DIJEMSS: Dinasti International Journal of Education Management and Social Science
E-ISSN: 2686-6331
P-ISSN: 2686-6358
DINASTI PUBLISHER
https://dinastipub.org/DIJEMSS dinasti.info@gmail.com +62 811 7404 455

DOI: https://doi.org/10.38035/dijemss.v8i1
https://creativecommons.org/licenses/by/4.0/
Received: 09 August 2026 | **Revised:** 19 September 2026 | **Accepted:** 30 October 2026

Effectiveness of STEM-Based Physics Learning in Enhancing 21st-Century Skills Among Senior High School Students

Mardia Hi Rahman[1], Fatma Hamid[2], Hutri Handayani Isra[3]
[1]Universitas Khairun, Ternate City, Indonesia, mardiah.rahman1@gmail.com
[2]Universitas Khairun, Ternate City, Indonesia, fatmahamid@unkhair.ac.id
[3]Universitas Khairun, Ternate City, Indonesia

Corresponding Author: mardiah.rahman1@gmail.com[1]

Abstract: Physics learning is expected to develop conceptual understanding and 21st-Century Skills, yet teacher-centered instruction may limit students' engagement in inquiry and authentic problem-solving. This study examined the effectiveness of STEM-based physics learning in improving physics learning achievement and four 21st-Century Skills. A quasi-experimental pretest-posttest control group design involved 60 eleventh-grade students, with 30 students in the experimental group and 30 in the control group. The experimental group received STEM-based instruction integrating scientific inquiry, technology, engineering design, mathematical reasoning, authentic problem-solving, and collaboration. Data were collected using achievement tests and structured observations of critical thinking, creativity, communication, and collaboration. Data were analyzed using descriptive statistics, independent samples t-tests, normalized gain, ANCOVA, and Cohen's d. The experimental group achieved a higher posttest mean (84.73 ± 3.99) than the control group (70.07 ± 3.10), with a significant difference, $t(58) = 15.906$, $p < 0.001$, and a very large effect (Cohen's $d = 4.11$). Significant differences were also found in critical thinking, creativity, communication, and collaboration ($p < 0.001$). These findings support the effectiveness of STEM-based physics learning in improving achievement and 21st-Century Skills.

Keywords: Physics learning, STEM education, 21st-Century Skills, physics education, senior high school.

INTRODUCTION

Twenty-first-century education requires students to develop not only academic knowledge but also competencies that support effective participation in complex social, technological, and professional environments. Among these competencies, critical thinking, creativity, communication, and collaboration are widely recognized as important dimensions of twenty-first-century learning. In science education, these competencies are particularly relevant because learning science involves investigating phenomena, analyzing evidence, solving problems, and communicating explanations (Dare et al., 2021; Kain et al., 2024). However, physics instruction at the secondary-school level may still rely heavily on teacher

explanations, conceptual transmission, and routine problem-solving, providing limited opportunities for students to engage actively in inquiry and authentic problem-solving.

Physics has considerable potential to support the development of higher-order thinking and twenty-first-century competencies because its concepts are closely connected to observable phenomena, mathematical reasoning, technological applications, and real-world problems. Effective physics instruction therefore needs to provide opportunities for students to connect theoretical concepts with practical applications while actively constructing knowledge. Such learning requires instructional approaches that integrate conceptual understanding with investigation, problem-solving, technology use, and collaborative learning (Bao & Koenig, 2019).

STEM (Science, Technology, Engineering, and Mathematics) provides an integrated framework for achieving these objectives. Integrated STEM learning connects scientific inquiry with technological applications, engineering design, and mathematical reasoning in the context of meaningful and authentic problems (Kelley & Knowles, 2016; Thibaut et al., 2018). Rather than learning physics concepts independently, students are encouraged to apply knowledge across disciplines, investigate problems, develop solutions, evaluate alternatives, and communicate their findings. These characteristics provide opportunities for students to engage in the cognitive and social processes associated with critical thinking, creativity, communication, and collaboration (Dare et al., 2021; Kain et al., 2024).

