



The Use of Bar Model Approach in Teaching Mathematics

Jiesson Maragana Lucero

Del Pilar National High School, Cagdianao, Dinagat Islands, Philippines

Email: jiesson.lucero@deped.gov.ph

ORCID: 0009-0000-8482-2199

Page | 1

Manuel Ballori Dayondon

Taft National High School, Nueva Ext., Brgy. Taft, Surigao City, Surigao del Norte, Philippines

Email: manuel.dayondon@deped.gov.ph

ORCID: [0009-0005-1333-4768](https://orcid.org/0009-0005-1333-4768)

Abstract

This study examined changes in Grade 7 students' mathematical performance, visualization skills, and problem-solving skills following instruction using the Bar Model Approach. A quantitative pre-experimental single-group pretest–posttest design was employed with 25 purposively selected students at Cagdianao National High School, Dinagat Islands, Philippines, during School Year 2025–2026. The intervention lasted one week, with one hour of instruction per day, and addressed contextualized word problems involving algebraic expressions and formulas. A researcher-developed 20-item test and a five-level skills rubric were used. The source reports expert validation, pilot testing, and a KR-20 coefficient of .75 for the achievement test. The reported achievement mean increased from 12.8 to 16.6, visualization scores from 20.0 to 65.4, and problem-solving scores from 31.6 to 80.1. Paired-samples tests were reported as significant at $p < .001$; however, the achievement-test t statistic is inconsistent with the reported standard deviations and requires verification. Item analysis identified comparison problems as a persistent area of difficulty, particularly Question 16, whose correct-response proportion increased from .16 to .32. The findings provide preliminary evidence of improved assessed performance after bar-model instruction, while the absence of a comparison group and use of the same test prevent causal attribution. The study recommends guided work on comparison relationships, explicit links between diagrams and equations, and more rigorous follow-up evaluation.

Keywords: Bar Model Approach, Mathematics Education, Visualization Skills, Problem-Solving Skills, Algebraic Word Problems

1. Introduction

Solving mathematical word problems requires learners to interpret a situation, identify relevant quantities, represent their relationships, and select suitable operations. A correct numerical answer does not necessarily show that a learner understands the structure of a problem. Students may perform familiar procedures successfully while finding it difficult to explain relationships or construct a useful representation. Instruction should therefore attend to both mathematical performance and the reasoning used to obtain a solution.

The Philippine context provides a reason to strengthen mathematical understanding and application. The Programme for International Student Assessment 2022 country note reports low mathematics proficiency among Filipino 15-year-olds (Organisation for Economic Co-operation and Development [OECD], 2023).

These national findings concern a different population from the Grade 7 participants in the present study. They provide educational context rather than a direct measure of achievement at the participating school.

The Bar Model Approach uses rectangular bars to represent quantities and their relationships. It can help learners organize a verbal problem before writing an equation. Equal parts may represent equal quantities, a whole can be partitioned into its components, and aligned bars can show a comparison. The instructional value lies in how the diagram expresses the mathematical relationship, rather than in drawing alone. A model must remain consistent with the information and the unknown quantity in the problem.

Recent research supports investigating visual representations while distinguishing their forms and instructional conditions. Schoenherr et al. (2024) synthesized visualization interventions in mathematics education and reported benefits for learning. Hadi and Csíkos (2025) reviewed visualization research in secondary-school mathematical problem-solving. Their work situates diagrams within a varied field of instructional approaches and learner characteristics, rather than treating every visual activity as equally effective.

Bar-model research is particularly relevant to the transition to algebra. Jahudin and Siew (2024) compared instruction combining Pólya's problem-solving steps and digital bar models, bar-model instruction, and conventional problem-solving among seventh graders in Malaysia. Morina and Vondrová (2026) examined how Grade 8 pupils in Kosovo developed bar-model representations following a brief intervention. These studies provide related evidence, but differences in grade, context, intervention, and design prevent direct equivalence with the present classroom study.

This study contributes a local assessment of a brief bar-model intervention among Grade 7 learners in a Philippine public school. It examines test performance and rubric-assessed visualization and problem-solving separately. The item analysis also identifies which competencies remained difficult after instruction. The objective is to document changes following the intervention and establish priorities for further teaching and evaluation, rather than infer an effect that the single-group design cannot isolate.

