Journal of Advanced Sciences and Mathematics Education
Volume 6, Issue 4, 240-249
e_ISSN: 2798-1606
DOI: 10.58524/jasme.v6i4.1691

GASING-based deep learning and grade IV students' multiplication concept mastery: A quasi-experimental comparison

Ana Kaprianie*
Universitas Pendidikan Indonesia, INDONESIA

Yunus Abidin
Universitas Pendidikan Indonesia, INDONESIA

Dinie Anggraeni Dewi
Universitas Pendidikan Indonesia, INDONESIA

Article Info

Article history:
Received: Aug 25, 2026
Revised: Oct 10, 2026
Accepted: Nov 02, 2026

Keywords:
Conceptual Understanding; Elementary Mathematics; GASING Method; Guided Inquiry; Multiplication; Quasi-Experiment.

Abstract

Background: Multiplication concept mastery requires connections among equal grouping, symbolic notation, and flexible calculation. The GASING method offers a three-stage sequence that may complement the mindful, meaningful, and joyful dimensions of Deep Learning in the Indonesian curriculum.
Aim: This study compared reported learning outcomes in GASING-based Deep Learning (GASING-DL) and inquiry-based Deep Learning (Inquiry-DL) for Grade IV students.
Methods: The source manuscript describes a four-week, nonequivalent pretest–posttest study at SDN 4 Kuala Kurun in January 2025 involving two intact classes reported as 32 students each. Multiplication concept mastery was assessed using eight constructed-response items scored from 0 to 4 (maximum 32). The original analysis plan included pretest-adjusted ANCOVA and gain-based comparisons.
Results: Reported descriptive scores rose from 10.81 to 25.30 in GASING-DL and from 12.28 to 19.39 in Inquiry-DL. The reported mean normalized gains were 0.683 and 0.353, respectively. Inferential outputs in the source record were not interpreted because their degrees of freedom and category counts conflict with the stated group sizes.
Conclusion: The available descriptive summaries favor GASING-DL within this classroom comparison, but the magnitude and statistical certainty of the between-group difference cannot be established reliably from the inconsistent inferential record. Replication and consistent participant-level analysis are needed before strong effectiveness claims can be made.

To cite this article: Kaprianie, A., Abidin, Y., & Dewi, D. A. (2026). GASING-based deep learning and grade IV students' multiplication concept mastery: A quasi-experimental comparison. *Journal of Advanced Sciences and Mathematics Education, 6*(4), 240-249.

INTRODUCTION

Multiplication is a gateway from counting-based arithmetic to relational and algebraic thinking, yet successful recall of products does not necessarily indicate an understanding of multiplicative structure. In Grade IV, learners are expected to interpret equal groups, arrays, operation properties, and contextual problems rather than reproduce multiplication tables alone (National Council of Teachers of Mathematics, 2000; Kemendikbud, 2022). Research on children's multiplicative reasoning has documented the difficulty of coordinating units and relationships across representations (Zwanch & Wilkins, 2021; Schulz, 2024). Studies of multiplication-table learning also distinguish memory interference from the construction of mathematical meaning (Dotan & Zviran-Ginat, 2022). These distinctions make conceptual mastery, rather than speed of calculation by itself, an appropriate focus for evaluating primary mathematics instruction (Hurrell, 2021; Hankeln & Prediger, 2025).

The pedagogical challenge is particularly relevant where students enter lessons with limited conceptual preparation and teachers must work within existing classroom resources. At SDN 4 Kuala Kurun, the source study describes two Grade IV classes whose mean pretest scores were far below the instrument maximum of 32. This local observation warrants attention, although it should not be generalized to all rural schools or to students sharing a particular cultural background. Broader research indicates that early numerical knowledge predicts later achievement but does not identify a single universally decisive skill (Ruiz et al., 2024). Similarly, language-sensitive materials and

* **Corresponding author:**
Ana Kaprianie, Universitas Pendidikan Indonesia, Indonesia
anakaprianiespd71@upi.edu

contextual interpretations matter when students must explain structures rather than merely transform symbols (Lenz et al., 2024; Purnomo et al., 2022).

The Merdeka Curriculum places conceptual understanding and meaningful engagement alongside the development of mathematical fluency (Kemendikbud, 2022; Kemendikbud, 2024). In this article, Deep Learning refers to the educational emphasis on mindful, meaningful, and joyful activity, not to artificial-intelligence algorithms. A useful instructional sequence should help students articulate what a multiplication expression represents before applying it to less familiar problems (Fullan & Langworthy, 2014; Hattie & Donoghue, 2016). Educational research on graphical representations and students' reasoning suggests that visible models can support, but do not automatically guarantee, conceptual connections (Sterner, 2025; Valenta et al., 2024). The instructional question is therefore how to organize representations, explanation, and fluency within lessons accessible to the students in this school.

