Solving Multistep Problems: What Will It Take?

Rebecca Seah
RMIT University
rebeca.seah@rmit.edu.au

Marj Horne
RMIT University
marj.horne@rmit.edu.au

Problem solving and reasoning are two key components of becoming numerate. Reports obtained from international assessments show that Australian students' problem solving ability is in a long-term decline. There is little evidence that teachers are embracing problem solving as part of the classroom routine. In this study, we analyse 598 Year 7 to 10 students' responses to a measurement task using Sfard's commognition framework. Four implications lead to recommendations on how to support curriculum, assessment and pedagogical alignment.

Problem solving has always been a key feature of mathematics learning. With the advent of STEM education, globalisation, and the continuing uncertainty caused by the pandemic, it has become central to many educational reform agendas. Weber and Leikin (2016) distinguish four broad traditions of problem solving based research: (i) problem solving as a research tool to investigate other constructs such as understanding of a concept; (ii) as an object of study in terms of the *resources* (knowledge and processes), *heuristics* (strategies), *metacognition*, and *beliefs* (about mathematics); (iii) the activity of problem posing; and (iv) as a didactical tool to teach conceptual understanding. Over the past two decades, problem solving research in Australasia evolved from being hidden by other research, to classroom practice in curricular reform and in the development of more general theoretical conceptions of problem solving as an activity, which is commonly termed as *working mathematically* (Clarke et al., 2007). It was hoped that subsuming problem solving to a general proficiency domain may help emphasise curricular, assessment and pedagogical alignment. The result of taking such a stand may explain the absence of a focus publication on problem solving as an object of study in Australasia since 2008 (Makar et al., 2020). Instead, problem solving appears peripherally in other studies such as STEM education, rich tasks, and learning progression research. Makar et al. also reported a 16-year absence of domain focus research in Algebra and Geometry and Measurement albeit a renewed interest in the latter appeared recently in the form of spatial reasoning.

The scarcity of focus research has dire consequences as there is little evidence that teachers are embracing problem solving effectively as part of their classroom routine. International assessments alert us about Australian students' weakness in geometry and algebra and that their problem solving ability is in a long-term decline (Thomson et al., 2016, 2017). It is estimated that 53% of young Australians did not possess the numeracy skills essential to work effectively in a modern economy (Lamb et al., 2020). Sfard (2021) rightly pointed out, the devil is in the detail. If problem solving is to be nurtured within a general label, attention must be placed in the inter-discursive gaps in the teaching-learning process.

In this paper, we analyse data collected from the *Reframing Mathematical Futures II* (RMF II) project (Siemon et al., 2018). Our purpose is to elicit evidence of students' problem solving and reasoning to highlight gaps when *working mathematically*. This leads to practical implications on how to support curriculum, assessment, and pedagogical alignment.

Theoretical Framework

The *Australian Curriculum: Mathematics* defines problem solving as the ability to make choices, interpret, formulate, model and investigate problem situations, and communicate solutions effectively (Australian Curriculum Assessment and Reporting Authority (ACARA),

2022. N. Fitzallen, C. Murphy, V. Hatisaru, & N. Maher (Eds.), *Mathematical confluences and journeys* (Proceedings of the 44th Annual Conference of the Mathematics Education Research Group of Australasia, July 3–7), pp. 490–497. Launceston: MERGA.

n.d). Together with understanding, fluency, and reasoning, problem solving is seen as an integral part of becoming proficient in mathematics across the three content strands.

The teaching of problem solving can be traced back to Polya's (1945) *How to solve it*, where students are taught to: 1) understand the problem; 2) make a plan; 3) carry out the plan; and 4) evaluate its effectiveness. Jonassen (2011) cautions on the fallacy of treating problem solving as a reproducible, algorithmic process, assuming that all problems are solved in pretty much the same way, and that generalisable processes can be applied in different contexts with different types of problems in order to yield similar results. He maintains that such views underestimate the role of domain knowledge and pattern recognition (analogical reasoning), resulting in the misrepresentation of knowledge and inhibiting transfer of skills learned. Successful problem solving needs two critical attributes, mental representation of the problem and manipulation and testing of the mental model of the problem to generate a solution.

