INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 7

Problem-Solving and Word-Problem Skills in Mathematics: Towards Essential Tools for Applying the Concepts to the Real-World Scenarios

DOI: 10.5281/zenodo.24190412

Kimme Yajaira C. Sayman
Western Colleges, Inc., Naic, Cavite
kycsayman@gmail.com

Abstract:

This study examined mathematics teachers' strategies in teaching problem-solving and word-problem skills as perceived by 120 junior high school students at Luis Y. Ferrer Jr. West National High School, Division of General Trias City, and translated the findings into essential tools for applying mathematical concepts to real-world scenarios. It used a descriptive-correlational design and a researcher-made, expert-validated questionnaire on a four-point scale. Frequency and percentage described the profile of the respondents; weighted mean assessed teachers' strategies; independent-samples t-test and one-way analysis of variance (ANOVA) tested differences by age, sex, and section; and Pearson correlation examined the relationship between the two strategy domains. Most respondents were 13 years old (72%) and female (59%), with exactly ten respondents from each of Sections A to L. Teachers' problem-solving strategies were rated as somewhat strategized (composite mean = 3.39, SD = 0.47), led by reflection and debriefing (3.66) and scaffolding (3.61), while encouraging multiple strategies was lowest (3.22). Word-problem strategies were rated as highly strategized (composite mean = 3.54, SD = 0.48), led by graphic organizers (3.78) and modeling/think-alouds (3.69). Sex differentiated both domains; age differentiated word-problem strategies only; section differentiated neither. The two domains were significantly and moderately negatively correlated ($r = -0.385$, $p < .001$). The study proposes two classroom tools, the Strategy Canvas Board and the Say-It-Back Protocol, to strengthen flexible strategy use and reading comprehension before computation.

Keywords: problem-solving skills, word-problem skills, mathematics teaching strategies, graphic organizers, scaffolding, real-world application

Introduction:

Problem-solving and word-problem skills are pivotal in mathematics education because they enable students to apply mathematical concepts to real-world scenarios. They do more than rehearse procedures; they require learners to interpret a situation, decide what is known and unknown, select a strategy, carry it out, and judge whether the answer makes sense. When these skills are well developed, students gain mathematical understanding, critical thinking, and analytical habits that transfer to daily decisions, work, and further study. When they are weak, students may compute correctly yet fail to begin, misread what a problem asks, or abandon a task at the first difficulty. For this reason, the quality of teachers' day-to-day strategies matters as much as the curriculum itself: the routines, questions, representations, and feedback that teachers use shape whether students experience mathematics as a set of rules to remember or as a way of thinking to use.

International and national discussions of mathematics reform have renewed attention to explicit, structured instruction in foundational skills. In New South Wales, for example, a streamlined back-to-basics curriculum was promoted to give all students equal access to essential mathematical knowledge and to reduce unnecessary complexity that can hinder comprehension (The Australian, 2025). The point is not to return to drill alone. It is to make the underlying structure of problems visible, to teach students how to reason with that structure, and to give them repeated, supported opportunities to practise. In Philippine classrooms, where word problems are a familiar source of difficulty, this balance is especially urgent. Students must cope simultaneously with mathematical demands and language demands, and teachers must therefore teach both the mathematics and the reading of the mathematics.

The problem addressed by this study is practical and local. At Luis Y. Ferrer Jr. West National High School, teachers were observed using a range of strategies for problem-solving and word problems, but it was not clear which strategies

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 7

students actually experienced most consistently, whether perceptions differed by age, sex, or section, or whether general problem-solving instruction and word-problem instruction were experienced as one coherent approach or as two different emphases. Without that evidence, efforts to improve instruction risk being generic: more activities, more worksheets, or more reminders to read carefully, without a clear diagnosis of what is already strong and what remains thin. This study therefore asked students themselves, as the direct recipients of instruction, how they perceived their mathematics teachers' strategies.

This study was guided by constructivism, Polya's (1945) four-stage problem-solving theory, and cognitive load theory (Sweller, 1988). Constructivism, as articulated by Vintere (2018), holds that learners actively construct knowledge from prior understanding through experience and reflection, rather than passively receiving it. Polya's stages, understanding the problem, devising a plan, carrying out the plan, and looking back, give that construction a usable sequence that students can internalise. Cognitive load theory adds a caution: if instruction overloads working memory with extraneous complexity, students have little capacity left for the intrinsic work of reasoning. Together, these lenses predict that effective teaching will make structure explicit, scaffold the difficult steps, use representations that offload memory, and create space for reflection, while gradually withdrawing support as students gain independence.

Accordingly, this study had seven objectives: (1) describe the profile of the student-respondents in terms of age, sex, and section; (2) assess their mathematics teachers' strategies in teaching problem-solving skills; (3) test whether that assessment differs significantly when grouped by profile; (4) assess teachers' strategies in teaching word-problem skills; (5) test whether that assessment differs significantly when grouped by profile; (6) determine the relationship between the assessments of problem-solving and word-problem strategies; and (7) use the findings to propose essential tools for applying mathematical concepts to real-world scenarios. The null hypotheses were that there is no significant difference in either assessment when grouped by profile and no significant relationship between the two domains. The study used a descriptive-correlational survey design, and its findings are intended for students, teachers, school administrators, curriculum developers, and future researchers who seek evidence-based, classroom-ready ways to strengthen mathematical problem-solving.

Figure 1. Conceptual Framework of the Study

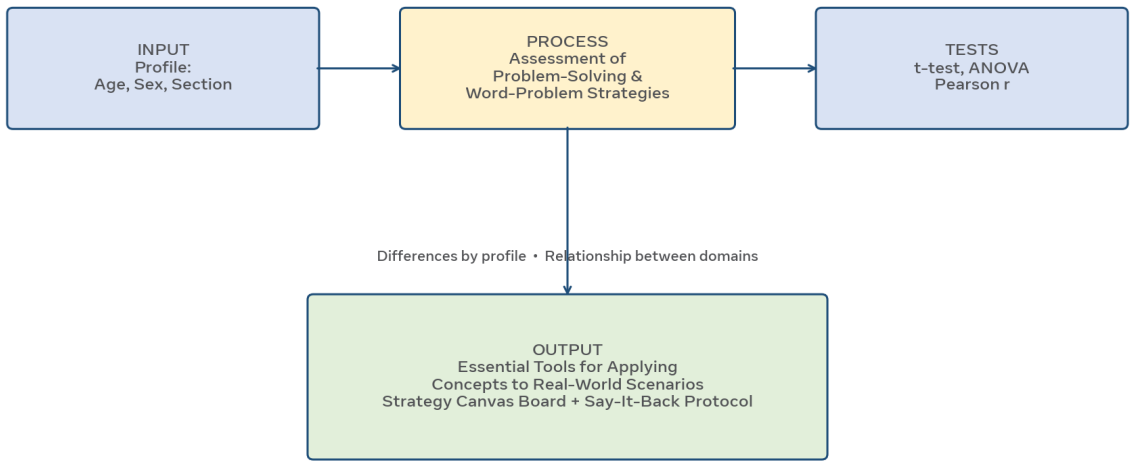


Figure 1. Conceptual framework: profile, assessed strategies, statistical tests, and proposed tools.

Literature Review:

Problem-Solving as Mathematical Thinking

Problem-solving is at the heart of mathematics because it requires learners to navigate complex situations by applying appropriate strategies and logical reasoning. Schoenfeld (2024) argued that effective problem-solving instruction must attend to students' cognitive and social behaviours, not only to whether they obtain a correct answer, because the way students frame a problem, choose resources, monitor progress, and recover from being stuck determines whether they can transfer what they know to unfamiliar tasks. This view aligns with the broader educational goal of preparing students for unforeseen challenges rather than for rehearsed item types. Teaching problem-solving well is difficult precisely because it draws on many skills at once: conceptual understanding, procedural fluency, strategic competence, adaptive reasoning, and a productive disposition (Smith & Brown, 2022).
Traditional instruction that emphasises the application of a specific algorithm to a particular problem type can limit flexibility and creativity, leaving students without a general toolkit for novel problems (Näveri & Leppäaho, 2011; Leppäaho, 2018). Dendane (2009) likewise noted that problem-solving is a complex topic to teach because it requires a multitude of skills, and that teachers need both subject-matter understanding and demonstrable teaching skill (Gozel, 2016, as cited in the original study; Guven & Ozcelik, 2017, as cited in the original study).

Polya's (1945) four-stage model remains the most widely used roadmap for making that complexity manageable. In the first stage, understanding the problem, students identify what is given, what is asked, and what conditions constrain the solution. In the second, devising a plan, they consider strategies such as drawing a diagram, making a table, working backward, or breaking the problem into smaller parts. In the third, carrying out the plan, they execute the steps methodically and perform the necessary calculations or constructions. In the fourth, looking back, they check reasonableness, consider alternative methods, and ask whether the method applies elsewhere. Carifio (2015) and Bachmann (2021) have revisited Polya's model and affirmed its value as a teachable heuristic, while more recent work has extended it: Cai and Hwang (2024) found in a meta-analysis that problem-posing, in which students formulate new problems, significantly improves their ability to interpret situations mathematically and develop meaningful strategies. Santos-Trigo (2024) similarly highlighted that teaching a repertoire of heuristic techniques fosters flexibility and adaptability, allowing learners to approach unfamiliar problems with confidence rather than waiting for a memorised procedure.