GASING (Gampang, Asyik, Menyenangkan, meaning Easy, Enjoyable, Fun) offers a sequence of concrete activity, abstract reasoning, and mencongak or mental calculation (Surya, 2018). Its initial stage invites learners to organize tangible objects into equal groups, followed by pictorial and symbolic explanations of the operation. The final stage develops fact retrieval through structured and engaging practice, but retrieval remains distinct from evidence of conceptual understanding. This distinction is consistent with research on concrete–representational–abstract mathematics instruction and the cognitive processes involved in learning multiplication facts (Ebner et al., 2025; Dotan & Zviran-Ginat, 2022). Other studies show that even younger learners can explore equal grouping in mathematically meaningful ways, while context selection influences the meaning attributed to multiplication (Levenson et al., 2024; Roan et al., 2025).

Evidence from mathematics education suggests that neither unguided exploration nor direct explanation should be treated as a universal instructional solution. A meta-analysis found benefits of guided play for some early mathematics outcomes, while its authors also noted variation across intervention types and measured domains (Skene et al., 2022). More recent syntheses identify beneficial effects from both intentional instruction and exploratory play, emphasizing the importance of alignment between pedagogy and target skills (Sun et al., 2026). Research on teacher learning in multiplicative thinking and work on primary mathematical reasoning likewise emphasizes the quality of tasks, scaffolds, and instructional enactment (Downton et al., 2023; Monteleone et al., 2023). An inquiry-based comparison is therefore meaningful only when the learning activities and omitted supports are described precisely rather than presenting inquiry as intrinsically deficient.

The literature also cautions against treating higher posttest scores as sufficient proof of an instructional mechanism. Reviews of mathematical word-problem interventions have found that implementation fidelity and research quality are associated with reported effects (Vessonen et al., 2025). Research on teacher feedback, language-responsive instruction, and representation construction demonstrates that instructional support can take different forms and may serve different students or tasks (Huang et al., 2024; Ponte et al., 2023; Tytler et al., 2023). Studies addressing mathematical reasoning, pattern recognition, and the relationship between curriculum and classroom mathematics further show why students' explanations are an important complement to summary achievement scores (Alderton & Pratt, 2025; Reuter & Reinhold, 2026; Višňovská & Cortina, 2025). Existing studies illuminate the broader components of structured mathematics instruction, but they do not by themselves establish the comparative performance of GASING-DL and the specific inquiry condition used at SDN 4 Kuala Kurun.

Against this background, the present study compared multiplication concept mastery in two intact Grade IV classes receiving GASING-DL and Inquiry-DL instruction. It examined reported baseline scores, descriptive pretest–posttest changes, and normalized gains, with an intended pretest-adjusted group comparison. The instructional contribution lies in specifying how concrete objects, representational bridges, and mental calculation were sequenced within a locally situated Deep Learning lesson. The empirical contribution is necessarily bounded by the nonequivalent-class design and the consistency of the underlying participant-level record. Framing the findings within these boundaries enables a meaningful comparison without claiming that a single school study establishes a general or uniquely GASING-specific learning mechanism.

METHOD

Research Design and Setting

The study used a quantitative nonequivalent-groups pretest–posttest quasi-experimental design because students were already assigned to intact classes before the intervention (Shadish et al., 2002). It was carried out in January 2025 at SDN 4 Kuala Kurun, Gunung Mas Regency, Central Kalimantan, Indonesia. The source manuscript identifies the experimental condition as GASING-based Deep Learning (GASING-DL) and the comparison condition as inquiry-based Deep Learning (Inquiry-DL). The study concerned one school and two existing classes, so class membership and instructional condition were inseparable; this feature constrains causal interpretation even when baseline scores are modeled statistically.

Participants and Selection

The manuscript states that two Grade IV classes participated, with 32 students per class and a total of 64 students. Eligibility required active enrollment, attendance at the pretest, instructional sessions and posttest, and written parental or guardian consent. The sample was drawn from existing classes rather than by individual random allocation. The study records the school as serving a resource-limited setting; wider demographic or regional representativeness cannot be inferred from the study as described. Numerical summaries and inferential test outputs retained in the source contain different implicit sample sizes; the inferential consequences are documented separately in the accompanying editorial audit.

