JPF | Volume 13 | Number 3 | 2025 | 461 - 478

*p* - ISSN: 2302-8939
*e* - ISSN: 2527-4015

**Jurnal Pendidikan Fisika**

**https://journal.unismuh.ac.id/index.php/jpf**

**DOI: 10.26618/h80x0b78**

# Phys'AR as a Learning Innovation: Strengthening Critical Thinking and Argumentation Skills in Applied Physics

**Faninda Novika Pertiwi[1)2)*], Udomsak Kitthawee[3)], Nurul Huda Ramadhan[1)], Izza Aliyatul Muna[1)]**

[1)]*Department of Science Education, Universitas Islam Negeri Kiai Ageng Muhammad Besari, Ponorogo, 63471, Indonesia*
[2]*Department of Physics Education, State University of Malang, Malang, 65145, Indonesia.*
[3)]*Suan Dusit Scientific Equipment Center, Suan Dusit University, Bangkok, 10300, Thailand.*

**Corresponding author: faninda@iainponorogo.ac.id*

Received: June 16, 2025; Accepted: August 19, 2025; Published: September 16, 2025

***Abstract*** **–** *Critical thinking and argumentation are essential twenty-first-century skills in physics education. Yet, conventional teaching methods often fail to provide students with sufficient opportunities to practice and develop these abilities. To address this challenge, this study introduced Phys'AR, an Augmented Reality (AR)–based learning medium designed not only to visualize abstract physics concepts but also to embed structured activities for constructing and evaluating arguments. The study employed an explanatory sequential mixed methods design involving 42 undergraduate students of physics education, divided into a 2023 control group and a 2024 experimental group. Quantitative data on critical thinking and argumentation skills were collected through standardized tests and analyzed using normality tests, homogeneity tests, and independent samples t-tests, while qualitative insights were gathered from classroom observations and student interviews. The findings showed that students in the experimental group significantly outperformed those in the control group in critical thinking, with higher post-test averages and more consistent score improvements. Analysis of argumentation revealed that students supported their claims with stronger data when learning with Phys'AR. However, most remained at a medium level in warrants and backings, and rebuttals were generally weak across both groups. These results indicate that Phys'AR is effective in strengthening evidence-based reasoning but requires complementary teaching strategies to promote deeper theoretical justification and counter-argumentation. The novelty of this study lies in extending the role of AR in physics education from a visualization tool to a platform for argumentation and critical reasoning. By highlighting which components of argumentation benefit most from AR and which require additional scaffolding, this research contributes both theoretical and practical insights for advancing critical thinking and argumentation in physics education.*

***Keywords:*** *applied physics; augmented reality; critical thinking; Phys' AR; scientific argumentation*



## I. INTRODUCTION

The twenty-first century is characterized by rapid scientific and technological advancements that have reshaped the way knowledge is generated, shared, and applied in everyday life. In this context, higher education faces the pressing task of preparing students not only with conceptual mastery but also with transferable skills that enable them to thrive in increasingly complex and uncertain environments. Critical thinking and argumentation are widely recognized as two essential skills that support lifelong learning, informed decision-making, and effective problem-solving in academic, professional, and civic spheres (Amani & Mkimbili, 2025; Cherian et al., 2013; Zhang et al., 2025). Critical thinking has been described as a multidimensional construct encompassing both cognitive and dispositional aspects. The cognitive dimension involves analyzing evidence, evaluating claims, identifying biases, and drawing logical conclusions, while the dispositional dimension reflects openness, ethical reasoning, and intellectual perseverance (Chen et al., 2024). Similarly, argumentation has emerged as a crucial component of scientific literacy, requiring individuals to construct, justify, and critically evaluate claims using valid evidence. It not only facilitates a deeper understanding of scientific concepts but also nurtures collaborative learning by encouraging students to engage with diverse perspectives (Mesa et al., 2025).

The growing body of literature underscores that critical thinking and argumentation are foundational for meaningful learning across educational levels. In science education in particular, these skills serve as cornerstones for inquiry-based practices, where students are expected to formulate hypotheses, test them against data, and communicate their findings persuasively. Scientific debates, peer discussions, and evidence-based reasoning tasks are consistently associated with enhanced higher-order thinking and problem-solving abilities (Chen et al., 2024; Gültepe & Kılıç, 2021). Moreover, the increasing complexity of global challenges, such as environmental sustainability, health crises, and technological ethics, demands that students be able to critically evaluate information and articulate coherent arguments in both academic and societal contexts. This demand aligns with the broader educational agenda, which positions twenty-first-century skills at the heart of curriculum reforms worldwide (Jegstad et al., 2025). Within this framework, fostering critical thinking and argumentation in physics education is particularly significant, as physics integrates abstract reasoning with real-world applications and requires students to interpret phenomena through evidence and models.

Despite this recognition, empirical evidence suggests that students frequently struggle to develop strong critical thinking and argumentation skills in higher education. Traditional teaching approaches that prioritize the transmission of factual knowledge tend to limit opportunities for active engagement, inquiry, and debate (Bati, 2019). In many classroom settings, discussions are dominated by instructors, leaving little room for students to articulate and defend their reasoning. Such teacher-centered practices not only reduce students' motivation to participate actively but also hinder the cultivation of argumentative discourse that is central to science learning

(Prabayanti et al., 2025). Pre-research conducted in the context of applied physics courses further supports this concern, showing that while multiple teaching methods such as discussions, presentations, and projects are employed, students rarely engage in structured debates or argumentation. As a result, their responses on written tests often consist of simple claims with weak or insufficient evidence. These shortcomings have been reported consistently in science classrooms, indicating a persistent gap between the importance of argumentation and its limited practice in instructional contexts (Chen et al., 2024; Mesa et al., 2025).

