Indonesian Journal of Educational Development (IJED)
Volume 7, Issue 2, 2026, pp. 643-657
ISSN: 2722-1059 (Online); ISSN: 2722-3671 (Print)
DOI: https://doi.org/10.59672/ijed.v7i2.6909

# Comparative effectiveness of STEAM and CTL using MUSI board for cognitive understanding

**Choirul Hadiyati[*)1], Achmad Buchori[2], Joko Siswanto[3]**
[1]Universitas PGRI Semarang, Semarang, Indonesia; larasatikinanthi37@gmail.com
[2]Universitas PGRI Semarang, Semarang, Indonesia; achmadbuchori@upgris.ac.id
[3]Universitas PGRI Semarang, Semarang, Indonesia; jokosiswanto@upgris.ac.id
[*)]Corresponding author: Choirul Hadiyati; E-mail addresses: larasatikinanthi37@gmail.com

**Article Info**

**Article history:**
Received July 08, 2026
Revised July 17, 2026
Accepted August 08, 2026
Available online August 17, 2026

**Keywords:** Cognitive understanding, CTL, Elementary mathematics, MUSI board, STEAM.



**Abstract.** Elementary students continue to struggle to understand Least Common Multiple (LCM) and Greatest Common Factor (GCF), highlighting the need for more effective instructional approaches that promote conceptual understanding. Although Science, Technology, Engineering, Arts, and Mathematics (STEAM) and Contextual Teaching and Learning (CTL) have shown positive outcomes, direct comparisons with the same instructional medium remain limited. This study's novelty is the direct comparison of STEAM and CTL using the same manipulative instructional medium (MUSI Board), isolating instructional effects from learning media. Therefore, this study aimed to compare the effectiveness of STEAM and CTL using the MUSI Board to improve cognitive understanding of LCM and GCF. A quasi-experimental pretest–posttest non-equivalent group design involved 30 fifth-grade students assigned through matched-group sampling to STEAM ($n = 15$) and CTL ($n = 15$). Data were analyzed using $t$-tests, normalized gain (N-Gain), and Cohen's $d$. Both groups improved significantly ($p < .001$), but STEAM achieved higher learning gains (0.487 vs. 0.352) with a very large educational effect (Cohen's $d = 1.975$). Conceptually, the findings strengthen evidence on the comparative effectiveness of STEAM and CTL. Practically, MUSI Board-assisted STEAM is recommended for elementary mathematics, while future studies should examine broader topics using larger samples.

## Introduction

Building students' conceptual understanding is a fundamental objective of elementary mathematics because it enables learners to interpret mathematical ideas, justify mathematical reasoning, and apply knowledge to unfamiliar situations rather than merely performing routine calculations. Among elementary mathematics topics, Least Common Multiple (LCM) and Greatest Common Factor (GCF) form essential conceptual foundations for fractions, proportional reasoning, algebraic thinking, and mathematical problem solving. Consequently, effective mathematics instruction should promote active knowledge construction through meaningful learning experiences instead of emphasizing procedural fluency alone. Recent studies consistently demonstrate that conceptual understanding develops more effectively when students actively engage in exploration, discussion, and collaborative knowledge construction (Asilevi et al., 2024; Brandl et al., 2024; F. Chen & Chen, 2025; Sormunen et al., 2023).

Despite these expectations, many elementary classrooms still prioritize procedural accuracy over conceptual understanding in mathematics instruction. Consequently, students frequently memorize procedures for determining LCM and GCF without understanding the relationships among multiples and factors. Preliminary observations conducted in Grade V of SDN 4 Ketringan revealed that 25 of 30 students (83.3%) demonstrated inadequate conceptual understanding of LCM and GCF. Most students scored below 60, whereas the highest score reached only 78. These findings indicate a substantial gap between expected learning outcomes and actual classroom practice, highlighting the urgency of implementing instructional approaches that facilitate conceptual learning.

Constructivist learning theory explains that learners construct knowledge through active interaction with ideas, learning materials, and social experiences. Within this perspective, Contextual Teaching and Learning (CTL) enables students to connect mathematical concepts with authentic real-life situations, whereas Science, Technology, Engineering, Arts, and Mathematics (STEAM) extends conceptual learning through interdisciplinary inquiry, engineering design, creativity, collaboration, and problem solving. Although both approaches are student-centered, STEAM theoretically provides richer opportunities for conceptual knowledge construction because it integrates multiple disciplinary perspectives within authentic problem-solving activities (Widana et al., 2021). Manipulative instructional media further enhance learning effectiveness. Accordingly, this study used the MUSI Board (Multifunctional Multiples and Factors Board), a manipulative medium designed to help students explore multiples, factors, LCM, and GCF through hands-on investigation and collaborative learning. Previous studies likewise demonstrate that manipulative media, collaborative learning, and game-based instruction improve students' engagement, motivation, and conceptual understanding (Anazifa & Djukri, 2023; Camacho-Sánchez et al., 2023; Tokac et al., 2023; Nadeem et al., 2023; Sailer, Murböck, et al., 2023).

Previous studies have consistently reported positive effects of STEAM, CTL, manipulative media, and collaborative learning. Nevertheless, important methodological limitations remain. Most previous studies investigated STEAM and CTL separately, employed different instructional media across comparison groups, focused on overall mathematics achievement rather than conceptual understanding, or were conducted in technology-rich learning environments. Consequently, existing evidence cannot determine whether improvements originate from the instructional approach itself or from the learning media employed during instruction. Direct comparative evidence using identical instructional media remains scarce, particularly in elementary mathematics learning involving LCM and GCF (Dabbous et al., 2023; Yang et al., 2025).

