M A T H L I N E
JURNAL MATEMATIKA DAN PENDIDIKAN MATEMATIKA
ISSN 2502-5872 (Print)
ISSN 2622-3627 (Elektronik)

This is an open access article under the CC BY-SA license
Volume 11 Number 4, November 2026, 1159-1174

THE EFFECT OF THE GASING METHOD USING CONCRETE MEDIA ON ELEMENTARY SCHOOL STUDENTS' MULTIPLICATION LEARNING OUTCOMES IN NORTH SULAWESI

Jein Anggreini Sarlota Waghe[1], Seni Apriliya[2*], Karlimah[3]
[1,2,3]Elementary School Teacher Education Study Program, Graduate School, Universitas Pendidikan Indonesia, Tasikmalaya Campus, Tasikmalaya, West Java 46115, Indonesia

*Correspondence: seni_apriliya@upi.edu

Received 27 October 2026; Revised 15 November 2026; Accepted 24 November 2026

ABSTRACT

Mastery of multiplication remains a barrier for many Indonesian primary school students, including fourth graders. The GASING (*Gampang, Asyik, dan Menyenangkan*) method moves from concrete objects to symbols to mental calculation, but evidence of its effectiveness comes largely from uncontrolled designs. This quasi-experiment (*non-equivalent pretest–posttest control group* design) compared GASING using concrete media with conventional instruction in two intact classes of a numeracy training programme at BPMP North Sulawesi ($n = 25$ per class; two 60-minute sessions). Outcomes were measured with a 10-item researcher-made test and analysed with ANCOVA and robustness checks (robust standard errors, bootstrapping, rank-based ANCOVA, normalised gain). Pretests did not differ significantly ($p = 0.316$). After controlling for the pretest, the experimental class scored 39.08 points higher (95% CI [33.14, 45.02]; $F(1, 47) = 175.34$; $p < 0.001$; partial $\eta^2 = 0.79$); mean normalised gains were 0.82 versus 0.29, and 88% versus 4% of students reached the operational mastery criterion (KKTP) of 70. Within this sample and setting, the findings provide initial evidence that a concrete–abstract–mental calculation sequence can improve multiplication outcomes in the short term. Because each condition was a single class taught by a different instructor, the intervention was brief, and the instrument was not formally validated, the effect should be confirmed through multi-site replication with independent instructors, validated instruments, and delayed tests.

Keywords: GASING Method, Concrete Media, Multiplication, Numeracy, Quasi-Experiment.

How To Cite: Waghe, J. A. S., Apriliya, S., & Karlimah, K. (2026). The Effect of the GASING Method Using Concrete Media on Elementary School Students' Multiplication Learning Outcomes in North Sulawesi. *Mathline: Jurnal Matematika dan Pendidikan Matematika, 11*(4), 1159-1174. http://doi.org/10.31943/mathline.v11i4.1246

PRELIMINARY

Fluent and meaningful multiplication is a cornerstone of primary mathematics: students who recall single-digit products and understand multiplication as combining equal groups can devote their limited working memory to multi-digit multiplication, division, and fractions, and instruction should develop conceptual and procedural knowledge together (Hurrell, 2021).

The problem is already visible in the primary years: Indonesian fourth graders scored 397 in TIMSS, far below the scale centrepoint of 500 (Mullis et al., 2016); they misapply the concept of multiplication and make computational errors with tens (Safitri et al., 2025) or process errors even on facts such as 6 × 8 (Sari et al., 2025); and in one class only 2 of 29 students (6.90%) had mastered the facts up to 10 × 10 (Pathiah et al., 2026). These gaps persist: Indonesian 15-year-olds scored 366 in PISA 2022 mathematics (OECD, 2023; Wijaya et al., 2024), and pandemic learning losses did not recover quickly (Betthäuser et al., 2023). Children with weak early numeracy rely on counting with fingers and objects (Mahligai et al., 2026), so numeracy requires deliberate strategies (Rahmania et al., 2024) and adaptive teaching for low performers (Epran et al., 2025); mathematics anxiety also lowers achievement across 90 countries ($r = -0.26$; Brunner et al., 2026).

The GASING method (*Gampang, Asyik, dan Menyenangkan*; easy, fun, and enjoyable), developed by Yohanes Surya, moves each operation from concrete manipulation through symbols to mental calculation (*mencongak*) until students reach a "critical point", the minimal facts and strategies needed for the next topic (Gesiradja et al., 2026; Langi et al., 2025). For multiplication, students form equal groups with objects, name them with symbols, and practise oral recall using patterns for ×1, ×10, ×5, and ×9.

This logic is shared by the concrete–representational–abstract (CRA) sequence and concreteness fading (Chin et al., 2022), which support transfer when each representation is explicitly linked to the next (Fyfe & Nathan, 2019). CRA is evidence-based for students with learning disabilities (Bouck et al., 2018; Ebner et al., 2025) and has outperformed comparison conditions in fraction learning (Flores et al., 2023, 2025; Zhang et al., 2022), although the benefits of manipulatives depend on engagement, materials, and teacher preparation (Angco & Angco, 2024; Ochugboju & Díez Palomar, 2025).

