

Bar Model Scaffolding for Percentage Word Problems in a Year Five Classroom

Hartini Ismail^{1*}, Mohd Azarul Mohd Mokhtar¹, Haryati Ismail²

¹ Department of Educational Studies, Faculty of Human Development, Universiti Pendidikan Sultan Idris, Perak, Malaysia

² Department of Science, Technology, Engineering and Mathematics, IPG Kampus Kota Bharu, Kelantan, Malaysia

*Corresponding Author

DOI: <https://doi.org/10.47772/IJRISS.2026.1026EDU0606>

Received: 17 September 2026; Accepted: 22 September 2026; Published: 05 October 2026

ABSTRACT

Many primary pupils can carry out a calculation but still struggle to turn a word problem into a useful mathematical representation. This paper revisits a small classroom intervention in which nine Year Five pupils from a public primary school in Sarawak, Malaysia, learned to use bar models for percentage problems. Over eight one-hour sessions, pupils identified the whole and the relevant part, drew and labelled a segmented bar, translated the relationship into a calculation, and checked whether the answer made sense. The class mean rose from 31.48 (SD = 5.55) on the pretest to 94.44 (SD = 8.34) on the posttest, and the median rose from 33.33 to 100.00. All nine pupils improved. Both the Wilcoxon signed-rank test and the exact sign test indicated a clear pre-post shift ($W = 0$, exact $p = .0039$). Every score moved from 33.33% or below at pretest to 83.33% or above at posttest. To make sense of this pattern, we introduce the Verbal-Visual-Symbolic Representation Bridge, which describes how pupils may move from understanding the problem, to constructing a diagram, to writing a mathematical expression, and finally to checking the result. The framework is offered as an interpretation of the teaching process, not as a tested causal model. The findings must also be read in light of the small sample, lack of a comparison group, limited information about the tests, and a marked ceiling effect at posttest. Even with these constraints, the study shows the value of examining each pupil's score pattern rather than relying only on a class mean. It also illustrates how the assessment itself can shape what can reasonably be claimed about learning.

Keywords: bar model, percentage, primary mathematics, representational scaffolding, word-problem solving

INTRODUCTION

Solving a mathematical word problem involves more than carrying out an operation. Pupils first need to decide which quantities matter, how those quantities are related, and what kind of mathematical representation fits the situation. Some difficulties therefore appear before any calculation begins. This is especially common when pupils depend on keywords or choose an operation without first making sense of the quantitative structure. Percentage problems add another layer of difficulty because pupils must coordinate the whole, the part, and a rate expressed per hundred.

A visual representation can give pupils something concrete to reason with as they move from words to mathematical notation. In the bar model, also known as the model method or strip diagram, rectangular bars and their partitions show the relationship between a whole, its parts, and an unknown value (Beckmann, 2004). The diagram is not simply an illustration placed beside the solution. Pupils must decide what each bar and segment represents, connect the labels to the information in the problem, and use the completed model to formulate a calculation.

Research on visual models is encouraging, although the results are not uniformly positive. Bar-model instruction has supported word-problem performance among learners with mathematics difficulties (Morin et al., 2017), while a percentage-focused intervention also reported improvement among secondary pupils (Abdul Gani et al., 2019). Yet a diagram helps only when pupils understand the relationships it is meant to show. De Koning et al. (2022), for example, found that model drawing could hinder performance when pupils followed superficial keyword cues. A recent meta-analysis similarly found that word-problem interventions can be effective, but that their effects vary with the quality and duration of instruction, the measures used, and the design of the study (Vessonen et al., 2025). Taken together, this evidence suggests that the bar model should be examined as part of an instructional process rather than treated as an effective technique in its own right.

This study began with a practical concern in a small primary school in Sarawak. Classroom diagnostic work showed that several Year Five pupils could perform basic percentage calculations but had difficulty deciding how to represent a percentage problem. The teaching sessions were designed to make the part-whole relationship visible before pupils began calculating. In revisiting the study, we focus on how much the pupils' scores changed, whether the pattern was shared across the class, and what the available evidence cannot establish.

