

The Effect of the Think Pair Share Learning Model on Fourth-Grade Students' Critical Thinking Skills in IPAS at Elementary School

Alya Firzanah¹, Chairul Amriyah², Ayu Reza Ningrum³

^{1,2,3}Universitas Islam Negeri Raden Intan Lampung

Jl. Letnan Kolonel H. Endro Suratmin, Sukarame Sub-district, Bandar Lampung, Lampung, Indonesia

¹alyafirzanah948@gmail.com

Abstract. Critical thinking is an essential competency in elementary science education, enabling students to analyze information, provide reasoned arguments, and make logical decisions. This study examined differences in students' critical thinking skills after instruction using the Think Pair Share (TPS) and Numbered Heads Together (NHT) models in Grade IV science at SDN 5 Jatimulyo, South Lampung, Indonesia. Conducted in August–September 2026, the study employed a quantitative approach with a quasi-experimental non-equivalent control group design. The participants were 72 students: 35 in the TPS group and 37 in the NHT group. Data were collected using a critical thinking skills test and analyzed through descriptive statistics, prerequisite tests, an independent-samples t-test, and Cohen's d. Grounded in Vygotsky's social constructivist theory, the learning models emphasize knowledge construction through social interaction and collaboration. The results showed higher mean posttest scores and N-Gain in the TPS group (59.71; 0.20) than in the NHT group (43.33; -0.04). Significant differences were found in *posttest* scores ($p=0.012$) and N-Gain ($p=0.037$), with a moderate effect size ($d=0.503$).

Keywords: Think Pair Share; Numbered Heads Together; Critical Thinking; Science; Elementary School.

1. Introduction

Critical thinking skills are one of the important competencies that need to be developed at the elementary school level because they are related to students' abilities in understanding information, analyzing problems, providing logical reasons, and making appropriate decisions (O'Reilly et al., 2022; Sasmita et al., 2023; Józsa et al., 2024; Novandi et al., 2025; Mametja et al., 2026). Dwyer et al. (2014) and Ennis (2011) explain that critical thinking is a reflective, rational process for determining actions in response to a problem. In the context of 21st-century education, critical thinking skills are part of the 4C skills (critical thinking, creativity, collaboration, and communication) that students need to meet the demands of learning and life. (Miller et al., 2023; Herlinawati et al., 2024; Nahar & Machado, 2025) Therefore, learning in elementary schools needs to be directed not only at mastering concepts, but also at the ability to use knowledge to analyze and solve problems.

One subject closely related to the development of critical thinking skills is Natural and Social Sciences (IPAS). Learning in IPAS is not only oriented towards understanding facts and concepts, but also encourages students to observe phenomena, connect information, provide arguments, and find solutions to various problems (Taylor, 2022; Nur'ariyani et al., 2023; Fatimah et al., 2026). However, the learning approach strongly influences these abilities. Learning that is still dominated

by the teacher can limit students' opportunities to explore ideas, discuss, and build understanding through thinking activities (Lombardi et al., 2021; Kozanitis & Nenciovici, 2023; Bhardwaj et al., 2025). Therefore, a learning model is needed that can increase student engagement through thinking activities and collaborative interactions.

These problems were found in the learning of science at SDN 5 Jatimulyo. Based on the analysis of the Mid-Semester Exam (UTS) scores for fourth-grade science, some students still had not achieved the Learning Objective Achievement Criteria (KKTP) of 66. In class IV A, out of 35 students, 15 students achieved mastery, and 20 students had not achieved mastery. Meanwhile, in class IV B, there were 21 out of 37 students who achieved mastery and 16 students who had not achieved mastery. Overall, out of 72 students, 36 students (50%) had not achieved mastery. The data were used as an initial picture of the achievement of science learning and were not used as a direct indicator that students' critical thinking skills were low. The results of observations and interviews with teachers showed that learning still tended to be teacher-centered. Students still had limitations in discussing, expressing opinions, providing reasons, analyzing information, and solving problems independently or in groups.

Table 1. Science Learning Completion Data for Grade IV SDN 5 Jatimulyo

Class	Number of Students	Completed	Not Completed	Percentage of Incomplete
IV A	35	15	20	57.14%
IV B	37	21	16	43.24%
Total	72	36	36	50%

One relevant learning model for addressing this problem is *Think Pair Share* (TPS). Developed by Frank Lyman in 1981, this model has three main stages: *think*, *pair*, and *share*. (Alsmadi et al., 2023; Iin Suriya Ningsih et al., 2025) In the *think* stage, students can understand the problem and organize their thoughts individually. Next, the *pair* stage allows students to discuss, compare ideas, and refine their understanding through interaction with a partner. In the *share* stage, students convey their thoughts to a wider group (Deore & Arora, 2022; Firja MKS, 2024). This stage supports the development of critical thinking skills because students are involved in the process of interpreting the problem, considering alternative answers, providing reasons, and communicating the results of their thinking.

Previous research shows that TPS is related to increased students' thinking activities and critical thinking skills. Ekamilasari et al. (2021) showed that critical thinking skills in science learning still need strengthening through learning models that give students space to analyze and build arguments. Alsaleh (2020) explained that collaborative learning can support critical thinking skills through the activity of exchanging ideas and evaluating various perspectives. In addition, Muhfahroyin et al. (2024) found that the application of TPS combined with *Problem-Based Learning* can increase students' critical thinking skills by 24.63%. Thamrin & Annisa Sulolipu (2023) also show that TPS can improve learning interactions at the elementary school level, while Azzahra & Purrohman (2024) found a difference in the critical thinking skills of fourth-grade students who learned using TPS compared to conventional learning, with a *t*-test significance value of 0.04. This finding is supported by Xu et al. (2023), who explain that TPS develops critical thinking skills through individual thinking, discussion, and evaluation of ideas. Huang et al. (2023) also found that TPS improves higher-order thinking skills in science learning because it encourages analysis and argumentation activities. At the elementary school level, Kasta et al. (2026) show that TPS influences critical thinking skills through more active learning interactions. Furthermore, Kilpeläinen-Pettersson et al. (2025) confirm that cooperative learning supports the development of critical thinking through communication activities and students' evaluation of ideas.