Evidence on GASING for multiplication is positive but methodologically limited. Armianti et al. (2016) reported a normalised gain of 0.59 in a one-group design with 17 matriculation students. Comparison-group studies reported better outcomes with KOBESI media (Halyadi et al., 2016), among sixth graders (Aprijon, 2020), fourth graders (normalised gains 0.56 versus 0.26; Maulida et al., 2024), and seventh graders in Bitung, North Sulawesi (Rumajar et al., 2024), and in Ambon (0.77 versus 0.35; Adam et al., 2026); uncontrolled studies agree (Anggraini et al., 2026; Bagus et al., 2026), although Puspitasari et al. (2026) found only a small, non-significant increase. These studies rarely controlled for prior ability, reported effect sizes with confidence intervals, or tested robustness. These gaps matter: when intact classes are compared, a pre-existing difference

in prior ability can masquerade as a treatment effect unless the pretest is controlled (Clifton & Clifton, 2019); without effect sizes and confidence intervals, readers cannot judge practical magnitude or precision (Kraft, 2020); and without robustness checks, conclusions from small samples with ceiling effects and unequal variances may rest on violated assumptions.

The novelty of this study lies in combining three elements. Substantively, it tests GASING as an explicit concrete–abstract–mental calculation sequence with a formative return to concrete media, whereas earlier comparison-group studies combined GASING with KOBESI media (Halyadi et al., 2016), TGT cooperative learning (Maulida et al., 2024), or dice games (Adam et al., 2026). Methodologically, it adds ANCOVA controlling for prior ability, effect sizes with 95% confidence intervals, and robustness checks, whereas previous studies relied mainly on posttest comparisons or normalised gains (Armianti et al., 2016; Maulida et al., 2024). Contextually, it provides evidence from a numeracy training programme in North Sulawesi, where the only previous comparison-group study involved seventh graders (Rumajar et al., 2024). Concrete media as such are not claimed as new.

This study therefore examined the effect of GASING using concrete media on fourth graders' multiplication learning outcomes compared with conventional instruction in a numeracy training programme in North Sulawesi. The research questions (RQ) were as follows.

RQ1. Do students taught with GASING using concrete media improve significantly from pretest to posttest, and by how much?

RQ2. After controlling for prior ability, do these students outperform students receiving conventional instruction?

RQ3. Is the difference consistent across alternative indicators (gain, normalised gain, KKTP attainment) and across analyses that relax normality and homoscedasticity assumptions?

METHODS

Research Design

A quasi-experiment with a non-equivalent pretest–posttest control group design was used: Class B received GASING instruction using concrete media (experimental), and Class C received conventional instruction (control). Because random assignment was not possible, baseline differences were tested explicitly and the pretest served as a covariate.

Setting and Participants

The study was conducted at the Education Quality Assurance Agency (*Balai Penjaminan Mutu Pendidikan*, BPMP) of North Sulawesi Province in a numeracy programme held outside regular lessons, with fourth graders from three schools in Manado ($N = 50$). Students were assigned by registration number (1–25 to Class B, experimental; 26–50 to Class C, control), irrespective of ability or school of origin. Because registration order is not random and may relate to motivation, parental support, or distance from the venue, selection bias cannot be excluded; the non-significant pretest difference and the pretest covariate reduce, but do not remove, this threat. The school composition of each class was not analysed, so school effects cannot be separated from the treatment effect. Both classes studied the same multiplication content from the Grade 4 national curriculum within the same schedule. The study was approved by BPMP North Sulawesi, and student names were coded before analysis. With 25 students per group, a two-tailed test at $\alpha = 0.05$ had 80% power to detect $d = 0.81$, adequate for large but not for small or moderate effects.

Treatment

Experimental class. The experimental class followed the GASING sequence (Table 1; Figure 1): students formed equal groups with coloured picture cards and boxes and described them orally (e.g., “three groups of four”), wrote the same situation as $3 \times 4 = 12$ and linked it to repeated addition, and then recalled products orally using the ×1, ×10, ×5, and ×9 patterns and commutativity. Students who could not yet recall a fact returned to the concrete media, so the sequence functioned as a formative cycle; each session ended with a short game-based exercise.

Control class. The teacher explained the multiplication table, students completed textbook and worksheet exercises individually, and errors were addressed through re-explanation. The number and duration of sessions were identical (2 × 60 minutes within one week), and concrete media were not used systematically.

Instructors and implementation consistency. The experimental class was taught by the first author, a GASING trainer who has delivered GASING training in about 12 regencies and cities in Indonesia, and the control class by a regular classroom teacher unfamiliar with GASING. Both classes covered the same facts and two-digit items with the same number and length of sessions, week, venue, and class size, and took the same test; they differed intentionally in media, representation, practice sequence, and response to errors (Table 1). Interaction intensity could not be equalised because GASING requires

more oral and game-based interaction. Implementation was documented with attendance lists; no independent observer or fidelity checklist was used.

