

Examining the effect of group investigation cooperative learning on student science achievement and collaboration skills

Erna Kusumawati^{*1}, Wahyu Edi Prasetyo, Ikah Salhah

1Program Studi Magister Administrasi Pendidikan, Sekolah Pascasarjana, Universitas Bina Bangsa, Serang, Banten, 42124, Indonesia

*Correspondence: ✉ ernakusumawati19@gmail.com

Abstract

Purpose: This study evaluates the pedagogical efficacy of the six-stage Group Investigation cooperative learning model on primary students' science achievement and collaboration skills, directly addressing cognitive and affective limitations associated with conventional teacher-directed instructional approaches in elementary science education.

Method: Employing a six-week quantitative, quasi-experimental pretest-posttest control group design, the research involved elementary students across intact classrooms. Data were gathered using a validated Science Achievement Test measuring higher-order cognitive domains alongside a dual-mode Collaboration Evaluation Framework comprising peer assessments and direct observations. Data were processed using multivariate and univariate analyses of covariance to control for baseline differences.

Findings: Multivariate analysis revealed a highly significant positive effect for Group Investigation on combined cognitive and interpersonal outcomes. Follow-up univariate testing demonstrated substantial, statistically significant gains in conceptual science mastery and peer collaboration skills compared to direct instruction, with high partial eta squared effect sizes confirming practical significance.

Practical implications: Findings provide elementary educators and curriculum designers with an actionable, scalable framework to simultaneously cultivate essential twenty-first-century collaboration competencies and deep conceptual science understanding.

Originality/value: This research offers novel empirical validation regarding scaffolded group investigation in primary education, demonstrating its capacity to bridge cognitive academic mastery and socio-emotional learning.

Article History

Received 20 January 2026

Revised 3 June 2026

Accepted 24 June 2026

Keyword

Group Investigation, cooperative learning, science achievement, collaboration skills, primary education.



This work is licensed under a [Creative Commons Attribution \(CC BY-SA\) 4.0 license](https://creativecommons.org/licenses/by-sa/4.0/) Internasional License.

How to cite:

Kusumawati, E., Prasetyo, W. E., & Salhah, I. (2026). Examining the effect of group investigation cooperative learning on student science achievement and collaboration skills. *Hamka Insight*, 6(1), 67–93. <https://doi.org/10.51574/hamka.v6i1.172>

INTRODUCTION

Global educational paradigms increasingly emphasize science literacy as a fundamental driver of 21st-century competence, prompting educational systems worldwide to evaluate and refine their pedagogical frameworks (Hu, 2022; Lau & Lam, 2017). International comparative assessments such as PISA and TIMSS demonstrate that instructional approaches significantly shape student performance across diverse cultural contexts (Aditomo & Klieme, 2020; Gao, 2014; Lau & Ho, 2022). In particular, the debate surrounding teacher-directed instruction versus inquiry-based learning highlights the complex mechanisms through which classroom practices influence science achievement and student dispositions (Areepattamannil et al., 2020). Effective inquiry-based learning environments rely heavily on structured instructional practices that guide young learners toward conceptual understanding rather than simple memorization (Cairns, 2019). Consequently, contemporary primary education demands instructional strategies that foster both cognitive domain mastery and social competence within inquiry-driven science learning contexts.

Enhancing science achievement during early childhood and primary education yields profound long-term social and academic benefits, serving as a foundational catalyst for lifelong scientific inquiry and intellectual growth (Burns et al., 2019; Chen, 2014). Student academic performance is deeply intertwined with non-cognitive factors, including self-concept, motivational beliefs, and intrinsic value attached to scientific learning tasks (Chang, 2014; Jansen et al., 2014; Lay et al., 2013). Furthermore, empirical evidence underscores that instructional quality can mediate the substantial impact of familial socioeconomic status and structural background disparities on student learning outcomes (Gustafsson & Yang Hansen, 2018; Korous et al., 2022; Tan, 2024). When classroom environments actively stimulate students' motivational beliefs, learners demonstrate heightened cognitive engagement and academic resilience across elementary grade levels (Liou & Ho, 2016). Investing in effective pedagogical

interventions thus holds significant multidimensional value, bridging socioeconomic divides while empowering young learners to develop essential scientific competencies.

Despite the widespread implementation of practical work and inquiry in school curricula, traditional hands-on activities often fail to deliver sustained affective engagement or deep conceptual mastery (Abrahams, 2009; Abrahams & Millar, 2008). Longitudinal studies reveal a worrying decline in student interest and attitudes toward science as they progress through school, largely due to rigid assessment policies and cookbook-style experiments (Childs & Baird, 2020; Gibson & Chase, 2002; Potvin & Hasni, 2014). Although guided-inquiry instruction offers potential benefits, many classroom implementations struggle to balance high student motivation with moderate conceptual achievement (Koksal & Berberoglu, 2012; Salchegger et al., 2021). Moreover, teachers frequently encounter pedagogical barriers when adopting student-centered practices, leaving practical tasks disconnected from genuine peer collaboration and scientific reasoning (Toplis, 2012; Yeung, 2009). Consequently, a critical gap persists in identifying structured cooperative learning models that simultaneously elevate academic achievement and cultivate deliberate interpersonal collaboration among elementary students.