Small classroom studies often report a mean gain and stop there. That summary can hide important differences: an apparent improvement may be shared by the class, driven by only a few pupils, or compressed because the test is too easy at posttest. We therefore examine the nine score pairs individually as well as at class level. We also use the score distribution to consider the sensitivity of the assessment. The study does not claim that the bar model is new. Its added value lies in bringing together the instructional movement from words to diagrams and symbols, the pattern of change for each learner, and the limits imposed by a near-ceiling posttest.

CONCEPTUAL FRAMEWORK

Word problem solving as representation

To solve a word problem, pupils need to build a mental account of the situation and connect it to a mathematical model. An external representation can reduce the amount of information that must be held in working memory and make relationships easier to inspect. This support depends on the accuracy of the mapping. Each segment, label, and comparison must correspond to a quantity or relationship in the problem. A neatly drawn diagram offers little help if it has been copied without understanding.

The bar model for percentages

In a percentage problem, a bar can represent the whole and equal segments can represent fractions or percentages of that whole. One example used during the intervention involved 50 books. The bar was divided into ten equal segments, so each segment represented 10%, or five books. Shading six segments for 30 English-language storybooks made the equivalence between 30/50 and 60/100 visible. In this way, the same problem could be read in words, seen as a proportional diagram, and expressed symbolically.

Figure 1: Example of a segmented bar representation for 30 of 50 books

10%	10%	10%	10%	10%	10%	10%	10%	10%	10%
5	5	5	5	5	5	5	5	5	5

Whole = 50 books; shaded part = 30 books = 6 of 10 equal segments = 60%.

The verbal visual symbolic representation bridge

We use the term Verbal-Visual-Symbolic Representation Bridge to describe the work pupils do between reading a problem and producing an answer. The bridge is not intended as a rigid sequence. A pupil may return to the wording, change a label, redraw the bar, or revise an equation at any point. Its purpose is to draw attention to the decisions involved in representing the problem and to distinguish those decisions from the simple act of drawing a rectangle.

Table 1: Stages and potential breakdowns in the proposed representation bridge

Stage	Central question	Observable instructional action	Potential breakdown
1 Verbal structuring	What quantities and relationships are described?	Identify the whole, parts, unknown and comparison stated in the problem	Selecting numbers without establishing their roles
2 Visual mapping	How can the relationship be displayed spatially?	Construct, partition and label a proportional bar	Drawing an unlabelled or structurally inaccurate bar
3 Symbolic formulation	Which operation expresses the displayed relationship?	Translate the bar into a number sentence or proportion	Choosing an operation from a keyword rather than the relation
4 Contextual verification	Does the result fit both the diagram and the situation?	Check magnitude, units and the relationship between whole and part	Reporting a calculation without evaluating its plausibility

The framework suggests several questions for later research. Bar models may be more useful when teachers connect the wording, diagram, and equation explicitly instead of presenting them as separate steps. Pupils' explanations of their labels and partitions may also reveal more than a completed diagram on its own. Finally, transfer may be more likely when pupils work with different problem structures and decide how the model should be drawn, rather than reproduce one familiar layout. The present score data cannot test these possibilities, but they provide a clear direction for subsequent studies.

Research Questions

1. To what extent did pupils' percentage word-problem scores change from pretest to posttest following the bar-model intervention?
2. How consistent were the direction and magnitude of score changes across the nine participating pupils?
3. What does the movement of the score distribution reveal about short-term performance transition and assessment sensitivity?
4. How can the Verbal-Visual-Symbolic Representation Bridge clarify the proposed instructional function of bar-model scaffolding?