Although various studies have shown that TPS can support the development of critical thinking skills, most previous studies have compared TPS with conventional learning. This condition indicates a research gap in the form of limited research comparing TPS with other cooperative learning models, especially Numbered Heads Together (NHT), in elementary school science learning. In fact, TPS and NHT differ in key ways, even though both are cooperative learning models. TPS emphasizes individual thinking processes, pair discussions, and the delivery

of thinking results (think-pair-share), while NHT emphasizes group responsibility through number assignments, group discussions, and random calling of members to submit answers (Mahmudah & Rasyid, 2022). These differences in mechanisms allow for different learning experiences in developing students' critical thinking skills.

Based on these conditions, the novelty of this research lies in the empirical testing of the differences in critical thinking abilities of students who learn using the Think Pair Share (TPS) and Numbered Heads Together (NHT) models in fourth-grade elementary school science learning. This research does not aim to determine the superiority of one model before testing, but rather to provide empirical evidence regarding the differences in critical thinking abilities through the application of the two cooperative learning models. SDN 5 Jatimulyo serves as the research context, based on problems of science learning identified through initial data, observations, and interviews. Thus, this study aims to analyze differences in fourth-grade students' critical thinking abilities in science learning using the Think Pair Share (TPS) and Numbered Heads Together (NHT) models at SDN 5 Jatimulyo.

2. Method

This study uses a quantitative approach with a quasi-experimental method to analyze the effect of the Think Pair Share (TPS) learning model on students' critical thinking skills. The research design is a nonequivalent control group design, which includes an experimental group and a control group with pre- and post-treatment measurements. This design was chosen because the division of students into classes has been determined by the school, so that individual randomization is not possible (Nurusofi et al., 2022; Mutmainnah et al., 2026). The experimental group learned using the TPS model, while the control group learned using the Numbered Heads Together (NHT) model. Both groups received equivalent materials, time allocation, and measurement instruments, while the difference in treatment lies in the learning model applied. The use of a control group with a cooperative learning model allows for a more specific test of the effect of TPS compared to learning without cooperative treatment (Sri Puji Retno et al., 2021).

Table 2. Nonequivalent Control Group Design

Group	Pretest	Treatment	Posttest
Experiment (IV A)	O ₁	X ₁ (TPS)	O ₂
Control (IV B)	O ₃	X ₂ (NHT)	O ₄

The research was conducted at SDN 5 Jatimulyo in the odd semester of the 2026/2027 academic year from August to September 2026. The research population included all 72 fourth-grade students, consisting of 35 students from class IV A and 37 students from class IV B. All members of the population were included as samples through a total sampling technique. Class IV A was designated as the experimental group and Class IV B as the control group based on the class division determined by the school. Thus, the research involved 72 students who all took part in the pretest, the learning process according to the treatment, and the posttest. Determination of groups based on the classes that had been formed was a characteristic of the nonequivalent control group design, so that initial ability measurements were carried out through a pretest to obtain information about the initial conditions of students.

The independent variable in this study is the TPS learning model, while the dependent variable is students' critical thinking skills. The TPS model is operationalized as a cooperative learning model implemented through three main stages: think, pair, and share (Huda et al., 2025). The think stage allows students to understand problems and develop their thinking individually; the pair stage allows them to discuss and clarify ideas with a partner, while the share stage allows them to communicate the results of their thinking. The implementation of learning is observed through observation sheets. Critical thinking skills are operationalized as students' abilities to analyze information, provide reasons, draw conclusions, clarify, and determine problem-solving

strategies logically (Li & Tu, 2024; Al-Malkil et al., 2025). Measurements are conducted using essay tests developed from indicators of critical thinking skills.

Table 3. Operational Definitions of Variables and Instruments

Variables	Operational Definition	Indicator	Instrument
Think Pair Share (TPS) Learning Model (X)	The cooperative learning model is implemented through the think, pair, and share stages to encourage students to think independently, discuss in pairs, and convey the results of their thinking.	1) Think; 2) Pair; 3) Share; 4) Student involvement in discussions.	Learning implementation observation sheet
Critical Thinking Skills (Y)	Students' ability to identify, analyze, provide reasons, evaluate information, draw conclusions, and determine problem-solving strategies logically.	1) Basic clarification; 2) providing reasons; 3) concluding; 4) further clarification; 5) setting strategies and tactics.	30-item essay test

The research was conducted in six meetings for each group. The first meeting was used for a pretest to obtain initial ability data. The second through fifth meetings provided treatment according to the learning model for each group, covering plant parts and their functions, continued plant parts, photosynthesis, and plant reproduction. The sixth meeting was used for a posttest to obtain data on critical thinking skills after the treatment. Both groups received equal materials and time, while the learning model was the main differentiator in the treatment.