Table 1. Comparison of Instruction in the Experimental and Control Classes

Aspect	Experimental class (GASING with concrete media)	Control class (conventional)
Starting point	Forming equal groups with concrete objects	Teacher's explanation of the multiplication table
Representation	Concrete → symbolic → mental calculation	Symbolic
Media	Coloured picture cards and boxes	Textbooks and student worksheets
Fact practice	Pattern-based oral recall (×1, ×10, ×5, ×9, commutativity) in game form	Individual written exercises and memorisation
Response to errors	Return to concrete media (formative cycle)	Re-explanation by the teacher
Instructor	First author (GASING trainer)	Classroom teacher (not trained in GASING)
Session 1 (60 min)	Concrete and abstract phases for ×1, ×2, ×5, and ×10 facts: equal groups with cards and boxes, oral naming, number sentences; game	Multiplication as repeated addition; the multiplication table; textbook exercises
Session 2 (60 min)	×9 pattern, commutativity, remaining facts, and two-digit items; oral recall with return to concrete media; game	Memorising the table; worksheet exercises including two-digit items; re-explanation
Sessions / duration	2 × 60 minutes within 1 week	2 × 60 minutes within 1 week

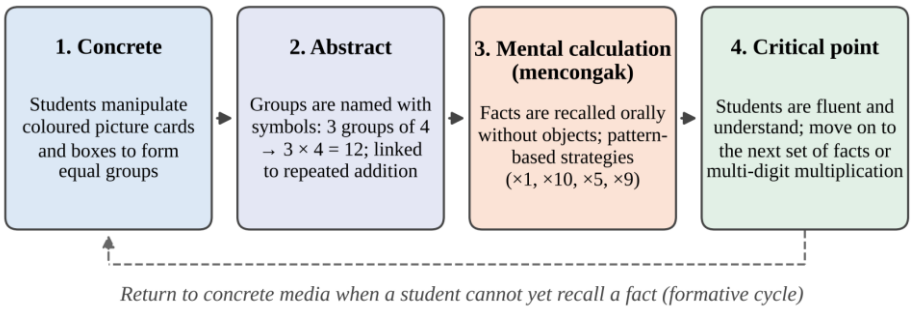

Figure 1. Sequence of multiplication learning with the GASING method using concrete media as implemented in the experimental class.

Instrument and Procedure

Outcomes were measured with a researcher-made test of 10 one- and two-digit multiplication items (10 points each; range 0–100), used as both pretest and posttest and written from a blueprint (Table 2) based on the Phase B (Grades 3–4) curriculum outcomes for multiplication of whole numbers. The instrument had not undergone formal expert validation or a reliability trial, and item-level data were not retained, so internal consistency, item difficulty, and discrimination could not be estimated. Identical items may produce a testing effect in both classes that favours the experimental class if its practice rehearsed the tested facts more closely; scores are therefore interpreted as performance on the content taught, not as general multiplication ability. The pretest preceded the first

session and the posttest followed completion of the material; answers were scored against an answer key by a scorer blind to class membership.

Table 2. Blueprint of the Multiplication Test

Indicator (Phase B: multiplication of whole numbers)	Content covered
Determine products using the ×1 and ×10 patterns	Basic facts with a factor of 1 or 10
Determine products using the ×5 and ×9 patterns	Basic facts with a factor of 5 or 9
Determine other basic facts using known facts and commutativity	Basic facts with both factors from 2 to 8
Multiply numbers involving a two-digit factor	Products with a two-digit factor
Total	**10 items, 10 points each (maximum 100)**

Note. Indicators were derived from the Phase B learning outcomes; each item was scored 0 or 10 against an answer key.

Data Analysis

Analyses were conducted in Python 3 with SciPy 1.18.1 and pandas 3.0.5 (two-tailed $\alpha = 0.05$). Normality, homogeneity of variance, and outliers were checked with Shapiro–Wilk tests, Levene tests, and the 1.5 × IQR rule. Baseline differences were tested with Welch's t, Mann–Whitney U, Hedges' g, and a two one-sided tests (TOST) equivalence procedure with bounds of ±0.5 pooled SD (Lakens, 2017). Within-group change (RQ1) was tested with paired t and Wilcoxon signed-rank tests, with Cohen's d_z, Cohen's d_{av}, and the rank-biserial correlation as effect sizes.

The between-group effect (RQ2) was estimated with ANCOVA (posttest as outcome, group as factor, pretest as covariate), which is more powerful than gain-score comparisons and corrects for chance baseline imbalance (Clifton & Clifton, 2019). Robustness (RQ3) was examined with HC3 robust standard errors, percentile bootstrap CIs (10,000 stratified resamples), Quade's rank ANCOVA, the group × time interaction in a 2 × 2 mixed ANOVA, and comparisons of gain scores and normalised gains, g = (posttest − pretest)/(100 − pretest), classified as high (≥ 0.70), medium (0.30–0.69), or low (< 0.30). Because KKTP is set by each school, the value used here is a study-specific benchmark rather than a universal standard: a score of 70, the KKTP applied by the participating schools for this content, served as the operational criterion, with 75 as a sensitivity check; attainment was compared with exact McNemar and Fisher tests.

RESULT AND DISCUSSION

Results

Descriptive Statistics and Data Screening

Table 3 presents the descriptive statistics. Pretest means were low in both classes (experimental 20.40, control 24.80); six experimental (24%) and three control students

(12%) scored zero, although whole-number multiplication belongs to earlier grades. This low baseline may reflect the participation of students with weak numeracy or previous instruction that relied mainly on rote memorisation; it matches the grade-4 difficulties reported in Indonesian classrooms (Pathiah et al., 2026; Safitri et al., 2025) and left considerable room for improvement. At posttest, the experimental mean was 84.80 and the control mean 47.60; nine experimental students (36%) scored 100, whereas the highest control score was 70. All experimental students improved (by 30–100 points); in the control class, 23 improved and two were unchanged. The 1.5 × IQR rule flagged only plausible values, so all data were retained.