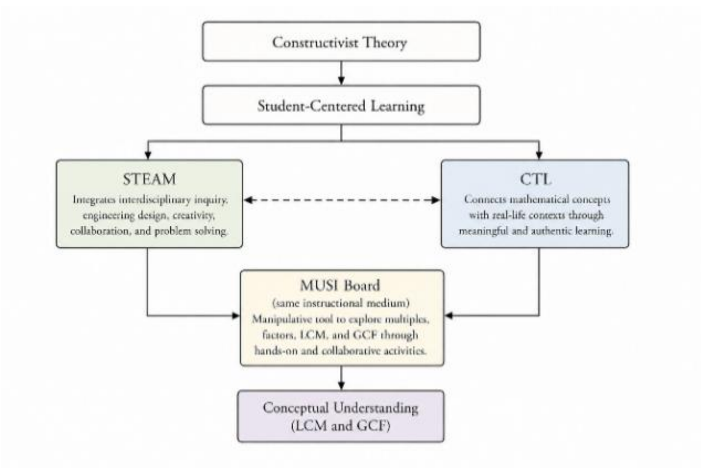


**Image 1.** Conceptual Framework of the Study

To address this gap, this study compares STEAM and CTL using the same manipulative instructional medium, the MUSI Board. Employing the same learning medium for both groups allows differences in students' conceptual understanding to be interpreted more confidently in relation to the instructional approach.

Based on the identified research gap, this study addressed the following research question: Does STEAM assisted by MUSI Board produce significantly higher cognitive understanding of LCM and GCF than CTL assisted by the same instructional medium? It was hypothesized that students learning through STEAM assisted by MUSI Board would achieve significantly higher cognitive understanding than those learning through CTL because STEAM integrates interdisciplinary inquiry, creativity, engineering design, and collaborative problem solving. Therefore, this study aimed to compare the effectiveness of STEAM and CTL using the MUSI Board in improving elementary students' cognitive understanding of LCM and GCF.

## Method

### *Research Method and Design*

We compared STEAM and CTL assisted by the MUSI Board using a quasi-experimental, non-equivalent pretest-posttest comparison group design. We selected this design because the fifth-grade classes had been naturally formed before the study, making random assignment of students to different groups impractical in the school setting. Compared with a true experimental design, the selected design allows researchers to maintain the existing classroom organization while minimizing disruption to the regular instructional process. To reduce potential selection bias and improve baseline equivalence, students were assigned to the experimental and control groups using matched-group sampling based on their pretest scores. Consequently, this design provided an appropriate and rigorous framework for comparing the effectiveness of STEAM and CTL while preserving the authenticity of classroom learning (Creswell & Creswell, 2023; Denny et al., 2023; Suryaningsih et al., 2025; Sandi, 2021).

**Table 1.** Non-equivalent Pretest–Posttest Comparison Group Design

| Group | Pretest | Treatment | Posttest |
|---|---|---|---|
| Experimental Group | $O_1$ | STEAM + MUSI Board. | $O_2$ |
| Control Group | $O_1$ | CTL + MUSI Board. | $O_2$ |

*Note.* $O_1$ = pretest; $O_2$ = posttest.
*Source:* Adapted from (Creswell & Creswell, 2023).

As summarized in Table 1, students in both experimental groups completed identical pretest and posttest assessments using the same evaluation instruments. To ensure that any observed differences in learning outcomes resulted from the instructional approaches rather than other classroom conditions, both groups received the same mathematical content, learning objectives, instructional duration, assessment procedures, and the MUSI Board as the manipulative learning medium. Consequently, the instructional approach remained the only experimental factor distinguishing the two groups. Maintaining these equivalent learning conditions provided a more rigorous basis for comparing the effectiveness of STEAM and CTL in improving students' cognitive understanding of LCM and GCF.

### *Participants and Sampling Technique*

The study was implemented during the 2025–2026 academic year at SD Negeri 3 Ketringan, Jiken District, Blora Regency, Central Java, Indonesia. Participants included all 30 fifth-grade students enrolled at the school, making total sampling appropriate because the target population was relatively small. Participants included all 30 fifth-grade students enrolled at SD Negeri 3 Ketringan; therefore, total sampling was employed because the target population was relatively small. To establish comparable baseline cognitive abilities, all students first completed a pretest on LCM and GCF. The pretest scores were ranked from the highest to the lowest, after which students with similar scores were paired. Each pair was then assigned alternately to the experimental group (STEAM assisted by MUSI Board) and the control group (CTL assisted by MUSI Board), resulting in two groups of 15 students with comparable initial cognitive abilities. This matched-group sampling procedure minimized selection bias and strengthened the comparability of the two groups before the instructional intervention.

### *Research Ethics*

This study followed ethical principles for educational research involving human participants. Before data collection, the principal of SD Negeri 3 Ketringan granted formal permission to conduct the research during regular mathematics lessons. The classroom teacher, students, and parents or guardians were informed about the purpose, procedures, and voluntary nature of the study. Participation did not affect students' academic assessment, and all participants could withdraw at any stage without consequences. To protect participants' privacy, all collected data were anonymized and used exclusively for research purposes.

### *Research Setting and Timeline*

The research was conducted during the 2025–2026 academic year at SD Negeri 3 Ketringan, Jiken District, Blora Regency, Central Java, Indonesia. Learning activities were integrated into regular mathematics lessons covering Least Common Multiple (LCM) and Greatest Common Factor (GCF)

### *Research Procedures*

Accurate measurement of students' cognitive understanding and the fidelity of classroom implementation required valid and reliable research instruments. The researchers developed the instruments before the intervention through expert validation and empirical testing to ensure each instrument appropriately represented the intended constructs. Validation procedures included expert judgment for the learning media and cognitive achievement test, followed by reliability and item quality analyses for the test instrument. Table 2 summarizes the research instruments and their validation procedures.

**Table 2.** Research Instruments

| Instrument | Purpose | Validation Method |
|---|---|---|
| MUSI Board. | Support mathematics learning on LCM and GCF through STEAM and CTL activities | Two instructional media experts (Aiken's V = 0.85, Highly Valid) |
| Essay Test (10 items) | Measure students' cognitive understanding of LCM and GCF (C1–C4) | Two mathematics education experts (Aiken's V = 0.88, Highly Valid; Cronbach's Alpha = 0.97) |
| Observation Sheet | Assess the fidelity of STEAM and CTL implementation during classroom learning | Expert judgment |

*Note.* Interpretation of Aiken's V followed the criteria: $\geq 0.80$ = Highly Valid; 0.60–0.79 = Valid; $< 0.60$ = Needs Revision.