Assessment Instrument

The outcome instrument was an eight-item constructed-response test of multiplication concept mastery aligned with the Phase B mathematics learning outcomes (Kemendikbud, 2022). Four indicators were addressed by two items each: interpretation, analysis, inference, and evaluation. Each item was scored from 0 to 4, giving a possible total score from 0 to 32. The source reports content review by two mathematics education lecturers with Aiken's V of at least 0.75 for each item, followed by a trial with 28 students in another school and Cronbach's alpha of 0.82 (Aiken, 1985). These reported quality indicators describe the original validation process; item-level ratings, scoring examples, and inter-rater agreement were not supplied.

The assessment blueprint is presented in Table 1 to show how each competency was operationalized and associated with the instructional sequence.

Table 1. Blueprint of the multiplication concept mastery assessment

Indicator	Target competency	Items	Max. score	Related instructional stage
Interpretation	Equal groups, repeated addition, arrays and contextual notation	Q1–Q2	8	Concrete
Analysis	Multiplication properties, strategies and patterns	Q3–Q4	8	Concrete–Abstract
Inference	Multiplicative reasoning in unfamiliar contexts	Q5–Q6	8	Abstract
Evaluation	Error detection, justification and strategy efficiency	Q7–Q8	8	Abstract–Mencongak
Total	Four conceptual indicators	8 items	32	Three-stage sequence

Table 1 shows that each indicator was assessed with two constructed-response items. Every indicator contributed a maximum of eight points. Interpretation focused on equal-group meanings and their representations. Analysis and inference placed greater weight on symbolic relationships and their use. Evaluation concerned the justification and assessment of proposed multiplication solutions.

Instructional Procedure

The two classes received eight 70-minute lessons over four weeks with the same classroom teacher. The teacher reportedly completed a two-day orientation and used standardized lesson plans with equivalent topic coverage and instructional time. In GASING-DL, the concrete phase lasted approximately 20 minutes and used seeds, marbles, or number cards to represent equal groups; the abstract phase lasted approximately 25 minutes and connected arrays and number lines to symbolic multiplication; and the mencongak phase lasted approximately 25 minutes and used rhythmic

practice, paired quizzes, and quick-response games. In Inquiry-DL, students explored prompted multiplication situations, discussed strategies, presented findings, and reflected on solutions, without the specified concrete–abstract–mencongak sequence. The researcher observed two lessons per group using a checklist, but numerical implementation-fidelity results were not supplied.

Figure 1 summarizes the implemented sequence and measurement points using only procedures described in the original manuscript.

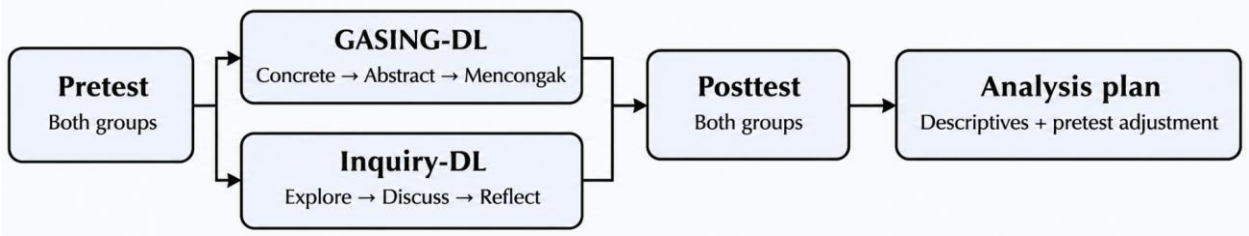


Figure 1. Instructional conditions and measurement sequence reported in the study

Figure 1 presents pretest and posttest measurement as shared features of both classes. The GASING-DL branch specifies three consecutive instructional phases. The Inquiry-DL branch retains classroom exploration, discussion, and reflection. Both groups were reported to receive the same total instructional time. The diagram represents the documented procedure and does not establish whether every lesson was implemented with identical fidelity.

Data Analysis and Research Ethics

The original statistical plan combined descriptive summaries, independent- and paired-samples t-tests, Shapiro–Wilk and Levene checks, normalized gain, ANCOVA with pretest as covariate, and a Mann–Whitney U sensitivity comparison (Hake, 1998; Field, 2018; Virtanen et al., 2020). Normalized gain was defined as (posttest − pretest)/(32 − pretest), with the formula applied at the student level before calculating the group mean. Because no participant-level dataset accompanied the article and the supplied summary output includes incompatible denominators, this revision retains the descriptive numbers as reported but does not treat the original inferential statistics as validated evidence. The source reports permission from the principal, informed parental consent, voluntary participation, and anonymized student data. No ethics committee approval identifier was supplied, so none has been created.