Table 4. Research Implementation Procedures

Meeting	Experimental Group	Control Group	Learning Materials/Activities	Objective
1	Pretest	Pretest	Measurement of early critical thinking skills	Identifying students' initial abilities
2	polling station	NHT	Parts of plants and their functions	Analyzing the structure and function of plant parts
3	polling station	NHT	Continued parts of plants	Identifying the relationship between plant structure and function
4	polling station	NHT	Photosynthesis	Analyze the processes and factors that influence photosynthesis
5	polling station	NHT	Photosynthesis and plant reproduction	Analyzing the relationship between the photosynthesis process and plant reproduction
6	Posttest	Posttest	Final critical thinking ability measurement	Identifying critical thinking skills after treatment

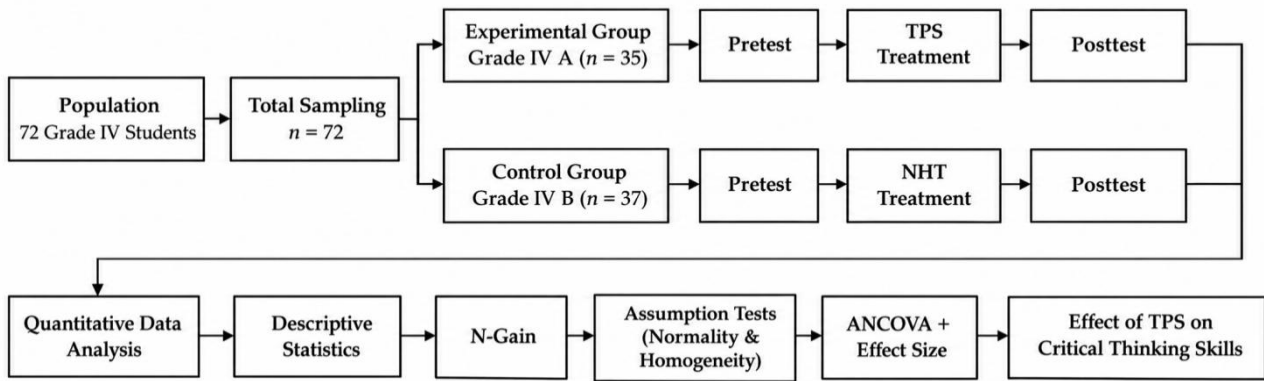


Figure 1. Research Implementation Flow

Data collection used test and documentation techniques. Tests were the primary instrument for measuring students' critical thinking skills, while documentation served as supporting data. The test instrument consisted of 30 descriptive questions administered during the pretest and posttest. The questions were developed based on five indicators of critical thinking skills: basic clarification, providing reasons, concluding, further clarification, and strategizing. The descriptive format was used to obtain information regarding students' reasoning processes, particularly in analyzing information, providing arguments, evaluating statements, and drawing conclusions. Documentation included a list of students, learning materials, and documentation of research activities.

Before being used in data collection, the test instrument was tested to determine the quality of the items through validity, reliability, difficulty level, and discriminatory power. Empirical validity was analyzed using *product-moment correlation*, while instrument reliability was analyzed using *Cronbach's Alpha*. The test items were declared valid if the empirical correlation value was greater than the critical value at the specified significance level. Reliability was used to determine the internal consistency of the instrument. Difficulty level analysis was used to identify easy, medium, and difficult items. In contrast, discriminatory power was used to determine the item's ability to differentiate students based on the level of ability being measured. Instrument quality testing was conducted prior to research data analysis to ensure that the instrument had adequate measurement characteristics (Sabrina & Arif, 2023).

Data analysis used descriptive statistics, improvement analysis, assumption testing, and hypothesis testing with the help of IBM SPSS Statistics Version 22. Descriptive statistics were used to obtain the minimum, maximum, average, and standard deviation of pretest and posttest scores in both groups. The improvement in thinking skills was analyzed using *normalized gain* (N-Gain) as a supporting analysis to describe the level of change in students' abilities after treatment. N-Gain values were interpreted into low, medium, and high categories based on criteria used in educational research. (Herawaty et al., 2023; Novianti & Haerudin, 2025; Sudirman & Marjanah, 2026).

The main hypothesis testing was conducted using Analysis of Covariance (ANCOVA) with posttest scores as the dependent variable, learning groups as factors, and pretest scores as covariates. ANCOVA compared critical thinking skills after treatment while accounting for differences in students' initial abilities. Before the analysis, assumption testing included normality, homogeneity of variance, linearity of the relationship between pretest and posttest, and homogeneity of regression slopes. The significance level was set at 0.05. The effect of the treatment was determined based on the significance of the group effect, while the effect size was reported to show the magnitude of the difference in practical terms. N-Gain was used as a supporting analysis to describe the level of improvement in critical thinking skills in the experimental and control groups (Jung et al., 2025; Schwarz, 2025).

3. Result

Instrument quality testing was conducted to ensure that the critical thinking skills instrument had adequate measurement characteristics before use in the research analysis. The testing included validity, reliability, difficulty level, and item discrimination. Table 5 presents the results of the instrument validity and reliability testing are presented in Table 5.

Table 5. Results of Instrument Validity and Reliability Tests

Test Stage	Valid Item Count	r count (Range)	r table	Cronbach's Alpha	Information
Pretest	30 of 30	0.270–0.709	0.2319	0.912	Very high
Posttest	30 of 30	0.351–0.696	0.2319	0.929	Very high

Based on Table 5, all 30 instrument items in the pretest and posttest met the empirical validity criteria. The calculated *r values in the pretest ranged from 0.270 to 0.709 and were all greater than the table r* of 0.2319. In the posttest, the calculated *r values* ranged from 0.351 to 0.696 and were also greater than the table *r*. These results indicate that all items had an adequate relationship with the instrument scores based on the empirical validity criteria used. Reliability testing produced Cronbach's Alpha values of 0.912 in the pretest and 0.929 in the posttest. Both values are in the very high category, so the instrument shows good internal consistency for use in measuring critical thinking skills in this study.

The characteristics of the test items were then analyzed based on their difficulty level. This analysis provided an overview of the distribution of item difficulty levels across the instrument. Table in presents the test results.