Metacognition is the thread that connects Polya's stages. Dulger and Ogan-Bekiroglu (2023) found that students who actively monitor their thinking perform better in mathematics because they are more aware of their strengths and weaknesses and more likely to adjust strategies when difficulties arise. Veenman and Kerseboom (2020) showed that metacognitive training in mathematics classrooms improves academic performance as students learn to plan, evaluate, and adjust their approaches. In practice, this means that reflection and debriefing are not optional extras after problem-solving; they are where students consolidate a method, name what worked, and build a toolbox they can reuse. Taylor and Roberts (2022) made the same point for reflection and metacognition, arguing that deliberate review transforms a single solved problem into a generalisable strategy. The present study therefore treats reflection, scaffolding, and strategic questioning not as isolated techniques but as mutually reinforcing moves that help students think about their own thinking.

Word Problems: Where Language Meets Mathematics

Word problems are a critical bridge between abstract mathematical concepts and practical applications, but they are also among the most challenging tasks for learners. Verschaffel et al. (2020) surveyed the field and emphasised that solving a word problem requires interpreting textual information, extracting relevant data, ignoring extraneous detail, and formulating a mathematical model. Jaffe and Bolger (2023) reviewed cognitive and linguistic factors in arithmetic word-problem solving and showed that reading comprehension, vocabulary, and the ability to build a mental representation of the situation all matter, especially for younger learners who are still consolidating both literacy and numeracy.

Boonen et al. (2016) found that visual representation type, spatial ability, and reading comprehension jointly shape word-problem performance at the item level, which helps explain why a student may succeed on a purely numerical

item yet struggle when the same mathematics is embedded in text. Chen and Jaffe (2023) added that task characteristics, such as the familiarity of the context and the complexity of the language, systematically affect performance, so word-problem difficulty is never purely mathematical.

Students frequently struggle to translate verbal descriptions into mathematical expressions, and teachers sometimes avoid word problems because of the perceived difficulty students have with them (Boonen et al., 2016). That avoidance is costly because it deprives students of the very practice they need. A more productive response is to teach the translation explicitly: how to identify the question, mark the numbers, paraphrase the situation, and choose a representation before computing. Bernardez (2021) and schema-based instruction offer one structured route. In schema-based instruction, students learn to categorise word problems into types or schemas, recognise the underlying structure, and retrieve an appropriate solution strategy, which improves efficiency and transfer. Powell and colleagues' work on schema-broadening instruction, cited in the original study, supports the same conclusion: when students learn the structure, they are less likely to rely on superficial keyword matching and more likely to reason about quantities and relationships. Vandewalle et al. (2022), as cited in the original study, warned that traditional keyword strategies alone can encourage surface-level processing unless they are combined with context-driven analysis and metacognitive modelling.

Visual and verbal scaffolds complement schema work. Drawing a simple picture, using a bar model, or completing a graphic organiser that separates knowns, unknowns, and the goal helps students externalise the problem and reduce working-memory load, consistent with cognitive load theory (Sweller, 1988; Gupta & Zheng, 2020). Moschkovich (2012) reminded educators that mathematical language itself must be taught: students need the precise meaning of terms and phrases if they are to interpret problems accurately and choose suitable strategies. This is especially important for learners with lower verbal or reading proficiency, for whom visual scaffolding can connect a situation to a sequence of operations without requiring them to hold every linguistic detail in mind at once. Storyboarding and other visual strategies that map out problem-solving steps have been reported to improve engagement and comprehension, particularly for students who struggle with abstract reasoning (Fadzil, 2025, as cited in the original study; Ghafar et al., 2023, as cited in the original study).

Teaching Strategies That Strengthen Both Domains

Several instructional strategies recur in the literature as effective across both general problem-solving and word problems. First, structured problem-solving routines give students a predictable sequence, such as identifying the unknown, selecting relevant information, choosing an operation, and checking the answer (Purcell, 2024). Second, modelling and think-alouds make expert thinking audible: the teacher articulates what they notice, what they are unsure about, and how they decide, showing patience and persistence when stuck. Third, scaffolding provides just enough support at first and gradually removes it as students gain independence, reducing initial cognitive overload while preserving productive struggle. Fourth, strategic questioning, including open-ended questions and brief Number Talks, prompts deeper reasoning and exposes students to multiple ways of seeing the same problem. Fifth, real-world contexts make mathematics relevant and give students a reason to care about the answer, while a growth mindset, praising effort and perseverance rather than only correct answers, helps students see challenges as opportunities (Boaler, 2023; Ramirez, 2022).

Collaborative learning, in which students discuss strategies, give feedback, and learn from peers, adds a social dimension: Siller and Ahmad (2024) and Johnson and Anderson (2023) reported benefits for achievement and attitude when collaboration is structured rather than merely grouping students together.

Two cautions qualify this optimistic picture. First, strategies interact. A teacher who models well but never asks students to generate their own strategies may produce dependence rather than independence. A teacher who encourages multiple strategies but never consolidates them through reflection may leave students with a collection of tricks rather than a coherent understanding. Second, context matters.

What counts as the right amount of scaffolding, the right moment to withdraw it, or the right representation depends on students' prior knowledge, language proficiency, and the mathematical demand of the task. This is why student perception is a useful, if partial, source of evidence: it captures whether strategies were experienced as consistent and

recognisable, not merely whether they were intended. The present study builds on this literature by measuring, in one school and with one instrument, how consistently students perceived ten problem-solving strategies and seven word-problem strategies, whether those perceptions varied by profile, and how the two domains related to each other, so that the proposed tools address the weakest, not merely the most familiar, links.

Together, the literature suggests a clear instructional logic: make the problem-solving process explicit through Polya's stages, manage cognitive load through scaffolds and representations, teach word-problem language and structure through schema work and graphic organisers, and consolidate learning through metacognitive reflection. What is less clear is how these elements cohere in a single classroom from the students' point of view, and whether teachers who are strong in general problem-solving routines are also experienced as strong in the language-heavy work of word problems.

Most prior studies examine one domain, one strategy, or one outcome; fewer test both domains with parallel items and ask whether perceptions differ by age, sex, or section within the same cohort. That gap is the contribution of this study. By documenting both domains, testing group differences, and examining their correlation, it provides a diagnostic basis for two targeted tools: one that makes multiple strategies visible and required, and one that makes reading and paraphrasing a non-negotiable first step before computation.

Methodology:

Research Design

This study used a descriptive-correlational research design. It was descriptive because it described the profile of the student-respondents and their assessment of their mathematics teachers' strategies in teaching problem-solving and word-problem skills as they existed at the time of the study. It was correlational because it tested whether a degree of relationship existed between the two strategy domains and whether assessments differed systematically by profile (Calmorin & Calmorin, 2013, as cited in the original study). A survey was appropriate because the aim was to gather data on present conditions, to establish the value of the facts, and to focus attention on the most important patterns to be reported. No intervention was implemented and no variables were manipulated; the study described and related naturally occurring perceptions.

Participants and Locale

The respondents were 120 learners at Luis Y. Ferrer Jr. West National High School, Division of General Trias City. Ten respondents were drawn from each of twelve sections, Sections A to L, giving an equal allocation across sections. The locale was chosen because the researcher observed a need to enhance learners' problem-solving and word-problem skills there. Participation was voluntary, and respondents could withdraw at any time. The study focused on students' perceptions of their mathematics teachers' strategies; teachers themselves were not surveyed, and no student achievement test was administered, so the findings describe perceived instruction rather than measured student performance.

Instrument

Data were gathered with a researcher-made survey questionnaire developed from a thorough reading of related literature and studies. Part 1 asked for the learner's profile: age, sex, and section. Part 2 asked students to assess their mathematics teachers' strategies using a four-point scale: 4 = Very True of My Math Teachers, 3 = True of My Math Teachers, 2 = Somewhat True of My Math Teachers, and 1 = Not True of My Teachers. Items 2.01 to 2.10 measured problem-solving strategies: structured approach, modelling and think-alouds, real-world context, multiple strategies, collaborative learning, growth mindset, strategic questioning, scaffolding, reflection and debriefing, and visual aids and tools. Items 2.11 to 2.17 measured word-problem strategies: read aloud and paraphrase, chunking and annotating, keyword analysis with caution, visualisation, graphic organisers, modelling/think-alouds, and reverse engineering. The questionnaire was validated by experts and by the researcher's former professors in graduate school, and it was revised in light of comments and suggestions from the proposal defence panel.