Table 2 shows that all research instruments met the required quality standards before classroom implementation. Expert validation confirmed that the MUSI Board and the cognitive achievement test possessed high content validity, while the reliability coefficient indicated excellent internal consistency of the essay test. Along with satisfactory item validity, difficulty level, and discrimination index obtained during the pilot testing stage, these findings support using the instruments to measure fifth-grade students' cognitive understanding of LCM and GCF while monitoring implementation fidelity across both instructional approaches.

**Table 3.** Interpretation Criteria of Aiken's V

| Aiken's V Value | Validity Category |
|---|---|
| 0.80 – 1.00 | Highly Valid |
| 0.60 – 0.79 | Valid |
| < 0.60 | Needs Revision |

The interpretation of Aiken's V coefficients followed the criteria presented in Table 3. Based on these criteria, the MUSI Board obtained an Aiken's V coefficient of 0.85 and the cognitive achievement test obtained a coefficient of 0.88, indicating that both instruments were classified as highly valid and suitable for classroom implementation.

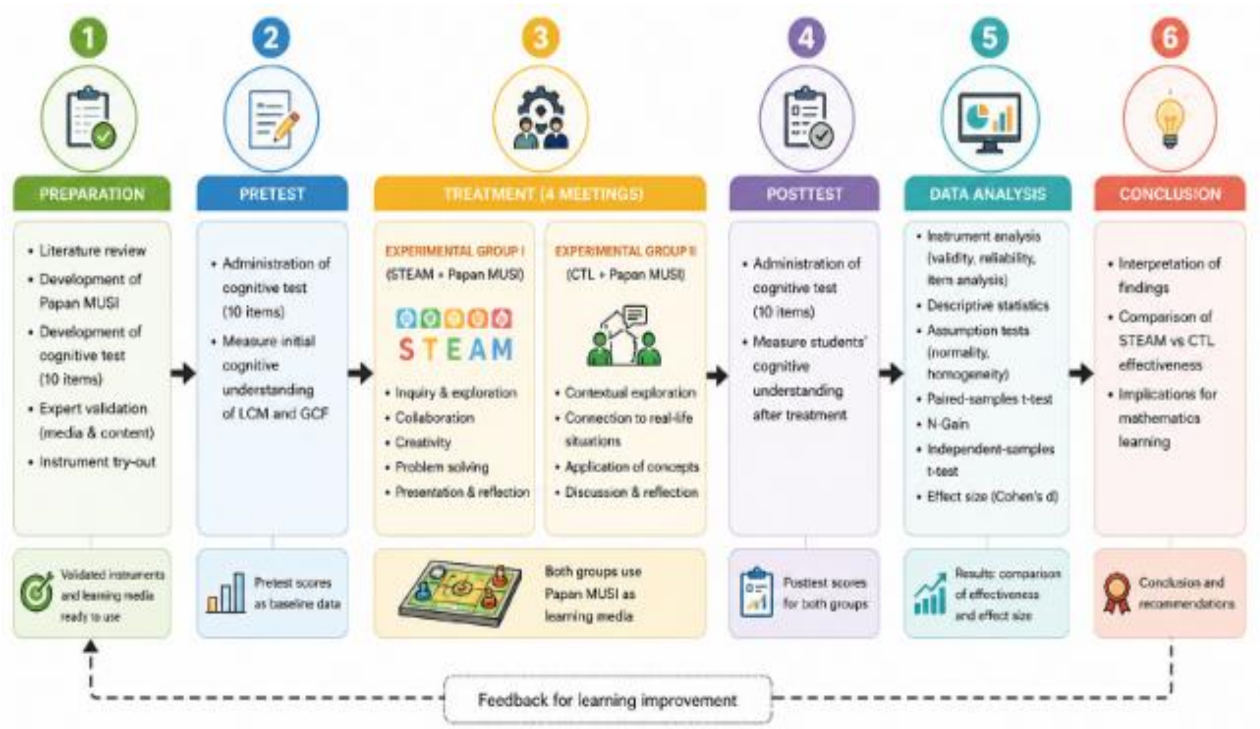


**Image 2.** Research Flow of the Study

Image 2 illustrates the overall research procedure, beginning with instrument preparation and validation, followed by pretest administration, implementation of learning interventions using STEAM and CTL assisted by the MUSI Board, posttest administration, and statistical analysis to compare students' cognitive understanding.

### *Implementation Fidelity*

To ensure the validity of the instructional intervention, the study monitored implementation fidelity throughout using structured classroom observation sheets. The observation instrument was developed based on the instructional syntax of STEAM and CTL and validated through expert judgment before classroom implementation. During each learning session, an independent observer evaluated the extent to which each stage of the lesson was implemented according to the lesson plan. The observation results were analyzed descriptively by calculating the percentage of implementation for each learning phase. High implementation fidelity indicated that both instructional approaches were delivered consistently and according to the planned procedures, strengthening the study's internal validity.