2. Literature Review

2.1 Theoretical foundation

The study's representational perspective is informed by Bruner's (1966) account of enactive, iconic, and symbolic representation. The concrete–pictorial–abstract sequence is a pedagogical application of these ideas, rather than the title of a distinct theory formulated by Bruner. Within this sequence, a bar diagram provides a pictorial representation that can be connected to symbolic expressions. The source documents bar-model instruction but does not specify whether concrete materials were used; a complete concrete–pictorial–abstract intervention should therefore not be assumed.

A constructivist perspective also supports asking learners to interpret situations and explain how a diagram represents a relationship. The source identifies constructivist learning as a foundation, but does not measure knowledge construction directly. Matzin and Mundia (2020) discuss a constructivist basis for bar-model teaching in their survey of teachers in Brunei. Their study concerns teachers' attitudes and reported use, so it should not be cited as an experiment demonstrating gains in student achievement.

2.2 Bar models and algebraic reasoning

Bar models can make part–whole, equal-group, and comparison structures available for examination before learners manipulate symbols. Jahudin and Siew (2024) reported stronger algebraic-thinking outcomes for their combined problem-solving and digital-model approach than for their other instructional groups. This suggests that the procedures surrounding a model deserve attention. The

current intervention used bar-model instruction, and digital technology or a formal Pólya sequence should not be attributed to it without documentation.

Morina and Vondrová (2026) reported improvement in the frequency and quality of bar-model representations after five lessons. Their exploratory work provides a relevant comparison for a short instructional intervention, while retaining the distinction between representation quality and problem-solving success. In the present study, separate skills scores allow these areas to be described, although the scoring criteria and relationships among the measures require clearer documentation.

2.3 Visualization and problem-solving assessment

Visual representations can assist learners in organizing information and making connections, but students still need to interpret relational language and connect diagrams with operations. The meta-analysis by Schoenherr et al. (2024) supports examining visual instruction systematically. Hadi and Csíkos (2025) further show that secondary-school visualization research covers different mathematical content and approaches. These findings justify a focused assessment of the particular problems and representations taught in a classroom.

A broader synthesis of word-problem interventions by Vessonen et al. (2025) emphasizes attention to intervention content, implementation, and methodological characteristics when evaluating outcomes. Their review concerns elementary-school learners and therefore provides a methodological perspective rather than direct evidence for Grade 7. The current study's small sample, short duration, and repeated assessment make transparent reporting especially important.

2.4 Study objectives and hypotheses

The study aimed to describe mathematical performance, visualization skills, and problem-solving skills before and after bar-model instruction; examine the reported pretest–posttest differences; and identify less-mastered competencies through item analysis. Separate null hypotheses specified no mean difference between pretest and posttest for each of the three outcomes. Visualization and problem-solving are treated as assessed outcomes, not demonstrated mediators of achievement change, because no mediation analysis was reported.

3. Methodology

3.1 Research design and participants

A quantitative pre-experimental single-group pretest–posttest design was used. The participants were 25 Grade 7 students enrolled at Cagdianao National High School, Cagdianao, Dinagat Islands, during School Year 2025–2026. Purposive selection focused on students who had already studied contextualized word problems involving algebraic expressions and formulas and were available for the intervention. The study site is reported separately from the first author's institutional affiliation.

Participants were regularly enrolled in Mathematics 7 and completed the intervention assessments. The source states that students absent during implementation or failing to complete either test were excluded. It does not report the initial eligible cohort, numbers excluded, or exact fieldwork dates. The resulting sample describes participating students rather than a representative population of Grade 7 learners.

3.2 Instruments and scoring

A researcher-developed mathematics test contained 20 contextualized problem-solving items on algebraic expressions and formulas. The source reports expert validation, pilot testing, and KR-20 reliability of .75. The term “standardized” has been removed because the supplied manuscript does not

document standardization, norming, or external test calibration. The number of validators, pilot sample, item-development procedures, and exact test specification were not supplied.

Visualization and problem-solving were assessed using a five-level rubric with a maximum score of 100. Achievement scores were reported on a 20-item scale. The source's qualitative classifications are retained in Tables 1 and 2, but the classification cutoffs and rubric descriptors are not available. Inter-rater reliability and separate psychometric evidence for the skills rubric were not reported. Thus, test reliability should not be interpreted as establishing the reliability of all three outcomes.

3.3 Intervention and data collection

Permission was secured from educational authorities and school administrators. The pretest established baseline performance. Students then received bar-model instruction for one week, with one hour of teaching per day, on contextualized word problems involving algebraic expressions and formulas. The same instrument was administered after the intervention. The source does not provide session-by-session lesson plans, attendance records, fidelity measures, or the precise interval between instruction and posttesting.