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 7

Data Gathering and Ethics

After securing approval from the Schools Division Superintendent, the researcher administered the questionnaire to the selected learners, explained its purpose, and assured confidentiality and voluntary participation. Informed consent was obtained, with separate copies for the respondent and the researcher. Responses were treated confidentially, names were optional, and no identifying information is reported in this paper. Completed questionnaires were checked for completeness before encoding.

Data Analysis

Frequency and percentage described the profile of the respondents. Weighted mean described the degree of students' responses on each indicator and on each domain as a whole, using the legend in Tables 2 and 4: 4.00–3.51 = Very True of My Math Teachers / Highly Strategized; 3.50–2.51 = True of My Math Teachers / Somewhat Strategized; 2.50–1.51 = Moderately True of My Math Teachers / Slightly Strategized; and 1.50–1.00 = Least True of My Math Teachers / Not Strategized. Independent-samples t-test tested differences by age (12 vs. 13) and by sex (male vs. female). One-way ANOVA tested differences across the twelve sections. Pearson correlation coefficient (r) tested the relationship between the problem-solving and word-problem assessments. All tests used a 0.05 level of significance, with the borderline sex difference for problem-solving strategies ($p = 0.051$) reported transparently as significant only at the 0.10 level in the original analysis.

Results and Discussion:

This section presents the results in the order of the research objectives: the profile of the respondents, the assessment of problem-solving strategies, differences in that assessment by profile, the assessment of word-problem strategies, differences in that assessment by profile, and the relationship between the two domains. Each table is followed by interpretation and discussion in light of the literature reviewed. A consistent caution applies throughout: the data are students' perceptions of teaching, not direct observations of lessons and not measures of student achievement. A high mean therefore means that students experienced a strategy as consistently present and recognisable; it does not prove that the strategy caused learning. Likewise, a significant difference by profile means that groups perceived instruction differently; it does not by itself prove that teachers treated groups differently, since age, sex, and section are also associated with developmental maturity, classroom composition, and prior experience. With that caution in mind, the patterns below are both coherent and practically useful.

Profile of the Student-Respondents

Table 1. Profile of the student-respondents in terms of age, sex, and section (N = 120)

Profile	Category	Frequency	Percentage
Age	12	34	28%
	13	86	72%
	Total	120	100%
Sex	Male	49	41%
	Female	71	59%
	Total	120	100%
Section	A	10	8%
	B	10	8%
	C	10	8%
	D	10	8%
	E	10	8%
	F	10	8%
	G	10	8%
	H	10	8%
	I	10	8%
	J	10	8%

	K	10	8%
	L	10	8%
	Total	120	100%

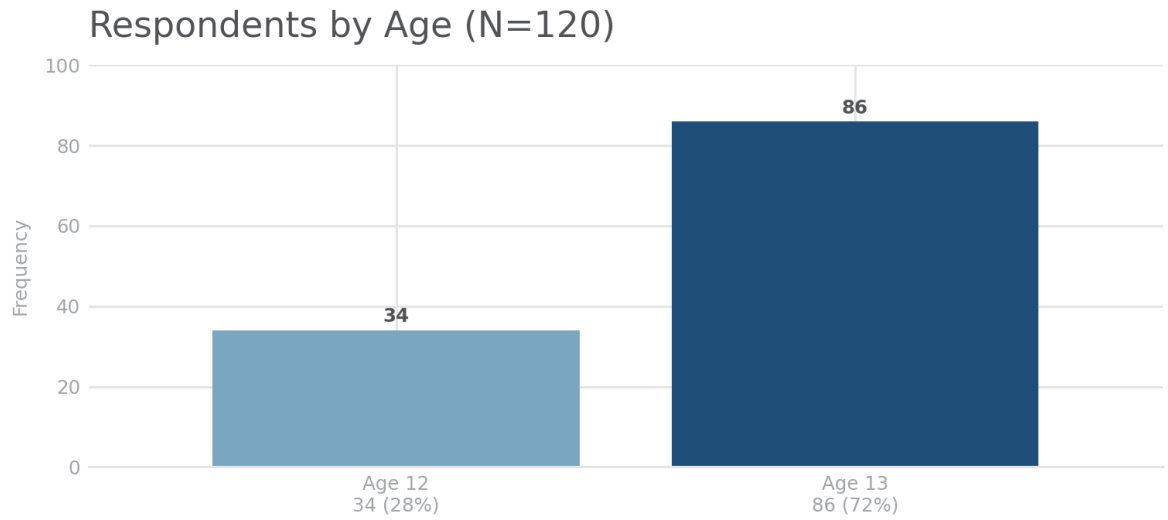


Figure 2. Distribution of respondents by age.

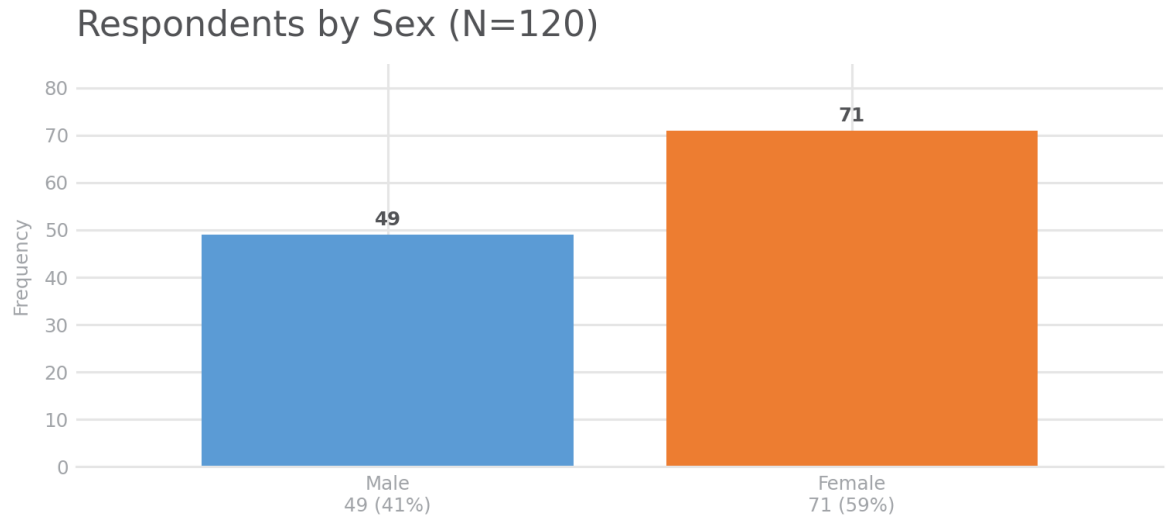


Figure 3. Distribution of respondents by sex.

Table 1 shows a sample that is young, female-skewed, and perfectly balanced by section. By age, 86 respondents (72%) were 13 years old and 34 (28%) were 12 years old. By sex, 71 respondents (59%) were female and 49 (41%) were male, an 18-percentage-point margin in favour of female representation. By section, each of Sections A to L contributed exactly 10 respondents (8% each), totalling 120. The design therefore achieved equal allocation across sections, which is a strength: any overall mean cannot be attributed to over-sampling one classroom, one teacher, or one timetable block. At the same time, the age and sex composition must be kept in view when interpreting group comparisons, because the 12-year-old group is less than half the size of the 13-year-old group and the male group is smaller than the female group, which affects statistical power and the precision of group means.

Taken together, the profile provides a balanced and reliable baseline for evaluating instructional strategies. The sample is large enough for the descriptive and correlational purposes of the study, the section design is disciplined, and the

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 7

age and sex distributions are transparent. The principal cautions are that the sample comes from one school, that it is cross-sectional, and that it relies on student report. Those cautions do not undermine the descriptive value of the results, but they do bound the generalisation: the findings describe how these 120 students experienced their mathematics teachers' strategies, and they generate hypotheses and tools for that school, rather than national estimates of teaching quality.

Assessment of Problem-Solving Strategies

Table 2. Assessment of mathematics teachers' strategies in teaching problem-solving skills

Indicator (Our Math teachers...)	Mean	SD	Qualitative Description	Interpretation	Rank
2.1 Structured Approach	3.32	0.47	True of my Math teachers	Somewhat strategized	6
2.2 Modelling & Think-Alouds	3.30	0.46	True of my Math teachers	Somewhat strategized	7
2.3 Real-World Context	3.38	0.49	True of my Math teachers	Somewhat strategized	4.5
2.4 Multiple Strategies	3.22	0.41	True of my Math teachers	Somewhat strategized	9
2.5 Collaborative Learning	3.24	0.43	True of my Math teachers	Somewhat strategized	8
2.6 Growth Mindset	3.38	0.49	True of my Math teachers	Somewhat strategized	4.5
2.7 Strategic Questioning	3.50	0.50	True of my Math teachers	Somewhat strategized	3
2.8 Scaffolding	3.61	0.49	Very true of my Math teachers	Highly strategized	2
2.9 Reflection & Debriefing	3.66	0.48	Very true of my Math teachers	Highly strategized	1
2.10 Visual Aids & Tools	3.34	0.48	True of my Math teachers	Somewhat strategized	5
Composite Mean	3.39	0.47	True of my Math teachers	Somewhat strategized	—

Legend: 4.00–3.51 Very True / Highly strategized; 3.50–2.51 True / Somewhat strategized; 2.50–1.51 Moderately True / Slightly strategized; 1.50–1.00 Least True / Not strategized.