To address these limitations, this study introduces the Group Investigation (GI) cooperative learning model as a highly structured, inquiry-driven framework tailored for young learners. While conventional inquiry can produce mixed achievement outcomes across countries, GI integrates scaffolded peer interaction with active scientific investigation to enhance overall instructional quality (Kang, 2020; Liou, 2021; Oliver et al., 2021). Unlike isolated practical tasks, the GI model leverages adaptive scaffolding strategies and iterative feedback to facilitate meaningful peer discourse and strategy uptake (Butterfuss et al., 2022; Kruiper et al., 2022). Novel focus is placed on overcoming multimodal and linguistic challenges in collaborative practical work, transforming group dynamics into rich environments for co-constructing knowledge (Pun & Cheung, 2023; Pun et al., 2023). By explicitly linking collaborative group inquiry

to both achievement and positive learning environments, this research offers a novel empirical evaluation of GI in early science education (Long et al., 2022; Rogers & Fraser, 2023).

The primary objective of this study is to examine the direct effect of Group Investigation cooperative learning on elementary students' science achievement and collaboration skills. Specifically, the study investigates how formative feedback loops within group investigations influence student intrinsic motivation and mitigate potential negative feedback effects (Fong et al., 2019; Gotwals, 2018). We examine the complex mechanisms linking instructional variables, academic self-concept, and cognitive activation to achievement outcomes in elementary classrooms (Jansen et al., 2022; Jo & Seo, 2021; Zhang et al., 2021). Additionally, the research tracks longitudinal changes in student engagement and self-efficacy throughout the implementation of the intervention (Maricuțoiu & Sulea, 2019). Rigorous structural equation modeling criteria are applied to validate the theoretical relationships between cooperative learning processes, collaboration skills, and academic performance gains (Hu & Bentler, 1999).

Establishing robust scientific literacy early in education has become exceptionally urgent in light of recent global challenges, where public understanding of science directly impacts societal well-being (Chan et al., 2023; Cheung et al., 2023). Elementary classrooms face heightened pressure to align instructional practice with rigorous accountability standards while maintaining student engagement (Hayes & Trexler, 2016). Aligning teaching practices with positive student attitudes toward scientific inquiry is critical, as scientific inquiry directly influences achievement through affective pathways (Wan et al., 2023). Furthermore, fostering a future time perspective and intrinsic valuation in young learners helps sustain motivation across developmental milestones (Simons et al., 2004). Addressing these contemporary demands requires operationalizing shared conceptions of effective teaching held by educators and learners alike (Yung et al., 2013).

Ultimately, this study offers substantial theoretical and practical contributions to primary science pedagogy and educational assessment. By demonstrating how Group Investigation builds both academic performance and teamwork, the findings provide actionable evidence for integrating formative assessment practices into elementary science classrooms (Yan & Chiu, 2023). Furthermore, the research illustrates how group-based inquiry fosters a growth mindset, helping students view collaborative problem-solving as a path toward mastery (Yan et al., 2021). From a policy perspective, these insights offer practical guidance for curriculum developers seeking to design cooperative learning environments that balance cognitive achievement with essential socio-emotional skills. Consequently, this study establishes a robust empirical foundation for scaling effective cooperative investigation strategies in early childhood and primary education.

METHODS

This study adopts a quantitative, quasi-experimental pretest-posttest control group design. A quasi-experimental approach was selected because educational institutions prohibit the random assignment of individual students to treatment conditions, as restructuring established classrooms would cause significant administrative and pedagogical disruption. This design minimizes internal validity threats—such as history and selection-maturation interactions—by establishing equivalence through pre-intervention baseline comparisons and statistical controls. The structural matrix of the research design is operationalized in Table 1:

Table 1: Quasi-Experimental Research Matrix

Group	Baseline Pretest (T1)	Experimental Intervention (X)	Post-Intervention Posttest (T2)
Experimental Group (<i>E</i>)	O_1, O_3	Group Investigation (X_{GI})	O_2, O_4
Control Group (<i>C</i>)	O_1, O_3	Direct Instruction (X_{DI})	O_2, O_4

Note: O_1 : Pre-intervention Science Achievement Test score, O_2 : Post-intervention Science Achievement Test score, O_3 : Baseline Collaboration Skills Assessment score, O_4 : Post-intervention Collaboration Skills Assessment score, X_{GI} : Implementation of the 6-phase Group Investigation Cooperative Learning model, X_{DI} : Implementation of teacher-centered Direct Instruction.

Research Participants and Sampling

Participant Characteristics and Sample Size Justification

The target population comprises **primary school students** enrolled in a standardized science curriculum. To prevent Type II statistical errors and meet high-impact publication criteria, the required sample size was calculated *a priori* using GPower 3.1.9.7. Parameter settings were derived from meta-analytic benchmarks in science education: a medium effect size ($f = 0.25$), significance threshold ($\alpha = .05$), statistical power ($1 - \beta = .85$), two groups, and one covariate (pretest baseline). GPower indicated a minimum requirement of $N = 128$. To accommodate full classroom enrollments and potential attrition resulting from student absenteeism or incomplete assessments, six intact classrooms comprising a final sample of $N = 174$ **students** were selected.

Sampling Rationale and Extraneous Controls Participant selection was conducted via Cluster Random Sampling at the classroom level across six intact classes. Baseline homogeneity among the selected classrooms was statistically verified prior to intervention using a one-way ANOVA on students' prior science GPA, showing no significant differences ($p > .05$).

Table 2. Participant selection

Group Allocation	Classrooms (k)	Sample Size (n)	Assigned Condition	Instructional
Experimental	3 Intact Classes	$n_1 = 86$	Group Investigation (X_{GI})	

Control	3	Intact	$n_2 = 88$	Direct	/	Conventional
Classes			Instruction (X_{DI})			

Inclusion criteria required participants to be enrolled in the same curriculum track, taught by instructors following standardized teaching guides, and present for all measurement sessions. Teacher-effect bias was controlled by delivering standardized facilitation training to instructors in the experimental condition prior to the study, ensuring faithful execution of the Group Investigation protocol while keeping learning objectives identical across both groups.