This gap highlights a significant challenge: while educators widely recognize the importance of integrating critical thinking and argumentation into science education, classroom practices often fail to cultivate these skills in meaningful ways. Students are rarely given opportunities to engage in tasks that demand critical reasoning and well-supported arguments, which limits their capacity to translate theoretical knowledge into sound explanations of real-world problems (Prabayanti et al., 2025). To overcome this issue, learning environments must be reimagined to actively involve students in activities that require both careful evaluation and structured argumentation. Approaches such as debates, collaborative inquiry, and technology-supported discussions have shown promise in narrowing this gap, as they encourage learners to present, defend, and refine their ideas through interaction with peers. Beyond strengthening understanding, these practices also nurture curiosity, collaboration, and reflective thinking (Mesa et al., 2025; Andrews, 2015).

Within the field of science education, technology-assisted learning has emerged as a promising pathway for addressing the limitations of conventional classroom practices. Augmented Reality (AR), in particular, has attracted increasing attention as an innovative tool that merges real and virtual environments. This tool allows learners to visualize complex phenomena, manipulate models, and engage in immersive experiences (Radu et al., 2023; Rebello et al., 2024). AR has been shown to enhance students' understanding of abstract concepts, improve motivation, and foster engagement in collaborative tasks (Demircioglu et al., 2022). Compared to other digital technologies, AR offers unique affordances in supporting inquiry, as it enables learners to interact directly with representations that connect theory and practice. Previous research has demonstrated that AR-based science learning can significantly enhance higher-order thinking skills, including problem-solving and creative reasoning (Lespita et al., 2023). However, while AR has often been employed to strengthen conceptual understanding, its potential in developing argumentation skills remains underexplored (Chen et al., 2024). This gap suggests a need for innovative AR applications that go beyond visualization, explicitly fostering discourse and critical reasoning.

The development of Phys'AR, an Augmented Reality–assisted learning medium for applied physics, represents a step toward addressing this need. Unlike many AR tools designed solely for conceptual visualization, Phys'AR integrates multiple features such as simulations, videos, quizzes, and interactive models to create a comprehensive learning platform. By embedding argumentation tasks into AR-based explorations, Phys'AR seeks to support students in constructing claims, providing evidence, and engaging with counter-arguments based on physics concepts. AR visualizations serve as scaffolds for students' investigations, making abstract phenomena observable and thereby encouraging evidence-based reasoning (Radu et al., 2023). Prior studies indicate that AR-based argumentation activities can enhance academic achievement and motivation, which are strongly linked to the development of both critical and argumentative skills (Demircioglu et al., 2022; Akbaş et al., 2019). By positioning argumentation at the center of AR-supported tasks, Phys'AR expands the role of technology in science education from a tool for knowledge delivery to a medium for cultivating essential cognitive and discursive competencies.

A closer examination of related literature reveals both opportunities and limitations of integrating AR into argumentation-focused science education. On one hand, AR has demonstrated effectiveness in engaging students and supporting conceptual understanding, with multiple studies reporting positive effects on motivation, collaboration, and learning outcomes (Wu et al., 2025; Alzahrani, 2025). On the other hand, evidence regarding its impact on argumentation is still fragmented. While some research highlights that AR can facilitate collaborative discourse by providing shared visual references (Rebello et al., 2024), systematic investigations into how AR supports Toulmin's argumentation components—claims, data, warrants, and rebuttals—remain scarce. Furthermore, studies show that students often perform better in producing claims and supporting data than in developing warrants or rebuttals, indicating uneven growth in argumentation skills even when supported by innovative media (Jon et al., 2023). This underscores a persistent research gap: although AR can enhance engagement and conceptual reasoning, its direct role in fostering comprehensive argumentation skills, particularly in physics contexts, has not been sufficiently addressed.

Therefore, the present study was conducted to explore how Phys'AR can be integrated into applied physics learning to foster both critical thinking and argumentation. Specifically, the research seeks to examine whether using Phys'AR enables students to construct stronger claims, provide evidence, build theoretical warrants, and engage in rebuttals. The novelty of this study lies in two aspects. First, it goes beyond the common use of AR as a visualization tool. Second, it investigates the interplay between critical thinking and argumentation within the specific context of applied physics education. By doing so, this study not only addresses a gap in the literature but also provides practical insights into how technology can be leveraged to enhance science education, making it more interactive, reflective, and aligned with the demands of the twenty-first century.

## II. METHODS

This study employed a mixed methods design using an explanatory sequential approach, which integrates quantitative and qualitative data to provide a more comprehensive understanding of the research problem. In this design, quantitative results are first collected and analyzed, followed by qualitative data that help explain or elaborate on the statistical findings (Creswell, 2022). The use of mixed methods is particularly appropriate in educational research, where complex skills, such as critical thinking and argumentation, are required. In this design, quantitative data were collected and analyzed to identify patterns and differences in students' skills across the two cohorts. These results then informed Phase 2, where qualitative data were gathered to explain and enrich the statistical outcomes. This design ensured that the study did not stop at identifying whether Phys'AR had an effect, but also explored how students engaged with the medium and why certain skills improved more than others. The research design is illustrated in Figure 1.

