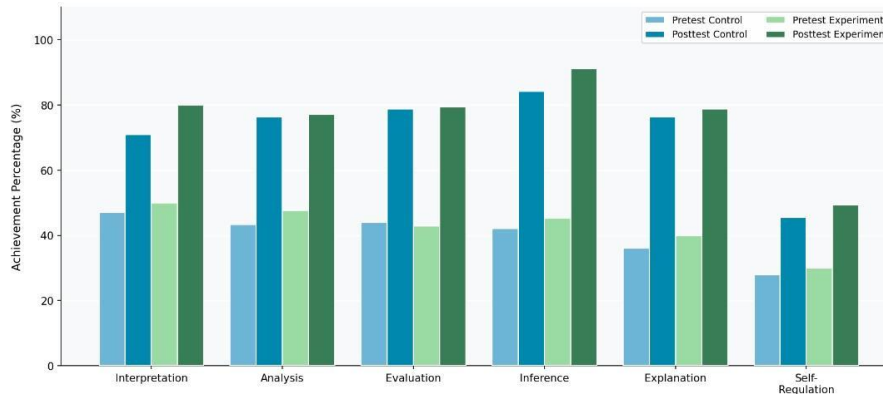


Figure 2. Percentage Achievement of Each Critical Thinking Indicator in Pretest and Posttest

Based on Table 2 and Figure 2, all indicators improved in both classes. The inference indicator recorded the highest N-Gain in the experimental class (from 45.04% to 91.18%, N-Gain = 0.84, high category). This significant improvement is attributable to the generalization stage embedded in the discovery learning-based worksheet, which explicitly requires students to formulate conclusions from self-collected and self-processed data a process that directly trains the capacity to identify patterns, test hypotheses, and construct valid generalizations (Dzikro et al., 2021). In contrast, the self-regulation indicator recorded the lowest N-Gain in both classes (0.27 for the experimental class and 0.25 for the control class). This is consistent with Nisa et al. (2022), who noted that self-regulation is the most challenging dimension of critical thinking to develop due to its reliance on high metacognitive abilities.

N-Gain Analysis and Distribution of Improvement Categories

To accurately measure the magnitude of learning improvement and account for differences in initial scores, a normalized gain (N-Gain) analysis was performed. Table 3 summarizes the N-Gain distribution for both classes.

Table 3. Distribution of N-Gain Categories for Experimental and Control Classes

N-Gain Category	Experimental Class (%)	Control Class (%)
High ($g > 0.70$)	8.82	0.00
Moderate ($0.30 \leq g \leq 0.70$)	88.23	100.00
Low ($g < 0.30$)	2.94	0.00
Average N-Gain	0.57	0.53
Standard Deviation	0.12	0.08

The experimental class obtained an average N-Gain of 0.57 (SD = 0.12), while the control class obtained 0.53 (SD = 0.08), both in the medium category. However, the experimental class showed a more varied distribution: 8.82% of students achieved the high category, 88.23% the medium, and 2.94% the low. All students in the control class fell into the medium category (100%). The presence of high N-Gain students only in the experimental class indicates that discovery learning-based worksheets are more effective in accelerating critical thinking development for high-performing students.

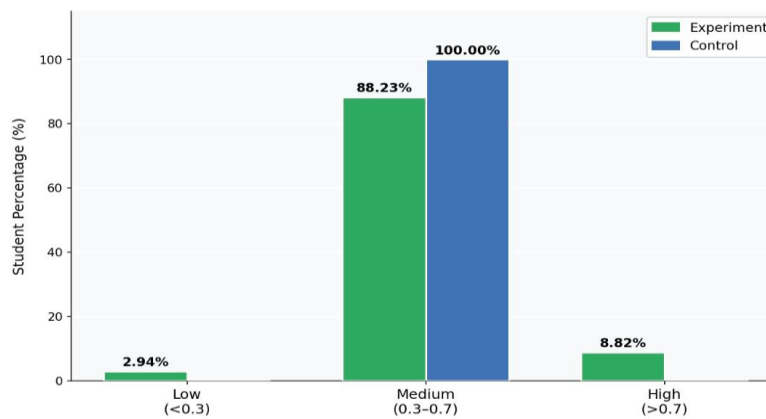


Figure 3. Distribution of N-Gain Categories for Experimental and Control Classes

Normality and Homogeneity Test Results

Before conducting hypothesis testing, prerequisite tests were performed to verify the assumptions required for parametric statistics. Table 4 presents the results of the normality test for all data.