RESULTS AND DISCUSSION

Table 2 consolidates the available descriptive statistics from the manuscript for the two Grade IV classes and separates observed scores from statistical inference.

Table 2. Reported pretest, posttest, and normalized-gain summaries

Measure	Inquiry-DL, control	GASING-DL, experimental
Reported class size	32 students	32 students
Pretest, M (SD)	12.28 (3.29)	10.81 (2.95)
Pretest range	6–18	6–16
Posttest, M (SD)	19.39 (2.98)	25.30 (3.29)
Posttest range	13–26	17–30
Absolute difference in reported means	7.11 points	14.49 points
Reported mean normalized gain (SD)	0.353 (0.147)	0.683 (0.144)

Table 2 indicates that the Inquiry-DL class had the higher mean pretest score. The GASING-DL class had the higher mean posttest score. Both groups showed positive changes between their reported mean scores. The reported normalized gain was higher for GASING-DL than for Inquiry-DL. These descriptive differences do not, by themselves, establish a valid pretest-adjusted treatment effect. Figure 2 visualizes the reported mean scores without reconstructing individual student observations or plotting unprovided confidence intervals.

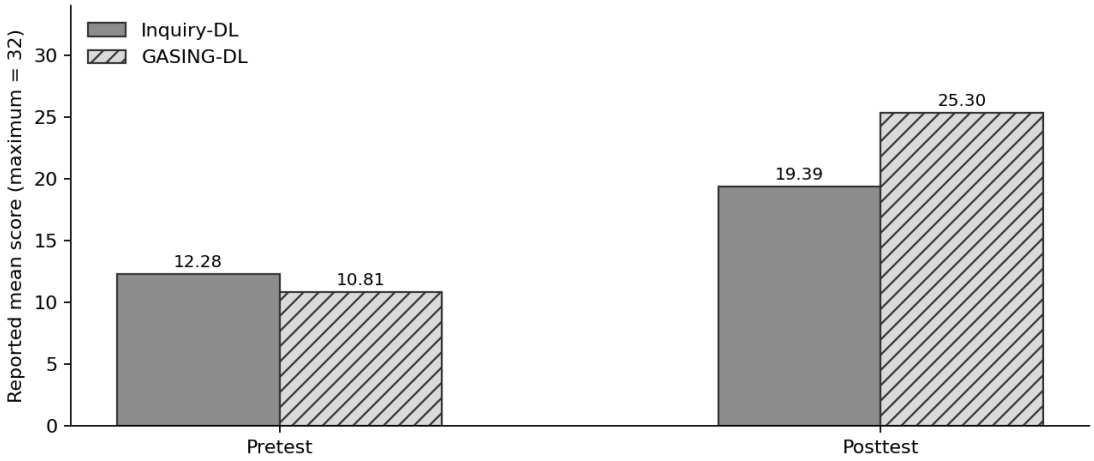


Figure 2. Reported mean pretest and posttest scores by instructional condition

Figure 2 makes the different starting points visible. Inquiry-DL began with a higher mean score than GASING-DL. Both groups had higher mean scores after instruction. GASING-DL had the higher posttest mean in the descriptive record. Bar heights represent reported group means only and should not be interpreted as estimates of statistical precision.

The originally supplied ANCOVA and other inferential outputs are not reproduced as confirmatory findings here. In the source record, the stated total of 64 participants conflicts with ANCOVA residual degrees of freedom of 125 and category frequencies summing to 64 within each group; the pretest t-test also uses test values incompatible with the stated group sizes. Accordingly, neither adjusted group means nor claims about the statistical significance of the treatment contrast can be accepted from these summaries without a coherent reanalysis of the actual observations. This is an evidentiary boundary, not a claim that the intervention failed or that its true effect was zero.

Discussion

The descriptive pattern favored GASING-DL at posttest despite its lower reported baseline mean, suggesting that the instructional sequence merits further investigation in this setting. The magnitude of the apparent difference is educationally interesting because the test targeted conceptual interpretation, analysis, inference, and evaluation rather than only the rapid recall of number facts. However, this pattern is not enough to establish an adjusted treatment effect when the inferential record contains incompatible sample denominators. Research on mathematics interventions has shown that fidelity and research design can influence apparent effect sizes (Vessonen et al., 2025; Fishstrom et al., 2022). The evidence therefore supports a cautious interpretation of observed scores, not a definitive causal claim about GASING-DL.