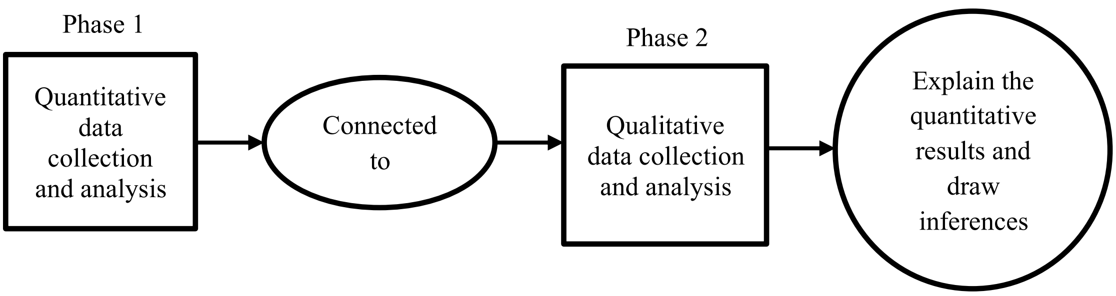


**Figure 1.** Explanatory research design

The research involved undergraduate students enrolled in the sixth semester of the Physics Education Program, who had chosen physics as their area of concentration. The participants were divided into two cohorts: one from 2023, which engaged in applied physics learning through conventional methods such as discussion and presentation, and one from 2024, which experienced applied physics learning assisted by Phys'AR media. In total, data were obtained from 42 students, representing a manageable yet meaningful sample size for both statistical and qualitative inquiry. This sampling choice was guided by the focus on depth of exploration in the explanatory phase, consistent with recommendations for mixed methods studies that balance breadth and depth (Creswell, 2022).

The instructional innovation in 2024 centered on the integration of Phys'AR, a learning medium that merges augmented reality with other multimedia features such as simulations, quizzes, and videos. The media was developed using the ADDIE model, which included systematic stages of analysis, design, development, implementation, and evaluation. To ensure quality, Phys'AR was validated by two experts in physics education before classroom implementation, followed by revisions and reliability testing until it met the standards for use in applied physics lectures. The incorporation of AR in this medium aligns with previous findings that highlight the capacity of augmented reality to enhance engagement and conceptual understanding in science education (Radu et al., 2023; Rebello et al., 2024). Moreover, AR-based environments are shown to provide authentic contexts for collaborative inquiry and argumentation (Demircioglu et al., 2022), making Phys'AR particularly relevant to the goals of this study.

Data collection focused on two primary constructs: critical thinking skills and argumentation skills. Critical thinking was measured through tests based on established indicators, including reasoning, hypothesis testing, argument analysis, probability and uncertainty evaluation, as well as problem-solving and decision-making (Tiruneh et al., 2017). Argumentation skills were assessed using Toulmin's model, which evaluates the quality of claims, data, warrants or backings, and rebuttals (Toulmin, 2015). These frameworks provided robust theoretical grounding for the measurement instruments, ensuring that the constructs were assessed in ways consistent with prior research in science education.

The quantitative phase of the study involved statistical testing of students' performance data. Tests for normality and homogeneity were applied to verify the assumptions of parametric testing, followed by independent sample t-tests to examine differences between the control and experimental groups. This analytic procedure allowed the researchers to determine whether the integration of Phys'AR had a statistically significant effect on students' critical thinking skills. The subsequent qualitative phase involved classroom observations and interviews, designed to provide deeper insight into students' experiences and to contextualize the numerical findings. The use of multiple data sources enhanced the validity of the study through triangulation (Creswell, 2022).

## III. RESULTS AND DISCUSSION

The effectiveness of Phys'AR-assisted learning in improving students' critical thinking skills was examined by comparing the performance of the 2024 experimental class with that of the 2023 control class, which did not use Phys'AR. Table 1 summarizes the descriptive statistics of students' critical thinking test scores. The pretest results show that both groups had comparable abilities before the intervention, with mean scores of 46.25 for the control group and 46.81 for the experimental group. This indicates that the two cohorts started from a similar baseline. However, after the intervention, clear differences emerged. The post-test average score of the experimental class reached 76.81, which was almost 10 points higher than the control group mean of 67.75. Moreover, the highest post-test score in the experimental group was 90, compared to 80 in the control group, while the lowest scores also shifted upward. These results suggest that the integration of Phys'AR not only raised the overall average performance but also improved the score range, reflecting both enhanced achievement and consistency across students.

**Table 1.** Critical thinking ability test results

| No | Value | Control class | Experiment class |
|---|---|---|---|
| 1 | Lowest pretest | 35 | 30 |
| 2 | Highest pretest | 60 | 60 |
| 3 | Average pretest | 46.25 | 46.81 |
| 4 | Lowest postest | 55 | 60 |
| 5 | Highest postest | 80 | 90 |
| 6 | Average post-test | 67.75 | 76.81 |

The next step was to ensure that the data met the assumptions for parametric testing. The results of the normality test are presented in Table 2. The Shapiro–Wilk statistics indicate that the significance values for all pretest and post-test data in both groups were greater than 0.05, ranging from 0.062 to 0.280. These values indicate that the distribution of scores did not deviate significantly from normality. Hence, the dataset fulfilled the assumption of normal distribution, which validates the application of further parametric tests.