Table 6. Results of the Test of the Level of Difficulty of Question Items

Test Stage	Difficult	Currently	Easy	Total
Pretest	2	28	0	30
Posttest	1	27	2	30
Total	3	55	2	60

Table 6 shows that the majority of test items fall into the moderate category. In the pretest, 28 items were moderate, and 2 were difficult, while in the posttest, 27 items were moderate, 1 was difficult, and 2 were easy. This distribution indicates that the instrument is dominated by items with a moderate level of difficulty, while items in the difficult and easy categories are found only in limited numbers. This composition suggests that the instrument is not concentrated at a particular level of difficulty.

In addition to difficulty level, we analyzed item quality based on discriminatory power. We used discriminatory power analysis to determine the item's ability to differentiate students based on differences in measured ability levels. Table 7 presents the results of the discriminatory power test are presented in Table 7.

Table 7. Results of the Test of Distinguishing Power of Question Items

Test Stage	Bad	Enough	Good	Very good	Total
Pretest	0	1	22	7	30
Posttest	0	0	17	13	30
Total	0	1	39	20	60

Based on Table 7, no items were categorized as poor in either the pretest or posttest. In the pretest, 22 items were categorized as good, 7 were very good, and 1 was fair. Meanwhile, in the posttest, 17 items were categorized as good, and 13 were categorized as very good. This distribution indicates that most items had good to excellent discriminatory power. Overall, the results of the validity, reliability, difficulty level, and discriminatory power tests indicate that the instrument meets the established measurement quality criteria and can be used in research data collection.

Descriptive statistics were then used to provide an overview of the critical thinking skills of students in the experimental and control groups before and after the learning process. The results of the descriptive statistical analysis are presented in Table 8.

Table 8. Descriptive Statistics of Critical Thinking Skills in the Experimental and Control Groups

Group	N	Range	Min	Max	Mean	Elementary School
Experiment Pretest	35	90.00	3.33	93.33	46.76	28.17
Experiment Posttest	35	100.00	0.00	100.00	59.71	27.48
Pretest Control	37	86.67	3.33	90.00	44.68	24.12
Posttest Control	37	100.00	0.00	100.00	43.33	26.57

Based on Table 8, the experimental group's average pretest score was 46.76, while the control group's was 44.68. Descriptively, these two values were relatively close, with a difference of 2.08 points, indicating that the initial critical thinking scores of the two groups were relatively similar. After the learning process, the experimental group's average score increased to 59.71, an increase of 12.95 points from the pretest. In contrast, the control group's average score decreased from 44.68 in the pretest to 43.33 in the posttest, representing a decrease of 1.35 points. Thus, the two groups showed different directions of change: the experimental group increased in average critical thinking score, whereas the control group decreased.

Further improvement analysis was conducted using N-Gain. The average N-Gain for the experimental group was 0.20, while the control group was -0.04. Based on the N-Gain interpretation criteria used, a value of 0.20 is considered low, while a value of -0.04 indicates a negative change in the average score of the control group. Therefore, the N-Gain results reinforce the descriptive finding that the control group not only showed a smaller increase than the experimental group, but also a negative direction of change. Although the improvement in the experimental group remained in the low category, the difference in the direction and magnitude of change between the two groups indicates different patterns of development in critical thinking skills following the learning process. The Discussion further examines the specific factors that may explain the decrease observed in the NHT group are examined further in the Discussion.

Before hypothesis testing, the research data were tested based on the assumptions of normality and homogeneity. These tests were conducted to determine the suitability of using parametric statistical analysis. Table 9 presents the results of the prerequisite tests are presented in Table 9.

Table 9. Results of the Normality and Homogeneity Tests

Data	Normality Test (Lilliefors) Sig.	Information	Homogeneity Test (Fisher) Sig.	Information
Pretest	Experiment = 0.092; Control = 0.269	Normal	0.360	Homogeneous
Posttest	Experiment = 0.674; Control = 0.282	Normal	0.841	Homogeneous

Based on Table 9, the significance values of the normality test for the pretest were 0.092 for the experimental group and 0.269 for the control group. For the posttest, the significance values were 0.674 for the experimental group and 0.282 for the control group. All significance values were greater than 0.05, indicating that the pretest and posttest data in both groups met the normality assumption. The homogeneity test also showed significance values of 0.360 for the pretest and 0.841 for the posttest. Both values were greater than 0.05, indicating that the variances of the experimental and control groups were homogeneous variances. Therefore, the pretest and posttest data met the assumptions required for subsequent parametric analysis.

To examine the magnitude and direction of improvement in students' critical thinking skills, an N-Gain analysis was conducted separately. Table in presents the results.

Table 10. N-Gain Results of Critical Thinking Skills

Group	Mean N-Gain	Interpretation
Experimental	0.20	Low
Control	-0.04	Negative

Based on Table 10, the experimental group obtained a mean N-Gain of 0.20, which was categorized as low based on the interpretation criteria used in this study. Meanwhile, the control group obtained a mean N-Gain of -0.04, indicating a negative change in the average critical thinking score. These results show that the experimental and control groups differed not only in the magnitude of change but also in the direction of change. The experimental group demonstrated a positive increase, whereas the control group demonstrated a negative change.

Hypothesis testing was conducted to determine differences in critical thinking skills between students who learned using the Think Pair Share model and students who learned using the Numbered Heads Together model. Testing was conducted on posttest and N-Gain scores. The results of the analysis are presented in Table 11.