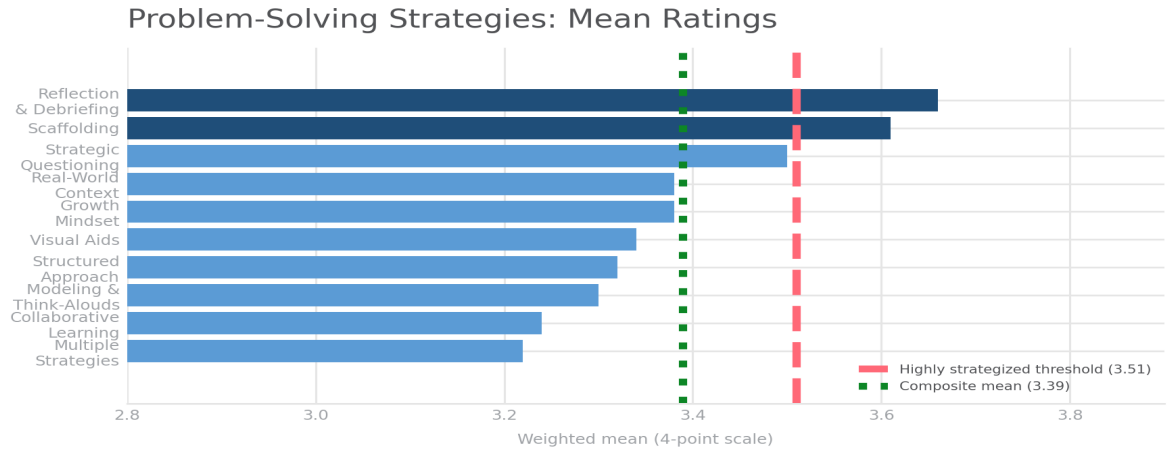


Figure 4. Mean ratings for problem-solving strategies. Only reflection and scaffolding cross the highly strategized threshold.

Table 2 and Figure 4 show that students perceived their teachers' problem-solving strategies as present but not yet fully consistent. The composite mean was 3.39 (SD = 0.47), which falls in the True of My Math Teachers band and is interpreted as somewhat strategized. In practical terms, students were saying that their teachers do use problem-solving strategies, and use them often enough to be recognisable, but not so consistently that students would describe them as a defining, always-present feature of every lesson. The standard deviation of 0.47 is moderate: there is meaningful variation around the mean, consistent with the group differences examined later, but not so much variation that the composite is meaningless. The distance from the highly strategized threshold (3.51) is small, only 0.12 points, which matters for improvement planning: the school does not need a wholesale change of approach; it needs to make already-present practices more deliberate, more frequent, and more evenly distributed across strategies.

The two highest-ranked indicators both crossed into the highly strategized band. Reflection and debriefing ranked first (M = 3.66, SD = 0.48), followed by scaffolding (M = 3.61, SD = 0.49). These are not the easiest strategies to implement, which makes their prominence encouraging. Scaffolding requires the teacher to judge how much support a class needs, to provide it in a form students can use, and then to withdraw it at the right moment. Reflection requires the teacher to protect time at the end of a problem, to ask what worked and what did not, and to help students name a transferable strategy rather than merely confirm an answer. That students rated both as very true suggests that their teachers were doing the two things that the literature identifies as most likely to build independence: giving temporary support and then making learning explicit through review. In Polya's terms, teachers were strong at the beginning of the cycle, providing a bridge into the problem, and at the end of the cycle, looking back. The middle of the cycle, where students must generate and compare alternative plans, is where the ratings fall, as discussed below.

Strategic questioning ranked third (M = 3.50, SD = 0.50), just below the highly strategized threshold. The item included open-ended questions and Number Talks, brief mental-mathematics discussions that prompt deeper reasoning. A mean of 3.50 on a four-point scale is high in absolute terms; it means that students regularly experienced questions that asked them to think, not only to recall. Its position, just short of the threshold, suggests that questioning was frequent but perhaps not yet systematically varied: for example, teachers may have asked good questions during whole-class discussion but asked fewer questions that required students to compare strategies, to justify a choice, or to explain why a tempting but incorrect approach fails. Strengthening questioning in those directions would connect directly to the lowest-ranked indicators, because a question such as 'Can you solve it another way, and which way is more efficient here?' simultaneously exercises strategic questioning, multiple strategies, and reflection.

The middle tier, real-world context and growth mindset tied at 3.38, visual aids and tools at 3.34, structured approach at 3.32, and modelling and think-alouds at 3.30, tells a more nuanced story than the composite alone. All are in the somewhat strategized band, and all are above the scale midpoint, so none should be read as absent. Yet their ordering is informative. Real-world context and growth mindset are dispositions and framings that can be communicated relatively briefly: connecting a problem to a familiar scenario, praising effort and perseverance, and framing a challenge as an opportunity. Visual aids and structured approach require materials, routines, and time, and modelling and think-alouds require the teacher to slow down and make private reasoning public, which can feel inefficient under time pressure. The fact that modelling and think-alouds ranked seventh in the general problem-solving domain, while the same kind of modelling ranked second in the word-problem domain (Section 4.4), is a key contrast. It suggests that teachers modelled thinking more visibly when the problem was textual and therefore obviously needed unpacking, and less visibly when the problem was presented as general problem-solving, where the mathematical steps may have seemed self-explanatory. If that interpretation is correct, the remedy is to make modelling a routine for all problem types, not only for word problems, by narrating decisions about representation, strategy choice, and checking in every lesson.

The lowest-ranked indicators are the most important for intervention. Encouraging multiple strategies ranked last (M = 3.22, SD = 0.41), and collaborative learning ranked eighth (M = 3.24, SD = 0.43). Both remain in the somewhat strategized band, so students were not saying that teachers never encouraged alternative methods or never used group work. They were saying that these practices were the least consistent of the ten. The gap between the top and bottom indicators is 0.44 points (3.66 minus 3.22), which is substantial on a four-point scale and larger than the gap between the composite and the highly strategized threshold. In classroom terms, a student who regularly hears 'What did we learn from this problem?' may less regularly hear 'Show me a second way,' or 'Compare your method with your partner's and decide which is more efficient and why.' The pattern points to a classroom culture that is teacher-

structured and reflective but still somewhat convergent: students are well supported to follow a good path and to review it, but less consistently invited to generate divergent paths or to learn from peers' paths. That is precisely the pattern that the Strategy Canvas Board, proposed in Section 4.7, is designed to change, by requiring at least two distinct pathways for a single real-world problem and by making comparison the normal closing routine.

Differences in Problem-Solving Assessment by Profile

Table 3. Differences in the assessment of problem-solving strategies by profile

Grouping	Test	Statistic	p-value	Decision	Interpretation
Age (12 vs. 13)	Independent-samples t-test	t = -1.89	0.061	Accept H0	Not significant
Sex (Male vs. Female)	Independent-samples t-test	t = -1.92	0.051	Reject H0 at 0.10*	Significant (borderline at 0.05)
Section (A to L)	One-way ANOVA	F = 1.321	0.247	Accept H0	Not significant

** The original table marks the sex difference as Reject H0 / Significant. At the conventional 0.05 level p = 0.051 is borderline; it is significant at the 0.10 level. It is reported transparently here and interpreted cautiously.*

Table 3 shows that age and section did not significantly differentiate students' assessment of problem-solving strategies, while sex did, albeit borderline at the conventional level. For age, the independent-samples t-test comparing 12- and 13-year-olds yielded t = -1.89, p = 0.061. Because p exceeds 0.05, the null hypothesis is accepted: there is no statistically significant difference in how the two age groups perceived their teachers' problem-solving strategies. For section, one-way ANOVA across Sections A to L yielded F = 1.321, p = 0.247, again leading to acceptance of the null hypothesis: group placement or classroom track did not produce significant variance in student evaluations. For sex, the t-test comparing male and female respondents yielded t = -1.92, p = 0.051. The original analysis marked this as Reject H0 / Significant. Strictly at alpha = 0.05, p = 0.051 is just above the threshold and is best described as borderline or marginally significant; at alpha = 0.10 it is significant. The substantive point is that male and female students did not perceive general problem-solving instruction in exactly the same way, and the difference is large enough to warrant attention even though it sits on the conventional boundary.

The absence of a section difference is also informative. With twelve sections and equal allocation, the ANOVA had a fair opportunity to detect differences if they were substantial. That it did not suggests that mathematics teachers' general problem-solving strategies were applied with reasonable consistency across classrooms, or at least that students in different sections experienced them similarly. This is consistent with a school in which teachers share a curriculum, plan together, or follow common routines for problem-solving. It does not prove that every teacher taught identically; it proves that section membership, as a proxy for classroom assignment, did not systematically shift perceptions. For school leaders, this is reassuring: an intervention to strengthen multiple strategies and collaboration can be designed as a school-wide routine rather than as a remedy for a few outlier sections, because the baseline is relatively uniform.