**Table 4.** Comparison of the Learning Syntax Used in the Experimental and Control Groups

| Learning Phase | Experimental Group (STEAM + MUSI Board) | Control Group (CTL + MUSI Board) |
|---|---|---|
| Orientation | Students were introduced to an authentic mathematical problem related to LCM and GCF using MUSI Board. | Students were introduced to a contextual problem related to daily life using MUSI Board. |
| Exploration | Students investigated multiples and factors through collaborative inquiry and experimentation. | Students explored mathematical concepts by connecting them with real-life situations through guided discussion. |
| Knowledge Construction | Students integrated science, technology, engineering, arts, and mathematics to develop solutions collaboratively. | Students constructed understanding through contextual activities, questioning, and authentic learning experiences. |
| Problem Solving | Students designed and tested solutions using MUSI Board and interdisciplinary reasoning. | Students solved contextual mathematical problems using MUSI Board based on real-life applications. |
| Reflection | Students presented, evaluated, and reflected on their solutions collaboratively. | Students reflected on their learning experiences and related mathematical concepts to everyday contexts. |

**Source:** Adapted from the STEAM and CTL learning models implemented in this study.

Table 4 summarizes the instructional syntax implemented in the experimental and control groups. Although both groups learned the same mathematical content using the same manipulative instructional medium (MUSI Board), the instructional procedures differed according to the characteristics of the STEAM and CTL approaches. This comparison clarifies that the instructional approach was the only experimental factor distinguishing the two groups.

### *Data Collection Techniques and Research Instruments*

Data collection relied on three research instruments: the MUSI Board. validation sheets, a cognitive achievement test, and classroom observation sheets. Two instructional media experts evaluated the quality of the instructional media, while two mathematics education specialists reviewed the cognitive assessment. High content validity was demonstrated by the coefficients of 0.85 for the MUSI Board. and 0.88 for the cognition exam, which were obtained using Aiken's V. A ten-item open-ended test spanning the cognitive domains of remembering (C1), comprehending (C2), applying (C3), and analyzing (C4) was created in accordance with the revised Bloom's taxonomy in order to assess students' cognitive understanding. Item validity, difficulty index, discrimination index, and Cronbach's Alpha reliability analysis were then used to assess instrument quality. The reliability coefficient of 0.973 confirmed excellent internal consistency, indicating that the instrument was suitable for evaluating elementary students' cognitive comprehension of LCM and GCF (Revelle, W., & Garner, 2024; Widana, 2022). The overall research sequence included instrument preparation, validation, implementation of the learning intervention, data collection, and statistical analysis.

Data analysis was conducted systematically to ensure that the research instruments met the required quality standards before evaluating the effectiveness of the instructional interventions. The evaluation of the research instruments preceded the main statistical analyses to ensure the collected data met the required quality standards. Subsequent analyses examined statistical assumptions before testing differences in the effectiveness of the STEAM and CTL approaches for improving students' cognitive understanding of LCM and GCF. Table 5 provides an overview of the analytical procedures used in this study.

**Table 5.** Data Analysis Procedures

| Analysis Stage | Purpose | Statistical Technique |
|---|---|---|
| Instrument validation | Evaluate content validity | Aiken's V |
| Item validity | Examine item consistency | Corrected Item–Total Correlation |
| Reliability analysis | Determine internal consistency | Cronbach's Alpha |
| Item quality | Evaluate item difficulty | Difficulty Index |
| Item quality | Evaluate item discrimination | Discrimination Index |
| Descriptive statistics | Summarize students' pretest and posttest performance | Mean, Standard Deviation, Minimum, Maximum |
| Assumption testing | Examine data normality | Shapiro–Wilk Test |
| Assumption testing | Examine variance homogeneity | Levene's Test |
| Within-group analysis | Compare pretest and posttest scores | Paired-Samples *t*-Test |
| Learning improvement | Measure improvement level | Normalized Gain (N-Gain) |
| Between-group analysis | Compare learning improvement between STEAM and CTL | Independent-Samples *t*-Test |
| Effect size | Determine the magnitude of treatment effectiveness | Cohen's *d* |

Table 5 shows that the analytical procedures were implemented sequentially, beginning with instrument evaluation, followed by descriptive and assumption testing, and concluding with hypothesis testing and effect size estimation. This sequence ensured that statistical analyses used valid and reliable instruments while meeting the assumptions required for parametric testing. As a result, the findings provide strong empirical support for comparing how well STEAM and CTL, using the MUSI Board, improve fifth-grade students' cognitive comprehension of LCM and GCF.

***Instrument Validity and Reliability Tests***

Instruction was delivered across four classroom sessions covering LCM and GCF. To ensure a fair comparison, both experimental groups learned with identical instructional objectives, mathematical content, lesson duration, assessment procedures, and the same manipulative learning medium, namely the MUSI Board. Differences between the groups were intentionally limited to the instructional approach. Students in the STEAM group engaged in interdisciplinary inquiry, collaborative investigation, creative design activities, and mathematical problem solving supported by the MUSI Board. In contrast, students in the CTL group explored the same mathematical concepts by relating them to authentic situations encountered in everyday life while using the identical learning medium. Controlling instructional materials and learning resources allowed differences in students' cognitive understanding to be interpreted primarily as the effect of the instructional approaches rather than variations in the learning environment.

***Data Analysis Techniques and Criteria***

The researchers evaluated the research instruments before examining the effectiveness of the instructional interventions to ensure all measurement tools met the required psychometric standards. Instrument quality was established through Aiken's V for content validity, item difficulty and discrimination index analyses, Cronbach's Alpha for reliability, and adjusted item–total correlation for item validity. Following instrument verification, students' pretest and posttest performance was summarized using descriptive statistics, including the mean, standard deviation, minimum, and maximum values. Next, assumptions for parametric analysis were examined using Levene's test to confirm homogeneity of variances and the Shapiro-Wilk test to assess data normality. The paired-samples t-test examined learning improvement within each instructional group, while normalized gain (N-Gain), computed using the following formula, quantified the amount of improvement.

$$\text{N-Gain}=\frac{\text{Posttest-Pretest}}{\text{Maximum Score-Pretest}} \quad (1)$$

In the N-Gain calculation, the maximum score represents the highest possible score that students could achieve. In contrast, students' performance before and after the instructional intervention is reflected in the pretest and posttest scores. An independent-samples t-test based on the N-Gain scores was then used to investigate differences in learning improvement between the STEAM and CTL groups. To complement the hypothesis testing, Cohen's *d* was calculated to determine the practical magnitude of the observed differences. Statistical decisions for all analyses were made using a significance level of $\alpha = 0.05$.