Table 11. Results of the Two Independent Samples t-Test

Data	t count	df	Sig. (2-tailed)	Information
Posttest	2,571	70	0.012	Significant
N-Gain	2,132	70	0.037	Significant

Based on Table 12, the results of *the independent-samples t-test* on posttest scores produced a calculated *t* value of 2.571 with 70 degrees of freedom (*df*) of 70 and a significance of 0.012. This significance value is smaller than 0.05, so there is a significant difference in the posttest scores between the experimental and control groups. Testing on N-Gain also produced a calculated *t* value of 2.132 with a *df* of 70 and a significance of 0.037. This value is smaller than 0.05, so there is a significant difference in the increase in critical thinking ability scores between the two groups.

These findings indicate that students in the Think Pair Share group experienced significantly different changes in critical thinking skill scores compared to those in the Numbered Heads Together group. These findings need to be understood within the context of the quasi-experimental design used, so that the statistical test results primarily demonstrate differences in outcomes and improvements between groups, rather than solely being the basis for attributing all score changes to the treatment.

To complement the interpretation of statistical significance, I performed an effect size analysis using Cohen's *d*. The calculation results showed a Cohen's *d* value of 0.503 for the N-Gain data. Based on Cohen's interpretation criteria, this value is in the moderate category. Thus, the differences in improvement were not only statistically significant but also showed a moderate level of practical significance. This finding provides an empirical basis that the differences in critical thinking skills development between the experimental and control groups have practical significance that needs to be considered in interpreting the research results.

4. Discussion

The results of the study indicate that there are differences in critical thinking skills between students who received learning using the Think Pair Share (TPS) model and students who received learning using the Numbered Head Together (NHT) model. The group that received learning with TPS showed a higher average *posttest* score than the NHT group. This finding indicates that the application of a cooperative learning model with a specific interaction structure can provide

different learning experiences in supporting the development of students' critical thinking skills in science learning.

The findings of this study align with those of Azzahra and Purrohman (2024), who demonstrated that implementing TPS (the T-Partnership) in fourth-grade science learning significantly improves critical thinking skills compared to conventional learning. Ramliyah et al. (2023) found similar results, explaining that TPS can increase student engagement through discussion and idea exchange. Furthermore, Muhfahroyin et al. (2024) demonstrated that implementing TPS can help improve critical thinking skills through learning that provides students with opportunities to analyze problems and develop arguments.

The differences in critical thinking skills between the TPS and NHT groups can be explained by the characteristics of the learning stages used. The TPS model gives students space to experience a gradual thinking process through the think, pair, and share stages. In the think stage, students can understand the problem independently and organize their thoughts before others' opinions influence them. This process is important in the development of critical thinking because students not only receive information but also make initial interpretations of the given problem.

The *pair* phase provides students with the opportunity to elaborate on their thinking through interaction with a partner. Through discussion, students can compare various possible answers, provide reasons, and clarify their understanding. This process aligns with Vygotsky's (1978) social constructivist perspective, which explains that students' cognitive development occurs through social interaction and the exchange of ideas with others. In the context of critical thinking, these interactions help students evaluate ideas and refine their understanding through discussion.

Next, the sharing stage provides students with the opportunity to communicate their thoughts to a wider group. The activity of expressing opinions openly encourages students to develop more systematic reasoning and be accountable for their answers. This process relates to critical thinking indicators according to Ennis (2011), specifically the ability to provide reasons, clarify, and draw conclusions based on available information.

Meanwhile, the control group using the NHT model also received cooperative learning through group discussions and individual accountability within the group. However, the main characteristic of NHT lies in the group accountability mechanism through the distribution of member numbers and the random calling of students to provide group answers. Unlike TPS, which explicitly provides individual thinking time before discussions, NHT emphasizes the readiness of group members to account for the results of the collective discussion. This difference in interaction structure can provide a different learning experience in the process of developing students' critical thinking skills.

An important finding that requires further consideration is the decrease in the average critical thinking score of the NHT group, from 44.68 in the pretest to 43.33 in the posttest, resulting in an N-Gain value of -0.04 . This result indicates that the NHT group not only showed a smaller increase than the TPS group but also experienced a slight negative change in the average score. This finding should not be interpreted as evidence that NHT inherently reduces students' critical thinking skills. Rather, it indicates that, under the conditions of this study, the NHT treatment was not accompanied by an increase in the measured critical thinking scores during the research period.

One possible explanation may lie in the interaction characteristics of NHT. Although NHT provides opportunities for group discussion and individual accountability, its mechanism emphasizes group responsibility through numbered members and random selection to present the group's answer. Consequently, students may have fewer opportunities to formulate and organize

individual reasoning before entering the group discussion than in the TPS *think* stage. For students who still require structured guidance in analyzing information, providing reasons, and formulating conclusions, the available discussion process may not necessarily result in an immediate improvement in individual critical thinking performance. This interpretation aligns with the distinction between TPS and NHT in terms of their interaction structures. However, the present study did not directly measure students' interaction processes or identify specific causes of the decline.

The negative N-Gain in the NHT group may also be related to the relatively short duration of the intervention and the characteristics of the students and learning materials. The research was conducted through a limited number of treatment meetings, while the development of critical thinking involves repeated opportunities to analyze information, evaluate ideas, provide reasons, and draw conclusions. Therefore, a short intervention may not be enough to produce a stable increase in individual critical thinking performance, particularly when students learn through a cooperative structure that emphasizes collective group responsibility. These considerations should be viewed as possible explanations rather than definitive causal factors because the study did not separately measure student engagement, individual participation during NHT activities, or other factors that could explain the change in scores.

The results also showed that the increase in critical thinking skills, based on the N-Gain value in both groups, remained in the low category. This indicates that changes in critical thinking skills require a more sustainable learning process and cannot depend solely on a single treatment period. The contrast between the two groups is therefore important: the TPS group showed a positive N-Gain of 0.20, whereas the NHT group showed a negative N-Gain of -0.04. Thus, although the magnitude of improvement in the TPS group was still categorized as low, the direction of change differed from that of the NHT group. This finding indicates that the TPS learning structure supports students' thinking activities. However, increases in critical thinking skills are still influenced by other factors such as learning duration, student characteristics, previous learning experiences, and the complexity of the material being studied.