Previous studies have reported positive relationships between STEM-based instruction and students' conceptual understanding, engagement, and twenty-first-century competencies. STEM-oriented physics learning has been shown to support conceptual understanding and 4C skills through technology-supported and inquiry-oriented activities (Asrizal et al., 2024). Meta-analytic and systematic evidence also indicates that STEM learning can contribute to the development of twenty-first-century competencies through interdisciplinary and authentic learning experiences (Hawur et al., 2025; Novianti et al., 2022; Yulianti et al., 2025). Project-based STEM learning further provides opportunities for students to work on complex tasks, communicate solutions, and engage in collaborative problem-solving (Patriot et al., 2025). At the same time, the implementation of STEM in secondary education continues to face challenges related to teacher readiness, instructional resources, curriculum integration, and assessment practices (Khalil et al., 2023; Margot & Kettler, 2019; Rizakhojayeva et al., 2025).

These challenges indicate that the effectiveness of STEM-based physics learning depends not only on the use of interdisciplinary activities but also on how systematically the learning process is designed and implemented. Instructional planning, authentic problem-solving, scientific inquiry, engineering design, and assessment of students' competencies need to be aligned within a coherent learning process. Although previous research has documented the benefits of STEM learning for conceptual achievement and selected twenty-first-century competencies, fewer studies have simultaneously examined physics learning achievement and the four core dimensions of critical thinking, creativity, communication, and collaboration within a systematically implemented STEM learning context.

Therefore, this study investigates the effectiveness of STEM-based physics learning in a senior high school setting by examining two complementary outcomes: physics learning achievement and students' observed 21st-Century Skills. Physics learning achievement is assessed through pretest and posttest measures, while critical thinking, creativity, communication, and collaboration are assessed separately through structured observation. This approach allows the study to distinguish academic achievement from broader twenty-first-century competencies while examining their relationship to an integrated STEM learning experience.

Accordingly, this study addresses the following research questions: (1) Does STEM-based physics learning result in higher physics learning achievement than conventional physics

instruction? and (2) Are there differences between the STEM and conventional groups in students' critical thinking, creativity, communication, and collaboration? The study is expected to provide empirical evidence regarding the implementation of STEM-based physics learning and practical implications for developing physics instruction that integrates conceptual learning with competencies required in twenty-first-century education.

METHODS

Research Design and Participants

This study employed a quantitative quasi-experimental approach using a pretest-posttest control group design. A quasi-experimental design was selected because the participants were members of intact classrooms and could not be randomly assigned to the experimental conditions. This design allows researchers to examine differences between groups under authentic educational conditions while maintaining the natural classroom setting (Creswell & Creswell, 2018; Fraenkel et al., 2019).

The study was conducted at SMA Negeri 7 Halmahera Selatan, Indonesia, during the 2026/2027 academic year. The participants consisted of 60 eleventh-grade science students from two intact classes. One class, consisting of 30 students, served as the experimental group and received STEM-based physics instruction, whereas the other class, also consisting of 30 students, served as the control group and received conventional physics instruction. Purposive sampling was used because the two classes had comparable academic characteristics, followed the same curriculum, were taught by the same physics teacher, and had similar learning facilities. However, because the participants were not randomly assigned, the findings are interpreted within the instructional context examined in this study.

Research Design

The research design is presented in Table 1.

Table 1. Pretest-Posttest Control Group Design

Group	Pretest	Treatment	Posttest
Experimental	O_1	X	O_2
Control	O_3	-	O_4

Source: Developed by the authors (2026).

Where O1 and O3 represent the pretests administered to the experimental and control groups, respectively; O2 and O4 represent the posttests; X represents STEM-based physics instruction.

The pretest and posttest were used to measure physics learning achievement, while 21st-Century Skills were assessed separately through structured classroom observation. This distinction was maintained because learning achievement and twenty-first-century competencies represent related but conceptually distinct outcome constructs (Kain et al., 2024).

STEM-Based Physics Learning Intervention

The experimental group received systematically organized STEM-based physics instruction integrating science, technology, engineering, and mathematics through authentic problem-solving and project-oriented learning activities. Integrated STEM learning emphasizes the connection of disciplinary knowledge with authentic problems and the integration of scientific inquiry, technological application, engineering design, and mathematical reasoning (Kelley & Knowles, 2016; Thibaut et al., 2018).

Students were engaged in contextual physics problems and were encouraged to identify problems, formulate scientific explanations, conduct investigations, analyze quantitative

information, develop possible solutions, evaluate alternative solutions, and communicate their findings. These activities provided opportunities for students to apply physics concepts in authentic contexts and engage in learning processes associated with higher-order thinking and twenty-first-century competencies (Dare et al., 2021; Kelley & Knowles, 2016).

The control group received conventional physics instruction primarily through teacher explanations, question-and-answer activities, textbook exercises, and individual problem-solving. Both groups followed the same curriculum and learning objectives and were taught by the same physics teacher. The principal distinction was the integration of interdisciplinary STEM activities in the experimental group.