The stated intervention involved visual representation through bars and relationships among known and unknown quantities. More detailed teaching procedures should be reported from the actual lesson records before submission. No additional activities, teaching materials, or assessment procedures have been assumed in this refinement. Reuse of the same test may have contributed practice or recall effects, which the design cannot separate from instructional learning.

3.4 Data analysis

Means and standard deviations summarized achievement and skills scores. Item proportions described the share of students answering each achievement item correctly. Paired-samples t tests were reported at $\alpha = .05$ because the same students completed both assessments. If all 25 students had complete pairs, the applicable degrees of freedom are 24. The normality of difference scores, outlier checks, confidence intervals, and original statistical output were not provided.

The reported achievement t statistic is retained in Table 3 but flagged because it is incompatible with the displayed means and standard deviations. In addition, the posttest item table omits Question 6. These discrepancies cannot be resolved without the original output or respondent-level records. Statistical significance, even when verified, would establish a within-group difference rather than demonstrate that bar-model instruction alone caused the change.

3.5 Ethical considerations

The source reports voluntary participation, informed consent, confidentiality, anonymity, data privacy, and the right to withdraw without penalty. School permission was also reported. Because the participants were Grade 7 students, the authors should specify the parental or guardian consent and student assent procedures actually used. An approving ethics body or documented exemption was not identified in the supplied manuscript and must be confirmed before submission.

4. Results and Discussion

4.1 Baseline performance

Table 1 presents the reported baseline scores for achievement, visualization, and problem-solving. The distinct scales and qualitative categories should be considered when interpreting their differences.

Table 1. Baseline mathematical performance and skills

Variable	Mean	SD	Qualitative Description
Pretest Score	12.8	4.67	High Performance
Visualization Skills	20.0	0.00	Very Low Skill
Problem-Solving Skills	31.6	6.31	Low Skill

Page | 5

Note. Achievement was measured using a 20-item test; skills were reported on a 100-point scale. Qualitative descriptions are reproduced from the source; cutoffs require author confirmation.

As shown in Table 1, achievement had a mean of 12.8 ($SD = 4.67$), while visualization and problem-solving had means of 20.0 and 31.6. The visualization SD of zero indicates identical reported baseline scores, which may reflect a rubric floor or uniform scoring and warrants examination. The pattern suggests a difference between numerical test performance and demonstrated skills; it does not, by itself, establish reliance on procedural knowledge. Assessing representations separately is consistent with the distinction between visual activity and mathematical reasoning in visualization research (Hadi & Csíkos, 2025).

4.2 Post-intervention performance

Table 2 summarizes the corresponding scores following the intervention. All three reported means were higher than at baseline.

Table 2. Post-intervention mathematical performance and skills

Variable	Mean	SD	Qualitative Description
Posttest Score	16.6	2.10	Very High Performance
Visualization Skills	65.4	15.57	High Skill
Problem-Solving Skills	80.1	8.59	High Skill

In Table 2, achievement increased to 16.6 ($SD = 2.10$), visualization to 65.4 ($SD = 15.57$), and problem-solving to 80.1 ($SD = 8.59$). The respective mean changes were 3.8 test points, 45.4 rubric points, and 48.5 rubric points. Achievement scores were less dispersed after instruction, whereas visualization scores became more varied from a uniform baseline. These descriptive changes are compatible with improved assessed performance following instruction, but they do not identify its sole cause. Morina and Vondrová (2026) provide related evidence of changes in representation after brief bar-model teaching, in a different population and setting.

4.3 Reported achievement comparison

Table 3 reproduces the source's paired comparison of achievement scores. The statistical inconsistency below must be resolved before relying on its inferential conclusion.

Table 3. Reported pretest–posttest achievement comparison

Comparison	<i>t</i> -value	<i>p</i> -value	Decision
Pretest vs Posttest	-13.4	< .001	Reject H_0

Note. Reported t and p are reproduced without alteration. For 25 complete pairs, $df = 24$. The achievement statistic requires verification against the original analysis.

As shown in Table 3, the source reports $t = -13.4$ and $p < .001$. With a mean change of 3.8 and $n = 25$, that t value implies a difference-score SD of approximately 1.42. However, the pretest and posttest SD s imply a minimum possible difference-score SD of $|4.67 - 2.10| = 2.57$. The figures therefore cannot describe the

same paired dataset, even allowing for the displayed rounding. The mean increase can be described, but the reported test should not be used as verified evidence. This finding is not a recalculation from raw scores; it is a consistency check of the supplied summaries.