Table 4. Normality Test Results for Pretest and Posttest Data

Parameter	Pretest Experiment	Posttest Experiment	Pretest Control	Posttest Control
Shapiro-Wilk	0.962	0.958	0.971	0.953
p-value (Sig.)	0.261	0.318	0.213	0.143
α	0.05	0.05	0.05	0.05
p-value (Asymp. Sig.)	0.261	0.318	0.213	0.143
Conclusion	Normal	Normal	Normal	Normal

Table 4 shows that the Shapiro-Wilk p-values for all data are greater than 0.05, confirming that all datasets are normally distributed. Table 5 presents the homogeneity test results.

Table 5. Homogeneity Test Results for Pretest and Posttest Data

Parameter	Pretest Control	Pretest Experiment	Posttest Control	Posttest Experiment
Number of Samples (N)	32	34	32	34
Fisher Statistic	0.009	–	1.565	–
p-value (Sig.)	0.925	–	0.216	–
α	0.05	0.05	0.05	0.05
Conclusion	Homogeneous	Homogeneous	Homogeneous	Homogeneous

Based on Table 5, Fisher's test p-values for both Pretest (0.925) and Posttest (0.216) data exceed α 0.05, confirming that the variance of both classes is homogeneous. Therefore, hypothesis testing using the independent samples t-test could proceed.

Hypothesis Testing

Table 6. Results of the Independent Samples T-Test

Parameter	Pretest	Posttest
Number of Samples (N Control/Experiment)	32 / 34	32 / 34
p-value (Sig.)	0.079	0.010
α	0.05	0.05
Decision	H_0 Accepted	H_0 Rejected
Conclusion	No significant difference	Significant difference exists
Cohen's d (Effect Size)	–	d = 0.74 (Large Effect)

The pretest data yielded a significance value of $0.079 > 0.05$, indicating no significant difference in initial critical thinking skills between the two classes and confirming the validity of the research design. In

contrast, the posttest significance value of $0.010 < 0.05$ confirms that H_0 is accepted: there is a significant difference in critical thinking skills between the experimental and control classes following the intervention. To assess practical significance beyond statistical significance, According to Cohen (1988), $d = 0.74$ represents a large effect, indicating that the difference between the groups is not merely statistically significant but also educationally meaningful and practically important.

The findings confirm that discovery learning-based worksheets are significantly more effective in improving students' critical thinking skills compared to the direct instruction model. This can be understood through Vygotsky's constructivist framework, which posits that knowledge is actively constructed through experience and social interaction (Gunawan et al., 2022). Through the stages of discovery learning, students are positioned as active agents constructing their own understanding rather than passive recipients of information.

As shown in Table 2, inference recorded the highest gain among all indicators, primarily driven by the generalization stage of the worksheet, which requires students to construct conclusions from data they collected and processed themselves a process that directly develops the capacity for evidence-based reasoning (Wahyudi, 2022). The remaining moderate-gain indicators (interpretation, analysis, evaluation, explanation; N-Gain range: 0.56–0.64) reflect the progressive scaffolding of the earlier discovery learning stages. Interpretation benefited from the stimulation stage through video-based phenomena, while analysis developed through problem statement and data collection activities. Evaluation improved via the verification stage, where students assessed findings against theory (Hosnan, 2022), and explanation improved through the data processing stage's requirement for written scientific justification (Dzikro et al., 2021). The moderate ceiling on these gains reflects the cognitive complexity of multi-variable physics analysis and the two-meeting duration of the intervention (Sapitri et al., 2021). The small absolute N-Gain gap ($\alpha = 0.04$) should be interpreted together with the Cohen's d of 0.74 (large effect), which confirms the practical significance of the intervention beyond statistical thresholds alone. This gap reflects the constraint of two instructional meetings rather than a weakness of the discovery learning approach itself (Apriani et al., 2023). Extended interventions would be expected to widen this gap, particularly for higher-order indicators requiring sustained metacognitive development (Wahyudi, 2022)(Field, 2018). It should be acknowledged, however, that this study does not directly test the causal contribution of each individual worksheet component to each indicator; the stage-indicator mapping represents a theoretically-grounded design inference supported by the pattern of differentiated per-indicator N-Gain outcomes (Dzikro et al., 2021). Future research employing component-level experimental designs would strengthen this causal argument. Despite this limitation, the practical effect confirmed by Cohen's d and the differentiated per-indicator gains provide meaningful evidence of the worksheet's instructional value beyond global improvement scores.