One plausible explanation for the observed pattern is the presence of explicit representational bridges in the GASING-DL lessons. The concrete phase made equal groups visible before students moved to arrays, symbols, and contextual problems, giving learners more than one form in which to express the same structure. Hankeln and Prediger (2025) document the significance of meaning-related explanations in translating among representations of multiplication, while Ebner et al. (2025) synthesize research on the concrete–representational–abstract approach. The GASING-DL sequence is consistent with these ideas, although the study did not measure each representational transition or compare separate instructional components. The interpretation is consequently a mechanism hypothesis grounded in the intervention description rather than an empirically isolated process.

The shift from concrete models to symbolic strategies may also have strengthened students' opportunities to reason about the properties of multiplication. Interpretations involving equal groups and arrays are not equivalent to simply choosing an operation from a verbal cue, a distinction emphasized in work on multiplicative situations (Roan et al., 2025). Studies of multiplicative thinking and children's equal-group constructions similarly show why the units and relationships being counted need explicit attention (Downton et al., 2023; Levenson et al., 2024). Bruno et al. (2024) illustrate how conceptual models may support the solution of multiplicative problems, albeit in a different population and study design. The present instrument's indicator structure is compatible

with those arguments, but without item-level scores it remains unknown which conceptual processes contributed most to the reported posttest difference.

Mencongak may have played a different pedagogical role by developing access to multiplication facts through structured, enjoyable repetition. Research on multiplication-table acquisition identifies interference among similar arithmetic facts as a genuine learning challenge (Dotan & Zviran-Ginat, 2022). Yet faster retrieval cannot be assumed to imply increased conceptual mastery, and this study did not independently assess fluency, enjoyment, or mathematical anxiety. Reviews of mathematics anxiety and classroom interventions emphasize that cognitive achievement and affective responses are related but distinct outcomes (Barroso et al., 2021; Robson et al., 2023; Donolato et al., 2026). Claims that mencongak freed working-memory resources or produced joyful learning therefore remain theoretically plausible but untested within the observed data.

The comparison with inquiry-based Deep Learning also requires a more precise interpretation than a simple contrast between guided and unguided instruction. Inquiry-DL in this study included prompting questions, small-group discussion, explanation, and reflection, so it was not an absence of instructional support. Research on collaborative reasoning in Grade IV and broader mathematical reasoning indicates that inquiry itself can be carefully scaffolded (Zeng et al., 2026; Valenta et al., 2024). Guided-play synthesis also suggests that exploration and structured teacher guidance can be complementary rather than competing approaches (Skene et al., 2022; Sun et al., 2026). Hence the contrast in this study pertains to two particular lesson packages, not to a general finding that inquiry is inferior to GASING.

The context of an existing school, one teacher, and two intact classes is central to assessing the transferability of the descriptive outcomes. A single teacher may reduce variation in instructional delivery but may also introduce enthusiasm, sequencing, or cross-condition contamination that cannot be separated from the method. The source notes that standardized lesson plans and limited classroom observations were used, although no numerical fidelity data were supplied. Studies on mathematics feedback, language-responsive instructional materials, and representation construction illustrate how teacher enactment and learner characteristics can modify instructional experiences (Huang et al., 2024; Lenz et al., 2024; Tytler et al., 2023). Future designs should record treatment differentiation and fidelity systematically alongside the outcome measures.

This study also invites consideration of what the assessment is capable of supporting. Its four indicators plausibly cover a richer domain than fact recall, but a total-score comparison cannot show whether students improved most in interpretation, inference, or evaluation. Investigations of early mathematical patterns, reasoning, and culturally meaningful problem settings show that achievement can change with task design and the opportunities students receive to explain their thinking (Reuter & Reinhold, 2026; Monteleone et al., 2023; Purnomo et al., 2022). Additional research highlights why curricular coherence and student perspectives should guide the selection of what counts as meaningful success (Višňovská & Cortina, 2025; Källberg & Roos, 2025). Item-level results, inter-rater scoring evidence, and delayed assessments would help distinguish conceptual development from familiarity with the test format.

The most important contribution of the study at its current stage is a specified, context-sensitive instructional comparison rather than a validated estimate of causal effectiveness. The three phases of GASING-DL are documented clearly enough to inform more rigorous classroom replication, while the inquiry condition provides a concrete comparator rather than an abstract pedagogical label. International scholarship on mathematical reasoning, word problems, and domain-specific skill development supports the value of studying both conceptual and procedural outcomes in authentic settings (English, 2023; Ruiz et al., 2024; Alderton & Pratt, 2025). At the same time, the unresolved statistical discrepancies and class-level assignment preclude the strongly causal interpretation originally attached to the reported ANCOVA and gain statistics. A transparent participant-level reanalysis, ideally followed by a multi-class replication and retention measurement, would provide a stronger foundation for claims about GASING-DL and comparable forms of instructional scaffolding. A broader methodological perspective similarly connects conceptual and procedural knowledge with classroom mindset and mathematical participation (Bui et al., 2023; Nordby et al., 2024; Maj-Tatsis & Tatsis, 2026; Ufer et al., 2024; Tang & Wong, 2026; Alibali et al., 2024).