4.4 Reported visualization comparison

Table 4 presents the paired comparison for visualization skills. The negative sign follows the source's pretest-minus-posttest comparison.

Table 4. Reported pretest–posttest visualization comparison

Comparison	<i>t-value</i>	<i>p-value</i>	Decision
PreVS vs PostVS	-14.6	< .001	Reject H_0

Note. If all 25 pairs were analyzed, $df = 24$. The source reports rejection of the no-difference hypothesis.

In Table 4, the reported t value of -14.6 and $p < .001$ indicate a higher posttest visualization score. Unlike the achievement comparison, this value is consistent with the displayed constant baseline and posttest SD after rounding. Nevertheless, the identical baseline score requires scrutiny of the rubric, scoring floor, and assessment procedure. The result concerns assessed representation skills rather than direct proof of conceptual understanding. Connecting bar diagrams with explicit reasoning is important in related Grade 7 research (Jahudin & Siew, 2024).

4.5 Reported problem-solving comparison

Table 5 presents the reported change in problem-solving rubric scores.

Table 5. Reported pretest–posttest problem-solving comparison

Comparison	<i>t-value</i>	<i>p-value</i>	Decision
PrePS vs PostPS	-26.8	< .001	Reject H_0

Note. If all 25 pairs were analyzed, $df = 24$. Original paired scores and diagnostics were not supplied.

As shown in Table 5, the source reports $t = -26.8$ and $p < .001$, alongside the increase from 31.6 to 80.1. These summaries are not demonstrably incompatible, but the test remains dependent on the original scoring and analysis procedures. A large change on an intervention-aligned rubric may reflect learning, familiarity, or scoring conditions. A comparison group and independent assessment would help distinguish these possibilities. Vessonen et al. (2025) similarly emphasize examining methodological and implementation factors when interpreting intervention outcomes.

4.6 Pre-intervention item analysis

Table 6 retains all 20 pretest item proportions and the source's mastery classifications. A proportion represents correct responses rather than a separate measure of reasoning quality.

Table 6. Pre-intervention item performance

Test Question	Sub-Topic/ Competency	Mean Proportion	Qualitative Description	Verbal Interpretation
Question 1	Division Problem	1.00	Mastered	Excellent
Question 2	Division Problem	1.00	Mastered	Excellent
Question 3	Part–Whole Relationship Problem	1.00	Mastered	Excellent
Question 4	Part–Whole Relationship Problem	0.76	Nearly Mastered	Very Satisfactory
Question 5	Multi-Step Problem	0.92	Mastered	Excellent
Question 6	Part–Whole Relationship Problem	0.92	Mastered	Excellent
Question 7	Part–Whole Relationship Problem	0.76	Nearly Mastered	Very Satisfactory
Question 8	Part–Whole Relationship Problem	0.60	Moderately Mastered	Satisfactory
Question 9	Multi-Step Problem	0.76	Nearly Mastered	Very Satisfactory
Question 10	Part–Whole Relationship Problem	0.60	Moderately Mastered	Satisfactory
Question 11	Division Problem	0.76	Nearly Mastered	Very Satisfactory
Question 12	Part–Whole Relationship Problem	0.64	Nearly Mastered	Very Satisfactory
Question 13	Part–Whole Relationship Problem	0.36	Low Mastery	Fairly Satisfactory
Question 14	Multi-Step Problem	0.72	Nearly Mastered	Very Satisfactory
Question 15	Comparison Problem	0.32	Low Mastery	Fairly Satisfactory
Question 16	Comparison Problem	0.16	Very Low Mastery	Needs Improvement
Question 17	Comparison Problem	0.24	Low Mastery	Fairly Satisfactory
Question 18	Comparison Problem	0.32	Low Mastery	Fairly Satisfactory

Test Question	Sub-Topic/ Competency	Mean Proportion	Qualitative Description	Verbal Interpretation
Question 19	Multi-Step Problem	0.24	Low Mastery	Fairly Satisfactory
Question 20	Multi-Step Problem	0.64	Nearly Mastered	Very Satisfactory

Note. For $n = 25$, proportions correspond to increments of .04. The displayed item proportions total 12.72 expected correct answers, which does not round to the reported mean of 12.8; confirm both summaries.

In Table 6, Questions 1–3 were answered correctly by all participants. The lowest proportion was .16 for comparison Question 16, equivalent to four of 25 students. Comparison Questions 15, 17, and 18 also had low proportions of .32, .24, and .32. Multi-step Question 19 had a proportion of .24. These item-specific findings identify comparison relationships and some multi-step tasks as priorities, without establishing why students answered incorrectly. The actual wording and student solutions would be needed to distinguish language interpretation, representation, and calculation errors.