**Table 2.** Results of the normality test of the critical thinking ability test results

**Tests of normality**

| | | Kolmogorov-Smirnov[a] | | | Shapiro-Wilk | | |
|---|---|---|---|---|---|---|---|
| | Class | Statistics | df | Sig. | Statistic | df | Sig. |
| Critical Thinking Ability | 1 | .162 | 20 | .176 | .943 | 20 | .274 |
| | 2 | .173 | 20 | .117 | .935 | 20 | .190 |
| | 3 | .173 | 22 | .087 | .916 | 22 | .062 |
| | 4 | .158 | 22 | .159 | .947 | 22 | .280 |

a. Liliefors Significance Correction

To further verify the suitability of the data, a homogeneity of variance test was performed. As shown in Table 3, the Levene's test results revealed significance values well above 0.05 across all measures, with the "Based on Mean" result yielding a value of 0.695. This indicates that the variances between the control and experimental groups were statistically equal. Together with the normality results, this outcome confirmed that the data met the requirements for independent sample t-tests.

**Table 3.** Homogeneity test results of the critical thinking ability test results

**Test of homogeneity of variance**

| | | Levene Statistic | df1 | df2 | Sig |
|---|---|---|---|---|---|
| Critical Thinking Ability | Based on Mean | .156 | 1 | 40 | .695 |
| | Based on Median | .222 | 1 | 40 | .640 |
| | Based on Median and with adjusted df | .222 | 1 | 39.823 | .640 |
| | Based on trimmed mean | .164 | 1 | 40 | .687 |

The results of the independent samples t-test are displayed in Table 4. The findings demonstrate a statistically significant difference between the control and experimental groups, with a t-value of –3.700 and a two-tailed significance level of 0.001 ($p < 0.05$). The negative sign reflects that the control group obtained lower scores than the experimental group. The mean difference was –8.841, with a 95% confidence interval ranging from –13.671 to –4.011. These results provide strong evidence that the use of Phys'AR in applied physics learning significantly improved students' critical thinking abilities compared to conventional methods.

**Table 4.** Critical thinking ability t test results

**Independent Samples Test**

| | | Levene's Test for Equality of Variances | | T-test for Equality of Means | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | 95% Confidence Interval of the Difference | |
| | | F | Sig | T | df | Sig (2-tailed) | Mean Difference | Std. Error Difference | Lower | Upper |
| Critical Thinking Ability | Equal variances assumed | .220 | .642 | -3.700 | 40 | .001 | -8.841 | 2.390 | -13.671 | -4.011 |
| | Equal variances not assumed | | | -3.724 | 39.935 | .001 | -8.841 | 2.374 | -13.639 | -4.043 |

In addition to examining students' critical thinking performance, this study also assessed their argumentation skills after participating in applied physics learning supported by Phys'AR. Argumentation was evaluated using Toulmin's framework, which breaks down arguments into claims, data, warrants or backings, and rebuttals. The quality of each structural component was rated on a three-point scale: high, medium, and low. The criteria employed in this study are summarized in Table 5, which outlines the standards used to judge the quality of each argumentation element

**Table 5.** Criteria for judging the quality of structural components of argumentation

| Structural Facet | Level | Description | Score |
|---|---|---|---|
| Claim | High | A claim without an opinion that includes background information | 3 |
| | Medium | Stating an opinion with background information or stating a stance on an issue, that is not stated as an opinion, but without background information | 2 |
| | Low | Simply stating an opinion | 1 |
| Data | High | Empirical: The use of specific data to back up the claim. The use of specific data to back up the claim. This evidence can include conceptual information as well. This is connected with evidence and data to the claim | 3 |
| | Medium | Conceptual: The use of conceptual information to back up a claim. This level may also include a personal opinion in linking the conceptual information to the claim. It does not rely on specific data to back up the claim, but includes more than a personal opinion | 2 |
| | Low | Opinion: The use of a personal opinion to back up a claim | 1 |
| Warrants, Backings, Qualifiers | High | Scientific: Data and reasoning that scientists use to investigate the phenomenon being argued, such as glaciers melting, sea levels, air temperature, water temperature, or species disturbance (McNeill & Pimentel, 2010) Data and theoretical groundings are connected in logical ways similar to ways in which scientists do this as well; Coordination of theory and evidence in the same ways that scientists use to connect data to hypotheses | 3 |
| | Medium | Rationale: Logical, attempts to use scientific understanding and language, is expressed through discussions of general scientific principles, possibly connected to personal experiences. (Dawson & Venville, 2010) | 2 |
| | Low | Personal: This is reasoning that relies on ideas from students' everyday lives, including, but not limited to, a student's opinion, personal feelings about the phenomena being studied, or expression of a student's expertise in an area to justify their claim. (McNeill & Pimentel, 2010) | 1 |
| Rebuttal | High | A counter-statement to the claim that uses empirical or conceptual evidence, as well as using scientific reasoning. A high-level rebuttal also refutes the counter-claim using scientific reasoning and empirical or conceptual evidence. This level of rebuttal is almost a complete argument within itself. | 3 |
| | Medium | A counter-statement to the claim that uses conceptual evidence, with a personal opinion, one possibly connected to refute the claim. The reasoning uses rational logic that makes an attempt to use scientific understanding and language. A medium-level rebuttal also refutes the counter-statement using personal opinion and/or conceptual evidence | 2 |
| | Low | A counter-statement to the claim that uses a personal opinion to refute the claim. It may or may not also refute the counter-statement to the claim, and if it does, it relies solely on a personal opinion. Conceptual information may be included in the personal opinion, but the overall effect of the statement is an opinion. The conceptual information is not the central focus of the statement | 1 |

The results of the argumentation assessment, presented in Table 6, reveal consistent differences between the control and experimental groups. For the claim component, most students reached the high category, with 12 students in the experimental group and 9 in the control group. In contrast, only one student in the control group fell into the low category. This indicates that students assisted by Phys'AR were better able to articulate claims supported by relevant background information rather than relying solely on opinion. A similar trend was observed in the data component, where 14 students in the experimental group reached the high category compared to 10 in the control group. These findings suggest that Phys'AR facilitated access to empirical and conceptual evidence, enabling students to construct claims with stronger and more specific support.