Data Collection Instruments and Procedures

Data collection relies on two validated, multi-faceted measurement instruments:

1. **Science Achievement Test (SAT):** A 25–30 item instrument combining multiple-choice and open-ended conceptual reasoning items mapped to higher-order cognitive domains (Analyzing, Evaluating, and Creating) of Revised Bloom’s Taxonomy. Content validity was evaluated by three science education experts and quantified using **Aiken’s V index**:

$$V = \frac{\sum s}{n(c - 1)}$$

Only items demonstrating an Aiken’s $V \geq .80$ were retained. Internal consistency reliability was confirmed using Kuder-Richardson 20 ($KR-20$) for dichotomous items and Cronbach’s alpha ($\alpha \geq .80$) for open-ended rubric scores.

2. **Collaboration Skills Evaluation Framework:** A dual-mode framework comprising a 20-item student Peer and Self-Assessment Survey (measured on a 4-point Likert scale: 1 = Strongly Disagree to 4 = Strongly Agree) and a structured Observation Checklist used by independent, trained classroom observers. Dimensions evaluated include Active Listening, Constructive Communication, Shared Responsibility, Respect for Diverse Perspectives, and Group Problem Solving. Survey internal consistency achieved Cronbach’s $\alpha \geq$

.80, while inter-rater reliability for observer checklists achieved high concordance via **Cohen's Kappa** ($\kappa \geq .75$).

Fieldwork Execution and Timeline

The intervention was conducted over a 6-week academic period covering a complete, standardized science unit across three systematic phases:

Table 3. 6-week academic period

Group / Condition	Operational Activities & Detailed Stages	Target Measures	Instruments /
Phase 1: Baseline Testing (Week 1)			
Both Groups (<i>N</i> = 174)	Administer baseline instruments to establish initial performance equivalence prior to intervention	•Pre-Science Test (<i>O</i> ₁) •Pre-Collaboration Assessment (<i>O</i> ₃)	Achievement Skills
Phase 2: Instructional Intervention (Weeks 2–7)			
Experimental Group (<i>X</i> _{GI})	6-Stage Group Investigation Model: 1. Topic Selection & Grouping: Form heterogeneous small groups (4–5 members). 2. Cooperative Planning: Define research questions and assign individual group roles. 3. Implementation: Carry out scientific investigations and process collected data. 4. Analysis & Synthesis: Structure and format findings into collaborative posters/models. 5. Presentation of Findings: Deliver peer-reviewed group presentations. 6. Evaluation: Conduct collaborative feedback and reflection sessions.	Implementation monitoring & group progress tracking	fidelity & continuous
Control Group	Direct / Conventional Instruction: • Teacher-led lectures and demonstrations	Standard curricular checks	progress

	• Individual textbook reading • Independent paper-and-pencil worksheet
(X_{DI})	exercises
Phase 3: Outcome Evaluation Week 8	
Both Groups $(N = 174)$	Administer post-intervention instruments to measure conceptual and skill gains • Post-Science Achievement Test (O_2) • Post-Collaboration Skills Assessment (O_4)

The Control Group completed the identical curriculum using conventional direct instruction, including instructor presentations, teacher demonstrations, and individual worksheet exercises.

Data Analysis Techniques

Data Preparation and Preliminary Diagnostics

Raw response data will be anonymized, coded, and analyzed using **SPSS (v28)** or **R**. Pre- and post-intervention scores will be converted into descriptive statistics (Means, M , and Standard Deviations, SD) alongside Normalized Gain Scores (g). Prior to inferential testing, data hygiene checks will screen for missing values, and the following statistical assumptions will be formally evaluated at $\alpha = .05$:

- *Normality*: Assessed via the **Shapiro-Wilk test** ($p > .05$).
- *Homogeneity of Variance*: Evaluated using **Levene's Test** ($p > .05$).
- *Homogeneity of Regression Slopes*: Tested to confirm the absence of covariate-by-treatment interactions.
- *Nested Cluster Effect Control*: Intraclass Correlation Coefficients (ICC) will be computed across intact classrooms to adjust for variance associated with group nesting.

Inferential Modeling and Hypothesis Testing

To evaluate the simultaneous impact of the instructional intervention across both dependent variables while controlling for initial baseline variations, a Multivariate Analysis of Covariance (MANCOVA) will be executed, entering pretest scores (O_1, O_3) as covariates: The primary multivariate null hypothesis will be evaluated using Pillai's Trace (or Wilks' Lambda). If the multivariate test yields statistical significance ($p < .05$), follow-up univariate ANCOVAs will be conducted for Science Achievement and Collaboration Skills independently to isolate specific intervention effects. Practical significance across all statistical tests will be quantified and reported using Partial Eta Squared (η_p^2).

RESULT AND DISCUSSION

Diagnostics, Baseline Equivalence, and Preliminary Checks

Prior to inferential modeling, raw response data ($N = 174$) were screened for missing entries, extreme outliers, and standard parametric assumptions. Normality was evaluated across both conditions using the Shapiro-Wilk test, confirming that pretest and posttest scores for both dependent variables conformed to normal distributions ($p > .05$).

Homogeneity of variance was established via Levene's Test, indicating non-significant variance differences between the experimental ($n = 88$) and control ($n = 86$) groups for post-intervention Science Achievement ($F(1,172) = 1.84, p = .177$) and Collaboration Skills ($F(1,172) = 2.11, p = .148$). Homogeneity of regression slopes was confirmed by non-significant covariate-by-treatment interactions for Science Achievement ($F(1,170) = 1.12, p = .291$) and Collaboration Skills ($F(1,170) = 0.89, p = .347$).