Hypotheses were tested at $\alpha = 0.05$ to determine whether meaningful differences existed in students' cognitive understanding between the two instructional approaches. The statistical hypotheses were formulated as follows: (1) $H_0$: No significant difference exists in the cognitive understanding of students who learn through the STEAM approach assisted by MUSI Board. and those who learn through CTL assisted by the same instructional medium; (2) $H_1$: A significant difference exists in the cognitive understanding of students who learn through the STEAM approach assisted by MUSI Board. and those who learn through CTL assisted by the same instructional medium. All statistical analyses were performed using IBM SPSS Statistics Version 27. We used the software to conduct descriptive statistics, assumption tests, hypothesis testing, normalized gain (N-Gain), and effect size analyses at a significance level of $\alpha = 0.05$.

## Results and Discussion

***Results***

***Descriptive Statistics of Students' Cognitive Understanding***

Examining students' cognitive knowledge both before and after the learning intervention was the first step in addressing the research objective of evaluating the efficacy of CTL with support from the MUSI Board and STEAM. Before inferential statistical testing, descriptive analysis of the pretest, posttest, and normalized gain (N-Gain) scores was conducted to provide a preliminary picture of each experimental group's pupils' learning progress. Table 6 displays the descriptive analysis's findings.

**Table 6.** Descriptive Statistics of Students' Cognitive Understanding

| Group | | N | Minimum | Maximum | Mean | Std. Deviation |
|---|---|---|---|---|---|---|
| STEAM | Pretes | 15 | 60.00 | 92.00 | 76.00 | 9.59 |
| | Postes | 15 | 78.00 | 97.00 | 87.26 | 5.96 |
| | N-Gain | 15 | .38 | .63 | 0.48 | 0.07 |

| Group | | N | Minimum | Maximum | Mean | Std. Deviation |
|---|---|---|---|---|---|---|
| | Valid N (listwise) | 15 | | | | |
| CTL | Pretes | 15 | 58.00 | 90.00 | 74.73 | 9.77 |
| | Postes | 15 | 69.00 | 95.00 | 83.13 | 7.68 |
| | N-Gain | 15 | .26 | .50 | 0.35 | 0.05 |
| | Valid N (listwise) | 15 | | | | |

a. No statistics are computed for one or more split files because there are no valid cases.

Table 6 presents the descriptive statistics of students' cognitive understanding before and after the intervention. Before the instructional treatment, the experimental group achieved a mean pretest score of 76.00 (SD = 9.60), whereas the control group obtained a mean score of 74.73 (SD = 9.78). Although the experimental group scored 1.27 points higher, the difference was not statistically significant, as confirmed by the independent-samples $t$-test, $t(28) = 0.358$, $p = .723$. These findings indicate that both groups had equivalent baseline cognitive understanding prior to the intervention, supporting the validity of subsequent comparisons between the two instructional approaches.

Following the intervention, both groups demonstrated improved cognitive understanding. The experimental group achieved a higher posttest mean score than the control group, indicating that students who learned through the STEAM approach assisted by the MUSI Board demonstrated greater conceptual improvement than those who learned through the CTL approach using the same instructional medium. Furthermore, the experimental group achieved a higher mean normalized gain (N-Gain = 0.487) than the control group (N-Gain = 0.352). According to Hake's classification, both values fall within the moderate improvement category; however, the experimental group's higher N-Gain indicates that STEAM assisted by the MUSI Board produced greater educational improvement in students' conceptual understanding of LCM and GCF. These descriptive findings were subsequently confirmed through inferential statistical analyses.

### *Assumption Testing*

We assessed data distribution and variance equality as an initial step to confirm the suitability of parametric statistical procedures. We evaluated the normality of the N-Gain scores with the Shapiro–Wilk test, and applied Levene's test to assess homogeneity of variance between the two experimental groups. Results of both assumption tests are presented in Table 7.

**Table 7.** Results of Normality and Homogeneity Tests

| | Kelompok | Kolmogorov-Smirnov[a] | | | Shapiro-Wilk | | |
|---|---|---|---|---|---|---|---|
| | | Statistic | df | Sig. | Statistic | df | Sig. |
| NGain | STEAM | 0.16 | 15 | 0.20* | 0.94 | 15 | 0.38 |
| | CTL | 0.12 | 15 | 0.20* | 0.94 | 15 | 0.51 |

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Results summarized in Table 7 confirmed that the assumptions required for parametric analysis were satisfied..The Shapiro–Wilk test showed that the pretest scores of both the experimental group ($p = .976$) and the control group ($p = .937$) were normally distributed because all significance values exceeded .05. Likewise, the posttest scores of the experimental group ($p = .976$) and the control group ($p = .956$) also satisfied the normality assumption. These findings indicate that the score distributions approximated normal distributions, supporting the use of parametric statistical procedures for hypothesis testing.

### *Hypothesis Testing*

The study then compared learning improvement between the two instructional approaches to determine whether meaningful differences existed after implementing the MUSI Board. Students'

N-Gain scores served as the basis for evaluating the research hypothesis through an independent-samples *t*-test. Results of the statistical analysis are presented in Table 8.

**Table 8.** Independent-Samples *t*-Test Results

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | F | Sig. | T | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
| | | | | | | | | | Lower | Upper |
| N-Gain | Equal variances assumed | 2.69 | 0.12 | 5.40 | 28 | .000 | 0.13 | 0.02 | 0.08 | 0.18 |
| | Equal variances not assumed | | | 5.40 | 26.22 | .000 | 0.13 | 0.02 | 0.08 | 0.18 |

A statistically significant difference in students' learning improvement between the two instructional modalities is shown by the results in Table 8 (t(28) = 5.40, p < 0.001). Students who learnt using the STEAM approach with MUSI Board's assistance made more progress in their cognitive comprehension than those who learned through CTL utilizing the same learning materials, according to the positive mean difference (0.13). As a result, the null hypothesis ($H_0$) was rejected, suggesting substantial differences between the two teaching strategies in enhancing elementary students' cognitive comprehension of LCM and GCF.