Research Instruments

Two principal types of instruments were used in the study: a physics achievement test and a structured observation instrument for 21st-Century Skills. The distinction between the two instruments was maintained to ensure that physics learning achievement and twenty-first-century competencies were assessed as separate constructs.

Physics Learning Achievement Test

Physics learning achievement was measured using pretest and posttest examinations consisting of physics questions aligned with the instructional objectives and competencies addressed during the intervention. The achievement test was designed to assess students' conceptual understanding and application of physics concepts. Before administration, the instrument was reviewed by physics education experts to assess its content relevance and alignment with the intended learning objectives.

The use of achievement tests to assess conceptual understanding and higher-order learning outcomes is consistent with the measurement of physics learning outcomes described in previous physics education research (Bao & Koenig, 2019). The present study therefore treated the achievement test as the primary instrument for measuring physics learning achievement rather than as a direct measure of all dimensions of 21st-Century Skills.

21st-Century Skills Observation Instrument

Students' 21st-Century Skills were assessed using a structured observation instrument consisting of four dimensions: critical thinking, creativity, communication, and collaboration. Each dimension consisted of five observable indicators, resulting in 20 indicators in total. The instrument used a four-point rating scale ranging from 1 to 4. A score of 4 indicated that the indicator was demonstrated fully, clearly, consistently, and as expected; a score of 3 indicated good performance with minor deficiencies; a score of 2 indicated partial performance with several deficiencies; and a score of 1 indicated that the indicator was not demonstrated or was demonstrated only to a very limited extent.

The critical thinking dimension consisted of five indicators: problem identification, information analysis, reasoning, evaluation, and conclusion drawing. Creativity was assessed through idea generation, flexibility, elaboration, originality, and modification. Communication included idea expression, concept explanation, argumentation, presentation, and response to questions. Collaboration included participation, task distribution, respect for peers' opinions, assistance to group members, and responsibility for group outcomes.

The observation was conducted during the instructional process in both the experimental and control groups. Scores for the four dimensions were recorded separately and subsequently used to describe and compare students' observed 21st-Century Skills. The conceptual dimensions of the instrument are consistent with the literature on 21st-century competencies and integrated STEM education (Dare et al., 2021; Kain et al., 2024).

Data Collection Procedure

Data collection was conducted in several stages. First, both groups completed the physics pretest to determine their initial achievement levels. Second, the experimental group participated in STEM-based physics instruction, while the control group received conventional instruction. During the learning process, structured observations were conducted to assess students' critical thinking, creativity, communication, and collaboration. Finally, both groups completed the physics posttest after the instructional intervention.

The achievement data and 21st-Century Skills observation data were recorded and analyzed separately according to their respective constructs. This procedure was intended to maintain alignment between the research questions, instruments, and outcome measures.

Data Analysis

The quantitative data were analyzed using descriptive and inferential statistical procedures with IBM SPSS Statistics. Means and standard deviations were calculated to summarize physics learning achievement and the four dimensions of 21st-Century Skills.

For physics learning achievement, the Shapiro–Wilk test was used to assess the normality assumption, and Levene's test was used to assess homogeneity of variance. These tests were used to evaluate the assumptions underlying parametric analysis (Field, 2018).

An independent samples t-test was first conducted on the pretest scores to determine whether the experimental and control groups differed significantly in initial physics achievement. An independent samples t-test was then conducted to compare posttest achievement between the two groups (Field, 2018; Sudjana, 2002).

Because the study used a pretest-posttest design, analysis of covariance (ANCOVA) was additionally conducted for posttest physics achievement, with pretest achievement entered as the covariate and instructional group as the independent variable. ANCOVA was used to determine whether the posttest difference between the groups remained statistically significant after accounting for measured baseline achievement.

Normalized gain (N-gain) was calculated to describe the magnitude of improvement in physics learning achievement from pretest to posttest:

$$\left[g = \frac{Posttest - Pretest}{Maximum\ Score - Pretest}\right] \quad (1)$$

Cohen's d was calculated to estimate the standardized magnitude of between-group differences. Reporting effect size alongside statistical significance provides information about the practical magnitude of an observed difference (Field, 2018).

For the four dimensions of 21st-Century Skills, descriptive statistics were calculated separately for critical thinking, creativity, communication, and collaboration. Independent samples t-tests were conducted to compare the observed scores of the experimental and control groups for each dimension. Cohen's d was also calculated for each comparison to estimate the magnitude of the between-group differences.