**Table 6.** Argumentation ability test results

| No | Claim | | | Data | | | Warrant/Backing/ Qualifier | | | Rebuttal | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | High | Medi-um | Low | High | Medi-um | Low | High | Medi-um | Low | High | Medi-um | Low |
| Control | 9 | 8 | 1 | 10 | 4 | 2 | 4 | 9 | 3 | 0 | 8 | 10 |
| Experi ment | 12 | 12 | 0 | 14 | 10 | 2 | 5 | 18 | 3 | 0 | 10 | 14 |
| Total | 21 | 20 | 1 | 24 | 14 | 4 | 9 | 27 | 6 | 0 | 18 | 24 |

For the warrants/backings/qualifiers component, most students remained in the medium category, with 18 students in the experimental group compared to 9 in the control group. While some students demonstrated high-level reasoning by integrating scientific theories with data, the majority relied on general principles without explicitly connecting evidence to formal frameworks, reflecting a common difficulty in argumentation (Jon et al., 2023). The greatest challenge emerged in the rebuttal component, where no student achieved a high score. In the experimental class, 14 students were categorized as low and only 10 as medium, while in the control group 10 were low and 8 mediums. This outcome indicates that although students could state claims and support them with data, they struggled to anticipate or counter opposing arguments using scientific reasoning. The absence of high-level rebuttals highlights that rebuttal construction remains an underdeveloped skill, even when supported by AR-based learning tools.

The detailed analysis of students' written responses provides further insight into the quality of their argumentation skills as structured through Toulmin's framework. For the claim component, many students successfully formulated high-quality claims, which were presented without personal opinions and supported by relevant background information. For example, one student explained, "*The refrigerator cools food by lowering the pressure of the refrigerant in the evaporator, so that the refrigerant evaporates and absorbs heat from the cooling room*". This statement reflects an informative claim anchored in physics concepts rather than subjective opinion. Another example was observed in students' explanations of wave applications: "*Microwaves heat food by emitting electromagnetic waves at a frequency that causes the water molecules in the food to vibrate and generate heat through molecular friction"*. The statement is also not an opinion, but is accompanied by a background, namely, physics concepts. A medium claim is characterized by stating an opinion with supporting background information. The following is an example of a student statement whose claim is at the medium level: "*I think microwaves are more efficient than electric stoves in heating food, because microwaves directly vibrate water molecules to speed up the heating process"*. When the statement includes the word "I think," it is an opinion that is supported, specifically regarding microwaves. Opinions also do not have to include the words "I think" or "in my opinion" when students compare something; it can also be stated as an opinion, for example, *"MRI is superior to CT Scan when used for soft tissue."* Claims in the low category are characterized by students only stating an opinion, for example *"MRI is better than CT scan"* or *"CT scan is better than MRI"*.

In the data component, high-quality responses contained specific empirical or conceptual evidence directly linked to the claim. For instance, one student argued: *"Microwaves heat water faster than conventional electric heaters (claim). Because based on simulations seen in Phys' AR media, 200 ml of water heated to 80°C takes 1 minute 30 seconds using a microwave, while when using an electric heater, it takes 3 minutes 10 seconds to reach the same temperature"*. This statement exemplifies how AR simulations provided concrete data to substantiate claims. Medium-level data were characterized by reliance on conceptual information without empirical evidence. An example is, *"Microwaves are more effective than conventional heaters, because microwaves directly affect the water molecules in food, so the heating process occurs from the inside out so that the process is faster than ordinary heaters"*. While scientifically plausible, the argument lacked quantitative support. In contrast, low-level responses relied on personal opinion, such as, *"food cooked using a microwave in my opinion tastes better."* These weaker responses underscore the ongoing challenge of guiding students beyond subjective statements to evidence-based reasoning.

The warrants, backings, and qualifiers component revealed the greatest variation in student performance. High-level responses demonstrated a strong connection between theoretical principles and empirical data, consistent with scientific reasoning. For example, a student wrote: *"The process of drying clothes in a washing machine is related to centrifugal force. Based on the simulation video on Phys' AR media, it appears that it is more efficient if the rotation speed is higher, as it will produce a greater centrifugal force and be able to remove more water. Based on theory, increasing rpm will increase* $\omega^2$, *which increases the centrifugal force on water"*. This statement shows how students linked simulation data with formal theoretical concepts. Medium-level responses often used logical reasoning or general scientific principles, but without deep theoretical grounding. For example, a student wrote: *"because the high rotation speed allows water to be pushed out of the clothing fibers faster, based on the basic principle that objects in circular motion experience an outward force of centrifugal force"*. These responses showed partial scientific reasoning but lacked explicit integration of data and theory. Finally, low-level responses relied mainly on personal experiences, for example, *"my washing machine can dry clothes very quickly. I know that because after every wash, the clothes just need to be dried in the sun for a while and they are already dry. Usually, I live in a humid area, but it still dries quickly with this machine"*. Such answers indicate reasoning rooted in everyday experience rather than formal scientific frameworks.