The sex difference, though borderline, should not be dismissed. In the original study's discussion, gender differences were noted in engagement with specific problem-solving steps and in responsiveness to collaborative versus independent techniques, with male and female students reacting differently to structured heuristics. In the present data, the direction of the t-value is negative, but without group means the study does not report which sex rated strategies higher for general problem-solving. The prudent interpretation is therefore that male and female learners perceived, processed, or responded differently to problem-solving strategies, and that teachers should check whether their strategies are equally accessible and engaging for both groups. Practical checks include whether questioning is distributed equitably, whether collaborative roles rotate so that both boys and girls practise explaining and listening, whether examples and contexts appeal across interests, and whether praise for effort and perseverance is given for the same behaviours in both groups. Future research should report group means, effect sizes, and confidence intervals for the sex comparison so that the practical magnitude of the difference, not only its p-value, can be judged.

Assessment of Word-Problem Strategies

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 7

Table 4. Assessment of mathematics teachers' strategies in teaching word-problem skills

Indicator (Our Math teachers...)	Mean	SD	Qualitative Description	Interpretation	Rank
2.11 Read Aloud & Paraphrase	3.33	0.47	True of my Math teachers	Somewhat strategized	6
2.12 Chunking & Annotating	3.44	0.50	True of my Math teachers	Somewhat strategized	5.5
2.13 Keyword Analysis (with caution)	3.44	0.50	True of my Math teachers	Somewhat strategized	5.5
2.14 Visualisation	3.54	0.50	Very true of my Math teachers	Highly strategized	4
2.15 Graphic Organisers	3.78	0.41	Very true of my Math teachers	Highly strategized	1
2.16 Modelling / Think-Alouds	3.69	0.46	Very true of my Math teachers	Highly strategized	2
2.17 Reverse Engineering	3.57	0.50	Very true of my Math teachers	Highly strategized	3
Composite Mean	3.54	0.48	Very true of my Math teachers	Highly strategized	—

Legend: as in Table 2. Note: Graphic Organisers is listed as Highly strategized in the original table despite the qualitative label True; its mean of 3.78 places it in the highly strategized band.

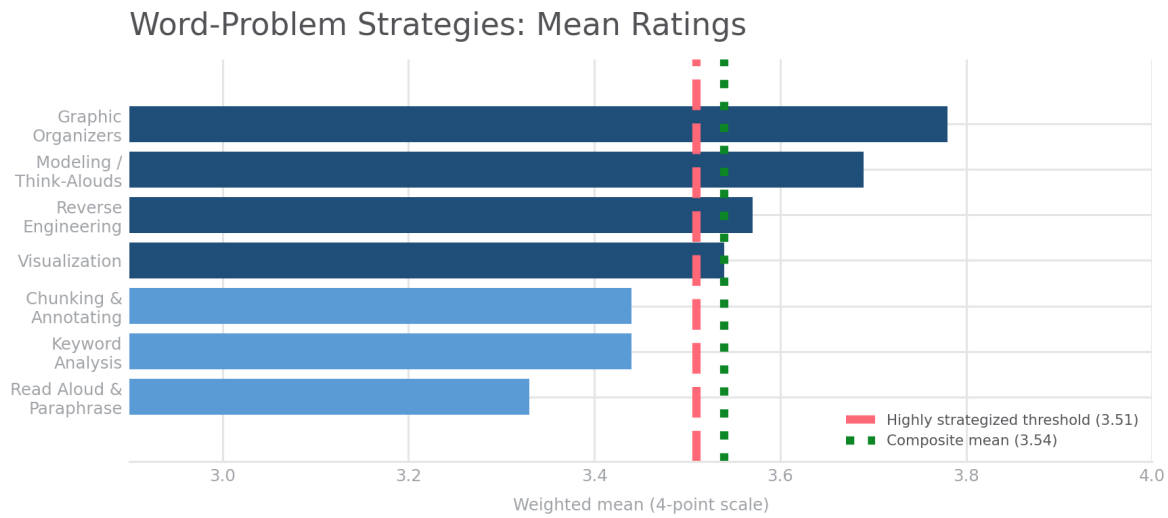


Figure 5. Mean ratings for word-problem strategies. Four of seven indicators cross the highly strategized threshold.

Table 4 and Figure 5 show a stronger and more sharply differentiated picture for word-problem strategies than for general problem-solving strategies. The composite mean was 3.54 (SD = 0.48), which falls in the Very True of My Math Teachers band and is interpreted as highly strategized. Four of the seven indicators crossed the highly strategized threshold, compared with only two of ten in the problem-solving domain. The gap between the two composite means is 0.15 points (3.54 minus 3.39), which is modest in absolute terms but meaningful in combination with the difference in the proportion of highly strategized indicators and with the negative correlation discussed in Section 4.6. Students experienced word-problem instruction as more consistently structured, more visibly modelled, and more reliant on concrete representations than general problem-solving instruction.

Graphic organisers ranked first by a clear margin (M = 3.78, SD = 0.41), the highest mean of any indicator in either domain. The item described using organisers to structure knowns, unknowns, and the goal. Its high rating and relatively low standard deviation (0.41, the lowest in the word-problem set) indicate broad agreement: most students recognised this tool and experienced it as a regular part of word-problem lessons. This is pedagogically important because a graphic organiser does several jobs at once. It externalises the problem structure, reducing working-memory load; it forces students to distinguish what is given from what is asked, which is the first step in Polya's model; it provides a common language for teacher and students to discuss a problem ('What goes in the unknown box?'); and it creates a visible product that the teacher can check quickly. The finding aligns with schema-based instruction and with evidence that visual organisers and schema-broadening instruction are highly effective for word-problem comprehension, because they help students see the underlying structure rather than chase keywords.

Modelling and think-alouds ranked second (M = 3.69, SD = 0.46). In the word-problem domain, this item included teacher modelling of thought processes and the use of Think-Aloud Pair Problem Solving (TAPPS), in which students verbalise reasoning in pairs. Its high ranking, in contrast to the seventh-place ranking of modelling in the general problem-solving domain, suggests that teachers made their thinking most audible when the problem was textual. That is a sensible adaptation: when a problem is presented as words, the teacher cannot assume that students will see the mathematical structure, so the teacher narrates the translation from words to quantities to operations. The implication is that the same modelling habit should be made explicit for non-textual problems as well, and that TAPPS should be protected as a routine, not treated as an occasional activity. When students verbalise reasoning in pairs, they practise the language of mathematics, hear alternative phrasings, and catch misinterpretations before they become calculation errors.

Reverse engineering ranked third (M = 3.57, SD = 0.50). The item described giving the answer first and asking students to create a problem that leads to it, fostering deeper understanding. This is a demanding strategy because it reverses the usual direction of work and requires students to think about the conditions that would produce a given result. Its high rating suggests that teachers in this school were willing to use non-routine, generative tasks for word problems, not only routine translation exercises. Reverse engineering is valuable because it forces students to consider the relationship between a situation and its mathematical model, to check plausibility, and to generate rather than merely solve. It also provides a natural bridge to problem-posing, which Cai and Hwang (2024) found to improve mathematical interpretation and strategy development. A teacher who regularly asks 'What problem could this answer come from?' is teaching students to see mathematics as a modelling activity, not only as a computational one.
Visualisation ranked fourth (M = 3.54, SD = 0.50), just above the highly strategized threshold. The item described drawing simple pictures or diagrams, such as squares for boxes, to represent the problem visually. Its position, below graphic organisers but above the text-processing strategies, suggests that teachers used visual representation as a core part of word-problem instruction, but that the more structured organiser was even more prominent. This is consistent with Boonen et al. (2016), who found that visual representation type matters and that drawing a diagram is not a generic skill but a specific act of representing quantities and relationships. A simple picture that does not encode the quantitative structure is less useful than a diagram that shows how quantities relate. Teachers can therefore strengthen visualisation by teaching a small repertoire of diagrams, such as bar models for part-whole and comparison problems, number lines for change problems, and tables for multi-step problems, and by asking students to explain what each part of the diagram represents before they compute.

The three lowest-ranked word-problem indicators were all text-processing strategies: chunking and annotating and keyword analysis tied at 3.44 (Rank 5.5), and read aloud and paraphrase at 3.33 (Rank 6). All remain in the somewhat strategized band, so they were present, but they were the least consistent of the seven. This pattern is the mirror image of the strength in graphic organisers. Teachers were strong at helping students represent a problem once it had been understood, but less consistent at helping students achieve that understanding through reading. Chunking and annotating, breaking a problem into smaller parts, identifying the question, circling numbers, and underlining key information, is a foundational reading strategy. Keyword analysis, used with caution, can hint at operations but must be subordinated to overall context, otherwise it encourages surface-level processing. Read aloud and paraphrase, having students read problems aloud and restate them in their own words, is the most direct check that a student has understood the situation before computing. That it ranked lowest is the single most actionable finding in the word-problem domain, and it is why the Say-It-Back Protocol is proposed as the second essential tool.