Implications and Limitations

The instructional sequence offers a practical design for teachers to consider: begin with tangible equal groups, develop representational and symbolic explanations, and use short structured practice for multiplication facts. This is a pedagogical implication rather than a validated prescription for increasing test scores in every school. The original study was limited to two intact classes in one school, a single teacher, a four-week intervention, and the absence of item-level and follow-up data. In addition, the summary statistics in the supplied manuscript cannot support confirmatory inference until the actual sample size and test outputs are reconciled. Future research should examine multiple teachers and classes, measure the fidelity of each instructional stage, and include delayed conceptual as well as fluency assessments.

CONCLUSION

The reported descriptive means showed a larger pretest–posttest change in the GASING-DL class than in the Inquiry-DL class during four weeks of Grade IV multiplication instruction. The instructional contribution of the study is the clear sequencing of concrete representations, abstract reasoning, and mencongak within a mindful, meaningful, and joyful lesson framework. Nevertheless, the source manuscript contains mutually incompatible participant counts and inferential statistics, preventing a reliable statement about the statistical significance or adjusted magnitude of the difference. The observed pattern should therefore be regarded as preliminary evidence within two existing classrooms rather than definitive proof of method effectiveness. Reanalysis of participant-level observations and independent replication are necessary before stronger generalizable conclusions are drawn.

ACKNOWLEDGMENT

The authors thank the school principal, teachers, and students who participated in the study. The original manuscript reports no external funding.

AUTHOR CONTRIBUTIONS STATEMENT

Ana Kaprianie: Conceptualization, investigation, data collection, data curation, formal analysis, and writing – original draft preparation. Yunus Abidin: Conceptualization, methodology, supervision, validation, and writing – review and editing. Dinie Anggraeni Dewi: Methodology, supervision, validation, interpretation of findings, and writing – review and editing.

REFERENCES

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement, 45*(1), 131–142.

Alderton, J., & Pratt, N. (2025). The construction of mathematical reasoning as a pedagogical object. *Educational Studies in Mathematics, 120*, 299–315. https://doi.org/10.1007/s10649-025-10420-1

Alibali, M. W., Matthews, P. G., Rodrigues, J., Meng, R., Vest, N. A., Jay, V., Menendez, D., Murray, J. O., Donovan, A. M., Anthony, L. E., & McNeil, N. M. (2024). A bird's-eye view of research practices in mathematical cognition, learning, and instruction: Reimagining the status quo. *Journal of Experimental Child Psychology, 248*, 106056. https://doi.org/10.1016/j.jecp.2024.106056

Baroody, A. J. (2006). Why children have difficulties mastering the basic number combinations and how to help them. *Teaching Children Mathematics, 13*(1), 22–31.

Barroso, C., Ganley, C. M., McGraw, A. L., Geer, E. A., Hart, S. A., & Daucourt, M. C. (2021). A meta-analysis of the relation between math anxiety and math achievement. *Psychological Bulletin, 147*(2), 134–168. https://doi.org/10.1037/bul0000307

Bruno, A., Polo-Blanco, I., Van Vaerenbergh, S., Fernández-Cobos, R., & González-López, M. J. (2024). Strategies for solving multiplicative problems using a conceptual model-based problem-solving approach: A case study with a student with autism spectrum disorder. *ZDM–Mathematics Education, 56*, 1239–1256. https://doi.org/10.1007/s11858-024-01568-w

Bui, P., Pongsakdi, N., McMullen, J., Lehtinen, E., & Hannula-Sormunen, M. M. (2023). A systematic review of mindset interventions in mathematics classrooms: What works and what does not?. *Educational Research Review, 40*, 100554. https://doi.org/10.1016/j.edurev.2023.100554

Donolato, E., Caviola, S., Masi, A., Lervåg, A., & Melby-Lervåg, M. (2026). Attitudes and emotions in the context of mathematical learning: A systematic review and meta-analysis of intervention studies. *Educational Research Review, 52*, 100815. https://doi.org/10.1016/j.edurev.2026.100815

Dotan, D., & Zviran-Ginat, S. (2022). Elementary math in elementary school: The effect of interference on learning the multiplication table. *Cognitive Research: Principles and Implications, 7*, 101. https://doi.org/10.1186/s41235-022-00451-0

Downton, A., Giumelli, K., McHugh, B., & Stenning, P. (2023). Impact of teachers' professional learning on students' learning of multiplicative thinking. *Mathematics Education Research Journal, 35*, 659–687. https://doi.org/10.1007/s13394-021-00408-4

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

English, L. D. (2023). Ways of thinking in STEM-based problem solving. *ZDM–Mathematics Education, 55*, 1219–1230. https://doi.org/10.1007/s11858-023-01474-7

Field, A. (2018). *Discovering statistics using IBM SPSS statistics (5th ed.).* SAGE Publications.