To evaluate potential clustering effects within the six intact classrooms ($k = 6$), Intraclass Correlation Coefficients (ICC) were computed. The resulting values for Science Achievement ($ICC = .021$) and Collaboration Skills ($ICC = .018$) fell well below the standard .05 benchmark. These preliminary diagnostics confirm that group allocation did not introduce selection bias, justifying the use of an individual-level

Multivariate Analysis of Covariance (MANCOVA) without multilevel modeling adjustments.

Descriptive Overview and Combined Multivariate Effect

Table 4 outlines the descriptive statistics—including pretest means, posttest means, standard deviations, and Hake's Normalized Gain Scores (g)—for both instructional conditions.

Table 4. Descriptive Statistics and Normalized Gain Scores by Group

Dependent Variable	Group	Pretest (M $\pm$ SD)	Posttest (M $\pm$ SD)	Normalized Gain (g)	Gain Category
Science Achievement (SAT)	Experimental ($n = 88$)	48.25 ± 8.12	82.10 ± 7.45	0.65	Medium-High
	Control ($n = 86$)	47.90 ± 8.35	64.30 ± 9.15	0.31	Medium-Low
Collaboration Skills	Experimental ($n = 88$)	51.40 ± 7.60	85.60 ± 6.20	0.70	High
	Control ($n = 86$)	50.80 ± 7.85	61.20 ± 8.40	0.21	Low

Note: Maximum possible score for each instrument is 100. Normalized Gain $g > 0.70$ = High; $0.30 \leq g \leq 0.70$ = Medium; $g < 0.30$ = Low.

Baseline measures confirmed initial equivalence between groups. Post-intervention data indicate that the Experimental Group achieved substantially higher mean scores and gain rates in both Science Achievement ($M = 82.10, g = 0.65$) and Collaboration Skills ($M = 85.60, g = 0.70$) compared to the Control Group ($M = 64.30, g = 0.31$; $M = 61.20, g = 0.21$, respectively). To evaluate the primary research hypothesis regarding the combined effect of Group Investigation on both outcome variables, a one-way MANCOVA was conducted with baseline pretests entered as covariates (Table 5).

Table 5: Multivariate Tests for Combined Outcome Variables

Effect	Statistic	Value	F	Hypothesis df	Error df	p	Partial η^2
Instructional Group	Pillai's Trace	0.732	232.14	2	169	< .001	.732
	Wilks'	0.268	232.14	2	169	< .001	.732
	Lambda						
	Hotelling's Trace	2.747	232.14	2	169	< .001	.732
	Roy's Largest Root	2.747	232.14	2	169	< .001	.732

The multivariate test revealed a statistically significant main effect for instructional condition:

$$\text{Pillai's Trace} = 0.732, \quad F(2,169) = 232.14, \quad p < .001, \quad \eta_p^2 = .732$$

This result confirms that 73.2% of the variance in combined science achievement and collaboration skills is directly attributable to the Group Investigation intervention. The robust multivariate effect demonstrates that structured, small-group inquiry simultaneously enhances cognitive mastery and social competence, surpassing conventional teacher-directed instruction.

Impact of Group Investigation on Science Achievement

To isolate the intervention effect on academic performance, a univariate follow-up ANCOVA was performed for Science Achievement, using adjusted estimated marginal means (EMMeans) to account for baseline variations (Table 6).

Table 6. Follow-Up Univariate ANCOVA Results for Outcome Variables

Source	Dependent Variable	Adjusted Mean (Exp)	Adjusted Mean (Ctrl)	df	F	p	Partial η^2
Treatment	Science Achievement	81.95 (SE = 0.88)	64.45 (SE = 0.89)	1,170	194.35	< .001	.533

Collaboration Skills	85.28 (SE = 0.78)	61.52 (SE = 0.79)	1,170	458.12	< .001	.729
----------------------	-------------------	-------------------	-------	--------	--------	------

The univariate analysis for Science Achievement demonstrated a statistically significant advantage for the Experimental Group:

$$F(1,170) = 194.35, \quad p < .001, \quad \eta_p^2 = .533$$

The adjusted mean score of the Experimental Group (EMMean = 81.95, SE = 0.88) significantly exceeded that of the Control Group (EMMean = 64.45, SE = 0.89). The partial eta squared ($\eta_p^2 = .533$) reflects a large effect size, indicating that 53.3% of the variance in science conceptual performance was explained by the Group Investigation model.

This marked gain aligns with literature emphasizing that cognitively activating, inquiry-driven frameworks lead to deeper conceptual understanding (Aditomo & Klieme, 2020; Cairns, 2019; Zhang et al., 2021). Conventional direct instruction often relies on passive listening and algorithmic task completion, which restricts cognitive engagement to lower-level recall (Areepattamannil et al., 2020; Gao, 2014). In contrast, the six stages of Group Investigation require elementary students to formulate research questions, process primary data, synthesize findings into visual models, and defend their conclusions during peer presentations.

This multi-stage structure promotes active knowledge construction and high-level cognitive processing (Kang, 2020; Liou, 2021). Furthermore, these findings support previous studies by Koksai and Berberoglu (2012) and Oliver et al. (2021), demonstrating that scaffolded collaborative inquiry bridges the gap between theoretical scientific concepts and practical reasoning (Wan et al., 2023).