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 7

Differences in Word-Problem Assessment by Profile

Table 5. Differences in the assessment of word-problem strategies by profile

Grouping	Test	Statistic	p-value	Decision	Interpretation
Age (12 vs. 13)	Independent-samples t-test	t = -2.98	0.003	Reject H0	Significant
Sex (Male vs. Female)	Independent-samples t-test	t = -2.18	0.031	Reject H0	Significant
Section (A to L)	One-way ANOVA	F = 0.729	0.705	Accept H0	Not significant

Table 5 shows a different pattern from Table 3: for word-problem strategies, both age and sex significantly differentiated students' assessments, while section did not. For age, the t-test comparing 12- and 13-year-olds yielded $t = -2.98$, $p = 0.003$, well below the 0.05 threshold, leading to rejection of the null hypothesis. The negative t-value indicates that 13-year-olds rated their teachers' word-problem strategies significantly higher than 12-year-olds did. For sex, the t-test yielded $t = -2.18$, $p = 0.031$, also leading to rejection of the null hypothesis: female students reported a significantly different, and higher, assessment than male students. For section, ANOVA yielded $F = 0.729$, $p = 0.705$, leading to acceptance of the null hypothesis: there was no significant difference across the twelve sections.

The age difference for word problems, contrasted with the absence of an age difference for general problem-solving, is the most developmentally informative result in the study. Word problems place a double demand on the learner: they require mathematical reasoning and linguistic interpretation. A 13-year-old, with an additional year of secondary schooling, is likely to have stronger reading comprehension, a larger mathematical vocabulary, more familiarity with word-problem formats, and more practice in translating situations into models. Those advantages would make the teacher's strategies more usable and therefore more noticeable: a graphic organiser is more helpful when the student can already parse the problem well enough to fill it in; a think-aloud is more intelligible when the student can follow the teacher's language; reverse engineering is more accessible when the student has a repertoire of problem types to draw on. A 12-year-old who is still consolidating those foundations may experience the same lesson as faster, denser, or less transparent, and may therefore rate the same strategies lower, not because the teaching was worse, but because the match between teaching and readiness was less secure. This interpretation aligns with research on developmental cognitive maturity and its influence on how students process teacher-guided word-problem interventions, and with the observation that higher grade levels exhibit distinct mathematical reasoning and problem-solving readiness.

The sex difference for word problems is clearer than the borderline sex difference for general problem-solving. At $p = 0.031$, it is significant at the conventional 0.05 level, and the direction indicates that female students rated word-problem strategies higher than male students did. This aligns with findings from secondary mathematics education in the Visayas region, where female students were reported to engage more with step-by-step problem-solving heuristics and to rate teacher instructional strategies higher than their male peers, and with research on gender-based academic self-efficacy in STEM classrooms across Asian secondary schools, which found that female students report higher awareness and appreciation of explicit schema-based strategies in word-problem instruction. Several mechanisms could produce this pattern. Female students may, on average, have stronger reading comprehension at this age, making word-problem strategies more usable. They may be more willing to use structured routines, such as graphic organisers, rather than attempting to solve mentally. They may be more responsive to modelling and think-alouds that emphasise careful, step-by-step reasoning. Or they may simply be more attuned to teacher strategies as a form of classroom engagement. The study cannot distinguish these mechanisms, but it can rule out the simplest dismissal: the difference is not a chance fluctuation in a small subgroup, because both groups are substantial and the p-value is comfortably below 0.05.

The absence of a section difference for word problems ($F = 0.729$, $p = 0.705$) is even stronger than the absence for general problem-solving. The F-value below 1.0 indicates that variation between sections was smaller than variation within sections, which is the opposite of what would be expected if section membership systematically shifted perceptions. This suggests that word-problem instruction, like general problem-solving instruction, was experienced consistently across the twelve sampled classrooms. For school leaders, this consistency is a foundation for school-wide intervention: a routine such as the Say-It-Back Protocol can be introduced across sections with a reasonable expectation that it will be implemented with similar fidelity, and its effects can be compared across sections without

worrying that baseline differences will confound the comparison. For researchers, the null section effect increases confidence that the age and sex effects are not artefacts of section composition, since sections were balanced and did not themselves produce differences.

Relationship Between Problem-Solving and Word-Problem Assessments

Table 6. Relationship between the assessments of problem-solving and word-problem strategies

Variables	Computed r	p-value	Decision	Interpretation
Problem-solving strategies and word-problem strategies	-0.385	< .001	Reject H0	Significant, moderate negative relationship

Table 6 shows a statistically significant, moderate negative relationship between students' assessments of their teachers' problem-solving strategies and word-problem strategies: Pearson $r = -0.385$, $p < .001$, leading to rejection of the null hypothesis. The negative sign means that students who rated general problem-solving strategies higher tended to rate word-problem strategies lower, and vice versa. The magnitude, 0.385, is conventionally described as moderate: it is large enough to be substantively interesting and to rule out the idea that the two domains are independent, but not so large that one domain determines the other. The squared correlation, r-squared = 0.148, indicates that about 15% of the variance in one assessment is linearly associated with variance in the other, leaving about 85% unshared. That balance is important: the two domains are related, but they are not the same construct, and a student's experience of one does not fully predict their experience of the other.

Whatever the mechanism, the practical implication is clear: teachers should deliberately integrate general problem-solving and word-problem instruction so that students experience them as one coherent approach, not as two competing emphases. In practice, integration means using the same language, the same Polya stages, and the same reflection routine for both kinds of problems, and explicitly naming the connection. A lesson might begin with a word problem, use the Say-It-Back Protocol to paraphrase it, use a graphic organiser to represent it, use the Strategy Canvas Board to generate two solutions, and close with the same reflection questions used for general problem-solving: What worked? What did not? What would you do differently next time? When the same routine is used for both domains, students are less likely to perceive a trade-off, and the correlation between the domains should move toward positive, reflecting a shared experience of coherent, well-structured instruction. Future research should test that prediction directly, by measuring both domains before and after an integrated intervention and by examining whether the correlation changes.

The negative correlation also has an implication for assessment and professional development. If teachers are evaluated separately on general problem-solving and on word problems, they may reasonably conclude that they must choose which to prioritise. If they are supported to see the two as one, with shared routines and shared tools, they can invest in practices that serve both. Professional development should therefore model integrated lessons, provide the two tools proposed below as a single package, and give teachers time to practise the transition from paraphrasing to representation to multiple strategies to reflection within a single lesson. That integrated package is the most direct response to the negative correlation: it replaces a perceived trade-off with a visible, teachable sequence that serves both domains at once.

Proposed Essential Tools: From Diagnosis to Classroom Routines

Table 7. Proposed essential tools for applying concepts to real-world scenarios

Tool	Target Indicator	Routine	Real-World Application
Strategy Canvas Board	Multiple Strategies (2.4) and Collaborative Learning (2.5)	Solve one problem in at least two quadrants: Visual/Drawing, Working Backward/Logical, Algebraic, and Estimation/Check; compare and justify.	Budgeting for a community fundraiser: buying supplies with a fixed amount and multiple discounts.

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 5, ISSUE NO. 7

Say-It-Back Protocol (3-Step Flipchart)	Read Aloud & Paraphrase (2.11), Chunking & Annotating (2.12)	1) I read... 2) It means... 3) I need to find... Partner covers numbers and restates goal before computing.	Calculating travel time, fuel costs, and stops for a family road trip from a word-problem scenario.

The findings lead directly to two essential tools, summarised in Table 7. They are called essential tools because they are not additional topics to cover; they are cognitive, procedural, and analytical routines that students use to interpret numbers, decode abstract concepts, and solve real-world problems. Each tool targets the weakest link in its domain, uses the strongest practices in that domain as its foundation, and is designed to be usable within ordinary lesson time without requiring new technology or extensive materials. Both tools were foreshadowed in the original study's output, and both are elaborated here so that a teacher can implement them tomorrow.

The Strategy Canvas Board addresses the lowest-ranked problem-solving indicators: encouraging multiple strategies (M = 3.22) and collaborative learning (M = 3.24). It is a structured template divided into quadrants: Visual/Drawing, Working Backward/Logical, Algebraic, and Estimation/Check. The routine is simple and non-negotiable: for a chosen real-world problem, each student or pair must solve the problem using at least two different quadrants, then compare the pathways and justify which is more efficient, more elegant, or more reliable for this problem. The teacher's role shifts from presenting the best method to orchestrating a comparison: 'Who used a drawing? Who worked backward? What did each method make visible? Where did each method risk an error?' The board makes multiple strategies visible, comparable, and discussable, which is exactly what a verbal exhortation to 'try another way' cannot do. It also structures collaboration: pairs can divide quadrants, then teach each other, so that collaboration has a product and a purpose, not merely a seating arrangement. A concrete application is budgeting for a community fundraiser, where students must figure out how to buy supplies with a fixed amount of money and multiple discounts. One quadrant might use a bar model to show the budget constraint; another might work backward from the total; a third might use an algebraic equation; a fourth might estimate and check. The mathematical content is the same; the thinking is multiplied. The expected outcome is cognitive flexibility: students learn that complex problems can be attacked from multiple angles, that different methods have different strengths, and that choosing a method is itself a mathematical decision.