Fishstrom, S., Wang, H.-H., Bhat, B. H., Daniel, J., Dille, J., Capin, P., & Vaughn, S. (2022). A meta-analysis of the effects of academic interventions on academic achievement and academic anxiety outcomes in elementary school children. *Journal of School Psychology, 92*, 265–284. https://doi.org/10.1016/j.jsp.2022.03.011

Fullan, M., & Langworthy, M. (2014). *A rich seam: How new pedagogies find deep learning.* Pearson.

Hake, R. R. (1998). Interactive-engagement versus traditional methods. *American Journal of Physics, 66*(1), 64–74. https://doi.org/10.1119/1.18809

Hankeln, C., & Prediger, S. (2025). Language is essential for avoiding surface translations: Associations of students' spontaneous use of meaning-related phrases for explicating structures with conceptual understanding of multiplication. *Educational Studies in Mathematics, 120*, 57–79. https://doi.org/10.1007/s10649-025-10414-z

Hattie, J., & Donoghue, G. M. (2016). Learning strategies: A synthesis and conceptual model. *npj Science of Learning, 1*, 16013. https://doi.org/10.1038/npjscilearn.2016.13

Huang, J., Cai, Y., Lv, Z., Huang, Y., & Zheng, X.-L. (2024). Toward self-regulated learning: Effects of different types of data-driven feedback on pupils' mathematics word problem-solving performance. *Frontiers in Psychology, 15*, 1356852. https://doi.org/10.3389/fpsyg.2024.1356852

Hurrell, D. P. (2021). Conceptual knowledge OR procedural knowledge OR conceptual knowledge AND procedural knowledge: Why the conjunction is important for teachers. *Australian Journal of Teacher Education, 46*(2), 57–71. https://eric.ed.gov/?id=EJ1296887

Källberg, P. S., & Roos, H. (2025). Meaning(s) of a student perspective in mathematics education research. *Educational Studies in Mathematics, 119*, 367–392. https://doi.org/10.1007/s10649-024-10374-w

Kemendikbud. (2022). *Capaian pembelajaran mata pelajaran matematika pada kurikulum merdeka.* Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia.

Kemendikbud. (2024). *Panduan implementasi pendekatan deep learning dalam pembelajaran kurikulum merdeka.* Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia.

Lenz, K., Obersteiner, A., & Wittmann, G. (2024). Who benefits most from language-responsive learning materials in mathematics? Investigating differential effects in heterogeneous classrooms. *Educational Studies in Mathematics, 116*, 185–211. https://doi.org/10.1007/s10649-024-10321-9

Levenson, E. S., Barkai, R., Mahamid, A., & Levy, S. (2024). Creating equal groups from n items: Mathematical creativity in kindergarten. *Educational Studies in Mathematics, 116*, 145–163. https://doi.org/10.1007/s10649-024-10322-8

Maj-Tatsis, B., & Tatsis, K. (2026). Exploring the use of a colouring task to elicit young students' mathematical reasoning. *Educational Studies in Mathematics, 122*, 291–309. https://doi.org/10.1007/s10649-025-10459-0

Monteleone, C., Miller, J., & Warren, E. (2023). Conceptualising critical mathematical thinking in young students. *Mathematics Education Research Journal, 35*, 339–359. https://doi.org/10.1007/s13394-023-00445-1

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.*

Nordby, S. K., Mifsud, L., & Bjerke, A. H. (2024). Computational thinking in primary mathematics classroom activities. *Frontiers in Education, 9*, 1414081. https://doi.org/10.3389/feduc.2024.1414081

Ponte, R., Viseu, F., Neto, T. B., & Aires, A. P. (2023). Revisiting manipulatives in the learning of geometric figures. *Frontiers in Education, 8*, 1217680. https://doi.org/10.3389/feduc.2023.1217680

Purnomo, Y. W., Pasri, Aziz, T. A., Shahrill, M., & Prananto, I. W. (2022). Students' failure to understand fraction multiplication as part of a quantity. *Journal on Mathematics Education, 13*(4), 681–702. https://eric.ed.gov/?id=EJ1379854

