SPEKTRA: Jurnal Kajian Pendidikan Sains
Vol. 12, No. 2, 2026, pp. 450-464
ISSN: 2548-642X, DOI: 10.32699/2tdkvx19

LINKING SCIENTIFIC METHOD AND ARGUMENTATION IN PHYSICS LEARNING: DEVELOPING STUDENTS' ARGUMENTATION SKILL

Siswanto Siswanto[1], Nuryunita Dewantari[1], Heni Safitri[3], Bambang Subali[2], Ellianawati[2]
[1]Department of Science Education, Faculty of Teacher Training and Education, Universitas Tidar, Magelang, Indonesia
[2]Department of Physics, Faculty of Teacher Training and Education, Universitas Terbuka, Tangerang Selatan, Indonesia
[3]Department of Physics Education, Faculty of Mathematics and Science, Universitas Negeri Semarang, Semarang, Indonesia

Article Info

Article history:
Received 08/05/2026
Accepted 30/10/26
Published 05/11/2026

Keywords:
Argumentation activity;
Argumentation skill;
Argumentation on science learning;
Physics learning;
Scientific method.

ABSTRACT

This study examined how integrating argumentation activities into the scientific method affects students' argumentation skills in scientific quality and structural completeness. A quasi-experimental design using a non-equivalent control group was employed. An essay test was used to measure students' argumentation skills. Pretest results showed no significant differences between the two classes in either scientific quality or structural completeness. Pretest and posttest comparisons revealed significant improvements in both classes ($p < 0.001$), with greater improvements in the experimental class. Regarding scientific quality, the experimental class showed a significant shift from the misconception and non-scientific categories toward the Scientifically Acceptable category, indicating conceptual restructuring. The N-gain score for the experiment class were 0.839 (high category), for control class were 0.557 (medium category) and the intervention had a large effect ($d = 0.80$). In terms of structural completeness, the experimental class achieved higher levels, particularly Level 4, demonstrating the ability to construct complete arguments comprising claims, data, explanations, and supporting evidence. The N-gain score for the experiment class were 0.703 (high category), for control class were 0.303 (medium category) and a very large effect on structural completeness ($d = 1.12$). These findings suggest that explicitly integrating argumentation activities into the scientific method can improve both the scientific quality and structural completeness of students' arguments. This study extends previous research on argumentation-based science learning by providing empirical evidence of the value of embedding structured argumentation activities within a scientific-method learning sequence. Furthermore, it highlights the complementary dimensions of argumentation skills, emphasizing both the scientific acceptability of students' reasoning and the completeness of their argument structures. These findings underscore the importance of integrating structured argumentation activities into physics learning to foster students' scientific reasoning and argumentation skills.

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

Corresponding Author:
Siswanto
Department of Science Education, Faculty of Teacher Training and Education, Universitas Tidar, Magelang, Indonesia
Email: siswanto@untidar.ac.id

1. INTRODUCTION

Developing students' argumentation skills is a key focus in contemporary science education because it reflects students' ability to understand, construct, and evaluate scientific knowledge. Argumentation skill involves not only formulating claims but also use scientific evidence and logical reasons that strengthen the claim, which is an integral part of epistemic practice in science (Jiménez-aleixandre & Erduran, 2007; Jimenez et al., 2024; J. Osborne, 2010). This skill is very important in learning physics considering that there are many abstract physics concepts that require strong scientific reasoning (Pols et al., 2023; Redish & Burciaga, 2013).

In line with this, the use of the scientific method has been adopted as a primary approach in science learning. The scientific method emphasizes the process of inquiry that uses systematic scientific stages such as the activities of the science process (observation, formulating problems, conducting experiments, analyzing data, and drawing conclusions) (Gerde et al., 2013; Gower, 1997; Vo & Simmie, 2025; Voit, 2019). However, findings indicate that the use of the scientific method is merely procedural and does not actually involve basic science practices such as the involvement of scientific argumentation activities (Driver, 2000; Tang, 2022).

Research findings have proven the effectiveness of inserting arguments into science learning on good science learning outcomes (Mcneill & Krajcik, 2008). Most of this research focuses on argumentation-based learning models without explicitly linking them to the structure of the scientific method commonly used in curricula (Berland & Reiser, 2008; Sampson et al., 2011). This marks a gap between theory and practice in the implementation of scientific methods with the development of scientific arguments.

Recent research has established the value of scientific argumentation in science education. A systematic review by (Li et al., 2022), identified 12 experimental studies examining different forms of scientific argumentation-based instruction and found evidence of its potential to promote multidimensional conceptual change. More recently, a meta-analysis of 17 experimental studies reported a positive overall effect of argumentation on students' academic achievement, science process skills, conceptual knowledge, and attitudes toward science (Fu & Zhan, 2026). Recent research has also moved beyond treating argumentation as a separate instructional activity by integrating it into inquiry-oriented learning. For example, Pols et al. (2023) developed a physics inquiry sequence in which argumentation was integrated into practical inquiry activities to help students understand scientific evidence and engage in inquiry more independently. These developments indicate that research is increasingly concerned not only with whether argumentation is beneficial, but also with how it can be deliberately incorporated into science learning practices.