Accordingly, the analysis distinguished between two complementary outcomes: physics learning achievement, measured using pretest and posttest achievement scores, and 21st-Century Skills, measured through structured observation of critical thinking, creativity, communication, and collaboration.

Ethical Considerations

Permission to conduct the study was obtained from the school administration before data collection. Participants were informed about the purpose of the study and the voluntary

nature of their participation. Students' identities and academic information were kept confidential and were used only for research purposes.

RESULTS AND DISCUSSION

Descriptive Results of Physics Learning Achievement

The study involved 60 eleventh-grade students, comprising 30 students in the experimental group and 30 students in the control group. The experimental group received STEM-based physics instruction, whereas the control group received conventional physics instruction. Physics learning achievement was assessed using pretest and posttest scores.

Table 2. Descriptive Statistics of Students' Achievement

Group	n	Pretest Mean ± SD	Posttest Mean ± SD	Mean Gain
Experimental	30	50.97 ± 2.59	84.73 ± 3.99	33.77
Control	30	49.90 ± 3.37	70.07 ± 3.10	20.17

Source: Developed by the authors (2026).

The descriptive results indicate that the two groups had relatively similar pretest performance. The experimental group obtained a mean pretest score of 50.97, whereas the control group obtained 49.90. The difference of 1.07 points indicates that the measured initial achievement of the two groups was relatively comparable.

Following the instructional intervention, both groups demonstrated improvement. The mean score of the experimental group increased from 50.97 to 84.73, corresponding to a gain of 33.77 points. Meanwhile, the control group increased from 49.90 to 70.07, corresponding to a gain of 20.17 points. Thus, the experimental group demonstrated a substantially greater absolute improvement.

The higher posttest performance in the experimental group is consistent with the characteristics of integrated STEM instruction, which connects scientific inquiry, technology, engineering design, and mathematical reasoning with authentic problems. Such learning allows students to apply physics concepts through investigation, experimentation, quantitative analysis, and solution development rather than relying primarily on direct explanation and routine exercises (Kelley & Knowles, 2016; Thibaut et al., 2018). Similar evidence has been reported in physics education research showing that STEM-based learning can strengthen conceptual understanding and engagement by connecting disciplinary concepts with meaningful applications (Asrizal et al., 2024; Bao & Koenig, 2019).

Normality and Homogeneity Testing

Before inferential analysis, the assumptions for parametric testing were examined. The Shapiro–Wilk test was applied to the posttest scores, while Levene's test was used to examine the equality of variances.

Table 3. Results of the Shapiro–Wilk Normality Test

Group	Statistic	Sig.	Interpretation
Experimental	0.967	0.472	Normal
Control	0.948	0.145	Normal

Source: Developed by the authors (2026).

The Shapiro–Wilk results produced p-values of 0.472 for the experimental group and 0.145 for the control group. Both values exceed 0.05, indicating that there was insufficient evidence to reject the normality assumption. Thus, the posttest scores were suitable for subsequent parametric analysis.

Table 4. Results of the Homogeneity Test

Test	Statistic	Sig.	Interpretation
Levene Test	2.326	0.133	Homogeneous

Source: Developed by the authors (2026).

The Levene's test produced a p-value of 0.133, which is greater than 0.05. Therefore, the two groups did not show statistically significant differences in score variance, and the homogeneity assumption was satisfied. These tests were used strictly to evaluate statistical assumptions and were not interpreted as evidence of instructional quality.

Equality of Initial Achievement

An independent samples t-test was conducted using the pretest scores to determine whether the experimental and control groups differed significantly before the intervention.

Table 5. Independent Samples t-Test of Pretest Scores

Outcome	Experimental Mean ± SD	Control Mean ± SD	Mean Difference	t	df	p-value	95% CI	Cohen's d
Pretest	50.97 ± 2.59	49.90 ± 3.37	1.07	1.375	58	0.174	[-0.49, 2.63]	0.36

Source: Developed by the authors (2026).

The independent samples t-test showed no statistically significant difference in pretest achievement between the experimental and control groups, $t(58) = 1.375$, $p = 0.174$. The mean difference was 1.07 points, with a 95% confidence interval ranging from −0.49 to 2.63. The standardized effect size was small (Cohen's $d = 0.36$). These results indicate that the two groups had statistically comparable measured physics achievement before the instructional intervention. However, because the participants were drawn from intact classrooms rather than randomly assigned, this baseline similarity does not establish equivalence on unmeasured characteristics.