***Learning Improvement***

To ascertain if students' cognitive comprehension of LCM and GCF significantly increased following the deployment of MUSI Board., learning improvement within each experimental group was assessed using the paired-samples t-test. The STEAM and CTL groups' pretest and posttest results were compared independently. Tables 8 and 9 display the paired-samples t-test results**.**

**Table 9.** Paired-Samples *t*-Test Results for the STEAM Group

| STEAM | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Paired Differences | | | | | t | Df | Sig. (2-tailed) |
| | | Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference | | | | |
| | | | | | Lower | Upper | | | |
| Pair 1 | Pretes - Postes | -11.26 | 4.00 | 1.03 | -13.48 | -9.04 | -10.88 | 14 | .000 |

Table 9 shows that students in the experimental group improved significantly in cognitive understanding following the STEAM intervention assisted by the MUSI Board. The mean difference between the pretest and posttest scores was −11.26 (SD = 4.00; SE = 1.03), with a 95% confidence interval ranging from −13.48 to −9.04. The paired-samples *t*-test confirmed that the improvement was statistically significant, $t(14) = -10.88, p < .001$. Because the analysis used pretest minus posttest scores, the negative mean difference indicates that students scored higher after the intervention. These findings demonstrate that the STEAM approach assisted by the MUSI Board produced substantial improvements in students' cognitive understanding of LCM and GCF.

**Table 10.** Paired-Samples *t*-Test Results for the CTL Group

| CTL | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Paired Differences | | | | | t | Df | Sig. (2-tailed) |
| | | Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference | | | | |
| | | | | | Lower | Upper | | | |
| Pair 1 | Pretes – Postes | -8.40 | 2.16 | 0.55 | -9.59 | -7.20 | -15.02 | 14 | .000 |

Table 10 shows that students in the CTL group improved significantly after the intervention. The mean difference between pretest and posttest scores was −8.40 (SD = 2.16; SE = 0.55), with a 95% confidence interval ranging from −9.60 to −7.20. The paired-samples t-test confirmed a statistically significant improvement, $t(14) = -15.02$, $p < .001$. Because the analysis compared pretest minus posttest scores, the negative mean difference indicates higher achievement after instruction. Similar significant improvements ($p < .001$) were also observed in the experimental group, demonstrating that both student-centered approaches enhanced students' cognitive understanding of LCM and GCF. Nevertheless, the larger gain in the experimental group indicates that STEAM supported by the MUSI Board promoted more meaningful conceptual learning than CTL using the same instructional medium.

### *Comparative Effectiveness*

To determine whether STEAM supported by the MUSI Board improved students' cognitive understanding more than CTL supported by the same learning materials, a comparative analysis was conducted. We used students' N-Gain scores in an independent-samples t-test to assess differences in learning improvement between the two instructional approaches after confirming that the assumptions for parametric analysis were met. Cohen's d was computed alongside the significance testing to estimate the instructional effect size and provide useful information about the relative efficacy of the two learning strategies. Table 11 displays the comparison analysis's findings.

**Table 11.** Comparative Effectiveness of STEAM and CTL Based on N-Gain Scores

| Independent Samples Effect Sizes | | | | | |
|---|---|---|---|---|---|
| | | Standardizer[a] | Point Estimate | 95% Confidence Interval | |
| | | | | Lower | Upper |
| NGainin | Cohen's d | 0.06 | 1.97 | 1.08 | 2.84 |
| | Hedges' correction | 0.07 | 1.92 | 1.05 | 2.76 |
| | Glass's delta | 0.05 | 2.29 | 1.17 | 3.38 |

a. The denominator used in estimating the effect sizes.
Cohen's d uses the pooled standard deviation.
Hedges' correction uses the pooled standard deviation, plus a correction factor.
Glass's delta uses the control group's sample standard deviation.

The effect size analysis revealed a very large practical impact of the intervention on students' learning gains (Table 11). Cohen's *d* was 1.97 (95% CI = 1.08–2.84), indicating a very large effect according to Cohen's criteria. Similar results were obtained for Hedges' *g* (1.92) and Glass's (2.29), confirming the robustness of the findings. These results indicate that the intervention was not only statistically significant but also educationally meaningful, demonstrating its high effectiveness in improving students' learning gains compared with the control group.

### *Observation Results of Learning Implementation*

To ensure that both instructional approaches were implemented consistently throughout the intervention, classroom observations were conducted during each instructional session using a structured implementation fidelity observation sheet. Table 12 presents the observation results.

**Table 12.** Observation Results of Learning Implementation

| | N | Mean | Std. Deviation | Minimum | Maximum |
|---|---|---|---|---|---|
| Experimental (STEAM + MUSI Board) | 4 | 36.56 | 4.72 | 30.00 | 40.00 |
| Control (CTL + MUSI Board) | 4 | 31.25 | 4.21 | 26.25 | 36.25 |

Table 12 shows that the experimental group achieved a higher implementation score (M = 36.56, SD = 4.72) than the control group (M = 31.25, SD = 4.21). These findings indicate that the MUSI Board was implemented more consistently across the four instructional sessions. This result is consistent with the higher posttest scores, N-Gain, and effect size observed in the experimental group, suggesting that consistent implementation of the STEAM approach contributed to improving students' conceptual understanding of LCM and GCF.

***Discussion***

Students who learned through STEAM assisted by MUSI Board achieved greater improvements in conceptual understanding of LCM and GCF than those receiving CTL with the same instructional medium. Higher N-Gain values and a very large effect size indicate that the observed improvement was educationally meaningful and statistically significant. Active engagement in interdisciplinary inquiry, collaborative problem solving, and knowledge construction likely contributed to deeper conceptual learning, highlighting the value of combining student-centered instruction with manipulative learning media (Almulla, 2023; Engelbrecht et al., 2024).