Despite the potential of argumentation-based learning, students may still experience difficulties in developing scientific argumentation competence, including identifying, evaluating, and generating the components of scientific arguments (Zhang & Browne, 2022). Explicit instruction can support the development of students' argumentation skills, while opportunities to apply argumentation in meaningful learning contexts can further support this development (Khishfe, 2022). In particular, clear guidance on the features of quality arguments, opportunities for individual and collaborative practice, and gradual scaffolding have been associated with improvements in students' argument quality (Dawson, 2024). These findings suggest that scientific argumentation needs to be deliberately supported through structured learning activities

rather than being expected to emerge automatically from students' participation in scientific practices.

Recent studies have explored different instructional configurations for integrating or supporting scientific argumentation in physics learning. (Pols et al., 2023), for example, integrated argumentation into physics inquiry to help students understand and use scientific evidence in conducting inquiry. Muslim et al. (2025) employed an Argument-Generating learning model supported by Toulmin's Argumentation Pattern to develop students' ability to construct and defend arguments in physics. More recently, Nurjannah et al. (2025), integrated the scientific method into an Argument-Driven Inquiry framework and examined its effects on students' cognitive and argumentation outcomes. However, these studies adopt different instructional configurations, with inquiry or argumentation-oriented models serving as the primary learning framework. Although these studies demonstrate the value of integrating argumentation into science learning, they employ different instructional frameworks. In ADI, argumentation is embedded within an inquiry-oriented instructional framework, whereas the present study examines explicit argumentation activities within a scientific-method learning sequence. Thus, the specific contribution of embedding explicit argumentation activities within scientific-method instruction, compared with scientific-method instruction without explicit argumentation, remains insufficiently examined. Moreover, limited attention has been given to examining this contribution in terms of both the scientific quality and structural completeness of students' arguments.

To address these gaps, this study embeds explicit scientific argumentation activities within a scientific-method learning sequence in physics instruction. Rather than proposing a new argumentation-oriented instructional model, this study investigates the contribution of explicitly incorporating argumentation activities into an established scientific-method framework. A quasi-experimental comparison is conducted between scientific-method instruction integrated with argumentation activities and scientific-method instruction without explicit argumentation. This design enables the study to examine how explicit argumentation activities contribute to two complementary dimensions of students' argumentation skills: the scientific quality of their arguments and the structural completeness of their argumentation. Thus, this study provides empirical evidence on the contribution of explicitly embedded argumentation to the development of students' scientific arguments within scientific-method instruction.

2. METHOD

Research Design

This study used a quantitative quasi-experimental method with a non-equivalent control group pretest–posttest design. This design was selected to compare the effectiveness of scientific method learning with and without integrated argumentation activities in developing students' argumentation skills. A quasi-experimental design is appropriate when participants are assigned to groups without individual random assignment, particularly in educational settings where existing classroom groups are maintained (Denny et al., 2023). In this study, two existing classes were used as the experimental and control groups. The experimental class received scientific method learning integrated with argumentation activities, while the control class received scientific method learning without argumentation activities. Pretests were given to identify students' initial argumentation skills before the intervention, while posttests were given to assess

their argumentation skills after the intervention. The use of pretest and posttest also allowed the initial status of the two groups to be considered when interpreting the results of the quasi-experimental comparison. The learning process was conducted over three meetings on the topic of temperature and heat.

Participants

The participants were selected using purposive sampling based on existing classroom groups. Two existing classes were purposively selected to participate in the study. The selection considered the requirements of the instructional intervention and the comparison of the two learning conditions. One class was assigned as the experimental group and the other as the control group. The study involved 68 students at a high school in Central Java Province, Indonesia, with 34 students in each class. All students in the selected classes participated in the learning activities, pretest, and posttest throughout the study.

Data Collection and Analysis

Data collection was conducted using a student argumentation skills test, consisting of three descriptive questions. The test was developed adopted by Cetin (2014). This test was developed to measure students' argumentation skills, both in terms of scientific aspects and the completeness of their argumentative structures. The argumentation skills indicators measured in this study refer to the Toulmin Argumentation Pattern (Jiménez-Aleixandre & Erduran, 2007; Muslim et al., 2025).

For the scientific aspects of their arguments, the researchers grouped them into five categories, as shown in Table 1. These five categories were then scored, as shown in Table 1. Then, test the difference in the pretest means of the experimental and control classes, then test the difference in the pretest and posttest means of each class. Then, calculating the value of the improvement in argumentation skills for each category is calculated using N-gain. Finally, test the effect size. Furthermore, the percentage of students in each category in each class, both before and after the treatment, was calculated. Specifically for the misconception category, the three questions used to measure argumentation skills refer to the topic of misconceptions on the temperature and heat (Kartal et al., 2011). The misconception points can be seen in Table 2.