Statistically, the test results show that the differences between the two groups are not only descriptively visible but also supported by the results of the hypothesis test. The significance values of the posttest and N-Gain indicate a significant difference between the TPS and NHT groups. In addition, Cohen's *d* of 0.503 indicates that the difference is moderate. The statistical findings therefore confirm that the observed difference between the two groups was unlikely to be explained solely by random variation under the statistical criteria applied in this study. However, statistical significance does not, by itself, establish that the NHT model caused the decrease observed in the NHT group. These results indicate that the application of TPS has a practical contribution in supporting the development of students' critical thinking skills. However, this interpretation needs to be understood in the context of a *quasi-experimental research design*, so that the results of the study show a relationship between the learning model and differences in critical thinking skills, not as the only factor determining changes in students' abilities.

This research contributes to the study of cooperative learning, particularly of comparing TPS with another cooperative model, namely NHT. Most previous studies have compared TPS with conventional learning. In contrast, this study provides an overview of how two cooperative learning models with different interaction characteristics can produce different developments in critical thinking skills. Thus, the results of this study broaden the understanding of the importance of learning interaction design in developing higher-order thinking skills in elementary school students. In particular, the findings show that comparison with another cooperative learning

model can reveal differences in the direction and magnitude of students' score changes that may not be apparent when TPS is compared only with conventional learning. The negative N-Gain observed in the NHT group also highlights the importance of examining not only whether cooperative learning is implemented, but also how the structure of interaction provides opportunities for individual reasoning and the development of critical thinking.

However, this study has several limitations. It was conducted in a single school with a limited number of students and specific science and science materials, so generalization of the results requires caution. Furthermore, the research group used pre-formed classes, so it did not randomize individuals. Although initial testing indicated that the data met the analysis assumptions, the possibility of differences in initial characteristics between groups may still exist. In addition, the study did not directly measure students' participation, engagement, or the quality of interaction during NHT learning. Therefore, the available data cannot conclusively explain the decrease in the NHT group's average score. Future research could expand the scope of schools, learning materials, and treatment duration to obtain a more comprehensive picture of the effectiveness of TPS in developing critical thinking skills. Future studies could also include systematic observations of individual participation and group interaction in NHT, as well as longer treatment periods, to examine whether the negative N-Gain observed in this study is consistent across different learning contexts or is specific to the conditions of the present study.

5. Conclusion

Based on the research results, there are differences in critical thinking skills between students who learned using the Think Pair Share (TPS) model and Numbered Heads Together (NHT) in science learning for grade IV of SDN 5 Jatimulyo. Based on the posttest and N-Gain results, the TPS group showed a more positive change in critical thinking skills scores than the NHT group. The results of the independent-samples t-test showed a significant difference between the two groups, while the Cohen's d value of 0.503 indicated a medium effect size. These findings indicate that the TPS structure, through the stages of think, pair, and share, provides space for students to develop understanding through individual thinking processes, pair interactions, and the delivery of ideas. TPS can be an alternative cooperative learning model to support student engagement and the development of critical thinking skills in science learning. However, these findings are limited to one school, certain materials, and a quasi-experimental design without individual randomization. Future research needs to involve more schools, longer treatment durations, and variations in materials to strengthen the generalizability of the findings.

6. References

- Al-Malkil, E.A.A., Gulnaz, F., Hagggar, A.I.B., & Alharbi, W. (2025). Investigating the Impact of Think-Pair-Share on Saudi EFL Learners' Descriptive Writing: A Study of Perceptions, Experiences, and Confidence. *Journal of Cultural Analysis and Social Change*, 935–951. <https://doi.org/10.64753/jcasc.v10i2.1716>
- Alsaleh, N. (2020). Teaching Critical Thinking Skills: Literature Review. In *TOJET: The Turkish Online Journal of Educational Technology* (Vol. 19, Number 1).
- Alsmadi, M. A., Tabieh, A.S., Alsaifi, R.M., & Al-Nawaiseh, S.J. (2023). The Effect of the Collaborative Discussion Strategy Think-Pair-Share on Developing Students' Skills in Solving Engineering Mathematical Problems. *European Journal of Educational Research*, volume-12-2023 (volume 12, issue 2, April 2023, 1123–1135. <https://doi.org/10.12973/eu-jer.12.2.1123>