Constructivist Learning Theory plausibly explains these outcomes by emphasizing that conceptual understanding develops through active exploration and social interaction. STEAM encourages students to investigate authentic problems from multiple disciplinary perspectives, while MUSI Board turns abstract concepts of multiples and factors into concrete, manipulable visual representations. During STEAM activities, students use MUSI Board to explore mathematical relationships, test ideas, discuss alternative solutions, and collaboratively verify their reasoning (Evi Yupani & Widana, 2023). The interaction between the manipulative features of MUSI Board and the inquiry-based STEAM learning process enables students to construct deeper conceptual understanding than CTL using the same instructional medium.

Similar improvements have been reported in studies demonstrating the benefits of STEAM, manipulative media, and collaborative learning for mathematics achievement and conceptual understanding (Hidayat et al., 2024; Lee & Lai, 2024). A distinguishing feature of the present research is the direct comparison of STEAM and CTL while controlling the instructional medium through the same MUSI Board. This approach provides stronger evidence that differences in learning outcomes were primarily associated with the instructional approach rather than the learning medium (Rahmadi et al., 2024).

Findings from the present study support previous evidence that student-centered and manipulative learning environments enhance mathematical understanding. Anazifa & Djukri (2023) reported that problem-based learning promoted higher-order thinking through active knowledge construction, whereas Tokac et al. (2023) and Camacho-Sánchez et al. (2023); Camacho-Sánchez et al. (2023) showed that game-based and manipulative learning improved students' conceptual understanding and engagement. The present study observed similar learning gains. However, previous studies generally examined a single instructional approach or employed different learning media, making it difficult to distinguish instructional effects from media effects. Directly comparing STEAM and CTL using the same MUSI Board extends previous research by providing stronger evidence that differences in conceptual understanding were primarily associated with the instructional approach rather than variations in instructional media (de Vries et al., 2024; Lin et al., 2024; Suh & Park, 2023). (Alt, 2023; Bakar et al., 2023; Hui et al., 2023).

Several limitations should be acknowledged, including the relatively small sample, implementation in a single elementary school, and the focus on LCM and GCF within a short intervention period. Future studies should involve larger samples, multiple schools, longer interventions, and broader mathematics topics to examine the generalizability and long-term effectiveness of MUSI Board-assisted STEAM.

## Conclusion

Both the experimental and control groups showed significant improvements in cognitive understanding after learning with MUSI Board. However, students in the experimental group achieved greater conceptual gains than those in the control group, indicating that STEAM provided richer opportunities for inquiry, collaboration, and conceptual reasoning while using the same instructional medium. Beyond demonstrating STEAM's effectiveness, these findings contribute to constructivist and STEAM-based mathematics education by showing that differences in students' conceptual understanding can be more confidently attributed to the instructional approach when the learning medium is controlled. For classroom practice, integrating MUSI Board into STEAM-based mathematics instruction offers teachers a practical way to promote students' conceptual understanding through inquiry, collaboration, and hands-on learning. Given the relatively small sample, single-school context, and short intervention period, further research with larger, more diverse samples, different mathematics topics, educational levels, and technology-supported learning environments is recommended to examine the broader applicability and long-term effectiveness of this approach.

## Acknowledgements

The authors express sincere gratitude to the principal, teachers, and students of SD Negeri 3 Ketringan, Jiken District, Blora Regency, Indonesia, for their valuable cooperation and active participation throughout this study. Appreciation is also extended to the mathematics education experts, instructional media experts, and all individuals whose support and constructive contributions facilitated the successful completion of this research.

## Bibliography

Almulla, M. A. (2023). Constructivism learning theory: A paradigm for students' critical thinking, creativity, and problem solving to affect academic performance in higher education. *Cogent Education, 10*(1), Article 2172929. https://doi.org/10.1080/2331186X.2023.2172929

Alt, D. (2023). Assessing the benefits of gamification in mathematics for student gameful experience and gaming motivation. *Computers & Education, 200*, Article 104806. https://doi.org/10.1016/j.compedu.2023.104806

Anazifa, R. D., & Djukri. (2023). Systematic review of problem-based learning research in fostering critical thinking skills. *Thinking Skills and Creativity, 49*, Article 101334. https://doi.org/10.1016/j.tsc.2023.101334

Asilevi, M. N., Kärkkäinen, S., Sormunen, K., & Havu-Nuutinen, S. (2024). A comparison of science learning skills in the teacher-centered approach and inquiry-based science fieldwork: Primary school students' perceptions. *International Journal of Education in Mathematics, Science and Technology, 12*(1), 1–19. https://doi.org/10.46328/ijemst.3146