Posttest Achievement and Effect Size

The posttest scores were compared using an independent samples t-test.

Table 6. Independent Samples t-Test of Posttest Achievement

Outcome	Experimental Mean ± SD	Control Mean ± SD	Mean Difference	t	df	p-value	95% CI	Cohen's d
Posttest	84.73 ± 3.99	70.07 ± 3.10	14.66	15.906	58	<0.001	[12.81, 16.51]	4.11

Source: Developed by the authors (2026).

The experimental group achieved a substantially higher posttest mean than the control group. The mean difference was 14.66 points, and the independent samples t-test indicated a statistically significant difference, $t(58) = 15.906$, $p < 0.001$. The 95% confidence interval for the mean difference ranged from approximately 12.81 to 16.51 points.

The standardized effect was also very large, with Cohen's $d \approx 4.11$. Therefore, the difference between the groups was not only statistically significant but also substantial in standardized magnitude.

These findings provide evidence supporting the effectiveness of STEM-based physics instruction in producing higher physics learning achievement under the conditions examined

in this study. The result is consistent with the integrated STEM perspective proposed by Kelley & Knowles (2016), in which disciplinary knowledge is applied through authentic problem-solving and interdisciplinary activity. It also aligns with previous studies reporting positive contributions of STEM-oriented physics learning to conceptual understanding and higher-order learning (Asrizal et al., 2024; Patriot et al., 2025; Yulianti et al., 2025).

Normalized Gain

Normalized gain was calculated to describe the proportion of possible improvement achieved by each group.

Table 7. Normalized Gain of Physics Learning Achievement

Group	Mean N-Gain	Interpretation
Experimental	0.689	High
Control	0.401	Medium

Source: Developed by the authors (2026).

The experimental group obtained a mean normalized gain of 0.689, whereas the control group obtained 0.401. Thus, the experimental group achieved a greater proportion of the maximum possible improvement during the instructional period.

The higher normalized gain is consistent with the greater absolute increase observed in the experimental group. This pattern suggests that the STEM learning experience provided more opportunities for students to apply physics concepts through inquiry, experimentation, engineering design, and problem-solving. Previous research similarly highlights authentic problem-solving and interdisciplinary activity as important characteristics of effective STEM learning (Dare et al., 2021; Thibaut et al., 2018).

ANCOVA Analysis

Because the research used a pretest-posttest design, ANCOVA was conducted as an additional analysis. Posttest achievement was used as the dependent variable, instructional group as the independent variable, and pretest achievement as the covariate.

Table 8. ANCOVA of Posttest Physics Learning Achievement

Source	F	df	p-value	Adjusted Mean Difference	95% CI
Pretest	5.20	1, 57	0.026	—	—
Instructional group	249.77	1, 57	<0.001	14.30	[12.49, 16.11]

Source: Developed by the authors (2026).

After controlling for pretest achievement, the instructional group remained statistically significant, $F(1,57) = 249.77$, $p < 0.001$. The adjusted mean difference was approximately 14.30 points, with a 95% confidence interval from 12.49 to 16.11.

The ANCOVA result strengthens the independent samples t-test finding because the difference in posttest achievement remained significant after controlling for measured initial achievement. Nevertheless, because the participants were members of intact classrooms and were not randomly assigned, the result is interpreted as evidence supporting the effectiveness of the intervention under the study conditions rather than definitive proof of causality.

Results of 21st-Century Skills

In addition to physics learning achievement, the study examined four dimensions of 21st-Century Skills through structured observation: critical thinking, creativity, communication, and collaboration. The four dimensions were analyzed separately as recommended for conceptually distinct competencies

Table 9. Descriptive Statistics of 21st-Century Skills

Dimension	Experimental Mean ± SD	Control Mean ± SD	Mean Difference
Critical Thinking	88.00 ± 3.32	72.07 ± 3.30	15.93
Creativity	88.07 ± 2.95	72.67 ± 2.75	15.40
Communication	88.93 ± 2.96	73.87 ± 3.19	15.07
Collaboration	91.60 ± 2.70	76.67 ± 2.59	14.93
Overall 4C	89.15 ± 2.52	73.82 ± 2.48	15.33

Source: Developed by the authors (2026).