- Azzahra, S., & Purrohman, P. S. (2024). The Effect of Think-Pair-Share (TPS) on the Critical Thinking Skills of Fourth Grade Science Students. *Ideguru: Journal of Teachers' Scientific Work*, 9(3), 1461–1467. <https://doi.org/10.51169/ideguru.v9i3.1058>
- Bhardwaj, V., Zhang, S., Tan, Y. Q., & Pandey, V. (2025). Redefining learning: student-centered strategies for academic and personal growth. *Frontiers in Education*, 10. <https://doi.org/10.3389/feduc.2025.1518602>
- Deore, M., & Arora, S. (2022). Effective Think-Pair-Share Pedagogical Strategy to Improve Inferential Statistics Concept Understanding. *Journal of Engineering Education Transformations*, 36 (S1), 25–32. <https://doi.org/10.16920/jeet/2022/v36is1/22170>
- Dwyer, C. P., Hogan, M. J., & Stewart, I. (2014). An integrated critical thinking framework for the 21st century. *Thinking Skills and Creativity*, 12, 43–52. <https://doi.org/10.1016/j.tsc.2013.12.004>
- Ekamilasari, E., PA, & PID (2021). Critical thinking skills and sustainability awareness for the implementation of education for sustainable development. In *Journal of Science Education Research* (Vol. 2021, Number 1). www.journal.uny.ac.id/jser
- Ennis, R. H. (nd). Critical Thinking. In *The Palgrave Handbook of Critical Thinking in Higher Education*. Palgrave Macmillan. <https://doi.org/10.1057/9781137378057.0005>
- Fatimah, S., Sarwi, S., Linuwih, S., Dewi, N. R., Hartono, H., & Bambang Subali. (2026). A systematic literature review on the integration of scientific reasoning skills into science education. *Journal of Turkish Science Education*, 23 (1), 70–90. <https://doi.org/10.36681/tused.2026.004>
- Firja MKS, M. (2024). Application of Think Pair Share Type Cooperative Learning Model to Improve Creative Thinking Ability. *Journal Corner of Education, Linguistics, and Literature*, 4 (001), 318–328. <https://doi.org/10.54012/jcell.v4i001.392>
- Herawaty, A., Rajagukguk, S., & Sitopu, J.W. (2023). Effect of Concept Sentence Learning Model on Students' Learning Outcomes. *Journal of Science Education Research*, 9(2), 843–849. <https://doi.org/10.29303/jppipa.v9i2.2988>
- Herlinawati, H., Marwa, M., Ismail, N., Junaidi, L., LO, & Situmorang, D. (2024). The integration of 21st-century skills in the curriculum of education. *Heliyon*, 10 (15), e35148. <https://doi.org/10.1016/j.heliyon.2024.e35148>
- Huang, T., Dewi, W. S., Gusnedi, G., & Sundari, P. D. (2023). Meta-Analysis of the Effect of Cooperative Learning Think Pair Share Type on Student Physics Learning Outcomes in Senior High School. *Journal of Science Education Research*, 9(11), 1148–1154. <https://doi.org/10.29303/jppipa.v9i11.4810>
- Huda, A. N., Hasibuan, P., Wati, S., & Mustafa, M. (2025). The Effect of the Think Pair Share (TPS) Learning Model on Students' Critical Thinking Skills in the Concept of the Qur'an and Hadith in Grade VII of the Diniyah Limo Jurai Islamic Boarding School. *ALSYS*, 5 (4), 1437–1454. <https://doi.org/10.58578/alsys.v5i4.6572>
- Iin Suriya Ningsih, Mahfud Iffendi, & Anggra Prima. (2025). The Use of the Think Pair Share Cooperative Learning Model to Improve Student Learning Outcomes in Arabic Language Learning at the Paqusatta Islamic Boarding School. *Hidayah : Scholars of Islamic Education and Sharia Law*, 2 (3), 26–41. <https://doi.org/10.61132/hidayah.v2i3.1255>
- Józsa, K., Oo, T. Z., Borbélyová, D., & Podráczky, J. (2024). Deductive Reasoning Skills in Children Aged 4–8 Years Old. *Journal of Intelligence*, 12(3), 33. <https://doi.org/10.3390/jintelligence12030033>

- Jung, W., Lee, K., Kim, H.-H., & Lim, C. (2025). How to construct analysis of covariance in clinical trials: ANCOVA with one covariate in a completely randomized design structure. *Korean Journal of Anesthesiology*, 78 (4), 315–320. <https://doi.org/10.4097/kja.24820>
- Kilpeläinen-Pettersson, J., Koskinen, P., Lehtinen, A., & Mäntylä, T. (2025). Cooperative Learning in Higher Education Physics – A Systematic Literature Review. *International Journal of Science and Mathematics Education*, 23(7), 2707–2729. <https://doi.org/10.1007/s10763-024-10538-3>
- Kozanitis, A., & Nenciovici, L. (2023). Effect of active learning versus traditional lecturing on the learning achievement of college students in humanities and social sciences: a meta-analysis. *Higher Education*, 86(6), 1377–1394. <https://doi.org/10.1007/s10734-022-00977-8>
- Li, M.-M., & Tu, C.-C. (2024). Developing a Project-Based Learning Course Model Combined with the Think–Pair–Share Strategy to Enhance Creative Thinking Skills in Education Students. *Education Sciences*, 14(3), 233. <https://doi.org/10.3390/educsci14030233>
- Lombardi, D., Shipley, T., Bailey, J., Bretones, P., Prather, E., Ballen, C., Knight, J., Smith, M., Stowe, R., Cooper, M., Prince, M., Atit, K., Uttal, D., LaDue, ND, McNeal, PM, Ryker, K., St. John, K., van der Hoeven Kraft, K. J., & Docktor, J. L. (2021). The Curious Construct of Active Learning. *Psychological Science in the Public Interest*, 22 (1), 8–43. <https://doi.org/10.1177/1529100620973974>
- Loves Yogen Ombi Kasta, Nuranisa, & Robert Budi Laksana. (2026). The Effect of the Think Pair Share Learning Model on the Critical Thinking Skills of Fifth Grade Students at SD Negeri 95 Palembang. *Pendas: Scientific Journal of Elementary Education*, 11 (03), 402–412. <https://doi.org/10.23969/jp.v11i03.56095>
- Mahmudah, H., & Rasyid, F. (2022). Engaging Students in a Cooperative Learning Model Of Reading Course Through Numbered Head Together. *Eternal (English Teaching Journal)*, 13 (1), 55–69. <https://doi.org/10.26877/eternal.v13i1.10339>
- Mametja, L. C., Mathebula, N. E., & Ntshangase, M. X. (2026). Exploring Primary School Learners' Understanding of Critical Thinking: A Case Study of Makhutswe Circuit, South Africa. *Asian Journal of Education and Social Studies*, 52 (1), 241–255. <https://doi.org/10.9734/ajess/2026/v52i12772>
- Muhfahroyin, M., Rachmadiarti, F., Mahanal, S., Zubaidah, S., & Siagiyanto, BE (2024). Improving Critical Thinking Skills of Low Academic Ability Students through TPS and PBL Integration in Biology Learning. *Journal of Turkish Science Education*, 20 (4), 606–618. <https://doi.org/10.36681/tused.2023.034>
- Mutmainnah, M., Nur, M., M., & Adawiyah, SR (2026). Think-pair-share learning model based on 3D media in developing students' critical thinking skills. *JPBIO (Jurnal Pendidikan Biologi)*, 11 (1), 198–205. <https://doi.org/10.31932/jpbio.v11i1.6234>
- Nahar, L., & Machado, C. (2025). Inquiry-based learning in Bangladesh: insights into middle and high school students' experiences and 21st-century skill development. *Disciplinary and Interdisciplinary Science Education Research*, 7 (1), 2. <https://doi.org/10.1186/s43031-025-00122-2>
- Novandi, M., Serani, G., Djudin, T., & Suratman, D. (2025). Critical Thinking Skills and Their Teaching in Elementary Schools. *Jurnal Pendidikan Dasar Perkhasa: Jurnal Penelitian Pendidikan Dasar*, 11 (1), 649–669. <https://doi.org/10.31932/jpdp.v11i1.4509>
- Novianti, S., & Haerudin, D. A. (2025). STAD (Student Team Achievement Division) Cooperative Learning Model with Kahoot Media to Improve Learning Outcomes of Grade XI Students in English Subject at SMAN 1 Darma. *Indo-MathEdu Intellectuals Journal*, 6 (5), 8441–8447. <https://doi.org/10.54373/imeij.v6i5.4073>