Table 3. Descriptive Statistics of Multiplication Learning Outcomes by Group

Group / measure	*M*	*SD*	95% CI	*Mdn*	Min–Max	Skewness	% score 0	% score 100
*Experimental (*n = *25)*								
Pretest	20.40	14.85	[14.27, 26.53]	20	0–50	0.01	24	0
Posttest	84.80	14.47	[78.83, 90.77]	90	60–100	−0.41	0	36
Gain	64.40	15.30	[58.09, 70.71]	70	30–100	−0.37	0	4
Normalised gain	0.82	0.17	[0.75, 0.89]	0.86	0.43–1.00	−0.54	0	—
*Control (*n = *25)*								
Pretest	24.80	15.84	[18.26, 31.34]	20	0–70	0.62	12	0
Posttest	47.60	9.26	[43.78, 51.42]	50	30–70	0.52	0	0
Gain	22.80	11.37	[18.11, 27.49]	20	0–50	0.32	8	0
Normalised gain	0.29	0.12	[0.24, 0.34]	0.29	0.00–0.50	−1.00	8	—

Note. Score range 0–100. Gain = posttest − pretest. Normalised gain = (posttest − pretest)/(100 − pretest). % score 0 and % score 100 = percentage of students with the minimum or maximum score; for gain scores, they denote no change and a gain of 100 points, respectively.

Posttest scores were non-normal in both classes (Shapiro–Wilk p = 0.003 and 0.009), as expected for a test scored in multiples of 10 with a ceiling, and variances were unequal for the posttest (Levene F = 5.88, p = 0.019), normalised gains (F = 7.42, p = 0.009), and ANCOVA residuals (F = 9.67, p = 0.003), although the residuals were normally distributed (W = 0.979, p = 0.500). Welch tests, HC3 errors, and nonparametric alternatives were therefore used alongside parametric analyses.

Baseline Differences

The pretest difference was not significant: −4.40 points (95% bootstrap CI [−13.20, 3.60]), Welch $t(47.80) = -1.01$, p = 0.316; Mann–Whitney U = 269.0, p = 0.393; Hedges' g = −0.28, 95% CI [−0.84, 0.28]. However, TOST with bounds of ±7.68 points did not confirm equivalence (90% CI [−11.69, 2.89]; p_{TOST} = 0.227), so pretest scores were controlled through ANCOVA. The control class scored slightly higher at baseline, so any imbalance worked against the hypothesis.

RQ1: Pretest-to-Posttest Change within Each Class

Both classes improved significantly (Table 3). The experimental mean gain was 64.40 points (95% CI [58.09, 70.71]), $t(24) = 21.05$, $p < 0.001$, $d_z = 4.21$, $d_{av} = 4.39$; Wilcoxon $W = 0$, $p < 0.001$, rank-biserial $r = 1.00$. The control mean gain was 22.80 points (95% CI [18.11, 27.49]), $t(24) = 10.02$, $p < 0.001$, $d_z = 2.00$, $d_{av} = 1.76$; Wilcoxon $W = 0$, $p < 0.001$, $r = 1.00$ for the 23 students whose scores changed. Pretest–posttest correlations were $r = 0.46$ ($p = 0.022$) and $r = 0.71$ ($p < 0.001$), respectively.

RQ2: Effect of GASING after Controlling for Pretest Scores

Regression slopes were nearly identical (experimental $b = 0.44$; control $b = 0.41$; group × pretest interaction $F(1, 46) = 0.03$, $p = 0.875$). As Table 4 shows, the pretest was a significant covariate, $F(1, 47) = 19.41$, $p < 0.001$, partial $\eta^2 = 0.29$, and after controlling for it, the group effect was large and significant, $F(1, 47) = 175.34$, $p < 0.001$, partial $\eta^2 = 0.79$. The adjusted means were 85.74 (experimental) and 46.66 (control), a difference of 39.08 points (95% CI [33.14, 45.02]); the model explained 79.4% of the posttest variance.

Table 4. Analysis of Covariance of Posttest Scores with the Pretest as Covariate

Source	*SS*	*df*	*MS*	*F*	*p*	Partial η^2
Pretest (covariate)	2,069.40	1	2,069.40	19.41	< 0.001	0.29
Group	18,692.22	1	18,692.22	175.34	< 0.001	0.79
Error	5,010.60	47	106.61			
Adjusted means at pretest = 22.60 (M; SE or 95% CI)						
Experimental	85.74		*SE* = 2.08			
Control	46.66		*SE* = 2.08			
Difference	39.08		95% CI [33.14, 45.02]			

Note. *SS* = sum of squares; *df* = degrees of freedom; *MS* = mean square. $R^2 = 0.794$. In the adjusted-means rows, the SS column gives the adjusted mean and the MS column its SE or 95% CI. Pooled pretest coefficient $b = 0.43$.

RQ3: Robustness of the Between-Group Effect

The group effect was stable across all robustness checks (Table 5): the HC3 robust test, $t(47) = 12.81$ (SE = 3.05; 95% CI [32.94, 45.22]), the bootstrap CI [33.28, 44.76], and Quade's rank ANCOVA, $F(1, 48) = 168.84$, all gave $p < 0.001$. The 2 × 2 mixed ANOVA showed a group × time interaction, $F(1, 48) = 119.08$, $p < 0.001$, partial $\eta^2 = 0.71$, confirming a larger improvement in the experimental class.