The descriptive results show consistently higher scores for the experimental group across all four dimensions. The highest experimental-group mean was observed for collaboration (91.60 ± 2.70), followed by communication (88.93 ± 2.96), creativity (88.07 ± 2.95), and critical thinking (88.00 ± 3.32). The control group obtained lower scores across all four dimensions.

These differences are consistent with the learning characteristics of integrated STEM instruction. STEM activities require students to analyze contextual problems, generate solutions, discuss alternative approaches, work collaboratively, and communicate evidence-based findings. These characteristics are closely related to the competencies emphasized in twenty-first-century learning and integrated STEM frameworks (Dare et al., 2021; Kain et al., 2024; Kelley & Knowles, 2016).

Comparison of the Four 4C Dimensions

Table 10. Independent Samples t-Test and Effect Size of 21st-Century Skills

Dimension	Mean Difference	t	df	p-value	95% CI	Cohen's d
Critical Thinking	15.93	18.637	58	<0.001	[14.22, 17.64]	4.81
Creativity	15.40	20.940	58	<0.001	[13.93, 16.87]	5.41
Communication	15.07	18.961	58	<0.001	[13.48, 16.66]	4.90
Collaboration	14.93	21.864	58	<0.001	[13.56, 16.30]	5.65
Overall 4C	15.33	23.760	58	<0.001	[14.04, 16.62]	6.13

Source: Developed by the authors (2026).

Independent samples t-tests indicated statistically significant differences between the experimental and control groups across all four dimensions. Critical thinking differed significantly between the two groups, $t(58) = 18.637$, $p < 0.001$, with a mean difference of 15.93 points and a very large effect size ($d = 4.81$).

Creativity also showed a statistically significant difference, $t(58) = 20.940$, $p < 0.001$. The experimental group exceeded the control group by 15.40 points, with a very large effect size ($d = 5.41$). This result is consistent with the characteristics of engineering design, which require students to generate, test, evaluate, and refine possible solutions (Khalil et al., 2023).

Communication showed a significant difference of 15.07 points, $t(58) = 18.961$, $p < 0.001$, with a very large effect size ($d = 4.90$). STEM activities require students to explain procedures, discuss evidence, present findings, and communicate decisions within collaborative learning situations. These activities provide repeated opportunities for scientific communication and argumentation (Patriot et al., 2025; Rizakhojayeva et al., 2025).

Collaboration showed the highest mean score in the experimental group, reaching 91.60 ± 2.70 compared with 76.67 ± 2.59 in the control group. The difference was statistically significant, $t(58) = 21.864$, $p < 0.001$, with a very large effect size ($d = 5.65$). The finding is consistent with the collaborative structure of STEM problem-solving, in which students work together to investigate problems, divide responsibilities, evaluate alternatives, and develop solutions.

For the overall 4C score, the experimental group obtained a mean of 89.15 ± 2.52 compared with 73.82 ± 2.48 in the control group. The difference was statistically significant, $t(58) = 23.760$, $p < 0.001$, with a very large standardized effect ($d = 6.13$). The consistently higher results across all four dimensions indicate that STEM-based instruction was associated with stronger observed performance in the four assessed twenty-first-century competencies.

Discussion

The combined findings demonstrate a consistent pattern across both academic achievement and the four observed dimensions of 21st-Century Skills. The experimental group not only achieved higher posttest physics scores but also obtained substantially higher scores in critical thinking, creativity, communication, and collaboration.

The improvement in critical thinking can be interpreted in relation to the inquiry-oriented nature of STEM learning. Students were required to examine contextual problems, consider evidence, analyze information, and formulate conclusions. Such activities move students beyond the reproduction of information toward analytical engagement with scientific problems. This mechanism is consistent with the literature describing inquiry and authentic problem-solving as important components of integrated STEM learning (Bao & Koenig, 2019; Dare et al., 2021).

The higher creativity score may be associated with engineering design processes. Students were required to generate possible solutions, test alternatives, and revise their designs or explanations. Such iterative activity creates opportunities for students to move beyond a single predetermined answer and consider multiple possible solutions. This interpretation is consistent with previous research showing positive relationships between STEM-oriented instruction, engineering design, and creative thinking in secondary education (Khalil et al., 2023; Novianti et al., 2022).

The communication findings can be explained by the collaborative and presentation-oriented characteristics of STEM learning. Students were expected to explain their reasoning, discuss evidence, communicate findings, and respond to peers. These activities create repeated opportunities to formulate and defend ideas using scientific reasoning. Similar mechanisms have been identified in STEM project-based learning, where communication is embedded in collaborative investigation and presentation (Patriot et al., 2025).