Table 1. The Scientific Categories

Category	Description	Score
Scientific acceptable (SA)	All answers are scientifically accepted	2
Partially Scientific acceptable (P-SA)	Some answers are scientifically correct	1
Non-Scientific acceptable (N-SA)	Incorrect answers	0
Misconception (Mis)	Misconception answers	0
Without explain (WE)	No answer/ only making claims without explaining	0

Table 2. Misconception Points in Argumentation Questions

No	Questions	Misconception Points
1	Two objects are heated using the same heater for the same amount of time. - First object: 1 kg of water (c = 4200 J/kg ºC) - Second object: 1 kg of oil (c = 2000 J/kg ºC) Based on this, both oil and water will show the same temperature when measured using the same thermometer. Do you agree with this statement? Explain!	- Temperature is assumed to depend only on mass - All objects will have the same temperature response to heat
2	Andi heats 500 grams of ice until it melts. Then, he also heats 500 grams of wax until it melts. Do you agree that the heat required to melt 500 grams of ice and 500 grams of wax is the same? Explain!	- Heat only depends on mass - The type of substance does not affect the heat requirement - All processes of change of state are considered the same
3	A child puts two different spoons into a glass of hot water. - Spoon A is made of iron - Spoon B is made of copper A few moments later, the child touches both ends of the spoons and feels that the copper end heats up faster than the iron end. Do you agree with the child? Explain!	- The rate of heat transfer is determined by the mass of the object

For the aspect of organ completeness, researchers grouped them into four levels, as shown in Table 3. The four levels were given scores as shown in Table 3. Then, test the difference in the pretest means of the experimental and control classes, the difference in the pretest and posttest means of each class, and test the effect size. Furthermore, the percentage of students in each category in each class, both before and after the treatment, was calculated.

Table 3. Level of Completeness of Argumentation Organs

Category	Description	Score
Level 1	Claim only	1
Level 2	Contains claim elements and data	2
Level 3	Contains claim elements, data, explanation	3
Level 4	Contains claim elements, data, explanation, and supporting evidence	4

Before being used in the study, the argumentation skills test instrument was validated by two experts in science education. Both experts hold doctoral degrees in science education and have more than five years of teaching experience. The validation was conducted to assess the feasibility

and content appropriateness of the test instrument for measuring students' argumentation skills. Four aspects were assessed: (1) the suitability of the test items with the indicators of argumentation skills, (2) the clarity and appropriateness of the language used in the test items, (3) the suitability of the questions with the relevant scientific concepts, and (4) the suitability of the expected answers with the questions. Based on the experts' assessment, the test instrument was declared feasible and could be used to measure students' argumentation skills, with a Content Validity Index (CVI) value of 1.0.

Inter-rater reliability was used to assess the consistency of scoring between raters. Two experts with experience in scientific argumentation independently evaluated 20% of the student responses to assess the reliability of the scoring process (Cetin, 2014). In this study, 14 student responses from both classes were evaluated by the two raters. The scientific aspect obtained a coefficient of 0.63, categorized as moderate agreement, with a 93% agreement percentage, while the completeness of argumentation obtained a coefficient of 1, categorized as perfect agreement, with a 100% agreement percentage.

Classroom Intervention

The learning intervention was implemented over three meetings on the topic of temperature and heat in both of groups. Both of classes were taught using the scientific method. The main difference between the two classes lies in the explicit integration of scientific argumentation practices (in the experiment class). In the control class, learning took place without a specific emphasis on argumentation development. Further differences in the interventions can be seen in Table 4. Table 5 contains the learning topics used during the intervention on the material of temperature and heat.

Table 4. Learning Intervention Framework

Control Class	Experimental Class
Observing	Observing
Asking	Asking
Reasoning	Reasoning
Trying	Trying
Presenting	Formulating arguments (claims, data, warrants, backing)
-	Presenting and revising arguments

Table 5. Learning topics

Meeting	Material
Meeting 1	- The effect of mass on the amount of heat required to raise the temperature of an object. - The effect of the type of substance on raising the temperature of an object.
Meeting 2	- The influence of mass on changes in the state of an object. - The effect of the type of substance on changes in the state of an object.
Meeting 3	- The effect of the type of substance on the heat rate

3. RESULTS AND DISCUSSION

3.1 Development of Argumentation Skills on the Scientific Aspect

Data analysis began with a data normality test for both classes. The test results showed that all data had a significance value of less than 0.05, thus concluding that the data were not normally distributed. Based on these findings, further inferential analysis was conducted using a non-parametric statistical approach.