The Say-It-Back Protocol addresses the lowest-ranked word-problem indicator: read aloud and paraphrase (M = 3.33), together with chunking and annotating (M = 3.44). It is a structured protocol card or flipchart with three sentence stems: (1) I read...; (2) It means...; and (3) I need to find.... The routine is paired: one student reads the problem aloud to a partner, the partner covers the numbers and restates the core goal in their own words, and together they identify the question, circle the numbers, and underline the key information before writing any mathematical symbols. The rule that comprehension precedes computation is made explicit and enforceable: no symbols until the paraphrase is agreed. A concrete application is calculating travel time, fuel costs, and stops for a family road trip based on a word-problem scenario. Before any calculation, students must say what the trip involves, what is known about distance, speed, and fuel, and what the problem is asking them to find. Only then do they choose a representation, such as a graphic organiser or a table, and compute. The expected outcome is a drastic reduction in misinterpretation errors and a habit of checking meaning before calculating. The protocol also directly addresses the age and sex differences reported for word problems: it gives younger students and students who find reading harder a structured, low-stakes way to practise paraphrasing, and it gives all students a shared language for discussing word problems.

Limitations

Four limitations bound the interpretation of these results. First, the study used student perceptions as its sole data source. Perceptions are valuable because they capture the lived experience of instruction, but they are not direct measures of teaching or learning. A student may rate a strategy highly because it was memorable, enjoyable, or recently used, rather than because it was frequent or effective. Future studies should triangulate student ratings with classroom observation, teacher self-report, and analysis of lesson plans and student work. Second, the study did not measure student achievement. It therefore cannot claim that highly rated strategies caused higher performance, nor that the negative correlation reflects a causal trade-off. A follow-up study that links perceived strategies to problem-solving test scores, with pre- and post-tests, would be needed to test causal claims. Third, the sample came from one

school and was cross-sectional. The findings describe 120 students at one point in time and cannot be generalised to other schools, grade levels, or regions without replication. The equal allocation across sections strengthens internal validity within this school, but it does not substitute for multi-school sampling. Fourth, the analysis reported only p-values and means, without group means for the t-tests, effect sizes, confidence intervals, or reliability coefficients for the scales. Those omissions limit judgments about practical magnitude and measurement quality. The borderline sex difference for general problem-solving ($p = 0.051$) is a case in point: without an effect size, it is impossible to know whether the difference is educationally meaningful or merely statistically borderline. Future research should report those statistics routinely, and should test whether the negative correlation replicates when the two domains are measured with independent methods.

Despite these limitations, the study has clear strengths. The sample was balanced across sections, the instrument covered both domains with parallel items, the statistical tests matched the research questions, and the findings are coherent: strong scaffolding and reflection, strong graphic organisers and modelling, weak multiple strategies and paraphrasing, age and sex differences for word problems, and a negative correlation that points to a need for integration. Those patterns are too systematic to be dismissed as noise, and they generate a focused, testable intervention in the form of two tools. The limitations therefore do not undermine the study; they define the next study.

Conclusion:

The 120 student-respondents were predominantly 13 years old (72%) and female (59%), with exactly ten respondents from each of Sections A to L. That balanced, section-stratified sample provides a credible baseline for evaluating instruction in this school. Mathematics teachers' problem-solving strategies were perceived as somewhat strategized (composite mean = 3.39), with reflection and debriefing (3.66) and scaffolding (3.61) as highly strategized strengths, and encouraging multiple strategies (3.22) and collaborative learning (3.24) as the least consistent practices. Word-problem strategies were perceived as highly strategized (composite mean = 3.54), with graphic organisers (3.78), modelling and think-alouds (3.69), reverse engineering (3.57), and visualisation (3.54) as strengths, and read aloud and paraphrase (3.33) as the least consistent practice. Together, these results describe teachers who are strong at supporting, structuring, modelling, and reviewing, and who now need to strengthen the practices that require students to generate ideas and language for themselves.

Tests of group differences showed that age and section did not differentiate perceptions of general problem-solving strategies, while sex showed a borderline difference ($t = -1.92$, $p = 0.051$) that warrants cautious attention and replication. For word-problem strategies, both age ($t = -2.98$, $p = 0.003$) and sex ($t = -2.18$, $p = 0.031$) were significant, with 13-year-olds and female students rating strategies higher, while section showed no significant difference ($F = 0.729$, $p = 0.705$). The pattern suggests that word-problem instruction, with its double demand of mathematics and language, is more sensitive to developmental maturity and to sex-related differences in reading and strategy use than general problem-solving routines are, while classroom assignment itself does not systematically shift perceptions. Teachers and school leaders should therefore design word-problem support that is differentiated by readiness and inclusive by sex, rather than assuming that a single, uniform routine will be experienced identically by all learners.

The two domains were significantly and moderately negatively correlated ($r = -0.385$, $p < .001$), meaning that students who rated one domain higher tended to rate the other lower. This finding should not be read as proof that teachers must choose between general problem-solving and word problems, but as evidence that students experienced the two as separate emphases rather than as one coherent approach. The remedy is integration: using the same Polya language, the same reflection routine, and the same tools for both kinds of problems, so that paraphrasing, representing, generating multiple strategies, and reflecting become a single, recognisable way of doing mathematics. The two proposed tools, the Strategy Canvas Board for multiple strategies and collaboration, and the Say-It-Back Protocol for reading and paraphrasing before computation, are designed to achieve that integration within ordinary lesson time. If implemented with deliberate comparison and reflection, and evaluated with pre- and post-measures of both perceptions and achievement, they offer a practical, evidence-based next step for strengthening problem-solving and word-problem skills and for helping students apply mathematical concepts to real-world scenarios.

References:

Anderson, P. (2022). Building procedural fluency through conceptual understanding. Mathematics Education Journal, 45(2), 78–92.
Arianto, F., & Hanif, M. (2024). Evaluating metacognitive strategies and self-regulated learning to predict primary school students' self-efficacy and problem-solving skills in science learning. Journal of Pedagogical Research, 8(3), 301–319. https://doi.org/10.33902/JPR.202428575
Bachmann, H. (2021). Polya's four-step problem solving method. https://henrikbachmann.com
Bernardez, K. (2021). The best word problem strategies that build strong mathematicians. https://www.kaylabernardez.com/word-problem-strategies
Boaler, J. (2023). Mathematical mindsets: Unleashing students' potential through creative math, inspiring messages, and innovative teaching. Jossey-Bass.
Boonen, A. J. H., van Wesel, F., Jolles, J., & van der Schoot, M. (2016). The role of visual representation type, spatial ability, and reading comprehension in word problem solving: An item-level analysis in elementary school children. International Journal of Educational Research, 77, 15–26.
Bos, D. O., Lariviere, M., & Lin, X. (2023). A meta-analysis of mathematics interventions: Examining the efficacy of interventions for students with mathematics difficulties. Educational Psychology Review. https://link.springer.com/article/10.1007/s10648-023-09843-0
Cai, J., & Hwang, S. (2024). Enhancing mathematical problem posing competence: A meta-analysis of intervention studies. International Journal of STEM Education, 11, Article 48. https://doi.org/10.1186/s40594-024-00507-1
Capone, R., De Angelis, V., & Petrillo, F. (2021). Enhancing problem-solving skills in primary school students: An intervention study. Journal on Mathematics Education, 12(1), 1–16.
Cardino, J. M., & Ortega-Dela Cruz, R. A. (2020). Understanding of learning styles and teaching strategies towards improving the teaching and learning of mathematics. LUMAT: International Journal on Math, Science and Technology Education.
Carifio, J. (2015). Updating, modernizing, and testing Polya's theory of mathematical problem solving. Journal of Educational and Human Development, 4(3), 102–109.
Chen, L., & Jaffe, E. (2023). Task characteristics associated with mathematical word problem-solving performance: A systematic review and meta-analysis. Educational Psychology Review, 35, 1–28. https://doi.org/10.1007/s10648-024-09954-2
Clark, D. (2020). Fostering adaptive reasoning in problem-solving tasks. International Journal of Mathematics Education, 38(4), 112–126.
De Vink, E., Schindler, M., & Lilienthal, A. J. (2022). Mathematics learning orientation: Mathematical creative thinking and its relation to problem-solving and reasoning. Journal of Educational Psychology, 114(3), 453–470.
Dendane, A. (2009). Mathematical problem-solving strategies among student teachers. https://files.eric.ed.gov/fulltext/EJ1208772.pdf
Dulger, M., & Ogan-Bekiroglu, F. (2023). Examination of the relationship between metacognitive awareness and problem-solving skills in mathematics. Journal of Educational Psychology, 115(4), 589–605.
Durgin, J. (n.d.). How to teach word problems: Strategies for elementary teachers. https://jodidurgin.com/ideas-strategies-how-to-teach-students-to-solve-word-problems/
Education Week. (2021). Four teacher-recommended instructional strategies for math. https://www.edweek.org/teaching-learning/opinion-four-teacher-recommended-instructional-strategies-for-math/2021/07
Edutopia. (2021). A strategy for teaching math word problems. https://www.edutopia.org/article/strategy-teaching-math-word-problems
Garcia, L. (2023). The role of strategy instruction in developing mathematical problem-solving skills. Educational Research Review, 50, 145–159.
Geiger, V., Goos, M., & Forgasz, H. (2015). A rich interpretation of numeracy for the 21st century: A survey of the state of the field. ZDM Mathematics Education, 47, 531–548.
Goldenberg, E. P., Mark, J., & Cuoco, A. (2003). Mathematical problem solving: A guide to developing strategies and skills. National Council of Teachers of Mathematics.
Gupta, U., & Zheng, R. Z. (2020). Cognitive load in solving mathematics problems: Validating the role of motivation and the interaction among prior knowledge, worked examples, and task difficulty. European Journal of STEM Education, 5(1), 05.
Harris, R., & King, T. (2024). Strengthening problem-solving skills in mathematics education. Journal of Educational Psychology, 60(1), 25–40.