Bakar, A. P., D. N. N., Shahrill, M., Zakariya, Y., & F. (2023). Digital escape game and students' learning outcomes in mathematics: Experience from Brunei. *SAGE Open, 13*(4), Article 21582440231216838. https://doi.org/10.1177/21582440231216838
Brandl, L., Stadler, M., Richters, C., Radkowitsch, A., Fischer, M. R., Schmidmaier, R., & Fischer, F. (2024). Collaborative problem-solving in knowledge-rich domains: A multi-study structural equation model. *International Journal of Computer-Supported Collaborative Learning, 19*(3), 341–368. https://doi.org/10.1007/s11412-024-09425-4
Camacho-Sánchez, R., Serna-Bardavío, J., Rillo-Albert, A., & Lavega-Burgués, P. (2024). Enhancing motivation and academic performance through gamified digital game-based learning methodology using the ARCS model. *Interactive Learning Environments, 32*(10), 6868–6885. https://doi.org/10.1080/10494820.2023.2294762
Chen, F., & Chen, G. (2025). Technology-enhanced collaborative inquiry in K–12 classrooms: A systematic review of empirical studies. *Science & Education, 34*(3), 1731–1773. https://doi.org/10.1007/s11191-024-00538-8
Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). SAGE Publications.
Dabbous, M., Sakr, F., Safwan, J., Akel, M., Malaeb, D., Rahal, M., & Kawtharani, A. (2023). Instructional educational games in pharmacy experiential education: A quasi-experimental assessment of learning outcomes, students' engagement and motivation. *BMC Medical Education, 23*(1), Article 753. https://doi.org/10.1186/s12909-023-04742-y
de Vries, M. J., Rooney, P., & Fletcher, S. (2024). STEAM education: Recent developments, challenges, and future directions. *International Journal of Technology and Design Education, 34*(2), 503–523. https://doi.org/10.1007/s10798-023-09831-8
Denny, M., Denieffe, S., & O'Sullivan, K. (2023). Non-equivalent control group pretest–posttest design in social and behavioral research. In A. L. Nichols & J. Edlund (Eds.), *The Cambridge handbook of research methods and statistics for the social and behavioral sciences* (Vol. 1, pp. 314–332). Cambridge University Press. https://doi.org/10.1017/9781009010054.016
Engelbrecht, J., & Borba, M. C. (2024). Recent developments in using digital technology in mathematics education. *ZDM–Mathematics Education, 56*(2), 281–292. https://doi.org/10.1007/s11858-023-01530-2
Evi Yupani & Widana, I. W. (2023). The impacts of the STEM-based inquiry learning models on critical thinking and concept mastery. *Indonesian Research Journal in Education, 7*(1), 171-184. https://doi.org/10.22437/irje.v7i1.24227
Hidayat, R., Qi, Y., Ariffin, T. F. P. N. B., Hadzri, M. H. B., Chin, L. M., Ning, J. L. X., & Nasir, N. (2024). Online game-based learning in mathematics education among Generation Z: A systematic review. *International Electronic Journal of Mathematics Education, 19*(1), Article em0763. https://doi.org/10.29333/iejme/14024
Hui, H. B., Mahmud, & S, M. (2023). Influence of game-based learning in mathematics education on the students' cognitive and affective domain: A systematic review. *Frontiers in Psychology, 14*, Article 1105806. https://doi.org/10.3389/fpsyg.2023.1105806
Lee, W.-C., & Lai, C.-L. (2024). Facilitating mathematical argumentation by gamification: A gamified mobile collaborative learning approach for math courses. *International Journal of Science and Mathematics Education, 22*(1), 11–35. https://doi.org/10.1007/s10763-024-10462-6
Lin, P.-Y., Hwang, G.-J., & Tu, Y.-F. (2024). Effects of contextualized technology-supported learning on students' engagement and academic achievement: A meta-analysis. *Educational Research Review, 42*, Article 100597. https://doi.org/10.1016/j.edurev.2024.100597
Nadeem, M., Oroszlanyova, M., & Farag, W. (2023). Effect of digital game-based learning on student engagement and motivation. *Computers, 12*(9), Article 177. https://doi.org/10.3390/computers12090177
Rahmadi, I. F., Lavicza, Z., Arkün Kocadere, S., Houghton, T., & Spector, J. M. (2024).

Investigating the design, participation and experience of teaching and learning facilitated by user-generated microgames on an open educational platform. *Educational Technology Research and Development, 72*, 1787–1811. https://doi.org/10.1007/s11423-024-10359-9

Revelle, W., & Garner, K. M. (2024). Measurement: Reliability, construct validation, and scale construction. In H. T. Reis, T. West, & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 471–501). Cambridge University Press. https://doi.org/10.1017/9781009170123.021

Sailer, M., Murböck, J., & Fischer, F. (2024). Digital learning in schools: What does it take beyond digital technology? *Teaching and Teacher Education, 153*, Article 104788. https://doi.org/10.1016/j.tate.2024.104788

Sandi, G. (2021). Pengaruh pendekatan STEM untuk meningkatkan pemahaman konsep elektroplating, keterampilan berpikir kritis dan bekerja sama (The effect of the STEM approach on enhancing understanding of electroplating concepts, critical thinking skills, and collaboration skills). *Indonesian Journal of Educational Development*, *1*(4), 579-585. https://doi.org/10.5281/zenodo.4559843

Sormunen, K., Vehmaa, S., Seitamaa-Hakkarainen, P., Lavonen, J., Hakkarainen, K., & Juuti, K. (2023). Learning science through a collaborative invention project in primary school. *Disciplinary and Interdisciplinary Science Education Research, 5*(1), Article 6. https://doi.org/10.1186/s43031-023-00074-5

Suh, J., & Park, S. (2023). Student-centered mathematics instruction for enhancing conceptual understanding: A systematic review. *International Journal of Science and Mathematics Education, 21*(8), 2479–2501. https://doi.org/10.1007/s10763-023-10373-y

Suryaningsih, N. M. A., Poerwati, C. E., Lestari, P. I., & Parwata, M. Y. (2025). Early childhood literacy skills: Implementation of the local genius-based STEM learning model. *Indonesian Journal of Educational Development (IJED)*, *6*(1), 160–172. https://doi.org/10.59672/ijed.v6i1.4627

Tokac, U., Novak, E., & Thompson, C. G. (2023). Effects of game-based learning on students' mathematics achievement: A meta-analysis. *Journal of Computer Assisted Learning, 39*(1), 1–18. https://doi.org/10.1111/jcal.12763

Widana, I. W., Sopandi, A. T., Suwardika, I. G. (2021). Development of an authentic assessment model in mathematics learning: A science, technology, engineering, and mathematics (STEM) approach. *Indonesian Research Journal in Education, 5*(1), 192-209. https://doi.org/10.22437/irje.v5i1.12992

Widana, I. W. (2022). Meta-analysis: The relationship between self-regulated learning and mathematical critical reasoning. *Education.Innovation.Diversity, 1*(4), 64-75. https://doi.org/10.17770/eid2022.1.6739

Yang, Y.-F., Lee, I.-C., Tseng, C.-C., Lai, S.-C. (2025). Developing students' self-regulated learning strategies to facilitate vocabulary development in a digital game-based learning environment. *Journal of Research on Technology in Education, 57*(5), 972–991. https://doi.org/10.1080/15391523.2024.2338092