To test for equivalence in initial abilities between the two groups, a Mann-Whitney U test was used to test for differences in pretest scores. The analysis results showed that there was no significant difference in pretest data between the experimental and control classes ($p = 0.732 > 0.05$). These test results indicate that the initial ability of students' argumentation skills in the experimental and control classes in the scientific aspect was the same.

Next, a test of mean differences between the pretest and posttest was conducted in each class using the Wilcoxon signed-rank test. The test results showed an increase in argumentation skills in the scientific aspect in both classes ($p < 0.001$). That's mean the learning intervention in the both of classes had a significant impact on improving argumentation skills.

However, Table 6 shows a difference in the quality of argumentation skills improvement in the scientific aspect both of classes. The improvement in the quality of argumentation in the scientific aspect in the experimental class showed better results than in the control class. Table 6 shows that in the experimental class, there was a more dominant shift to the Scientifically Acceptable category, while in the control class, the shift was predominantly to the Partially Acceptable category. These results demonstrate that the improvement in the experimental class was not only quantitative, but indeed reflected a better quality of argumentation than in the control class in the scientific aspect.

Furthermore, the improvement in the quality of argumentation in the experimental class, which was better than the control class for the scientific aspect, was also proven by N-gain score and the effect size test. The N-gain score for the experiment class were 0.839 (high category) and for control class were 0.557 (medium category). The results of the Cohen's d test for the difference in effect between the experiment and control classes showed a score of $d = 0.8$, in the large effect category. This value indicates that the learning intervention had a significant practical impact on improving the quality of students' scientific argumentation. Therefore, it can be concluded that the integration of argumentation activities into scientific method-based learning is not only effective in improving argumentation skill scores but also substantially improves the scientific quality of students' argumentation. Table 7 contains the examples of student responses.

Table 6. Scientific Aspect on Argumentation Skill Achievement in the Control and Experimental Classes

Category	Control Class (%)		Experimental Class (%)	
	Pretest	Posttest	Pretest	Posttest
Scientific acceptable (SA)	0	38,24	0	76,47
Partially Scientific acceptable (P-SA)	20,59	44,12	17,65	17,65
Non-Scientific acceptable (N-SA)	44,12	0	47,06	0
Misconception (Mis)	35,29	17,65	35,29	5,88
Without explain (WE)	0	0	0	0

Table 7. The Examples of Student Answers for Scientific Aspects

Category	Example of Student Answers
Scientific Acceptable (SA)	In my opinion, temperature increase depends not only on the mass of an object but also on the specific heat of the substance and the amount of heat added. Even if they have the same mass, two substances can experience different temperature increases because their respective abilities to absorb heat are different. Experimental results show that the substance with the lower specific heat increases temperature more quickly. So, mass does have an effect, but it's not the only factor determining temperature change.
Partially Scientifically Acceptable (P-SA)	In my opinion, the temperature of an object isn't solely determined by mass. While a larger mass usually takes longer to heat up, the type of substance also plays a role. Experiments show that even with the same mass, the temperature increase can be different. So, besides mass, there are other factors that influence temperature; I'm just not sure what the concept is.
Non-Scientific Acceptable (N-SA)	In my opinion, the temperature of an object is more influenced by its size. If the object is smaller, its temperature rises more quickly, while if it's larger, its temperature rises more slowly. So, the temperature difference is due to the shape and size of the object, not anything else.
Misconceptions (Mis)	In my opinion, temperature depends solely on the object's mass. If the mass is greater, its temperature will be lower because the heat is distributed among more objects. Conversely, if the mass is small, the temperature will be higher because the heat is concentrated in a small area. So, the only thing that determines temperature is the mass of an object.

The results of the study indicate that the integration of argumentation activities into scientific method-based learning significantly improved students' scientific argumentation. The lack of significant differences in the pretest between the experimental and control classes indicates that both groups had equivalent initial abilities, and therefore, the changes observed in the posttest can be attributed to the learning intervention. This is in accordance with the basic principle in quasi-experimental research that equality of initial abilities in both classes is very important to estimate the effects of the intervention carried out (Cohen et al., 2011; Creswell & Creswell, 2017).

In the experimental class, the improvement in the quality of scientific argumentation in the scientific aspect, as indicated by the dominant shift to the Scientifically Acceptable category, proves that the explicit integration of argumentation in the scientific method is able to build better conceptual knowledge of students. This is in accordance with research by Osborne (2010), that argumentation is a basic scientific practice that must be presented in science learning because it is able to train students to build and construct scientific concepts as a whole. In addition, according to Jiménez-Aleixandre & Erduran (2007), the quality of scientific argumentation built by students reflects the quality of students' understanding of a scientific concept.

Another important finding concerns misconceptions. In the experimental class, misconceptions were reduced better than in the control class. According to (Potvin et al., 2020),

conceptual change can occur in students when they experience cognitive conflict in their minds, which will encourage the process of conceptual restructuring. In the context of this study, the argumentation activity intervened in the experimental class was able to trigger cognitive conflict in students' minds through discussion and structuring of claims, data, warrants, and support. The research findings indicate that misconceptions cannot be changed simply by presenting conceptual information; rather, an active process involving evaluation and reflection is needed, one of which is through argumentation practice (Siswanto et al., 2022).