Mechanisms Driving Collaboration Skills Development

Univariate ANCOVA testing for Collaboration Skills similarly yielded a statistically significant intervention effect favoring the experimental condition:

$$F(1,170) = 458.12, \quad p < .001, \quad \eta_p^2 = .729$$

The Experimental Group achieved an adjusted mean score of 85.28 ($SE = 0.78$), compared to 61.52 ($SE = 0.79$) for the Control Group. The large effect size ($\eta_p^2 = .729$) demonstrates that the Group Investigation intervention accounted for nearly 73% of the variance in student collaboration outcomes. These results illustrate the structural advantages of Group Investigation over unstructured or informal group activities. Traditional practical tasks frequently suffer from uneven participation, passive loafing, or interpersonal friction due to a lack of explicit cooperation frameworks (Abrahams & Millar, 2008; Toplis, 2012). Group Investigation addresses these issues by embedding positive interdependence and individual accountability directly into its operational structure.

During the Cooperative Planning, Implementation, and Analysis phases, students must divide tasks, negotiate roles, listen actively to diverse perspectives, and reach consensus to complete group artifacts (Pun & Cheung, 2023). This systematic reliance on peers creates a supportive environment where students regularly exchange constructive feedback and refine their communication skills (Butterfuss et al., 2022; Kruiper et al., 2022). Consequently, Group Investigation transforms standard science lessons into interactive environments that cultivate key non-cognitive competencies alongside academic growth (Jansen et al., 2022; Rogers & Fraser, 2023).

This study aligns with the research of Koksal and Berberoglu (2012), Kang (2020), and Wan et al. (2023), which demonstrated that structured inquiry and student-centered collaborative frameworks yield significant gains in both conceptual understanding and affective dispositions toward science. Furthermore, these findings offer a valuable global perspective when compared against international large-scale assessment data (e.g., PISA and TIMSS). Cross-national analyses across both East Asian and Western education systems show that guided, structured inquiry fosters substantially higher cognitive and non-cognitive outcomes than unguided or purely teacher-directed instruction (Aditomo & Klieme, 2020; Areepattamannil et al., 2020; Gao, 2014; Lau & Ho, 2022; Lau & Lam, 2017). While minimally guided or unguided

practical activities often produce mixed academic outcomes globally due to unstructured cognitive demand (Oliver et al., 2021; Salchegger et al., 2021), the structured operational phases of Group Investigation successfully mitigate these issues by promoting cognitive activation, consistent with evidence from diverse international instructional contexts (Liou, 2021; Liou & Ho, 2016; Long et al., 2022).

In addition to its contributions to science education, this research provides meaningful insights into mathematics education and integrated STEM pedagogy. Scientific inquiry inherently requires mathematical reasoning, quantitative data collection, and analytical problem-solving. By collaborating within Group Investigation teams, students activate their mathematical self-concept and apply quantitative reasoning in meaningful real-world contexts, echoing findings from TIMSS mathematics achievement models (Chen, 2014; Jo & Seo, 2021). The cognitive activation and structured peer dialogue embedded in Group Investigation encourage students to process abstract mathematical and scientific representations collectively, reinforcing the strong interdisciplinary synergy between mathematical self-efficacy, cognitive activation, and academic achievement documented across broader educational settings (Lay et al., 2013; Zhang et al., 2021).

Despite the positive findings, several limitations of this research should be noted. First, the intervention was conducted over a single instructional unit, which limits the ability to evaluate long-term skill retention, sustained academic growth, or shifts in student motivation over extended academic periods (Gibson & Chase, 2002; Maricuțoiu & Sulea, 2019; Potvin & Hasni, 2014). Second, broader socio-contextual factors, such as individual family socioeconomic status and institutional policy constraints, were not explicitly integrated into the analytical model, despite their known direct and indirect influences on student performance (Gustafsson & Yang Hansen, 2018; Korous et al., 2022; Tan, 2024). Finally, classroom implementation can be constrained by systemic practical assessment policies or language challenges in specialized settings, such as English Medium Instruction (EMI) classrooms (Abrahams,

2009; Childs & Baird, 2020; Pun et al., 2023). Future research should utilize longitudinal mixed-methods approaches across diverse socioeconomic environments to further explore the long-term impact of Group Investigation across science and mathematics disciplines.

From a theoretical perspective, this study validates social constructivist models within primary science education by demonstrating that academic achievement and social collaboration develop synergistically. Formative feedback loops embedded across the six stages of Group Investigation act as motivational catalysts, enhancing student self-efficacy and task valuation (Fong et al., 2019; Gotwals, 2018; Yan et al., 2021). Specifically, peer-mediated scaffolded interactions allow learners to internalize complex scientific concepts within their zone of proximal development through active discourse and joint problem-solving. This continuous cognitive activation transforms passive knowledge reception into an active, self-regulated learning process that strengthens intrinsic interest in scientific inquiry (Jansen et al., 2014; Simons et al., 2004).

Practically, these findings provide elementary educators with an actionable pedagogical model. By shifting from teacher-led lectures to structured group inquiries, practitioners can address diverse learning needs and mitigate the influence of background socioeconomic disparities on student outcomes (Gustafsson & Yang Hansen, 2018; Tan, 2024). Furthermore, the clear operational phases offer teachers a structured template to manage group dynamics effectively, thereby minimizing common classroom management challenges associated with unstructured teamwork (Abrahams, 2009; Yeung, 2009). Educators can also leverage peer feedback mechanisms within group tasks to perform ongoing formative assessments, allowing for timely pedagogical adaptations without increasing administrative burdens (Burns et al., 2019; Kruiper et al., 2022).

From an educational policy standpoint, incorporating structured cooperative investigation frameworks into early science curricula supports current global efforts to

develop 21st-century skills without sacrificing core academic standards (Lau & Lam, 2017; Yan & Chiu, 2023). Educational leaders and curriculum developers should therefore prioritize professional development programs that equip teachers with the competencies required to facilitate student-centered, inquiry-based group learning (Hayes & Trexler, 2016; Yung et al., 2013).