Collaboration showed the highest mean score among the four dimensions in the experimental class. This result is consistent with the structure of STEM activities, which commonly involve collaborative investigation, division of responsibilities, collective decision-making, and shared development of solutions. STEM therefore provides a learning environment in which collaboration is not simply an additional activity but an integral component of completing authentic tasks (Rizakhojayeva et al., 2025; Thibaut et al., 2018).

The large effect sizes across all four dimensions indicate that the observed differences were substantial in standardized terms. However, these effect sizes should be interpreted within the context of the study design. The participants were drawn from two intact classes in a single school, and the 4C measures were obtained through observational assessment. Therefore, the findings provide strong evidence of between-group differences under the specific instructional conditions examined, but they should not be generalized uncritically to all senior high school students.

The achievement findings and 4C findings also provide complementary evidence for the theoretical pathway underlying integrated STEM learning. STEM activities expose students to authentic problems, requiring them to construct knowledge actively through inquiry, design, quantitative reasoning, collaboration, and communication. These processes can simultaneously support conceptual learning and the development of competencies associated with twenty-first-century education (Kain et al., 2024; Kelley & Knowles, 2016).

Overall, the results suggest that systematic STEM-based physics instruction was associated with higher physics learning achievement and stronger observed performance across critical thinking, creativity, communication, and collaboration than conventional instruction. The findings therefore provide empirical support for integrating interdisciplinary STEM activities into secondary-school physics learning while recognizing the contextual limitations of the quasi-experimental design.

CONCLUSIONS AND RECOMMENDATIONS

Conclusion

This study investigated the implementation of STEM-based physics learning and its contribution to improving senior high school students' 21st-Century Skills. The findings demonstrated that students who participated in STEM-based learning achieved significantly higher posttest scores than those who received conventional instruction. The implementation of STEM learning effectively promoted active student engagement through scientific inquiry, engineering design, collaborative problem-solving, and authentic learning experiences, resulting in improvements in conceptual understanding as well as critical thinking, creativity, collaboration, and communication skills. These findings confirm that STEM-based physics learning represents an effective instructional strategy for enhancing both academic achievement and twenty-first-century competencies in secondary education. The study contributes to science education by providing empirical evidence that a systematically implemented STEM approach can improve the quality of physics instruction while supporting the development of competencies required to address increasingly complex scientific and technological challenges.

Limitations

This study has several limitations that should be considered when interpreting the findings. First, the study was conducted in a single senior high school and involved only two intact classes with 60 students. The participants were not randomly assigned, which limits the generalizability of the findings and prevents definitive causal conclusions beyond the study context. Second, the intervention was conducted over a relatively short instructional period; therefore, the study could not determine whether the observed improvements in physics learning achievement and 21st-Century Skills would be sustained over a longer period. Third, the study assessed critical thinking, creativity, communication, and collaboration through structured classroom observation, while other relevant competencies, such as scientific literacy, digital competence, learning motivation, and self-efficacy, were not examined comprehensively. Finally, the findings should be interpreted within the specific instructional, curricular, and school conditions in which the STEM-based learning intervention was implemented.

Recommendations

Future research should involve larger samples from multiple schools and diverse educational contexts to improve the generalizability of the findings. Longer intervention periods and longitudinal assessments are also recommended to examine the sustainability of changes in physics learning achievement and 21st-Century Skills. Future studies should additionally employ well-validated 4C assessment instruments and, where applicable, establish inter-rater reliability to strengthen the measurement of critical thinking, creativity, communication, and collaboration. Further research may also examine broader competencies, including scientific literacy, digital competence, learning motivation, and self-efficacy, to provide a more comprehensive understanding of the effects of integrated STEM-based physics learning.

Author Contributions

Mardia Hi Rahman conceived the study, designed the research methodology, supervised the research process, interpreted the findings, and prepared the manuscript. Fatma Hamid contributed to the development of the research framework, data collection, statistical analysis, interpretation of the results, and critical revision of the manuscript. Hutri Handayani Isra participated in data collection, classroom implementation, data validation, literature review, manuscript editing, and final proofreading. All authors reviewed, approved, and agreed to the published version of the manuscript and accept responsibility for the integrity of the work.

Acknowledgment

The authors would like to express their sincere gratitude to Universitas Khairun for providing academic support throughout this research. The authors also extend their appreciation to the principal, physics teachers, and students of SMA Negeri 7 Halmahera Selatan for their valuable cooperation and participation during the implementation of this study. Their support and commitment were essential to the successful completion of this research.