Furthermore, the findings of misconceptions, such as "Heat only depends on mass," indicate that students' conceptions are intuitive and unscientific. However, after the learning process, there was a significant decrease as students began to reconstruct their concepts scientifically through the learning activities. This aligns with findings in science learning that evidence-based argumentation can help students connect concepts more coherently (J. F. Osborne et al., 2015; Siswanto et al., 2022)

In this study, the development of scientific argumentation skills also provides strong evidence that the implementation of basic science practices places scientific argumentation at the center of constructing holistic knowledge. Through argumentation practice, students not only memorize scientific concepts but also construct holistic and meaningful concepts. This is supported by research by (Sampson et al., 2010), which showed that engaging in argumentation improves the quality of students' scientific reasoning.

The large effect size ($d = 0.80$) indicates that the intervention has a strong practical impact. This aligns with previous research showing that argumentation-based learning has a significant impact on conceptual understanding and the quality of scientific argumentation (Cetin, 2014; Siswanto et al., 2022). Thus, the integration of argumentation in learning not only improves learning outcomes, but also develops deeper scientific thinking skills.

3.2 Development of Argumentation Skills on the Completeness of Organs

Data analysis began with a data normality test for both classes. The test results showed that all data had a significance value of less than 0.05, thus concluding that the data were not normally distributed. Thus, non-parametric statistics was used.

Same with the scientific aspect, to test for equivalence in initial abilities between the two groups, a Mann-Whitney U test was used to test for differences in pretest scores. The analysis results showed that there was no significant difference in pretest data both of classes ($p = 0.41 > 0.05$). These test results indicate that the initial ability of students' argumentation skills on the both of classes in the organ completeness aspect was the same. Next, a test of mean differences between the pretest and posttest was conducted in each class using the Wilcoxon signed-rank test. The test results showed an increase in argumentation skills in the organ completeness aspect in both classes ($p < 0.001$). This result indicates that the learning intervention in the both of classes had a significant impact on improving argumentation skills on the organ completeness.

However, the development of argumentative skills in the aspect of completeness of organs in the experimental class showed a trend of results that were much better when compared to the control class (Table 8). In the experimental class, the trend of shifting was dominated by level 4, which means developing towards the preparation of complete and perfect arguments including claims, data, explanations, and complete support. This is different from the control class, where the trend of shifting was dominated by levels 2 and 3 (the shift to level 4 was much less). This

finding is strong evidence that the improvement in argumentative skills in the experimental class is not only quantitative, but shows a better and more comprehensive quality of development.

Same with the scientific aspect, the improvement in the quality of argumentation in the experimental class, which was better than the control class for the organ completenest aspect, was also proven by N-gain score and the effect size test. The N-gain score for the experiment class were 0.703 (high category) and for control class were 0.303 (medium category). The results of the Cohen's d test for the difference in effect between the experimental and control classes showed a score of d = 1.12, in the very large effect category.

These results demonstrate that explicit argumentation involvement has a significant impact on building a comprehensive and systematic argumentative structure. This finding indicates that integrating argumentative activities and scientific methods can develop students' argumentative skills, not only in terms of the scientific quality of the text, but also in terms of the structure and completeness of the text. Examples of student responses for each level can be seen in Table 9.

Table 8. Achievement of Argumentation Skills in the Completeness of Organs Aspect in the Control and Experimental Classes

Level	Control Class (%)		Experimental Class (%)	
	Pretest	Posttest	Pretest	Posttest
Level 1	0	0	0	0
Level 2	94,12	44,12	88,24	11,76
Level 3	5,88	47,06	11,76	32,35
Level 4	0	8,82	0	55,88

Table 9. The Example of Student Answers for the Aspect of Organ Completeness

Level	Example of Student Answers
Level 2	In my opinion, metal heats up more quickly than wood (Claim). When touched, metal will definitely feel hotter than wood, even though both are heated in the same place (Data).
Level 3	In my opinion, metal will be hotter than wood (claim). When touching metal, your hand will feel hotter than when touching wood (Data). This is because heat will transfer through metal more easily than wood when we touch both, so the heat will reach our hand more easily when touching metal (Warrant).
Level 4	In my opinion, metal will feel hotter than wood (claim). Metal has better heat conductivity than wood, so when touched metal will feel hotter than wood even though both are in the same place (Data). Metal is a good conductor, unlike wood, so the heat from metal travels faster to our hands than when we touch wood, resulting in heat transfer from wood to hands being slower (warrant). Based on the concept of heat, objects that have a high conductivity value will definitely transfer heat faster (backing).