The most challenging component for students was the rebuttal. Notably, no student achieved a high-level rebuttal, which would require refuting counterclaims with scientific reasoning and empirical or conceptual evidence. Medium-level rebuttals were somewhat more frequent, with students combining personal opinions with some conceptual reasoning. For example, one student argued: "*Indeed, there are some who feel that it tastes different when food is cooked using a microwave, but I think it depends on the type of food. In principle, microwaves only vibrate water molecules, so they do not damage the chemical structure of food directly*." This reflects an attempt to counter a claim using scientific concepts, albeit with reliance on opinion. However, the majority of responses remained at the low level, where rebuttals were based solely on personal judgment. An example is: "*I do not agree that microwaves are more effective than ordinary stoves, because I feel that food from microwaves is not as warm as if it is heated using a stove*." Such rebuttals lacked theoretical or empirical support, showing that while students were able to present claims and data, they struggled to engage in higher-level argumentation that involves anticipating and countering opposing views.

The results of this study indicate that Phys'AR-assisted learning produced measurable gains in students' critical thinking and, more specifically, in the claim and data components of scientific argumentation. Post-test means were significantly higher for the experimental cohort than for the control group, as confirmed by the independent-samples t-test. This pattern aligns with prior research showing that technology-supported inquiry promotes deeper analytical processing and problem solving in science, particularly when learners can manipulate representations and observe cause–and–effect relations (Demircioglu et al., 2022; Lespita et al., 2023). In our setting, AR visualizations appear to have functioned as concrete anchors for reasoning: students were not merely recalling facts but mobilizing evidence from simulations and multimedia resources to justify their conclusions. That shift is consistent with the conceptualization of critical thinking as both a set of cognitive operations (analysis, evaluation, inference) and a disposition toward evidence-based judgment (Chen et al., 2024; Amani & Mkimbili, 2025). The improvement in critical thinking thus coheres with the literature that positions argument-evidence coordination as a key lever for higher-order learning in physics (Gültepe & Kılıç, 2021; Zhang et al., 2025).

The argumentation analysis helps explain how those gains emerged. Relative to the control group, more students in the Phys'AR cohort produced high-quality claims and data. In both components, learners moved beyond opinion statements and generic examples to articulate claims embedded in a disciplinary context and supported by empirical or conceptual evidence. Students frequently referenced measurable outputs from Phys'AR simulations such as time–temperature comparisons for microwave heating to warrant their conclusions (Glassner & Schwarz, 2007). This mirrors findings that AR can scaffold observation and measurement in ways that are difficult to realize with static media, thereby enriching the evidentiary basis of students' arguments (Radu et al., 2023; Rebello et al., 2024). In effect, Phys'AR lowered the cognitive cost of accessing relevant evidence, allowing students to dedicate more effort to selecting, linking, and interpreting information germane to the claim. Such movement from assertion to substantiation reflects a central goal of scientific literacy: making and defending knowledge claims with data and reasons rather than preference or authority alone (Mesa et al., 2025; Andrews, 2015; Joiner & Jones, 2003).

At the same time, the warrants/backings/qualifiers component largely remained at a medium level, and rebuttals were predominantly low across both groups, with no student reaching the highest category. This uneven profile of argumentation quality stronger in claims and data, weaker in warrants and rebuttals is well documented in the literature. Studies repeatedly show that students can state what they believe and even provide supporting information, yet struggle to link that information to theory through explicit warrants or to anticipate and refute counterclaims (Jon et al., 2023). Our results echo that pattern: learners readily drew on simulation outputs to support their positions, but only a minority connected those outputs to formal principles (e.g., relating increased spin rate to $\omega^2$ and centrifugal force). Even fewer constructed scientifically grounded rebuttals. This suggests that while AR supplies rich evidentiary material, the inferential bridges that connect data to theory require targeted scaffolding beyond access to visualization.

These findings bear directly on the theorized relationship between critical thinking and argumentation. On one reading, the improved post-test performance and the stronger claim data pairs indicate growth in analytic reasoning consistent with the cognitive dimension of critical thinking (Chen et al., 2024). Yet the persistent difficulties with warrants and rebuttals point to limits in students' epistemic cognition that is, how they understand the status of evidence, the role of theory, and the standards for justifying knowledge claims. Prior work has documented positive associations between argumentation and critical thinking dispositions, with argument skill predicting critical thinking in educational contexts (Akbaş et al., 2019; Chowning et al., 2012). Our results refine that association: AR-enhanced inquiry may accelerate progress on the evidentiary (claim–data) axis of argumentation and the performance indicators of critical thinking, but without explicit instruction in warrant construction and counter-argument, growth on more demanding argumentative moves remains constrained (Hasnunidah et al., 2015; Gültepe & Kılıç, 2021).