REFERENCES

Asrizal, Nazifah, N., Effendi, H., & Helma. (2024). STEM-Smart Physics E-Module to Promote Conceptual Understanding and 4C Skills of Students. *International Journal of Information and Education Technology*, *14*(2), 279–286. https://doi.org/10.18178/ijiet.2024.14.2.2049

Bao, L., & Koenig, K. (2019). Physics Education Research for 21st Century Learning. *Disciplinary and Interdisciplinary Science Education Research*, *1*, 2. https://doi.org/10.1186/s43031-019-0007-8

Creswell, J. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th ed.). SAGE Publications.

Dare, E. A., Keratithamkul, K., Hiwatig, B. M., & Li, F. (2021). Beyond Content: The Role of STEM Disciplines, Real-World Problems, 21st Century Skills, and STEM Careers within Science Teachers' Conceptions of Integrated STEM Education. *Education Sciences*, *11*(11), 737. https://doi.org/10.3390/educsci11110737

Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics* (5th ed.). SAGE Publications.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2019). *How to Design and Evaluate Research in Education* (10th ed.). McGraw-Hill Education.

Hawur, P. K., Supahar, Rasyid, F., & Gunawan, C. W. (2025). Global Trends in Innovation and Evaluation of 21st Century Physics Learning: A Bibliometric and Systematic Review. *Kasuari: Physics Education Journal (KPEJ)*, *8*(2), 542–560. https://doi.org/10.37891/kpej.v8i2.1000

Kain, C., Koschmieder, C., Matischek-Jauk, M., & Bergner, S. (2024). Mapping the Landscape: A Scoping Review of 21st Century Skills Literature in Secondary Education. *Teaching and Teacher Education*, *151*, 104739. https://doi.org/10.1016/j.tate.2024.104739

Kelley, T. R., & Knowles, J. G. (2016). A Conceptual Framework for Integrated STEM Education. *International Journal of STEM Education*, *3*(1), 11. https://doi.org/10.1186/s40594-016-0046-z

Khalil, R. Y., Tairab, H., Qablan, A., Alarabi, K., & Mansour, Y. (2023). STEM-Based Curriculum and Creative Thinking in High School Students. *Education Sciences*, *13*(12), 1195. https://doi.org/10.3390/educsci13121195

Margot, K. C., & Kettler, T. (2019). Teachers' Perception of STEM Integration and Education: A Systematic Literature Review. *International Journal of STEM Education*, *6*(1), 2. https://doi.org/10.1186/s40594-018-0151-2

Novianti, B. A., Lasmawan, I. W., & Suharta, I. G. P. (2022). Efek STEM dalam Pembelajaran Sains terhadap Keterampilan Abad 21: Meta Analisis. *Kappa Journal: Physics & Physics Education*, *6*(2). https://doi.org/10.29408/kpj.v6i2.18830

Patriot, E. A., Ismet, I., Wiyono, K., Rizka, M. R., & Melati, P. (2025). STEM-Project Based Learning: Implementation to Optimize the 21st Century Skills in Pre-Service Teacher of Physics Education. *Proceedings of the 6th Sriwijaya University Learning and Education International Conference (SULE-IC 2024)*, 622–639. https://doi.org/10.2991/978-2-38476-390-0_51

Rizakhojayeva, G., Ramankulov, S., Akeshova, M., Nurizinova, M., Tuyakov, Y., & Abdrakhmanov, R. (2025). STEM-Based Approaches to Soft Skills Development: A Synthesis of Meta-Analytic Findings and Empirical Evidence. *Frontiers in Education*, *10*, 1663155. https://doi.org/10.3389/feduc.2025.1663155

Sudjana. (2002). *Metode Statistika* (6th ed.). Tarsito.

Thibaut, L., Ceuppens, S., De Loof, H., De Meester, J., Goovaerts, L., Struyf, A., & Petegem, P. Van. (2018). Integrated STEM Education: A Systematic Review of Instructional Practices in Secondary Education. *European Journal of STEM Education*, *3*(1), 2. https://doi.org/10.20897/ejsteme/85525

Yulianti, D., Shafira, F. F., Marwoto, P., & Noviyatun, S. (2025). A Systematic Review of Trend STEM Education Research on Physics Learning in Indonesia. *Jurnal Pendidikan Fisika Indonesia*, *21*(2), 192–208. https://doi.org/10.15294/jpfi.v21i2.23548