Table 5. Robustness of the Between-Group Effect across Analyses and Outcome Indicators

Analysis / indicator	Estimate (E − C)	95% CI	Test statistic	*p*	Effect size
Controlling for pretest					
ANCOVA, conventional *SE*	39.08	[33.14, 45.02]	$F(1, 47) = 175.34$	< 0.001	$\eta_p^2 = 0.79$
ANCOVA, HC3 robust *SE*	39.08	[32.94, 45.22]	$t(47) = 12.81$	< 0.001	—
ANCOVA, bootstrap (10,000)	39.08	[33.28, 44.76]	—	—	—

Analysis / indicator	Estimate (E − C)	95% CI	Test statistic	*p*	Effect size
Quade rank ANCOVA	—	—	$F(1, 48) = 168.84$	< 0.001	—
Mixed ANOVA, group × time	—	—	$F(1, 48) = 119.08$	< 0.001	$\eta_p^2 = 0.71$
Unadjusted indicators (Welch tests)					
Pretest (baseline)	−4.40	[−13.20, 3.60][a]	$t(47.80) = -1.01$	0.316	$g = -0.28$ [−0.84, 0.28]
Posttest	37.20	[30.40, 43.60][a]	$t(40.83) = 10.83$	< 0.001	$g = 3.02$ [2.21, 3.83]
Gain score	41.60	[34.00, 48.80][a]	$t(44.32) = 10.91$	< 0.001	$g = 3.04$ [2.22, 3.85]
Normalised gain	0.53	[0.45, 0.61][a]	$t(41.54) = 12.64$	< 0.001	$g = 3.52$ [2.63, 4.40]

Note. E = experimental; C = control; g = Hedges' g with 95% CI. [a]Percentile bootstrap CI for the mean difference (10,000 resamples). Mann–Whitney rank-biserial correlations were −0.14 (pretest), 0.96 (posttest), 0.96 (gain), and 0.99 (normalised gain). HC3 = heteroscedasticity-consistent standard error (SE = 3.05); its CI is estimate ± $t(47)$ × SE. Mann–Whitney $U = 269.0$ (pretest; $p = 0.393$), 614.0 (posttest), 612.5 (gain), and 622.5 (normalised gain; all other $p < 0.001$).

Alternative indicators led to the same conclusion: the experimental class had a 41.60-point larger mean gain (Hedges' $g = 3.04$), a higher mean normalised gain (0.82, high, versus 0.29, low; $g = 3.52$), and rank-biserial correlations of 0.96–0.99.

Gain categories differed between the classes, $\chi^2(2, N = 50) = 38.27$, $p < 0.001$: 21 experimental students (84%) achieved high and 4 (16%) medium gains, whereas no control student achieved a high gain (11 medium, 14 low). KKTP attainment in the experimental class rose from 0 to 22 students (88%; exact McNemar $p < 0.001$), whereas only one control student (4%) reached it at both time points (posttest Fisher's exact $p < 0.001$); at the stricter threshold of 75, 18 experimental students (72%) and no control students reached the criterion.

Discussion

After controlling for pretest scores, the experimental class outperformed the control class by about 39 points on a 100-point scale, consistently across analyses and indicators; every experimental student gained at least 30 points, whereas two control students did not improve.

Possible Mechanisms

Manipulating equal groups gives the symbol $a \times b$ a referent: in three boxes of four cards, the factors take distinct roles (number of groups and group size), the total can be checked by counting, and rearranging the boxes shows that $3 \times 4 = 4 \times 3$. Writing each arrangement as a number sentence bridges concrete and symbolic knowledge (Chin et al., 2022; Fyfe & Nathan, 2019), and oral practice builds automatic recall, which is most productive when combined with conceptual understanding (Hurrell, 2021).

The patterns also reduce the learning burden: with commutativity and the ×1, ×10, ×5, and ×9 patterns, the 100 facts of the 10 × 10 table reduce to 21 facts (products of 2, 3,

4, 6, 7, and 8) that must be recalled separately. In cognitive load terms, rules that organise many facts into a few schemas reduce the elements that novices must hold in working memory, and manipulating objects is less demanding than following a verbal explanation of an abstract table (Sweller et al., 2019).

Which stage contributed most cannot be determined, because the stages were delivered as one package and only total scores were available. However, experimental students who started at zero reached 60–100, none had a low gain, and the pretest–posttest correlation was lower in the experimental class (0.46 versus 0.71), suggesting that the concrete phase and formative return to concrete media mainly helped students with little initial understanding, while students who knew some facts gained mainly from pattern-based mental calculation, in line with the benefits of concrete-to-abstract sequences for low achievers (Bouck et al., 2018; Ebner et al., 2025); this hypothesis requires stage-level data.

Games and oral recall add retrieval practice with immediate feedback in a low-threat format that may reduce anxiety (Angco & Angco, 2024; Brunner et al., 2026), and GASING involves more teacher–student interaction than individual worksheet practice. These components belong to GASING as designed, but affective variables and interaction were not measured, and the contribution of concrete media cannot be isolated.

Comparison with Previous GASING Studies

These findings agree with earlier comparison-group studies of GASING (Adam et al., 2026; Aprijon, 2020; Halyadi et al., 2016; Maulida et al., 2024; Rumajar et al., 2024): the experimental normalised gain (0.82) resembles those of Adam et al. (2026) (0.77) and Gesiradja et al. (2026) (0.74) and exceeds those of Armianti et al. (2016) and Maulida et al. (2024), whose control gains (0.26 and 0.35 with Adam et al.) are close to ours (0.29). The non-significant result of Puspitasari et al. (2026) among students with relatively high pretests is consistent with GASING benefiting low-ability students most, a hypothesis that requires direct testing.