METHODOLOGY

Design and setting

The study used a one-group pretest–posttest classroom design within a teacher-inquiry setting. The earlier proceedings paper used both “action research” and “quasi-experiment”; the one-group designation is more precise for the surviving record. There was no comparison group, random assignment, independent assessor, or documented control of concurrent teaching. The analysis therefore documents change within this class; repeated testing, teacher influence, maturation, and other events remain plausible explanations.

The study was conducted in a small public primary school in the Meradong district of Sarawak, Malaysia. The intervention addressed percentage word problems in the Year Five mathematics curriculum.

Participants

The participants were the nine pupils in one Year Five class, comprising five boys and four girls. Classroom observations, interviews, diagnostic tasks, drill exercises, and the pretest had indicated that they needed additional support with percentage word problems. As the study involved the whole accessible class, the sample was purposive, small, and closely tied to its school context.

Intervention

The intervention consisted of eight one-hour lessons over one month. During each lesson, pupils identified the whole and the relevant part or unknown, drew and labelled a rectangular bar, divided it into equal units, and placed the numerical information from the problem onto the diagram. They then selected an operation, calculated the answer, and checked it against the model. The teacher first demonstrated the process, then guided pupils through similar tasks before asking them to work more independently. Calculation was deliberately delayed until the pupils had represented the relationship between the quantities.

Table 2: Intervention sequence and expected pupil actions

Intervention phase	Instructional emphasis	Expected pupil action
Problem comprehension	Identify the question, whole, part and known quantities	Underline and restate relevant information
Diagram construction	Represent the whole with a labelled rectangular bar	Draw, partition and label the bar
Relational mapping	Connect text quantities to bar segments	Explain what each segment represents
Symbolic solution	Translate the displayed relation into an operation	Calculate the unknown value
Verification	Compare the numerical answer with the diagram and context	Check magnitude and state the answer in context

Measure and procedure

The archived report describes teacher-developed pretest and posttest assessments of percentage word problems, with scores retained as percentages. The original question papers, marking rubrics, item-level responses, item specifications, and documentation of their development were not available for this reanalysis. Consequently, the number and wording of items, scoring weights, content coverage, difficulty and discrimination indices, expert review, reliability, and equivalence of the two forms cannot be established. Score steps of approximately 16.67 percentage points are visible in the surviving totals, but they do not establish that either test comprised six equally weighted items: different scoring schemes could produce the same steps. No sample item is presented as an original test item because its provenance cannot be verified. The 50-book example in Figure 1 illustrates the instruction; it must not be mistaken for a recovered test question. These missing details restrict interpretation of the score change, especially given the posttest ceiling.

Data analysis

We summarised the scores using the mean, standard deviation, median, range, and the change recorded for each pupil. We then examined whether all pupils changed in the same direction, how widely the gains varied, and how the class moved between the observed pretest and posttest score bands. In this paper, high immediate performance refers only to the observed posttest range of 83.33%-100%; it is not a validated mastery category. The paired differences were markedly non-normal (Shapiro-Wilk $W = 0.536$, $p < .001$), partly because the scores occurred at a small number of test increments and clustered near the posttest ceiling. We therefore used the Wilcoxon signed-rank test and reported the exact sign test as a simple check of directional consistency. We did not rely on a paired-samples t test or Cohen's d_z because the normality assumption was untenable and score compression could exaggerate the standardised gain. The resulting p values indicate a shift within this group; they do not show that the intervention caused it.

Ethics

The school administration gave permission for the classroom study. Parents or guardians provided consent, and the pupils provided assent before data collection. Participation was voluntary. Names and other direct identifiers

were removed from the dataset used for this reanalysis, and the findings are reported anonymously. The surviving records do not include an institutional ethics approval reference number, so none is reported here.

RESULTS

Descriptive change

Every pupil scored higher on the posttest than on the pretest. The class mean increased by 62.96 percentage points, from 31.48 (SD = 5.55) to 94.44 (SD = 8.34), while the median increased from 33.33 to 100.00. Pretest scores ranged from 16.67 to 33.33; posttest scores ranged from 83.33 to 100.00.