4.7 Post-intervention item analysis

Table 7 presents the available posttest item results. Question 6 was absent from the supplied table and has not been reconstructed or imputed.

Table 7. Post-intervention item performance

Test Question	Sub-Topic/ Competency	Mean Proportion	Qualitative Description	Verbal Interpretation
Question 1	Division Problem	1.00	Mastered	Excellent
Question 2	Division Problem	1.00	Mastered	Excellent
Question 3	Part–Whole Relationship Problem	1.00	Mastered	Excellent
Question 4	Part–Whole Relationship Problem	0.96	Mastered	Excellent
Question 5	Multi-Step Problem	0.96	Mastered	Excellent
Question 7	Part–Whole Relationship Problem	0.96	Mastered	Excellent
Question 8	Part–Whole Relationship Problem	0.92	Mastered	Excellent
Question 9	Multi-Step Problem	0.96	Mastered	Excellent
Question 10	Part–Whole Relationship Problem	0.96	Mastered	Excellent
Question 11	Division Problem	0.88	Mastered	Excellent
Question 12	Part–Whole Relationship Problem	0.84	Mastered	Excellent
Question 13	Part–Whole Relationship Problem	0.60	Moderately Mastered	Satisfactory
Question 14	Multi-Step Problem	0.88	Mastered	Excellent

Test Question	Sub-Topic/ Competency	Mean Proportion	Qualitative Description	Verbal Interpretation
Question 15	Comparison Problem	0.48	Moderately Mastered	Satisfactory
Question 16	Comparison Problem	0.32	Low Mastery	Fairly Satisfactory
Question 17	Comparison Problem	0.48	Moderately Mastered	Satisfactory
Question 18	Comparison Problem	0.84	Mastered	Excellent
Question 19	Multi-Step Problem	0.76	Nearly Mastered	Very Satisfactory
Question 20	Multi-Step Problem	0.80	Nearly Mastered	Very Satisfactory

Note. Question 6 is missing from the source table. The 19 displayed proportions total 15.60; the full item analysis cannot be reconciled with the reported mean until the missing item is verified.

As shown in Table 7, comparison Question 16 remained the lowest-performing item at .32, equivalent to eight of 25 students. Its proportion increased by .16, or 16 percentage points, from baseline. Questions 15 and 17 each rose to .48, while Question 18 rose to .84. Multi-step Question 19 increased from .24 to .76. The pattern shows uneven improvement within competencies, making item-level follow-up more useful than assuming that every comparison problem remained equally difficult.

The available results support additional teaching focused on the lower-performing comparison items. Teachers could ask students to explain which quantity is larger, what the difference represents, and how the equation follows from the bars. These are instructional recommendations, not procedures documented as part of the intervention. Bar-model research supports attention to representation quality and mathematical relationships (Morina & Vondrová, 2026), while the present data cannot isolate the contribution of those specific features.

5. Educational Implications and Limitations

The descriptive findings suggest that bar-model instruction merits further classroom evaluation as a means of supporting representations and word-problem solutions. Teachers should connect diagrams with the quantities, unknowns, and equations they represent, rather than assess diagram production alone. Persistent difficulty with some comparison items indicates a need for targeted guided practice and discussion of relational language. Staff development may also be useful because teachers' experience and use of bar-model methods vary (Matzin & Mundia, 2020).

The study does not demonstrate that visualization or problem-solving mediated the change in achievement. All outcomes were measured before and after the same intervention, and no causal pathway was tested. The proposed representational explanation is theoretically plausible, but should remain an interpretation. Broader visualization findings support further investigation without substituting for a controlled evaluation of this particular approach (Schoenherr et al., 2024).

Limitations include the 25-student purposive sample, absence of a comparison group, brief intervention, use of the same test, incomplete rubric documentation, and lack of follow-up assessment. Changes could reflect instruction, repeated testing, concurrent learning, or other conditions. The achievement t inconsistency, discrepancy between pretest item proportions and the mean, and missing posttest item further limit inference. Outcomes describe this participating group and cannot establish effectiveness across Philippine schools.

6. Conclusion and Recommendations

The source reports higher achievement, visualization, and problem-solving scores after one week of bar-model instruction. Comparison Question 16 remained a notable difficulty, while several other comparison and multi-step items improved. These results provide preliminary descriptive support for continued classroom use and evaluation, rather than conclusive evidence of causal effectiveness. The achievement significance test requires correction or verification before submission.