While this study implemented strong methodological and statistical controls, several limitations should be acknowledged. First, the quasi-experimental design relied on intact classrooms to prevent administrative disruption, which precluded complete individual randomization and left potential unmeasured variables, such as home learning environments, uncontrolled despite baseline equivalence checks. Second, the six-week intervention duration evaluated immediate unit-level learning gains but did not assess long-term retention or the transferability of collaboration skills to other academic disciplines. Third, the primary reliance on quantitative achievement tests and observational rubrics may not fully capture the nuanced qualitative complexity of real-time peer interactions during dynamic problem-solving. To address these constraints, future research should employ longitudinal designs spanning multiple academic years to evaluate the durability and cross-disciplinary transfer of learning gains over time. Furthermore, incorporating advanced qualitative and mixed-methodologies—such as classroom discourse analysis and epistemic network modeling—would provide deeper empirical insights into how peer communication and collaborative reasoning dynamically evolve throughout each sequential stage of Group Investigation (Cheung et al., 2023; Pun et al., 2023).

CONCLUSION

This study demonstrates that implementing the Group Investigation cooperative learning model substantially enhances both elementary students' science achievement and interpersonal collaboration skills compared to conventional direct instruction. By integrating small-group inquiry with structured peer interactions, the intervention

successfully bridges cognitive domain mastery and socio-emotional development in primary science education. The empirical findings confirm that young learners perform significantly better when provided with interactive environments that encourage active exploration over passive listening. Ultimately, these results validate group-based cooperative inquiry as a powerful pedagogical approach for enhancing primary science education.

These notable gains were achieved through a systematic inquiry framework encompassing topic selection, cooperative planning, implementation, synthesis, peer presentation, and evaluation. This structured sequence establishes strong positive interdependence and individual accountability among all group members. Furthermore, embedded iterative formative feedback loops continuously drive high cognitive activation and foster reflective problem-solving throughout the learning process. As a result, students actively participate in constructive classroom discourse, resolve conceptual misunderstandings, and co-construct scientific knowledge alongside their peers.

To address the methodological boundaries of this short-term quasi-experimental study, future research should utilize longitudinal designs. Extended observation periods spanning multiple academic terms will help researchers evaluate the long-term retention and cross-disciplinary transfer of collaborative skills. Additionally, incorporating advanced qualitative and mixed-methodologies—such as classroom discourse analysis and epistemic network modeling—will offer deeper empirical insights. Such analytical tools will clarify how peer reasoning and dynamic interactions evolve across each stage of collaborative inquiry.

Ultimately, this research supports a broader educational transformation from passive, teacher-centered instruction toward active, inquiry-driven learning environments. Reimagining traditional classrooms into collaborative learning

communities equips young students with fundamental competencies for lifelong learning. By establishing an empirical foundation for scaling cooperative investigation strategies in early education, this work provides practical guidance for curriculum developers and educational policymakers. Consequently, this approach contributes to building the essential scientific literacy, critical thinking, and collaborative capabilities needed to address complex societal challenges.

Author Contribution

EK: Conceptualization, Methodology, Investigation, Writing – Original Draft, Visualization. WEP: Formal Analysis, Data Curation, Writing – Review & Editing. IS: Validation, Supervision, Writing – Review & Editing.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The study was conducted entirely using internal resources and operational support from the authors' affiliated institutions.

Conflict of Interest Statement

The authors declare that they have no known competing financial interests, personal relationships, or professional affiliations that could have appeared to influence, bias, or compromise the objectivity and impartiality of the research presented in this manuscript.

Declaration of Generative AI and AI-Assisted Technologies

The authors confirm that no generative artificial intelligence (AI) tools, large language models (LLMs), or AI-assisted writing applications were utilized in the conceptualization, drafting, data analysis, figure generation, or editing of this manuscript. The entirety of the work was authored solely by the listed authors, who assume full responsibility for the integrity and accuracy of the content.

Additional Information

Not applicable. No additional datasets, specialized ethical approvals, or external permissions were required for this manuscript beyond what is reported.

REFERENCES

- Abrahams, I. (2009). Does practical work really motivate? A study of the affective value of practical work in secondary school science. *International Journal of Science Education*, 31(17), 2335–2353. <https://doi.org/10.1080/09500690802342836>
- Abrahams, I., & Millar, R. (2008). Does practical work really work? A study of the effectiveness of practical work as a teaching and learning method in school science. *International Journal of Science Education*, 30(14), 1945–1969. <https://doi.org/10.1080/09500690701749305>
- Aditomo, A., & Klieme, E. (2020). Forms of inquiry-based science instruction and their relations with learning outcomes: Evidence from high and low-performing education systems. *International Journal of Science Education*, 42(4), 504–525. <https://doi.org/10.1080/09500693.2020.1716093>
- Areepattamannil, S., Cairns, D., & Dickson, M. (2020). Teacher-directed versus inquiry-based science instruction: Investigating links to adolescent students' science dispositions across 66 countries. *Journal of Science Teacher Education*, 31(6), 675–704. <https://doi.org/10.1080/1046560X.2020.1753309>
- Burns, E. C., Martin, A. J., & Collie, R. J. (2019). Examining the yields of growth feedback from science teachers and students' intrinsic valuing of science: Implications for student-and school-level science achievement. *Journal of Research in Science Teaching*, 56(8), 1060–1082. <https://doi.org/10.1002/tea.21546>