- Nur'ariyani, S., Jumyati, J., Yuliyanti, Y., Nulhakim, L., & Leksono, S.M. (2023). Scientific Approach to Learning Science in Elementary Schools. *Journal of Science Education Research*, 9 (8), 6659–6666. <https://doi.org/10.29303/jppipa.v9i8.3680>
- Nurusofei, I., Mahardika, I., & Budiarmo, AS (2022). Impact of the Think Pair Share Cooperative Learning Model with Multi-Representation Based Worksheet on Students' Critical Thinking. *Journal of Science and Science Education*, 3 (2), 110–114. <https://doi.org/10.29303/jossed.v3i2.1804>
- O'Reilly, C., Devitt, A., & Hayes, N. (2022). Critical thinking in the preschool classroom: A systematic literature review. *Thinking Skills and Creativity*, 46, 101110. <https://doi.org/10.1016/j.tsc.2022.101110>
- Sabrina, N., & Arif, K. (2023). The Effect Of Numbered Head Together (NHT) Cooperative Learning Model On Students' Learning Outcomes On Human Respiratory System Materials At SMPN 2 Solok City. *Universe*, 4 (1), 26–31. <https://doi.org/10.24036/universe.v4i1.341>
- Sasmita, R.N., Sapriya, S., & Maryani, E. (2023). Critical thinking in social studies learning for Elementary School Students. *Nazhruna: Journal of Islamic Education*, 5 (3), 1377–1387. <https://doi.org/10.31538/nzh.v5i3.2355>
- Schwarz, W. (2025). The ANCOVA model for comparing two groups: a tutorial emphasizing statistical distribution theory. *Frontiers in Psychology*, 16. <https://doi.org/10.3389/fpsyg.2025.1600764>
- Sri Puji Retno, A., Khoiri, N., & Ihda Norra, B. (2021). The Effectiveness of Cooperative Learning Model in the Type of Numbered Heads Together (NHT) Assisted with Picture Puzzle Media on Self-Efficacy and Student Learning Outcomes in the Coordination System Material. *Journal of Physics: Conference Series*, 1779 (1), 012058. <https://doi.org/10.1088/1742-6596/1779/1/012058>
- Sudirman, R., & Marjanah, M. (2026). The Effectiveness of the Quiz Team Learning Model on Islamic Education Learning Outcomes of Eighth-Grade Students at SMP Negeri 2 Tinambung. *Journal of Multidisciplinary Science: MIKAILALSYS*, 4 (2), 3180–3190. <https://doi.org/10.58578/mikailalsys.v4i2.11121>
- Taylor, C. (2022). *Phenomenon-Based Instruction in the Elementary Classroom* [Boise State University]. <https://doi.org/10.18122/td.1956.boisestate>
- Thamrin, I., & Annisa Sulolipu, A. (nd). Implementation of Think Pair Share in Improving Learning Outcomes and Student Interaction Patterns at the UPT SPf SD Inpres Campus IKIP, Rappocini District, Makassar City – *Journal of Governance and Politics (JGP)*, 3 (2).
- Thornhill-Miller, B., Camarda, A., Mercier, M., Burkhardt, J.-M., Morisseau, T., Bourgeois-Bougrine, S., Vinchon, F., El Hayek, S., Augereau-Landais, M., Mourey, F., Feybesse, C., Sundquist, D., & Lubart, T. (2023). Creativity, Critical Thinking, Communication, and Collaboration: Assessment, Certification, and Promotion of 21st Century Skills for the Future of Work and Education. *Journal of Intelligence*, 11 (3), 54. <https://doi.org/10.3390/jintelligence11030054>
- Xu, E., Wang, W., & Wang, Q. (2023). The effectiveness of collaborative problem solving in promoting students' critical thinking: A meta-analysis based on empirical literature. *Humanities and Social Sciences Communications*, 10 (1), 16. <https://doi.org/10.1057/s41599-023-01508-1>