Reuter, D., & Reinhold, F. (2026). Investigating the complexity of pattern recognition in primary students. *Educational Studies in Mathematics, 121*, 229–260. https://doi.org/10.1007/s10649-025-10464-3

Roan, E., Kularajan, S., & Czocher, J. A. (2025). Reconsidering task classification systems for situations that can be modelled with multiplication. *ZDM–Mathematics Education, 57*, 411–424. https://doi.org/10.1007/s11858-025-01657-4

Robson, D. A., Johnstone, S. J., Putwain, D. W., & Howard, S. (2023). Test anxiety in primary school children: A 20-year systematic review and meta-analysis. *Journal of School Psychology, 98*, 39–60. https://doi.org/10.1016/j.jsp.2023.02.003

Ruiz, C., Kohnen, S., von Hagen, A., Kwok, F. Y., & Bull, R. (2024). Which domain-specific skills at the beginning of formal schooling predict later mathematical achievement? A systematic review and meta-analysis. *Educational Research Review, 42*, 100583. https://doi.org/10.1016/j.edurev.2023.100583

Schulz, A. (2024). How do multiplicative and relational thinking contribute to multidigit division?. *Journal für Mathematik-Didaktik, 45*, Article 21. https://doi.org/10.1007/s13138-024-00247-3

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Houghton Mifflin.

Skene, K., O'Farrelly, C. M., Byrne, E. M., Kirby, N., Stevens, E. C., & Ramchandani, P. G. (2022). Can guidance during play enhance children's learning and development in educational contexts? A systematic review and meta-analysis. *Child Development, 93*(4), 1162–1180. https://doi.org/10.1111/cdev.13730

Sterner, H. (2025). Using graphical representations to develop students' correspondence relationships and covariational thinking in pattern generalizations in primary school. *International Journal of Science and Mathematics Education, 23*, 1573–1597. https://doi.org/10.1007/s10763-024-10520-z

Sun, J., Xie, S., Li, J.-B., Li, J., & Mu, F. (2026). Early mathematics interventions in pre-primary education: A three-level meta-analysis of intentional teaching and exploratory play approaches. *Educational Research Review, 52*, 100830. https://doi.org/10.1016/j.edurev.2026.100830

Surya, Y. (2018). *Matematika GASING: Gampang, asyik, menyenangkan.* Yohanes Surya Foundation.

Tang, E., & Wong, T. T.-Y. (2026). Potential underlying mechanisms of spatial-mathematical associations among primary school children. *Journal of Experimental Child Psychology, 261*, 106369. https://doi.org/10.1016/j.jecp.2025.106369

Tytler, R., Prain, V., Kirk, M., Mulligan, J., Nielsen, C., Speldewinde, C., White, P., & Xu, L. (2023). Characterising a representation construction pedagogy for integrating science and mathematics in the primary school. *International Journal of Science and Mathematics Education, 21*, 1153–1175. https://doi.org/10.1007/s10763-022-10284-4

Ufer, S., Kaiser, A., Niklas, F., & Gabler, L. (2024). I have three more than you, you have three less than me? Levels of flexibility in dealing with additive situations. *Frontiers in Education, 9*, 1340322. https://doi.org/10.3389/feduc.2024.1340322

Valenta, A., Rø, K., & Klock, S. I. (2024). A framework for reasoning in school mathematics: Analyzing the development of mathematical claims. *Educational Studies in Mathematics, 116*, 91–111. https://doi.org/10.1007/s10649-024-10309-5

Vessonen, T., Hellstrand, H., Kurkela, M., Aunio, P., & Laine, A. (2025). The effectiveness of mathematical word problem-solving interventions among elementary schoolers: A systematic review and meta-analysis. *International Journal of Educational Research, 132*, 102642. https://doi.org/10.1016/j.ijer.2025.102642

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., & SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods, 17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2

Višňovská, J., & Cortina, J. L. (2025). Curriculum and mathematical coherence: Exploring the tensions in teaching and learning number. *ZDM–Mathematics Education, 57*, 1051–1063. https://doi.org/10.1007/s11858-025-01706-y

Zeng, H., Feng, J., & Dai, M. (2026). A structured five-step collaborative inquiry approach to mathematical reasoning in elementary school: Effects and implementation evidence. *Frontiers in Psychology, 17*, 1813634. https://doi.org/10.3389/fpsyg.2026.1813634

Zwanch, K., & Wilkins, J. L. M. (2021). Releasing the conceptual spring to construct multiplicative reasoning. *Educational Studies in Mathematics, 106*, 151–170. https://doi.org/10.1007/s10649-020-09999-4