Based on Table 8, the findings indicate that the experimental class experienced a more significant improvement, with a shift in argumentation skills from the dominant aspect of completeness to level 4. This indicates that the experimental class was able to develop argumentation skills with complete argumentation structures according to Toulmin's

Argumentation Framework (Magalhães, 2020). This finding provides evidence that the intervention, in addition to developing scientific aspects, also developed students' skills in organizing systematic and structured arguments.

The results of this study also support the claim that argumentation skills cannot develop naturally in students through conventional science learning but must be directly cultivated through argumentation practices explicitly and systematically presented in the science learning process. This aligns with several research findings that argue that argumentation training must be explicitly integrated into learning so that students can develop complete and systematic argumentation structures appropriate for constructing coherent concepts (Mcneill & Krajcik, 2008). (J. Osborne et al., 2004) also emphasized the importance of direct argumentation training in the science learning process.

Upon closer examination, the findings across the two indicators suggest that direct argumentation training supported complementary dimensions of students' scientific argumentation. Table 6 shows that students in the experimental class increasingly produced Scientifically Acceptable arguments, indicating an improvement in the scientific quality of their reasoning. Complementing this result, Table 8 shows a greater shift toward Level 4 in the experimental class, while Table 9 illustrates arguments that include claims, data, explanations, and supporting evidence. Taken together, these patterns suggest that students not only developed more scientifically acceptable ideas but also became better able to organize and justify those ideas in more complete argument structures. This interpretation is consistent with scaffolding theory, which emphasizes the role of structured instructional support in developing complex thinking skills (Lin et al., 2012; Tabak & Kyza, 2018). It also aligns with research suggesting that argumentation as a scientific practice can support meaningful knowledge construction (Clark et al., 2009) and help students construct more complete and logical arguments (Berland & Neill, 2010; Berland & Reiser, 2008).

The very large effect size ($d = 1.12$) indicates a very large difference between the experimental and control classes in the completeness of argumentation structures. When this result is considered together with the large effect size in the scientific aspect ($d = 0.80$), it indicates that the integration of argumentation activities supported the development of students' argumentation skills in two complementary aspects. In the scientific aspect, students in the experimental class were able to construct more Scientifically Acceptable arguments, while in the completeness aspect, they were better able to organize their arguments using claims, data, explanations, and supporting evidence. These results indicate that the development of argumentation skills involves not only the scientific quality of students' reasoning but also the completeness of the structure used to express and justify that reasoning. This is in line with the view that scientific argumentation is a multidimensional construct in which scientific content and argument structure need to be considered together (Haudek & Zhai, 2024).

Thus, the results of this study indicate that integrating argumentation activities into scientific method–based learning is an effective approach for developing students' scientific argumentation skills comprehensively. The intervention was associated not only with improvement in students' ability to construct scientifically acceptable reasoning but also with improvement in their ability to express and justify that reasoning in complete argument structures. This finding is consistent with research showing that argumentation instruction can improve students' use of evidence, reasoning, and justification in solving scientific problems

(Iwuanyanwu, 2023). These findings also align with research emphasizing the importance of argumentation in developing scientific literacy (J. Osborne, 2010; Schwarz et al., 2003; Siswanto et al., 2022).

4. CONCLUSION

This study demonstrates that integrating argumentation activities into scientific method-based learning is effective in improving students' argumentation skills. In terms of scientific aspects, the learning intervention was able to encourage a shift from misconceptions to a more valid scientific understanding, thus improving not only the accuracy of answers but also the quality of scientific justification. In terms of the completeness of argumentation, students in the experimental class demonstrated better ability to construct complete and systematic arguments, reaching Level 4. The findings of this study provide empirical evidence that argumentative skills cannot develop automatically in students during science learning without direct argumentative practice. Developing argumentative skills requires explicit and systematic training directly integrated into science learning. Effect size analysis, which showed large to very large categories, reinforces that the intervention has a significant practical impact. Therefore, integrating argumentation activities into scientific methods is an effective way to develop students' argumentation skills, both in terms of scientific quality and the completeness of argumentation organs.

REFERENCES

Berland, L. K., & Neill, K. L. M. C. (2010). A Learning Progression for Scientific Argumentation : Understanding Student Work and Designing Supportive. *ScienceEducation*, 765–793. https://doi.org/10.1002/sce.20402

Berland, L. K., & Reiser, B. J. (2008). Making Sense of Argumentation and Explanation. *ScienceEducation*. https://doi.org/10.1002/sce.20286

Cetin, P. S. (2014). Explicit argumentation instruction to facilitate conceptual understanding and argumentation skills. *Research in Science & Technological Education*, *32*(1), 1–20. https://doi.org/10.1080/02635143.2013.850071

Clark, D., Sampson, V., Stegmann, K., Marttunen, M., Kollar, I., & Janssen, J. (2009). Scaffolding Scientific Argumentation Between Multiple Students in Online Learning Environments to Support the Development of 21 st Century Skills. *In the National Academies' Board on Science Education Workshop on Exploring the Intersection of Science Education and 21st Century Skills*, 1–44.