Teachers may incorporate bar diagrams into contextualized problem-solving and provide additional support for comparison relationships. School leaders may support teacher learning and the development of appropriate instructional examples. Future research should use comparison groups, document lesson implementation, employ distinct assessment forms where feasible, report rubric reliability, and assess retention. The immediate publication priority is to reconcile the statistical and item-level results and confirm the missing methodological and author details.

Funding

The study was supported by the researchers' personal funds. No external funding agency provided financial support.

CRedit Authorship Contribution Statement

Jiesson Maragana Lucero: Conceptualization, Methodology, Investigation, Formal analysis, Writing – original draft, Writing – review and editing, and Project administration. Manuel Ballori Dayondon: Validation, Supervision, and Writing – review and editing. Roles follow the contributions stated in the source and require final author approval.

Ethical Statement

The source reports informed consent, voluntary participation, confidentiality, anonymity, and school permission. The authors must confirm the institutional ethics approval or applicable exemption and the parental or guardian consent and student assent procedures used for minors before submission.

Declaration of Interests

The authors declare no competing financial, personal, institutional, or professional interests that influenced the research, as stated in the supplied manuscript.

Data Availability

The source states that data supporting the findings are available from the corresponding authors upon reasonable request. Any access must respect participant confidentiality and the applicable consent conditions.

AI Usage

AI-assisted tools were used to refine language, organize and format the manuscript, verify references, and check consistency among reported summaries. No participant data were generated or reanalyzed. The authors must review and approve the final text, references, statistical reporting, and disclosure before submission.

Page | 11

References

Bruner, J. S. (1966). *Toward a theory of instruction*. Harvard University Press.

Hadi, W., & Csikos, C. (2025). A systematic review on visualization in mathematical problem-solving in secondary schools. *European Journal of Science and Mathematics Education*, 13(4), 352–367. <https://doi.org/10.30935/scimath/17410>

Jahudin, J., & Siew, N. M. (2024). The effects of Polya's problem solving with digital bar model on the algebraic thinking skills of seventh graders. *Problems of Education in the 21st Century*, 82(3), 390–409. <https://doi.org/10.33225/pec/24.82.390>

Matzin, E. S., & Mundia, L. (2020). Efficacy of the bar model method of teaching mathematics to Year 7 students: Case study of teachers in Brunei Darussalam. *Journal of Educational Issues*, 6(1), 402–421. <https://doi.org/10.5296/jei.v6i1.17154>

Morina, Q., & Vondrová, N. (2026). Efficacy of the bar model method in solving algebraic word problems: An exploratory study with grade 8 Kosovo pupils. *Eurasia Journal of Mathematics, Science and Technology Education*, 22(5), Article em2822. <https://doi.org/10.29333/ejmste/18341>

Organisation for Economic Co-operation and Development. (2023). PISA 2022 results (Volume I and II)—Country notes: Philippines. https://www.oecd.org/en/publications/pisa-2022-results-volume-i-and-ii-country-notes_ed6fbcc5-en/philippines_a0882a2d-en.html

Schoenherr, J., Strohmaier, A. R., & Schukajlow, S. (2024). Learning with visualizations helps: A meta-analysis of visualization interventions in mathematics education. *Educational Research Review*, 45, Article 100639. <https://doi.org/10.1016/j.edurev.2024.100639>

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, Article 102642. <https://doi.org/10.1016/j.ijer.2025.102642>

Author Verification Before Submission

Remove this editorial completion section once the following points are resolved. This refinement retains the original title and data; unresolved source inconsistencies prevent a fully submission-ready version.

- Verify the achievement pretest and posttest means, SDs, paired t value, p value, and valid sample against the original dataset or statistical output. The reported $t = -13.4$ is incompatible with the supplied summaries.
- Reconcile the pretest item-proportion sum of 12.72 with the reported mean of 12.8. Supply posttest Question 6 and verify all 20 item proportions against the reported posttest mean.
- Supply the rubric descriptors and classification cutoffs, test specification, validation and pilot details, scoring procedures, and any inter-rater agreement evidence. Explain the identical visualization baseline scores.
- Confirm school and ethics permissions, parental or guardian consent and student assent, fieldwork dates, participant exclusions, and actual instructional procedures.

- Confirm the second author's ORCID. The supplied manuscript duplicates the first author's identifier for both authors; the second has been flagged rather than assigned an invented identifier.