Instructor and Selection Effects

The most serious alternative explanation is the instructor. The first author, who designed the lessons, taught the experimental class, so instructor expertise, enthusiasm, and expectations are fully confounded with the treatment, and evaluations by those who develop or deliver an intervention tend to report larger effects than independent evaluations (Wolf et al., 2020). Identical content and time, a blind-scored objective test, and a substantial control-class gain (22.80 points) reduce, but cannot remove, this concern; likewise, registration-order assignment may have produced differences in unmeasured

characteristics such as motivation. The estimate is therefore best read as the effect of GASING delivered by an experienced GASING trainer, not by regular teachers.

Magnitude of the Effect

The standardised effects were very large (Hedges' g = 3.0–3.5; partial η^2 = 0.79) relative to the benchmarks of Kraft (2020), under which 0.20 *SD* is already large. Small samples, quasi-experimental designs, and researcher-made tests inflate effect sizes (Cheung & Slavin, 2016), as do close test–treatment alignment, the small control *SD* (9.26), and coarse scoring. The large gain after only two sessions also reflects the very low baseline, a narrow test of directly practised facts, and an immediate posttest, so it represents short-term mastery of the taught content rather than durable competence. Conversely, the ceiling effect (36% scored 100) probably underestimated the raw difference, so the raw difference (about 39 points) and KKTP attainment (88% versus 4%) are the more interpretable results.

Limitations and Implications

With one class per condition, class, school-composition, and instructor differences cannot be ruled out, clustering makes the reported standard errors too small, and registration-order assignment is open to selection bias. Testing, novelty, and instructor expectancy effects and greater interaction intensity in the experimental class threaten internal validity, so the estimates should be treated as upper bounds. The instrument was not validated and its reliability and item properties are unknown; error in the pretest weakens the covariate adjustment, and the results describe performance on these 10 items rather than multiplication ability in general. The intervention was brief with an immediate posttest, fidelity and affective outcomes were not measured, and the sample came from one training programme in one province.

The findings are therefore promising initial evidence rather than a basis for policy. For practice, they suggest beginning multiplication with meaning rather than memorisation: teachers can (1) provide concrete media for every student to form equal groups, (2) have students write and say their manipulations as number sentences, (3) teach the ×1, ×10, ×5, and ×9 patterns and commutativity explicitly, (4) conduct short, frequent mental calculation practice, for example as games, and (5) use it to decide who returns to concrete media, supported by training in managing transitions between phases (Ochugboju & Díez Palomar, 2025). Future research should use multi-site designs with several classes per condition, random assignment, independent instructors, multilevel models, validated

parallel-form instruments, fidelity observations, delayed tests, and comparisons of GASING with and without the concrete phase.

CONCLUSION

Within the limits of its sample and implementation, this quasi-experiment provides initial evidence that GASING instruction using concrete media was associated with substantially greater short-term gains in fourth graders' multiplication learning outcomes than conventional instruction in a numeracy training programme in North Sulawesi: after controlling for the pretest, the experimental class scored about 39 points higher, and 88% versus 4% of students reached the operational KKTP, consistently across robust, bootstrap, and nonparametric analyses. The findings should not be generalised to all primary school students. Given the single class per condition, the instructor confound, registration-order assignment, the brief intervention, unvalidated instrument, identical tests, and immediate posttest, the effect should be confirmed by multi-site studies with independent instructors, validated parallel-form instruments, and delayed assessments.

DECLARATIONS

Author Contribution : Jein Anggreini Sarlota Waghe: Conceptualization, Investigation, Formal Analysis, and Writing – Original Draft.
Seni Apriliya: Supervision and Writing – Review & Editing.
Karlimah: Supervision and Writing – Review & Editing.

Funding Statement : No funding. This research was self-funded by the authors.

Conflict of Interest : The authors declare no conflict of interest.

Additional Information : The study was conducted with permission from BPMP North Sulawesi Province, and student identities were anonymised before analysis. The anonymised data and analysis scripts are available from the corresponding author upon reasonable request.

REFERENCES

Adam, D., Muslihin, H. Y., & Apriani, I. F. (2026). Effectiveness of the GASING method based on dice games in improving multiplication skills of primary school students. *Indonesian Journal of Primary Education, 10*(1), 167–186. https://doi.org/10.17509/ijpe.v10i1.98405

Angco, R. J., & Angco, L. B. (2024). A meta-synthesis of effective practices and outcomes in the use of manipulatives for teaching mathematics. *Human Behavior, Development and Society, 25*(2), 50–59. https://doi.org/10.62370/hbds.v25i2.274848

Anggraini, N. M., Ayu, H. D., & Tiningrum, A. (2026). Peningkatan keterampilan aritmatika dasar melalui metode GASING interaktif menggunakan kartu angka pada siswa kelas 1 SD [Improving basic arithmetic skills through the interactive GASING method using number cards among first-grade elementary students]. *Pendas: Jurnal Ilmiah Pendidikan Dasar, 11*(2), 276–285. https://doi.org/10.23969/jp.v11i2.52105