Pedagogically, the implication is not to replace AR with other tools, but to wrap AR within instructional designs that make the missing moves unavoidable. Structured debate protocols (e.g., assigning roles for claim, evidence, warrant, and rebuttal), explicit modeling of how to construct warrants from canonical laws and models, and checklists that prompt students to anticipate and test counterclaims can convert rich AR experiences into full Toulmin-style arguments. Such designs resonate with broader recommendations to integrate critical thinking across curricula via tasks that require evaluation, justification, and revision of ideas (Jegstad et al., 2025; Bavlı & Özdemir, 2025; Lv et al., 2025). They also align with evidence that cooperative inquiry and teacher facilitation can move students from plausible narratives to theory-consistent explanations in STEM learning (Fonseca et al., 2025; Nurilma et al., 2023). In short, AR can supply the "data-rich" substrate; instruction must engineer the argumentative work of connecting, qualifying, and contesting that data.

Finally, the study's contribution lies in extending AR research from conceptual understanding to a more granular analysis of argument quality within applied physics. Whereas much prior AR scholarship reports gain in motivation and comprehension (Wu et al., 2025; Alzahrani, 2025), our results differentiate which argumentation components benefit most from AR (claims and data) and which require additional scaffolds (warrants and rebuttals). This differentiation helps reconcile mixed findings in the literature by suggesting that AR's primary value is evidentiary rather than dialogic: it makes phenomena visible and measurable, thereby strengthening the support side of arguments; to strengthen the challenge side, instruction must deliberately cultivate counter-argument and theory–evidence coordination. Future studies could compare AR-only conditions with AR plus structured debate or embed metacognitive prompts that require students to state their warrants and generate counterexamples, to test whether these additions shift the distribution from medium to high in the warrant/backing category and elevate rebuttals beyond the low band. Such work would further clarify how technology and pedagogy can be integrated to cultivate the full arc of scientific argumentation and the critical thinking it is meant to serve.

## IV. CONCLUSION AND SUGGESTION

The study examined the effectiveness of Phys'AR, an augmented reality–based learning medium, in enhancing students' critical thinking and argumentation skills in applied physics courses. The findings demonstrated that students in the experimental class who engaged with Phys'AR achieved significantly higher post-test scores in critical thinking compared to their peers in the control class. Moreover, analysis of argumentation revealed notable improvements in the quality of claims and supporting data, with more students in the Phys'AR group providing scientifically grounded and evidence-based reasoning. However, most students remained at a medium level in warrants and backings, while rebuttals were predominantly low across both groups, indicating that although AR integration enriched evidence generation, it did not by itself foster deeper theoretical justification or counter-argumentation.

Despite these promising results, the study has several limitations. The sample size was relatively small and drawn from a single institutional context, which may constrain the generalizability of findings. The absence of high-level rebuttals also suggests that additional pedagogical scaffolding is required to complement AR tools. Future research should explore how Phys'AR can be integrated with structured debate formats, teacher facilitation, and metacognitive prompts to strengthen warrants and rebuttals, as well as replicate the study across larger and more diverse student populations. The contribution of this study lies in extending AR research from conceptual understanding to the domain of argumentation and critical thinking in physics education. By differentiating which components of argumentation benefit most from AR and which require further instructional design, this research offers both theoretical and practical insights for curriculum innovation and for advancing twenty-first-century skill development in science learning.

## REFERENCES

Akbaş, Y., Şahin, İ. F., & Meral, E. (2019). Implementing argumentation-based science learning approach in social studies: Academic achievement and students' views. *Review of International Geographical Education Online, 9*(1), 209–245. https://doi.org/10.33403/rigeo.529139

Alzahrani, A. (2025). A systematic review of the use of information communication technology, including augmented reality, in the teaching of science to preschool children. *International Journal of Educational Research Open, 9*, 1-11. https://doi.org/10.1016/j.ijedro.2025.100453

Amani, J., & Mkimbili, S. T. (2025). Developing critical thinking skills among secondary schools students: The role of research project. *Thinking Skills and Creativity*, *58*, 1-10 https://doi.org/10.1016/j.tsc.2025.101883

Andrews, R. (2015). Critical thinking and/or argumentation in higher education. In M. Davies & R. Barnett (Eds.), *The Palgrave handbook of critical thinking in higher education,* 49-62. Palgrave Macmillan. https://doi.org/10.1057/9781137378057_3

Bati, K. (2019). Are we ready for argumentation in science classrooms? An investigation into the scientific discussion climate in a Turkish elementary school. *Critical Questions in Education, 10*(1), 29–43.

Bavlı, B., & Özdemir, K. (2025). The conceptualization and enactment of critical thinking across diverse curricula in higher education. *Thinking Skills and Creativity, 58*, 1-16. https://doi.org/10.1016/j.tsc.2025.101914

Chen, X., Zhao, H., Jin, H., & Li, Y. (2024). Exploring college students' depth and processing patterns of critical thinking skills and their perception in argument map (AM)-supported online group debate activities. *Thinking Skills and Creativity, 51*, 1-13. https://doi.org/10.1016/j.tsc.2024.101467

Chowning, J. T., Griswold, J. C., Kovarik, D. N., & Collins, L. J. (2012). Fostering critical thinking, reasoning, and argumentation skills through bioethics education. *PLoS ONE, 7*(5), 1-8. https://doi.org/10.1371/journal.pone.0036791

Creswell, J. W. (2022). *A concise introduction to mixed methods research* (2nd ed.). SAGE Publications.