Hattie, J. (2021). Visible learning for mathematics: What works best to optimize student learning. Corwin Press.
Hernandez, M., & White, S. (2021). Strategic competence in mathematical problem-solving: A comparative study. Mathematics Teaching Journal, 29(3), 67–80.
Jaffe, E., & Bolger, M. (2023). Cognitive processes, linguistic factors, and arithmetic word problem solving: A review. Educational Psychology Review, 35, 1–25. https://doi.org/10.1007/s10648-023-09821-6
Johnson, D., & Anderson, R. (2023). The impact of collaborative problem-solving in mathematics classrooms. Journal of Mathematics Education, 15(2), 45–63.
Jones, C. (2023). Mathematical problem-solving and cognitive development. Journal of Cognitive Mathematics, 55(2), 90–105.
Karabulut, R., & Özkubat, U. (2019). Mathematical problem-solving processes of students with special needs: A cognitive strategy instruction. International Electronic Journal of Elementary Education, 12(1), 1–10.
Kola, A. J., & Sunday, O. S. (2020). Assessment strategies for enhancing students' mathematical problem-solving skills. Eurasia Journal of Mathematics, Science and Technology Education, 16(12), em1919. https://doi.org/10.29333/ejmste/9239
Leppäaho, H. (2018). Teaching mathematics through problem solving: A review of the literature. Journal of Mathematical Behavior, 52, 379–399.
Lewis, G., & Patel, N. (2023). Encouraging mathematical discourse to enhance adaptive reasoning skills. Mathematics Teacher Education, 41(2), 74–89.
Li, X., & Ma, Y. (2020). The effectiveness of digital tools in enhancing mathematics problem-solving skills. Computers & Education, 150, 103843.
Liang, Z., Zhang, J., & Zhang, X. (2022). Analogical math word problems solving with enhanced problem-solution association. arXiv:2212.00837. https://doi.org/10.48550/arXiv.2212.00837
Martinez, J., & Lopez, K. (2021). The importance of conceptual understanding in mathematical problem-solving. Advances in Mathematics Education, 27(4), 102–118.
Miller, S. (2024). Word problems and procedural fluency: The intersection of skills in mathematics learning. Educational Studies in Mathematics, 58(3), 130–145.
Morrison, K., & Sullivan, P. (2022). The role of practice in developing problem-solving skills in mathematics. Mathematics Teaching Research Journal, 18(1), 23–40.
Moschkovich, J. N. (2012). Mathematics, the Common Core, and language: Recommendations for mathematics instruction for ELs aligned with the Common Core. Understanding Language, Stanford University.
Näveri, L., & Leppäaho, H. (2011). Developing mathematical problem-solving skills in primary school. https://files.eric.ed.gov/fulltext/EJ1358674.pdf
Nguyen, T., & Davis, R. (2023). Integrating procedural and conceptual knowledge in mathematics instruction. Mathematics Pedagogy Review, 35(2), 88–101.
Old Dominion University. (2021). Doing what works: Five evidence-based strategies to specially design mathematics instruction. https://ttac.odu.edu/curriculum-and-instruction/math/doing-what-works-five-evidence-based-strategies-to-specially-design-mathematics-instruction/
Parker, H., & Evans, J. (2021). Fostering a productive disposition towards mathematics: Strategies for educators. Journal of Mathematical Attitudes, 19(3), 56–72.
Polya, G. (1945). How to solve it: A new aspect of mathematical method. Princeton University Press.
Prakitipong, N., & Nakamura, S. (2006). Analysis of mathematics performance of grade five students in Thailand using Newman procedure. Journal of International Cooperation in Education, 9(1), 111–122.
Purcell, J. (2024). Teaching problem-solving strategies in math: A step-by-step guide. https://purcellmath.com/techniques-of-instruction/effective-classroom-techniques/teaching-problem-solving-strategies-in-math-a-step-by-step-guide/
Ramirez, O. (2022). Developing a growth mindset in mathematics classrooms. Educational Psychology Review, 49(1), 78–92.
Santos-Trigo, M. (2024). Mathematical problem solving: Current themes, trends, and research perspectives. In M. Santos-Trigo & L. Moreno-Armella (Eds.), Problem solving in mathematics education (pp. 1–20). Springer.
Schoenfeld, A. H. (2024). Problem solving in mathematics education. ZDM–Mathematics Education. https://doi.org/10.1007/s11858-024-01578-8
Sidenvall, J. (2019). Literature review of mathematics teaching design for problem solving and reasoning. Nordic Studies in Mathematics Education, 24(1), 51–74.

Siller, H.-S., & Ahmad, S. (2024). Analyzing the impact of collaborative learning approach on grade six students' mathematics achievement and attitude towards mathematics. EURASIA Journal of Mathematics, Science and Technology Education.
Smith, A., & Brown, L. (2022). Enhancing problem-solving abilities through cognitive development strategies. Mathematics Education Research Journal, 36(2), 67–83.
Smith, L., & Brown, T. (2023). The effectiveness of lecture-based versus inquiry-based learning in secondary mathematics education. Educational Research Quarterly, 47(1), 78–96.
Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257–285.
Taylor, B., & Roberts, M. (2022). Reflection and metacognition in mathematical problem-solving. International Journal of STEM Education, 46(1), 97–112.
The Australian. (2025). Back-to-basics curriculum on offer to all. https://www.theaustralian.com.au/education/nsw-offers-other-states-its-slimmeddown-school-syllabus/news-story/187c58ab6dd50aabdd5fa44cf4a0eea4
Toh, T. L., Kaur, B., & Toh, P. C. (2023). Problem solving in mathematics education: Tracing its development in Singapore. ZDM–Mathematics Education.
Turel, Y., & Johnson, D. (2021). Step-by-step algorithmic instruction and its impact on mathematical problem-solving proficiency. International Journal of Mathematics Teaching, 12(3), 112–127.
Veenman, M. V. J., & Kerseboom, J. (2020). Enhancing metacognitive awareness in mathematics classrooms: Effects on problem-solving performance. European Journal of Psychology of Education, 35(3), 487–505.
Verschaffel, L., Depaepe, F., & Van Dooren, W. (2020). Word problems in mathematics education: A survey. ZDM–Mathematics Education, 52, 1–12. https://doi.org/10.1007/s11858-020-01130-4
Vintere, A. (2018). A constructivist approach to the teaching of mathematics to boost competences needed for sustainable development. Proceedings of the Latvian Academy of Sciences, 72(1), 1–9.
Vula, E., & Berisha, V. (2020). Mathematics word problem solving through collaborative action research. Practical Teacher Research, 2(1), 34–46.
Watson, D. (2023). Mathematical confidence and problem-solving persistence: The role of attitude in student success. Journal of Mathematics Education, 50(2), 120–135.
We Are Teachers. (2021). 14 effective ways to help your students conquer math word problems. https://www.weareteachers.com/math-word-problems
Wijayanti, P., Herman, T., & Usdiyana, D. (2017). The effect of CORE learning model on students' mathematical problem-solving ability. Journal of Physics: Conference Series, 895(1), 012061.
Wilson, K. (2024). Critical thinking and adaptive reasoning in mathematical problem-solving. Teaching Mathematics for Understanding, 22(4), 140–158.
Zhang, Y., & Lin, C. (2023). Real-world applications of mathematical problem-solving: Enhancing student engagement and learning outcomes. Mathematics Education Research Journal, 35(1), 67–82.