Cohen, L., Manion, L., & Morrison, K. R. B. (2011). *Research Methods in Education*. Routledge. https://books.google.co.id/books?id=p7oifuW1A6gC

Creswell, J. W., & Creswell, J. D. (2017). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications. https://books.google.co.id/books?id=KGNADwAAQBAJ

Dawson, V. (2024). Teachers' support in developing year 7 students' argumentation skills about water-based socioscientific issues. *International Journal of Science Education*, *46*(3), 222–239. https://doi.org/10.1080/09500693.2023.2226334

Denny, M., Denieffe, S., & O'Sullivan, K. (2023). Non-equivalent Control Group Pretest–Posttest Design in Social and Behavioral Research. In A. L. Nichols & J. Edlund (Eds.), *The Cambridge Handbook of Research Methods and Statistics for the Social and Behavioral Sciences: Volume 1: Building a Program of Research* (pp. 314–332). Cambridge University Press. https://doi.org/DOI: 10.1017/9781009010054.016

Driver, R. (2000). Establishing the Norms of Scientific Argumentation in Classrooms. *Science Education*, *84*(3), 287–312.

Fu, X., & Zhan, Q. (2026). Effectiveness of argument in science education: a meta-analysis based on the results of 17 experiments. *Research in Science & Technological Education*, *44*(2), 469–489. https://doi.org/10.1080/02635143.2025.2527603

Gerde, H. K., Schachter, R. E., & Wasik, B. A. (2013). Using the Scientific Method to Guide Learning: An Integrated Approach to Early Childhood Curriculum. *Early Childhood Education Journal*, *41*(5), 315–323. https://doi.org/10.1007/s10643-013-0579-4

Gower, B. (1997). *Scientific Method: A Historical and Philosophical Introduction (1st ed.).* Routledge. https://doi.org/https://doi.org/10.4324/9780203046128

Haudek, K. C., & Zhai, X. (2024). Examining the Effect of Assessment Construct Characteristics on Machine Learning Scoring of Scientific Argumentation. *International Journal of Artificial Intelligence in Education*, *34*(4), 1482–1509. https://doi.org/https://doi.org/10.1007/s40593-023-00385-8

Iwuanyanwu, P. N. (2023). When Science Is Taught This Way , Students Become Critical Friends : Setting the Stage for Student Teachers. *Research in Science Education*, *53*(6), 1063–1079. https://doi.org/10.1007/s11165-023-10122-9

Jiménez-aleixandre, M. P., & Erduran, S. (2007). *Chapter 1 Argumentation in Science Education : An Overview* (pp. 3–4).

Jiménez-Aleixandre, M. P., & Erduran, S. (2007). *Argumentation in Science Education: An Overview BT - Argumentation in Science Education: Perspectives from Classroom-Based Research* (S. Erduran & M. P. Jiménez-Aleixandre (eds.); pp. 3–27). Springer Netherlands. https://doi.org/10.1007/978-1-4020-6670-2_1

Jimenez, P. C., Alred, A. R., & Dauer, J. M. (2024). Describing undergraduate students' reasoning and use of evidence during argumentation about socioscientific issues systems. *Frontiers in Education*, *Volume 9-2024*. https://doi.org/10.3389/feduc.2024.1371095

Kartal, T., Öztürk, N., & Yalvaç, H. G. (2011). Misconceptions of science teacher candidates about heat and temperature. *Procedia - Social and Behavioral Sciences*, *15*, 2758–2763. https://doi.org/https://doi.org/10.1016/j.sbspro.2011.04.184

Khishfe, R. (2022). Nature of Science and Argumentation Instruction in socioscientific and scientific contexts. *International Journal of Science Education*, *44*(4), 647–673. https://doi.org/10.1080/09500693.2022.2050488

Li, X., Wang, W., & Li, Y. (2022). Systematically reviewing the potential of scientific argumentation to promote multidimensional conceptual change in science education. *International Journal of Science Education*, *44*(7), 1165–1185. https://doi.org/10.1080/09500693.2022.2070787

Lin, T.-C., Hsu, Y.-S., Lin, S.-S., Changlai, M.-L., Yang, K.-Y., & Lai, T.-L. (2012). *A review of empirical evidence on scaffolding for science education. International Journal of Science and Mathematics Education*, *10*(2), 437–455. https://doi.org/10.1007/s10763-011-9322-z

Magalhães, A. L. (2020). Teaching How to Develop an Argument Using the Toulmin Model. *International Journal of Multidisciplinary and Current Educational Research (IJMCER)*, *2*(3), 1–7.