Aprijon, A. (2020). Pelatihan matematika Gasing pada materi penjumlahan dan perkalian dua digit dengan dua digit untuk siswa kelas VI sekolah dasar negeri [GASING mathematics training on two-digit by two-digit addition and multiplication for sixth-grade public elementary school students]. *Menara Riau, 14*(1), 45–55. https://doi.org/10.24014/menara.v14i1.12520

Armianti, A., Yani, I., Widuri, K., & Sulistiawati, S. (2016). Pengaruh matematika GASING (Gampang, ASyIk, dan menyenaNGkan) pada materi perkalian bilangan bulat terhadap hasil belajar peserta matrikulasi STKIP Surya [The effect of GASING mathematics on integer multiplication on the learning outcomes of STKIP Surya matriculation students]. *Kreano, Jurnal Matematika Kreatif-Inovatif, 7*(1), 74–81. https://doi.org/10.15294/kreano.v7i1.5012

Bagus, M. B., Yuniarti, Y., & Kurniawan, D. T. (2026). Implementation of in-depth learning with the GASING method based on mathematical games for elementary school student. *International Journal of Education, Vocational and Social Science, 5*(4), 1162–1176. https://doi.org/10.63922/ijevss.v5i04.5951

Betthäuser, B. A., Bach-Mortensen, A. M., & Engzell, P. (2023). A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. *Nature Human Behaviour, 7*(3), 375–385. https://doi.org/10.1038/s41562-022-01506-4

Bouck, E. C., Satsangi, R., & Park, J. (2018). The concrete–representational–abstract approach for students with learning disabilities: An evidence-based practice synthesis. *Remedial and Special Education, 39*(4), 211–228. https://doi.org/10.1177/0741932517721712

Brunner, M., Preckel, F., Götz, T., Lüdtke, O., & Keller, L. (2026). High math anxiety is associated with lower math achievement across 90 countries: An individual participant data meta-analysis of representative student and adult samples. *Psychological Bulletin, 152*(2), 207–253. https://doi.org/10.1037/bul0000514

Cheung, A. C. K., & Slavin, R. E. (2016). How methodological features affect effect sizes in education. *Educational Researcher, 45*(5), 283–292. https://doi.org/10.3102/0013189X16656615

Chin, H., Chew, C. M., & Suseelan, M. (2022). Mathematics learning from concrete to abstract (1968–2021): A bibliometric analysis. *Participatory Educational Research, 9*(4), 445–468. https://doi.org/10.17275/per.22.99.9.4

Clifton, L., & Clifton, D. A. (2019). The correlation between baseline score and post-intervention score, and its implications for statistical analysis. *Trials, 20*, Article 43. https://doi.org/10.1186/s13063-018-3108-3

Ebner, S., MacDonald, M. K., Grekov, P., & Aspiranti, K. B. (2025). A meta-analytic review of the concrete-representational-abstract math approach. *Learning Disabilities Research & Practice, 40*(1), 31–42. https://doi.org/10.1177/09388982241292299

Epran, E., Kusumah, Y. S., & Herman, T. (2025). Analysis of mathematics teacher's pedagogical competency and students's numeracy in solving numeracy problems.

Mathline: Jurnal Matematika dan Pendidikan Matematika, 10(1), 179–192. https://doi.org/10.31943/mathline.v10i1.796

Flores, M. M., Hinton, V. M., Sanchez, V. M., Kohnke, S., Lee, J., Podemski, M. O., & Burdette, M. G. (2025). Teaching rational number concepts to fifth graders using the concrete-representational-abstract-integrated sequence. *The Journal of Educational Research, 118*(3), 268–280. https://doi.org/10.1080/00220671.2025.2469852

Flores, M. M., Hinton, V. M., Shadoan, L., & Monroe, C. (2023). Teaching rational number concepts and skills to students receiving tier two interventions. *The Journal of Educational Research, 116*(4), 206–215. https://doi.org/10.1080/00220671.2023.2248046

Fyfe, E. R., & Nathan, M. J. (2019). Making "concreteness fading" more concrete as a theory of instruction for promoting transfer. *Educational Review, 71*(4), 403–422. https://doi.org/10.1080/00131911.2018.1424116

Gesiradja, M. C., Sudiapermana, E., Mamonto, F., Lutfiansyah, D. Y., Shantini, Y., & Sebastian, E. J. V. (2026). Effectiveness of GASING method in improving elementary students' basic arithmetic ability: West Java, Indonesia and Philippines. *Jurnal Math-UMB.EDU, 13*(3), 382–395. https://doi.org/10.36085/math-umb.edu.v13i3.10360

Halyadi, H., Agustianie, D., Handayani, T., & Windria, H. (2016). Penggunaan KOBESI dalam matematika GASING untuk meningkatkan pemahaman materi perkalian siswa SD [Using KOBESI in GASING mathematics to improve elementary students' understanding of multiplication]. *Suska Journal of Mathematics Education, 2*(2), 81–88. https://doi.org/10.24014/sjme.v2i2.2149

Hurrell, D. (2021). Conceptual knowledge OR procedural knowledge or conceptual knowledge AND procedural knowledge: Why the conjunction is important to teachers. *Australian Journal of Teacher Education, 46*(2), 57–71. https://doi.org/10.14221/ajte.2021v46n2.4

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher, 49*(4), 241–253. https://doi.org/10.3102/0013189X20912798

Lakens, D. (2017). Equivalence tests: A practical primer for *t* tests, correlations, and meta-analyses. *Social Psychological and Personality Science, 8*(4), 355–362. https://doi.org/10.1177/1948550617697177