Table 3: Descriptive statistics for pretest, posttest, and change scores

Measure	n	Mean	SD	Median	Minimum	Maximum
Pretest	9	31.48	5.55	33.33	16.67	33.33
Posttest	9	94.44	8.34	100.00	83.33	100.00
Change	9	62.96	7.35	66.67	50.00	66.67

Note. Values are percentages or percentage-point changes. Figures are calculated from the individual scores reported in the source manuscript.

Individual score patterns

Table 4: Individual pretest, posttest, and change scores

Participant	Pretest	Posttest	Change
P1	33.33	100.00	66.67
P2	33.33	100.00	66.67
P3	33.33	100.00	66.67
P4	33.33	100.00	66.67
P5	33.33	100.00	66.67
P6	33.33	83.33	50.00
P7	16.67	83.33	66.66
P8	33.33	83.33	50.00
P9	33.33	100.00	66.67

At posttest, six pupils scored 100% and the remaining three scored 83.33%. Their individual gains ranged from 50.00 to 66.67 percentage points. Improvement was therefore not confined to a few members of the class. At the same time, the clustering of scores at 83.33% and 100% suggests that the test had limited capacity to distinguish performance at the upper end.

Paired inferential analysis

All nine paired changes were positive. The exact two-sided sign test yields $p = .0039$ under the null model of equally likely improvement or decline in independent pairs. A zero sum of negative ranks is consistent with a Wilcoxon signed-rank statistic of $W = 0$, although tied change magnitudes require care in calculating an exact

Wilcoxon probability. We rely on the sign test for the exact directional probability and show all scores in Table 4. Neither a small p value nor the large observed gain isolates an intervention effect in this one-group design.

Distributional transition and assessment sensitivity

The pretest scores took only two values: eight pupils scored 33.33% and one scored 16.67%. At posttest, six pupils scored 100% and three scored 83.33%. In other words, every pupil moved from the observed pretest band of 16.67%-33.33% to the posttest band of 83.33%-100%. The smallest gain was 50.00 percentage points and the median gain was 66.67 points. Within the material covered by the test, the short-term change was substantial.

This pattern also makes the weakness of the posttest visible. Once two thirds of the class reached the maximum score, the test could no longer distinguish a sound performance from a more advanced understanding. Nor can these scores show whether pupils could use the model with unfamiliar wording, a less routine percentage structure, or after a delay. The ceiling effect is therefore part of the finding, not merely a statistical inconvenience. The pupils performed well on the sampled items immediately after teaching, but the test tells us much less about the depth, durability, or transfer of their learning.

Table 5: Analytical interpretation of the observed score transition

Analytical lens	Observed evidence	Defensible interpretation	Interpretive boundary
Group change	Mean increase of 62.96 percentage points	Substantial pre-post shift in this class	No causal attribution without a comparison condition
Directional consistency	9 of 9 pupils improved	Improvement was not driven by a minority of pupils	Small sample and repeated testing remain competing explanations
Transition pattern	All pupils moved from 16.67%-33.33% to 83.33%-100%	High immediate performance on the assessed domain	The bands are descriptive, not validated mastery categories
Assessment sensitivity	6 of 9 pupils obtained 100%	Posttest captured successful performance	Ceiling limits discrimination, transfer claims and effect estimation

DISCUSSION

Scores improved sharply and consistently after the eight bar-model lessons. This pattern is consistent with the instructional rationale for the approach: a segmented bar makes the whole-part relationship visible and gives pupils an intermediate representation between the wording of a problem and its symbolic solution. Similar improvements have been reported in earlier studies of bar-model instruction (Abdul Gani et al., 2019; Morin et al., 2017).

The findings should not be read as a definitive estimate of effectiveness. They show what happened in one small class whose pupils had struggled to represent percentage problems. All nine pupils improved, including the pupil who began with the lowest score. For classroom practice, the useful point is not simply that a bar was drawn. The labels and partitions gave pupils a way to work through the relationship described in the problem before they selected an operation.