Dawson, V. M., & Venville, G. (2010). Teaching strategies for developing students' argumentation skills about socioscientific issues in high school genetics. *Research in Science Education, 40*(2), 133–148. https://doi.org/10.1007/s11165-008-9104-y

Demircioglu, T., Karakus, M., & Ucar, S. (2022). The impact of augmented reality-based argumentation activities on middle school students' academic achievement and motivation in science classes. *Education Quarterly Reviews, 5*(1), 22-35. https://doi.org/10.31014/AIOR.1993.05.02.464

Fonseca, X., Spangenberger, P., Baer, M., Schmidt, R., & Söbke, H. (2025). Location-based augmented reality in education: A systematic literature review. *Computers and Education Open*, *9*, 1-24. https://doi.org/10.1016/j.caeo.2025.100277

Glassner, A., & Schwarz, B. B. (2007). What stands and develops between creative and critical thinking? Argumentation? *Thinking Skills and Creativity, 2*(1), 10–18. https://doi.org/10.1016/j.tsc.2006.10.001

Gültepe, N., & Kılıç, Z. (2021). The effects of scientific argumentation on high school students' critical thinking skills. *International Journal of Progressive Education, 17*(6), 183–200. https://doi.org/10.29329/ijpe.2021.382.13

Hasnunidah, N., Susilo, H., Irawati, H. M., & Sutomo, H. (2015). Argument-driven inquiry with scaffolding as the development strategy of argumentation and critical thinking skills of students in Lampung, Indonesia. *American Journal of Educational Research, 3*(9), 1185–1192.

Jegstad, K. M., Heggernes, S. L., Jøsok, E., Ryen, E., Svanes, I. K., & Tørnby, H. (2025). Approaches to critical thinking in primary education classrooms: A systematic review. *Educational Research Review, 48*, 1-26. https://doi.org/10.1016/j.edurev.2025.100711

Joiner, R., & Jones, S. (2003). The effects of communication medium on argumentation and the development of critical thinking. *International Journal of Educational Research, 39*(8), 861–871. https://doi.org/10.1016/j.ijer.2004.11.008

Jon, E., Asrial, A., Hasibuan, M. H. E., & Hariyadi, B. (2023). Interaction of the reflective ethnobiology (REBI) model and creative thinking argumentation in biodiversity material. *JPBI (Jurnal Pendidikan Biologi Indonesia), 9*(3), 540–551. https://doi.org/10.22219/jpbi.v9i3.29279

Lespita, E., Purwanto, A., & Syarkowi, A. (2023). Application of problem-based learning model assisted by augmented reality media to improve students' high-order thinking skills. *Jurnal Pendidikan Fisika, 11*(1), 1–12. https://doi.org/10.26618/jpf.v11i1.9069

Lv, X., Zhou, J., & Ren, X. (2025). The bidirectional relationship between critical thinking and academic achievement is independent of general cognitive ability: A three-year longitudinal study on elementary school children. *Learning and Individual Differences, 120,* 1-12. https://doi.org/10.1016/j.lindif.2025.102666

McNeill, K. L., & Pimentel, D. S. (2010). Scientific discourse in three urban classrooms: The role of the teacher in engaging high school students in argumentation. *Science Education, 94*(2), 203–229. https://doi.org/10.1002/sce.20364

Mesa, M. del M., Valor, D., Montes, A., & García-Casas, I. (2025). Energy under debate: Fostering critical thinking in chemical engineering through in-class debates. *Education for Chemical Engineers, 51*, 121–132. https://doi.org/10.1016/j.ece.2025.03.002

Nurilma, F. R., Supriana, E., & Diantoro, M. (2023). Using STEM-based 3D-multimedia to improve students' critical thinking skills in uniform circular motion. *Jurnal Pendidikan Fisika, 11*(2), 193–201. https://doi.org/10.26618/jpf.v11i2.10785

Prabayanti, E., Usman, U., Khaeruddin, K., & Setiawan, T. (2025). Analysis of students' critical thinking abilities in physics learning: A case study at SMAN 5 Sidrap. *Jurnal Pendidikan Fisika, 12*(3), 141–151. https://doi.org/10.26618/jpf.v12i3.15317

Radu, I., Huang, X., Kestin, G., & Schneider, B. (2023). How augmented reality influences student learning and inquiry styles: A study of 1-1 physics remote AR tutoring. *Computers & Education: X Reality, 2*, 1-15. https://doi.org/10.1016/j.cexr.2023.100011

Rebello, C. M., Deiró, G. F., Knuutila, H. K., Moreira, L. C. de S., & Nogueira, I. B. R. (2024). Augmented reality for chemical engineering education. *Education for Chemical Engineers, 47*, 30–44. https://doi.org/10.1016/j.ece.2024.04.001

Tiruneh, D. T., De Cock, M., Weldeslassie, A. G., Elen, J., & Janssen, R. (2017). Measuring critical thinking in physics: Development and validation of a critical thinking test in electricity and magnetism. *International Journal of Science and Mathematics Education, 15*(4), 663–682. https://doi.org/10.1007/s10763-016-9723-0

Toulmin, S. E. (2015). *The uses of argument* (Updated ed.). Cambridge University Press.

Wu, J. F., Zhang, Z., Wu, Y., & Tang, X. J. (2025). The effectiveness of immersive media in promoting consumer products: Augmented vs. virtual reality. *Journal of Retailing and Consumer Services, 86*, 1-12. https://doi.org/10.1016/j.jretconser.2025.104323

Zhang, Y., Guo, W., Ji, B., & Shi, H. (2025). An exploratory case study of Chinese college students' critical thinking skills through interactive independent courses: Evidence from behavioral and brain imaging. *Thinking Skills and Creativity, 57*, 1-15. https://doi.org/10.1016/j.tsc.2025.101867