Langi, E. L., Kondolayuk, M. L., & Domassang, K. D. (2025). Construction of integer operations of the Gasing-mathematics method from the perspective of APOS theory. *Qalamuna: Jurnal Pendidikan, Sosial, dan Agama, 17*(2), 1551–1566. https://doi.org/10.37680/qalamuna.v17i2.8110

Mahligai, F. H., Huda, N., & Siburian, J. (2026). Early numeracy skills as predictors of numeracy assessment performance: A mixed-methods study using the Early Grade Mathematics Assessment framework. *Kreano, Jurnal Matematika Kreatif-Inovatif, 17*(1), 124–142. https://doi.org/10.15294/kreano.v17i1.34517

Maulida, S. A., Isrokatun, I., & Julia, J. (2024). Pengaruh pembelajaran kooperatif TGT dengan metode GASING terhadap pemahaman konsep matematis perkalian dua angka di SD [The effect of TGT cooperative learning with the GASING method on elementary students' conceptual understanding of two-digit multiplication]. *SJME (Supremum Journal of Mathematics Education), 8*(1), 101–113. https://doi.org/10.35706/sjme.v8i1.10789

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS 2015 international results in mathematics*. TIMSS & PIRLS International Study Center, Boston College. https://timssandpirls.bc.edu/timss2015/international-results/

Ochugboju, A. O., & Díez Palomar, F. J. (2025). Systematic review on the use of virtual and physical manipulatives by primary school teachers. *Social Sciences & Humanities Open, 12*, Article 102188. https://doi.org/10.1016/j.ssaho.2025.102188

OECD. (2023). *PISA 2022 results (Volume I): The state of learning and equity in education*. OECD Publishing. https://doi.org/10.1787/53f23881-en

Pathiah, Makki, M., & Wahyuningsih, B. Y. (2026). Pengembangan media PEKAN (Perkalian Menyenangkan) terhadap pemahaman konsep siswa kelas IV SDN 45 Mataram [Development of PEKAN (fun multiplication) media for fourth-grade students' conceptual understanding at SDN 45 Mataram]. *Pendas: Jurnal Ilmiah Pendidikan Dasar, 11*(2), 204–216. https://journal.unpas.ac.id/index.php/pendas/article/view/49751

Puspitasari, C. D., Nayif, M. A. S., Hidayati, P. A., Anggraeni, S. E. N., Azizah, N. A. N., & Setyawan, A. (2026). Upaya meningkatkan hasil belajar matematika melalui penerapan metode Gasing dalam perkalian dua digit pada siswa kelas V SDN Bilaporah 1 [Improving mathematics learning outcomes through the GASING method in two-digit multiplication among fifth-grade students at SDN Bilaporah 1]. *Pendas: Jurnal Ilmiah Pendidikan Dasar, 11*(2), 500–514. https://doi.org/10.23969/jp.v11i2.47607

Rahmania, U. G., Safitri, R. R., Putri, A. F., Nurohman, S., & Salehudin, A. (2024). Systematic literature review: How important are literacy and numeracy for students, and how to improve it? *Indonesian Journal of Educational Research and Review, 7*(2), 416–429. https://doi.org/10.23887/ijerr.v7i2.79797

Rumajar, J. R., Pulukadang, R. J., & Manurung, O. (2024). Pembelajaran Gasing pada materi perkalian bilangan bulat: Pengaruh & dampaknya terhadap hasil belajar siswa [GASING learning on integer multiplication: Its effect and impact on student learning outcomes]. *Jurnal Riset dan Inovasi Pembelajaran, 4*(3), 2127–2139. https://doi.org/10.51574/jrip.v4i3.2356

Safitri, K., Disurya, R., & Tanzimah. (2025). Analisis kesulitan belajar matematika operasi hitung perkalian siswa kelas IV SD Negeri 4 Pulau Rimau [Analysis of fourth-grade students' learning difficulties in multiplication at SD Negeri 4 Pulau Rimau]. *Pendas: Jurnal Ilmiah Pendidikan Dasar, 10*(3), 498–508. https://journal.unpas.ac.id/index.php/pendas/article/view/31721

Sari, D. P., Riswari, L. A., & Santoso, D. A. (2025). Analysis of students' errors in solving mathematics story problems on multiplication and division. *Jurnal Kiprah Pendidikan, 4*(4), 672–683. https://doi.org/10.33578/kpd.v4i4.p672-683

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Wijaya, T. T., Hidayat, W., Hermita, N., Alim, J. A., & Talib, C. A. (2024). Exploring contributing factors to PISA 2022 mathematics achievement: Insights from Indonesian teachers. *Infinity Journal, 13*(1), 139–156. https://doi.org/10.22460/infinity.v13i1.p139-156

Wolf, R., Morrison, J., Inns, A., Slavin, R., & Risman, K. (2020). Average effect sizes in developer-commissioned and independent evaluations. *Journal of Research on Educational Effectiveness, 13*(2), 428–447. https://doi.org/10.1080/19345747.2020.1726537

Zhang, S., Yu, S., Xiao, J., Liu, Y., & Jiang, T. (2022). The effects of concrete-representational-abstract sequence instruction on fractions for Chinese elementary students with mathematics learning disabilities. *International Journal of Science and*

Mathematics Education, 20(7), 1481–1498. https://doi.org/10.1007/s10763-021-10215-9