Looking at the individual scores changes the interpretation. A mean gain of 62.96 points is impressive, but it does not show that every pupil improved or that the lowest posttest score was higher than the highest pretest score. That complete separation between the observed score bands is noteworthy. It is not conclusive, however, because practice effects, close teaching-test alignment, and limited item difficulty remain possible explanations. For a sample this small, showing every score is more honest and more useful than presenting a significance test in isolation.

The Verbal-Visual-Symbolic Representation Bridge helps explain what the bar model may have contributed to the lessons. Pupils had to organise the verbal information, map the relationships onto a diagram, formulate an operation, and check the answer in context. Framing the process in this way separates meaningful modelling from copying a diagram. It also points to evidence that future studies could collect, such as whether pupils identify the whole correctly, partition the bar proportionally, write an equation that matches the relationship, and revise an answer that does not fit the situation.

The near-ceiling posttest scores nevertheless call for restraint. They may represent genuine success with the taught problems, but they may also reflect an undemanding posttest, close similarity between practice and test items, or the effect of repeated exposure. The safest interpretation is therefore limited to immediate performance on the assessed percentage problems. The study does not show that the learning was retained or transferred to other kinds of problems.

Drawing a bar is not automatically helpful. De Koning et al. (2022) showed that diagrams can reinforce shallow reasoning when pupils map keywords mechanically. Teachers therefore need to ask what each bar or segment means, why it has been partitioned in a particular way, and whether the resulting answer is reasonable. The discussion around the representation is at least as important as the drawing itself.

Practical implications

- Introduce the bar model through explicit links among the problem text, the diagram, and the symbolic equation.
- Ask pupils to label the whole, the known part, the unknown, and the value of each segment before calculating.
- Use contrasting problems so that pupils cannot solve by copying one fixed bar layout or following keywords.
- Include transfer items with unfamiliar wording and delayed assessment to test whether understanding endures.
- Treat pupils' diagrams and explanations as diagnostic evidence, not only their final numerical answers.

Limitations and directions for research

The study has six consequential limitations. Nine pupils in one class offer little basis for population generalisation. With no comparison condition, other causes of the gain cannot be separated from the lessons. Original test forms and item-level records are unavailable, so content validity, reliability, form equivalence, and item properties cannot be evaluated retrospectively. Six pupils scored at the posttest maximum, limiting discrimination at the upper end. The immediate posttest provides no evidence of retention or transfer. Finally, no archived pupil explanations or diagrams were available to test the proposed representation bridge. These are limitations of this study, not additional measurements to be inferred from the score totals. A prospective replication should preserve both forms and their marking guides; obtain expert review against an explicit content blueprint; pilot items and report their difficulty and discrimination; assess scoring agreement and reliability when appropriate; establish form comparability; include a comparison group across schools; and administer delayed and unfamiliar transfer tasks. Collecting dated drawings, explanations, and think-aloud records would allow the framework's stages to be coded and tested directly.

Inclusive instructional value

We did not collect subgroup evidence for pupils with identified learning needs, so the study cannot establish differential effects for those learners. The design may still have inclusive potential. A labelled external model gives teachers and pupils a shared object to discuss, makes breakdowns in reasoning easier to locate, and offers an alternative to instruction that depends heavily on verbal explanation or memorised procedures. These are plausible teaching advantages, not outcomes demonstrated by the present data. Future work should examine accessibility, language demands, pupils' own diagrams, and what happens when the scaffold is gradually withdrawn.

Research agenda generated by the framework

The next step is to test the proposed framework with process evidence rather than infer its stages from total scores. A prospective study could compare bar-model teaching with another explicit representation strategy while holding teaching time and assessment conditions as similar as possible. Pre-specified coding of pupils' drawings and explanations could track correct identification of the whole, proportional partitions, correspondence between labels and quantities, equation choice, and contextual checking. Independent coders and reported agreement would strengthen these observations. Delayed and unfamiliar percentage problems could then test retention and transfer. These proposals describe future data collection; they are not findings from the nine pupils.