Mcneill, K. L., & Krajcik, J. (2008). Scientific Explanations : Characterizing and Evaluating the Effects of Teachers ' Instructional Practices on Student Learning. *Journal of Research in Science Teaching*, *45*(1), 53–78. https://doi.org/10.1002/tea

Muslim, Sihombing, R. A., & Putra, I. H. (2025). The Power of Argumentation: A Systematic Literature Review on Science Learning and Multiple Skills Development. *Asian Journal of University Education*, *21*(3), 814–832. https://doi.org/10.24191/ajue.v21i3.55

Nurjannah, Sani, N. K., Kade, A., & Khuzaimah, A. U. (2025). Jurnal Pendidikan Fisika Enhancing Cognitive and Argumentation Skills through Integration of Argument-Driven Inquiry and the Scientific Method in Physics Education. *Jurnal Pendidikan Fisika*, *13*(3), 291–307. https://doi.org/10.26618/9h25w918

Osborne, J. (2010). Arguing to Learn in Science : The Role. *Science*, *463*. https://doi.org/10.1126/science.1183944

Osborne, J., Erduran, S., & Simon, S. (2004). ENHANCING THE QUALITY OF ARGUMENTATION IN SCHOOL SCIENCE. *Journal of Research in Science Teaching*, *41*(10), 994–1020.

Osborne, J. F., Henderson, J. B., Macpherson, A., Szu, E., & Wild, A. (2015). The Development and Validation of a Learning Progression for Argumentation in Science. *Journal of Research in Science Teaching*, *53*(6), 821-846. https://doi.org/10.1002/tea.21316

Pols, C. F. J., Dekkers, P. J. J. M., & Vries, M. J. De. (2023). Integrating argumentation in physics inquiry : A design and evaluation study. *Physical Review Physics Education Research*, *19*(2), 20170. https://doi.org/10.1103/PhysRevPhysEducRes.19.020170

Potvin, P., Nenciovici, L., Malenfant-Robichaud, G., Thibault, F., Sy, O., Mahhou, M. A., Bernard, A., Allaire-Duquette, G., Blanchette Sarrasin, J., Brault Foisy, L.-M., Brouillette, N., St-Aubin, A.-A., Charland, P., Masson, S., Riopel, M., Tsai, C.-C., Bélanger, M., & Chastenay, P. (2020). Models of conceptual change in science learning: establishing an exhaustive inventory based on support given by articles published in major journals. *Studies in Science Education*, *56*(2), 157–211. https://doi.org/10.1080/03057267.2020.1744796

Redish, E. F., & Burciaga, J. R. (2013). *Teaching Physics with the Physics Suite* (Vol. 414, Issue 2004). Hoboken, NJ: John Wiley & Sons.

Sampson, V., Grooms, J., & Walker, J. O. I. P. (2010). Argument-Driven Inquiry as a Way to Help Students Learn How to Participate in Scientific Argumentation and Craft Written Arguments : An Exploratory Study. *ScienceEducation*, *95*(2), 217–257. https://doi.org/10.1002/sce.20421

Sampson, V., Grooms, J., & Walker, J. P. (2011). Argument-Driven Inquiry as a way to help students learn how to participate in scientific argumentation and craft written arguments: An exploratory study. *Science Education*, *95*(2), 217–257. https://doi.org/10.1002/sce.20421

Schwarz, B. B., Neuman, Y., Gil, J., & Ilya, M. (2003). Construction of Collective and Individual Knowledge in Argumentative Activity. *Journal of the Learning Sciences*, *12*(2), 219–256. https://doi.org/10.1207/S15327809JLS1202_3

Siswanto, Hartono, Subali, B., & Mastruri, M. (2022). Infusing Explicit Argumentation in Science Reading Activities : Helping Prospective Science Teachers Reduce Misconception and Foster

Argumentation Skills. *Pegem Journal of Education and Instruction*, *12*(3), 177–189. https://doi.org/10.47750/pegegog.12.03.19

Tabak, I., & Kyza, E. A. (2018). Research on Scaffolding in the Learning Sciences. In *international Handbook of the Learning Sciences* (1 st, p. 10). Routledge.

Tang, K. (2022). *Material inquiry and transformation as prerequisite processes of scientific argumentation : Toward a social-material theory of argumentation*. *January*, 969–1009. https://doi.org/10.1002/tea.21749

Vo, D. Van, & Simmie, G. M. (2025). Assessing Scientific Inquiry: A Systematic Literature Review of Tasks, Tools and Techniques. *International Journal of Science and Mathematics Education*, *23*(4), 871–906. https://doi.org/10.1007/s10763-024-10498-8

Voit, E. O. (2019). Perspective : Dimensions of the scientific method. *PLOS Computational Biology*, *15*(9), 1–14. https://doi.org/https://doi.org/10.1371/journal.pcbi.1007279 September

Zhang, J., & Browne, W. J. (2022). Exploring Chinese high school students ' performance and perceptions of scientific argumentation by understanding it as a three-component progression of competencies. *Journal of Research in Science Teaching*, *December 2021*, 847–884. https://doi.org/10.1002/tea.21819