In contrast, the self-regulation indicator showed the lowest improvement in both classes (experimental N-Gain = 0.27, control = 0.25). This indicates that independently reflecting on and evaluating one's own thinking processes requires more intensive and sustained practice beyond two learning sessions. As argued by Dewi & Wardani (2022), developing self-regulation requires structured metacognitive scaffolding and extended exposure time. Flavell (as cited in Rahmadani, 2022) similarly emphasized that metacognitive development is an ongoing process that cannot be achieved through brief interventions.

Overall, these findings align with the meta-analysis by Apriani et al. (2023), which reported that discovery learning produces an average effect size of 0.68 (medium-large category) on higher-order thinking skills in high school students. The superiority of the experimental class is due to the worksheet structure that systematically activates higher-order cognitive processes at each discovery learning stage. At the problem statement stage, students formulate hypotheses, fostering analytical and evaluation skills; data collection develops explanatory skills; and the generalization stage directly targets inference.

These results contrast with the direct instruction approach in the control class, where students passively received teacher-delivered information without engaging in concept discovery (Rahmadani & Anugraheni, 2022). This confirms Hosnan's (2022) view that discovery learning is structurally more conducive to developing critical thinking skills than expository learning. Evidence from various contexts further corroborates this finding: Marzuki et al. (2022) found that discovery learning significantly increased student activeness in physics; Utami et al. (2022) demonstrated its effectiveness in simultaneously developing critical thinking and

conceptual understanding; Yuli Ekawati et al. (2022) confirmed its benefits when integrated with virtual laboratory activities; and Wulandari & Setyawan (2022) showed that STEM-integrated discovery learning worksheets effectively fostered critical thinking across multiple indicators. Collectively, the per-indicator N-Gain pattern across the six Facione dimensions confirms that the worksheet's stage-by-stage structure produces a differentiated rather than uniform improvement profile, which has direct diagnostic value: teachers can identify which specific thinking competencies each student still needs to develop, an advantage that holistic scores alone cannot provide (Hermawan et al., 2025).

Theoretically, this study strengthens the argument that constructivist, student-centered learning models are significantly more effective in developing higher-order thinking skills than expository models (Wahyu et al., 2021). It provides empirical evidence that integrating Facione's six indicators into a discovery learning-based worksheet structure results in differentiated improvement per indicator, giving teachers richer diagnostic information than holistic assessments alone.

Practically, this study offers a worksheet model adaptable by high school physics teachers, particularly for topics involving directly observable or simulation-based phenomena. Teachers are advised to allocate additional time to the verification and generalization stages, as these directly target evaluation, inference, and self-regulation indicators that require more intensive cultivation. This study also supports the implementation of Indonesia's Independent Curriculum, which emphasizes project-based and discovery-based learning. The findings are further corroborated by Widyaningrum et al. (2024), who demonstrated that a STEM-integrated discovery learning approach significantly enhanced critical thinking in the physics classroom, underscoring the enduring relevance of discovery-based instructional design across evolving curriculum frameworks. Furthermore, Fadillah et al. (2025) confirmed that integrating structured digital worksheets with inquiry-based physics learning produces significant gains in critical thinking, while Hermawan et al. (2025) showed that worksheets explicitly aligned with Facione's framework yield measurably differentiated indicator-level improvements in secondary science students.

CONCLUSION

This study demonstrates that integrating Facione's six critical thinking indicators into a discovery learning-based worksheet is an effective instructional strategy for eleventh-grade physics, producing both statistically significant and practically meaningful gains. The experimental class outperformed the control class on posttest scores and average normalized gain, with the intervention yielding a large practical effect as confirmed by Cohen's *d*. The most educationally important finding is the differentiated pattern of improvement across indicators: inference benefited most, directly reflecting the generalization stage of the worksheet, while self-regulation showed the lowest development in both classes. This pattern underscores that metacognitive skills require more extended scaffolding than a two-meeting intervention can provide. The practical implication is clear: teachers should embed discovery learning-based worksheets within a longer instructional sequence incorporating dedicated self-reflection checkpoints, rather than treating them as a one-time activity. The worksheet's dual function as both an instructional scaffold and a diagnostic instrument offers indicator-level feedback that enables more targeted and differentiated teaching decisions than holistic assessments alone. This study acknowledges three key limitations: the two-meeting duration constrains the magnitude of gains; the causal contribution of each individual worksheet stage to each specific indicator has not been directly tested; and generalizability is bounded by the high school physics linear motion context. Future research should replicate this framework across different physics topics, school levels, and cultural contexts, and employ component-level experimental designs to establish clearer causal relationships between specific discovery learning stages and individual critical thinking indicators.