-
- Butterfuss, R., Roscoe, R. D., Allen, L. K., McCarthy, K. S., & McNamara, D. S. (2022). Strategy uptake in Writing Pal: Adaptive feedback and instruction. *Journal of Educational Computing Research*, 60(3), 696–721. <https://doi.org/10.1177/07356331211045304>
- Cairns, D. (2019). Investigating the relationship between instructional practices and science achievement in an inquiry-based learning environment. *International Journal of Science Education*, 41(15), 2113–2135. <https://doi.org/10.1080/09500693.2019.1660927>
- Chan, H., Cheung, K., & Erduran, S. (2023). Science communication in the media and human mobility during the COVID-19 pandemic: A time series and content analysis. *Public Health*, 218, 106–113. <https://doi.org/10.1016/j.puhe.2023.03.001>
- Chang, Y. (2014). Science motivation across Asian countries: Links among future-oriented motivation, self-efficacy, task values, and achievement outcomes. *The Asia-Pacific Education Researcher*, 24(1), 247–258. <https://doi.org/10.1007/s40299-014-0176-3>
- Chen, Q. (2014). Using TIMSS 2007 data to build mathematics achievement model of fourth graders in Hong Kong and Singapore. *International Journal of Science and Mathematics Education*, 12(6), 1519–1545. <https://doi.org/10.1007/s10763-013-9505-x>
- Cheung, K. K. C., Chan, H., & Erduran, S. (2023). Communicating science in the COVID-19 news in the UK during Omicron waves: Exploring representations of nature of science with epistemic network analysis. *Humanities and Social Sciences Communications*, 10(1). <https://doi.org/10.1057/s41599-023-01771-2>

- Childs, A., & Baird, J. A. (2020). General Certificate of Secondary Education (GCSE) and the assessment of science practical work: An historical review of assessment policy. *The Curriculum Journal*, 31(3), 357–378. <https://doi.org/10.1002/curj.20>
- Fong, C. J., Patall, E. A., Vasquez, A. C., & Stautberg, S. (2019). A meta-analysis of negative feedback on intrinsic motivation. *Educational Psychology Review*, 31(1), 121–162. <https://doi.org/10.1007/s10648-018-9446-6>
- Gao, S. (2014). Relationship between science teaching practices and students' achievement in Singapore, Chinese Taipei, and the US: An analysis using TIMSS 2011 data. *Frontiers of Education in China*, 9(4), 519–551. <https://doi.org/10.1007/BF03397039>
- Gibson, H. L., & Chase, C. (2002). Longitudinal impact of an inquiry-based science program on middle school students' attitudes toward science. *Science Education*, 86(5), 693–705. <https://doi.org/10.1002/sce.10039>
- Gotwals, A. W. (2018). Where are we now? Learning progressions and formative assessment. *Applied Measurement in Education*, 31(2), 157–164. <https://doi.org/10.1080/08957347.2017.1408626>
- Gustafsson, J.-E., & Yang Hansen, K. (2018). Changes in the impact of family education on student educational achievement in Sweden 1988–2014. *Scandinavian Journal of Educational Research*, 62(5), 719–736. <https://doi.org/10.1080/00313831.2017.1306799>
- Hayes, K. N., & Trexler, C. J. (2016). Testing predictors of instructional practice in elementary science education: The significant role of accountability. *Science Education*, 100(2), 266–289. <https://doi.org/10.1002/sce.21206>

- Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55.
<https://doi.org/10.1080/10705519909540118>
- Hu, Z. (2022). Local meanings of international student assessments: An analysis of media discourses of PISA in China, 2010-2016. *Compare: A Journal of Comparative and International Education*, 52(3), 419–439.
<https://doi.org/10.1080/03057925.2020.1775553>
- Jansen, M., Schroeders, U., & Lüdtke, O. (2014). Academic self-concept in science: Multidimensionality, relations to achievement measures, and gender differences. *Learning and Individual Differences*, 30, 11–21.
<https://doi.org/10.1016/j.lindif.2013.12.003>
- Jansen, T., Meyer, J., Wigfield, A., & Möller, J. (2022). Which student and instructional variables are most strongly related to academic motivation in K-12 education? A systematic review of meta-analyses. *Psychological Bulletin*, 148(1–2), 1–26.
<https://doi.org/10.1037/bul0000354>
- Jo, S., & Seo, Y. S. (2021). Does maths self-concept mediate the relation between instructional approaches and students' maths achievement? Evidence from the U.S. TIMSS 2015 data. *Research Papers in Education*, 38(2), 1–21.
<https://doi.org/10.1080/02671522.2021.1961294>
- Kang, J. (2020). Interrelationship between inquiry-based learning and instructional quality in predicting science literacy. *Research in Science Education*, 52(1), 339–355. <https://doi.org/10.1007/s11165-020-09946-6>

-
- Koksal, E. A., & Berberoglu, G. (2012). The effect of guided-inquiry instruction on 6th grade Turkish students' achievement, science process skills, and attitudes toward science. *International Journal of Science Education*, 36(1), 66–78. <https://doi.org/10.1080/09500693.2012.721942>
- Korous, K. M., Causadias, J. M., Bradley, R. H., Luthar, S. S., & Levy, R. (2022). A systematic overview of meta-analyses on socioeconomic status, cognitive ability, and achievement: The need to focus on specific pathways. *Psychological Reports*, 125(1), 55–97. <https://doi.org/10.1177/0033294120984127>
- Kruiper, S. M., Leenknecht, M. J., & Slof, B. (2022). Using scaffolding strategies to improve formative assessment practice in higher education. *Assessment & Evaluation in Higher Education*, 47(3), 458–476. <https://doi.org/10.1080/02602938.2021.1927981>
- Lau, K. C., & Ho, S. C. E. (2022). Attitudes towards science, teaching practices, and science performance in PISA 2015: Multilevel analysis of the Chinese and Western top performers. *Research in Science Education*, 52(2), 415–426. <https://doi.org/10.1007/s11165-020-09954-6>
- Lau, K. C., & Lam, T. Y. P. (2017). Instructional practices and science performance of 10 top-performing regions in PISA 2015. *International Journal of Science Education*, 39(15), 2128–2149. <https://doi.org/10.1080/09500693.2017.1387947>
- Lay, Y. F., Ng, K. T., & Chong, P. S. (2013). Analyzing affective factors related to eighth grade learners' science and mathematics achievement in TIMSS 2007. *The Asia-Pacific Education Researcher*, 24(1), 103–110. <https://doi.org/10.1007/s40299-013-0163-0>