CONCLUSION

After eight one-hour bar-model lessons, all nine pupils scored higher on the immediate percentage word-problem posttest. The gain was consistent across the class, but six pupils reached the test maximum and the original test forms and item-level records could not be recovered. The proposed Verbal–Visual–Symbolic Representation Bridge describes a plausible instructional sequence; pupil process data would be needed to test it. These results support a narrow conclusion of improved immediate scores in this classroom. The design and surviving records do not establish an intervention effect, the quality or equivalence of the tests, retention, or transfer.

Declarations

Ethics approval and consent to participate. The archived manuscript reports school permission, parental or guardian consent, and pupil assent. The authors must verify the wording against the available permissions and consent records before submission. No institutional ethics approval identifier was available in the source materials, and approval is not claimed here.

Funding. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

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

Data availability. The anonymised pupil-level scores used in the analysis are reported in this article. The remaining archived classroom records are not publicly available because they concern minors and consent was not obtained for open deposition. Requests for further information may be directed to the corresponding author and will be considered within the relevant ethical and institutional restrictions.

Author contributions. The contributions of Hartini Ismail, Mohd Azarul Mohd Mokhtar, and Haryati Ismail require confirmation by all three listed authors before submission.

REFERENCES

1. Abdul Gani, M., Tengah, K. A., & Said, H. (2019). Bar model as intervention in solving word problem involving percentage. *International Journal on Emerging Mathematics Education*, 3(1), 69–76. <https://doi.org/10.12928/ijeme.v3i1.11093>
2. Beckmann, S. (2004). Solving algebra and other story problems with simple diagrams: A method demonstrated in grade 4–6 texts used in Singapore. *The Mathematics Educator*, 14(1), 42–46.
3. Bruner, J. S. (1966). *Toward a theory of instruction*. Harvard University Press.
4. de Koning, B. B., Boonen, A. J. H., Jongerling, J., van Wesel, F., & van der Schoot, M. (2022). Model method drawing acts as a double-edged sword for solving inconsistent word problems. *Educational Studies in Mathematics*, 111, 29–45. <https://doi.org/10.1007/s10649-022-10150-8>
5. Hegarty, M., & Kozhevnikov, M. (1999). Types of visual-spatial representations and mathematical problem solving. *Journal of Educational Psychology*, 91(4), 684–689. <https://doi.org/10.1037/0022-0663.91.4.684>

6. Ismail, H., & Maat, S. M. (2016). Keberkesanan model bar dalam penyelesaian masalah matematik berayat murid Tahun Lima. In *E-Proceedings of the International Conference on Education 2016 (ICE2016)* (pp. 496–502). Sarawak Association for Development of Professionalism in Education.
7. Morin, L. L., Watson, S. M. R., Hester, P., & Raver, S. A. (2017). The use of a bar model drawing to teach word problem solving to students with mathematics difficulties. *Learning Disability Quarterly*, 40(2), 91–104. <https://doi.org/10.1177/0731948717690116>
8. Vessonen, T., Dahlberg, M., Hellstrand, H., Widlund, A., Söderberg, P., Korhonen, J., Aunio, P., & Laine, A. (2024). Task characteristics associated with mathematical word problem-solving performance among elementary school-aged children: A systematic review and meta-analysis. *Educational Psychology Review*, 36, 117. <https://doi.org/10.1007/s10648-024-09954-2>
9. 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>
10. Xin, Y. P., & Jitendra, A. K. (1999). The effects of instruction in solving mathematical word problems for students with learning problems: A meta-analysis. *The Journal of Special Education*, 32(4), 207–225. <https://doi.org/10.1177/002246699903200402>