References

- Aditya, D. Y. (2021). The influence of the discovery learning model on students' critical thinking skills in physics. *Journal of Physics Education*, 9(2), 102–113. <https://doi.org/10.21009/1.09210>
- Amaniyah, R., & Sugiarto, B. (2021). The effectiveness of guided discovery-based student worksheets in science learning to improve critical thinking skills. *Journal of Educational Science and Technology*, 7(1), 45–54. <https://doi.org/10.26858/est.v7i1.15561>

- Anjarsari, P., & Prasetyo, Z. K. (2022). The importance of training critical thinking skills in science learning in secondary schools. *Indonesian Journal of Educational Science*, 6(1), 1–12. <https://doi.org/10.31605/ijes.v6i1.1825>
- Apriani, D., Setiawan, A., & Purwanto, A. (2023). Meta-analysis of the effect of discovery learning on higher-order thinking skills of high school students. *Journal of Physics Education Research*, 10(1), 55–67. <https://doi.org/10.20961/jper.v10i1.70123>
- Ardiansyah, A., Ishafit, I., & Toifur, M. (2021). The effect of discovery learning-based student worksheets on critical thinking skills and physics learning activities. *Journal of Science Education Research*, 7(2), 91–100. <https://doi.org/10.30997/jser.v7i2.4876>
- Azizah, R., Yuliati, L., & Latifah, E. (2021). Difficulties in solving physics problems in high school students. *Journal of Physics Research and Its Applications*, 6(2), 55–67. <https://doi.org/10.17977/um024v6i12021p055>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Dewi, N. R., & Wardani, S. (2022). Discovery-based learning innovation to improve students' higher-order thinking skills. *Phenomenon: Journal of Mathematics and Natural Sciences Education*, 12(1), 25–37. <https://doi.org/10.21580/phen.2022.12.1.10965>
- Dzikro, W., Zulfiani, Z., & Juanengsih, N. (2021). The influence of the discovery learning model on critical thinking skills and student learning outcomes. *Bioedukasi: Journal of Biology Education*, 14(1), 48–58. <https://doi.org/10.20961/bioedukasi-uns.v14i1.44256>
- Ekawati, Y. I., & Ilyas, A. (2022). The effectiveness of scientific-based student worksheets on high school students' critical thinking skills. *Indonesian Journal of Science and Mathematics Education*, 5(2), 180–191. <https://doi.org/10.24042/ij sme.v5i2.12308>
- Facione, P. A. (2015). *Critical thinking: What it is and why it counts*. Insight Assessment, 1–30.
- Fadillah, A., Harjono, A., & Gunawan, G. (2025). Enhancing students' critical thinking skills through inquiry-based learning integrated with digital worksheets in high school physics. *Jurnal Pendidikan Fisika Indonesia*, 21(1), 1–12. <https://doi.org/10.15294/jpfi.v21i1.45123>
- Fahaluddin, I., Halim, A., & Khaldun, I. (2021). The effect of discovery-based worksheet models on students' activities and critical thinking skills in physics learning. *Indonesian Journal of Physics Education*, 17(1), 99–109. <https://doi.org/10.15294/jpfi.v17i1.22685>
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.
- Fuadi, H., Robbia, A. Z., Jamaluddin, J., & Jufri, A. W. (2022). Analysis of the causes of low science literacy of students and alternative solutions. *Jurnal Ilmiah Profesi Pendidikan*, 5(2), 108–116. <https://doi.org/10.29303/jipp.v5i2.127>
- Gunawan, R., Sahidu, H., & Harjono, A. (2022). The effect of the discovery learning model assisted by PhET simulation media on the critical thinking of prospective teacher students. *Jurnal Kumparan Fisika*, 5(1), 11–18. <https://doi.org/10.33059/jkf.v5i1.4521>
- Hermawan, D., Setiawan, A., & Rustaman, N. Y. (2025). The role of structured student worksheets in developing Facione's critical thinking competencies in secondary science education. *International Journal of Science Education*, 47(2), 215–234. <https://doi.org/10.1080/09500693.2025.2187456>
- Hosnan, M. (2022). *Scientific and contextual approaches in 21st century learning*. Ghalia Indonesia.
- Kusprayanto, H. N., Santoso, A., & Akbar, S. (2022). Development of student worksheets based on guided discovery to improve critical thinking skills and science concept understanding in elementary school students. *Journal of Education*, 6(4), 61–72. <https://doi.org/10.31004/joe.v6i4.3874>
- Laili, N., Budi, S. H., & Supriyadi, B. (2022). Analysis of high school students' critical thinking skills on static electricity. *Proceedings of the National Seminar on Physics Education*, 7(1), 28–40.