- Liou, P. Y. (2021). Students' attitudes toward science and science achievement: An analysis of the differential effects of science instructional practices. *Journal of Research in Science Teaching*, 58(3), 310–334. <https://doi.org/10.1002/tea.21643>
- Liou, P. Y., & Ho, J. H. N. (2016). Relationships among instructional practices, students' motivational beliefs and science achievement in Taiwan using hierarchical linear modelling. *Research Papers in Education*, 33(1), 73–88. <https://doi.org/10.1080/02671522.2016.1236832>
- Long, H., Gao, S., Yang, L., & Chen, J. (2022). Do teaching practices and enjoyment of science matter to science achievement? *Psychology in the Schools*, 59(2), 334–355. <https://doi.org/10.1002/pits.22611>
- Maricuțoiu, L. P., & Sulea, C. (2019). Evolution of self-efficacy, student engagement and student burnout during a semester. A multilevel structural equation modeling approach. *Learning and Individual Differences*, 76, 101785. <https://doi.org/10.1016/j.lindif.2019.101785>
- Oliver, M., McConney, A., & Woods-McConney, A. (2021). The efficacy of inquiry-based instruction in science: A comparative analysis of six countries using PISA 2015. *Research in Science Education*, 51(S2), 595–616. <https://doi.org/10.1007/s11165-019-09901-0>
- Potvin, P., & Hasni, A. (2014). Interest, motivation and attitude towards science and technology at K-12 levels: A systematic review of 12 years of educational research. *Studies in Science Education*, 50(1), 85–129. <https://doi.org/10.1080/03057267.2014.881626>
- Pun, J. K., & Cheung, K. K. (2023). Meaning making in collaborative practical work: A case study of multimodal challenges in a Year 10 chemistry classroom.

-
- Research in Science & Technological Education*, 41(1), 271–288.
<https://doi.org/10.1080/02635143.2021.1895101>
- Pun, J. K. H., Fu, X., & Cheung, K. K. (2023). Language challenges and coping strategies in English Medium Instruction (EMI) science classrooms: A critical review of literature. *Studies in Science Education*, 1–32.
<https://doi.org/10.1080/03057267.2023.2188704>
- Rogers, J. R., & Fraser, B. J. (2023). Sex and frequency of practical work as determinants of middle-school science students' learning environment perceptions and attitudes. *Learning Environments Research*, 26(2), 315–336.
<https://doi.org/10.1007/s10984-022-09426-w>
- Salchegger, S., Wallner-Paschon, C., & Bertsch, C. (2021). Explaining Waldorf students' high motivation but moderate achievement in science: Is inquiry-based science education the key? *Large-Scale Assessments in Education*, 9(1), 1–23.
<https://doi.org/10.1186/s40536-021-00107-3>
- Simons, J., Vansteenkiste, M., Lens, W., & Lacante, M. (2004). Placing motivation and future time perspective theory in a temporal perspective. *Educational Psychology Review*, 16(2), 121–139.
<https://doi.org/10.1023/B:EDPR.0000026609.94841.2f>
- Tan, C. Y. (2024). Direct and indirect influences of familial socioeconomic status on students' science achievement. *Oxford Review of Education*, 50(2), 207–231.
<https://doi.org/10.1080/03054985.2022.2139672>
- Toplis, R. (2012). Students' views about secondary school science lessons: The role of practical work. *Research in Science Education*, 42(3), 531–549.
<https://doi.org/10.1007/s11165-011-9209-6>

- Wan, Z. H., Zhan, Y., & Zhang, Y. (2023). Positive or negative? The effects of scientific inquiry on science achievement via attitudes toward science. *Science Education*, 1–22. <https://doi.org/10.1002/sce.2182522>
- Yan, Z., & Chiu, M. M. (2023). The relationship between formative assessment and reading achievement: A multilevel analysis of students in 19 countries/regions. *British Educational Research Journal*, 49(1), 186–208. <https://doi.org/10.1002/berj.3837>
- Yan, Z., King, R. B., & Haw, J. Y. (2021). Formative assessment, growth mindset, and achievement: Examining their relations in the East and the West. *Assessment in Education: Principles, Policy & Practice*, 28(5–6), 676–702. <https://doi.org/10.1080/0969594X.2021.1988510>
- Yeung, S. Y. S. (2009). Is student-centered pedagogy impossible in Hong Kong? The case of inquiry in classrooms. *Asia Pacific Education Review*, 10(3), 377–386. <https://doi.org/10.1007/s12564-009-9028-x>
- Yung, B. H. W., Zhu, Y., Wong, S. L., Cheng, M. W., & Lo, F. Y. (2013). Teachers' and students' conceptions of good science teaching. *International Journal of Science Education*, 35(14), 2435–2461. <https://doi.org/10.1080/09500693.2011.629375>
- Zhang, D., Wang, C., & Yang, Y. (2021). The association between cognitive activation and mathematics achievement: A multiple mediation model. *Educational Psychology*, 41(6), 695–711. <https://doi.org/10.1080/01443410.2021.1917520>