- Marzuki, M., Sadrina, S., & Rizqi, I. (2022). Application of the problem-discovery learning model to increase student activeness in physics subjects. *CIRCUIT: Scientific Journal of Electrical Engineering Education*, 6(1), 19–30. <https://doi.org/10.22373/crc.v6i1.12209>
- Nisa, S., Utaya, S., & Utomo, D. H. (2022). The influence of the discovery learning model on the critical thinking skills of geography education students. *Journal of Education*, 7(10), 1308–1318. <https://doi.org/10.31004/joe.v7i10.3678>
- Novayani, W. (2021). The effect of the discovery learning model on junior high school students' critical thinking skills on environmental pollution. *Scientific Journal of Biology Education*, 7(1), 29–38. <https://doi.org/10.22437/bio.v7i1.12096>
- Nurrohmi, Y., Harjono, A., & Gunawan, G. (2022). The effect of discovery learning-based worksheets on students' critical thinking skills in physics. *Jurnal Pendidikan Fisika Indonesia*, 18(1), 45–56. <https://doi.org/10.15294/jpfi.v18i1.32561>
- OECD. (2023). *PISA 2022 results (volume I): The state of learning and equity in education*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- Rahmadani, N., & Anugraheni, I. (2022). Increasing students' learning activities through problem-based learning in elementary school. *Scholaria: Journal of Education and Culture*, 7(3), 241–250. <https://doi.org/10.24246/j.scholaria.2022.v7.i3.p241-250>
- Rahmawati, I., Suyanto, S., & Mahmudi, A. (2026). Discovery learning-based instructional materials and their impact on higher-order thinking skills: A systematic review. *Jurnal Penelitian Pendidikan IPA*, 12(1), 88–101. <https://doi.org/10.29303/jppipa.v12i1.51234>
- Ristasari, T., Priyono, B., & Sukaesih, S. (2023). The effect of problem solving learning model with mind mapping on students' critical thinking skills. *Journal of Biology and Biology Education*, 6(1), 18–28. <https://doi.org/10.15294/jbe.v6i1.38712>
- Sapitri, E., Mujasam, M., & Yusuf, I. (2021). Improving students' critical thinking skills through the implementation of discovery learning-based student worksheets. *Tadulako Online Journal of Physics Education*, 9(1), 35–44. <https://doi.org/10.22487/tojtiptpe.v9i1.1521>
- Utami, M. F. L. B., Widodo, A., & Sopandi, W. (2022). Application of discovery learning strategy to improve critical thinking skills and understanding of science concepts in junior high school students. *Jinop: Journal of Learning Innovation*, 8(1), 83–95. <https://doi.org/10.22219/jinop.v8i1.19234>
- Wahyu, Y., Suastra, I. W., Sadia, I. W., & Suarni, N. K. (2021). The effectiveness of mobile-augmented reality using a scientific approach to improve science literacy. *International Journal of Instruction*, 13(3), 77–91. <https://doi.org/10.29333/iji.2020.1336a>
- Wahyudi, E. (2022). Application of discovery learning in physics learning to improve learning outcomes and critical thinking skills. *Indonesian Journal of Science and Mathematics Education*, 5(1), 45–57. <https://doi.org/10.24042/ijsme.v5i1.11542>
- Widyaningrum, R., Sarwanto, S., & Karyanto, P. (2024). Discovery learning with STEM approach to improve critical thinking skills in physics. *Jurnal Pendidikan IPA Indonesia*, 13(1), 89–99. <https://doi.org/10.15294/jpii.v13i1.43217>
- Wulandari, R., & Setyawan, A. (2022). Development of STEM-based student worksheets integrated with discovery learning to improve students' critical thinking. *Indonesian Journal of Science and Mathematics Education*, 5(3), 270–283. <https://doi.org/10.24042/ijsme.v5i3.13491>
- Yuli Ekawati, I., Hikmawati, H., & Verawati, N. N. S. P. (2022). The effect of laboratory-assisted discovery learning model virtual learning on students' critical thinking skills. *Jurnal Pijar MIPA*, 17(4), 420–428. <https://doi.org/10.29303/jpm.v17i4.3271>
- Zulyusri, Z., Santosa, T. A., Festiyed, F., Yerimadesi, Y., & Yohandri, Y. (2023). Effectiveness of discovery learning in improving critical thinking skills: A meta-analysis. *Jurnal Penelitian Pendidikan IPA*, 9(3), 1376–1385. <https://doi.org/10.29303/jppipa.v9i3.3474>