For Polya, a problem worth solving is one where the solution is unknown to the solver. Consider the Drink Bottles task (coded as GSODA to mean Geometry Soda) in Figure 1 where the context should be familiar to most students. The solution for Part a is fairly straightforward, involving the application of decimal fraction knowledge and proportional reasoning. We acknowledge the difficulties students face when learning fraction and decimal arithmetic (Lortie-Forgues et al., 2015). Part b has several possible solutions and requires an understanding of array, dimensionality, units, and multiplication. In Part c, while the conversion between capacity and mass is given, understandings of magnitudes of measures, that 1L is equivalent to 1000g, and the processes needed to obtain a correct result are needed. To effectively address this task, a solver needs a robust understanding of mathematical concepts and the ability to carry out the processes flexibly, accurately, efficiently, and appropriately. Further, since it is possible to solve the problem by stating the answer without explanation, the task requires students to explain their reasoning. Reasoning is observed when:

> ... students explain their thinking, when they deduce and justify strategies used and conclusions reached, when they adapt the known to the unknown, when they transfer learning from one context to another, and when they compare and contrast related ideas and explain their choices (Australian Curriculum Assessment and Reporting Authority [ACARA], n.d).

Drink bottles

A 1.25-liter bottle of soda water is 285 millimeters high and has a diameter of 85 millimeters.

a. [SODA1]
   How many bottles would be needed to fill a 10-liter container with soda water? Explain your reasoning.
b. [SODA2]
   What are the dimensions of a carton that would firmly hold 12 bottles of soda water? Explain your reasoning.
c. [SODA3]
   One milliliter of water weighs 1 gram and each empty bottle weighs 80 grams. The cardboard in the box weighs 750 grams. How heavy would the full carton of 12 bottles of soda water be?

Figure 1. Drink bottles task.

Following Sfard's (2021) commognition, reasoning about the process of solving a problem is a form of discourse that requires an understanding of the syntactic and sematic structure of its schemas and is characterised by four components. First, a mathematics discourse is *endorsed narratives* accepted by its participants as faithful accounts of the situation using communicational tools that make the discourse distinguishable from others. In the drink bottles task, the endorsed narratives are about measurement concepts, units, and dimensionality. The communication tools that distinguish the endorsed narratives from other stories hinge on the *keywords* (e.g., litre, millimetres, gram, dimensions) used to explain the focal objects and actions of the discourse. By nature of its abstraction, the third characteristics of a mathematical discourse is the use of *visual mediators* (e.g., symbols, diagrams, and words) to support effective communications. In this instant, a student may use a combination of linguistic, symbolic/algebraic, and diagrammatic tools to explain their reasoning. Indeed, while not necessary, a drawing may help the student to comprehend and clarify their solution for the GSODA task. Lastly, mathematics discourse is made distinct by the *routines*, the recurrent ways of performing different kinds of tasks in obtaining solutions. Routines guide our response to an expectation. They are task specific and depend on their interpreters. As learning is a process of routinisation of our actions, exposing the discursive gaps that threaten the process of learning is a critical step in turning obstacles into opportunities for learning. Since it is not possible to conduct large-scale observations of students' problem solving abilities, the GSODA task served to provide a practical context to help students see the mathematical relevance, and to determine how student coordinate and connect various information that is the hallmark of solving real-life problems faced by the world today. Analysing students' solutions can further help to infer the discursive routines enacted in problem solving situations.

Methodology

The data analysed here are taken from the RMF II project. Using a design-based research method, we applied an iterative cycle of designing, testing, and re-designing assessment tasks and scoring rubrics. Tasks were compiled into multiple assessment forms, both to validate the forms and to test the Learning Progressions, which was the aim of the project. Figure 2 shows the marking rubric for the GSODA task. Note that GSODA3 was included on some forms and not others, thus accounting for the difference in the total sample collected (see Table 1).

SCORE	DESCRIPTION for GSODA1
0	No response or irrelevant response
1	incorrect with no clear reasoning or working
2	Incorrect but with clear attempt to calculate, may use addition and make an error
3	Correct (8 or 8 bottles) but no reasoning or calculations shown
4	Correct, reasoning or working to justify (e.g., $8 \times 1.25 = 10$ litres)
SCORE	DESCRIPTION for GSODA2
0	No response or irrelevant response
1	All dimensions incorrect
2	Height dimensions correct, others not correct
3	Dimensions recognised for array chosen (e.g., 3×4; 2×6; or 1×12) but calculation error in one dimension (e.g., for a 3×4 array correctly calculates 4×85 mm but incorrectly calculates 3 by 85 mm
4	Correct for array chosen (e.g., 340 mm by 255 mm by 285 mm for a 3 by 4 array). Also correct if a small amount added for width of cardboard
SCORE	DESCRIPTION for GSODA3
0	No response or irrelevant response
1	Partially correct working with at least one correct component
2	Error in calculating liquid mass or one component missing
3	Correct (16.71 kg or 16710 g), that is 12×1250 g + 12×80 g + 750 g = 16710 g or 16.71 kg

Figure 2. GSODA task marking rubric.

The participants were Year 7 to 10 students from across Australia States and Territories. Two groups of cohorts were involved. The first set of data – the trial data, was taken from 214 students from five high schools across social strata in New South Wales, Queensland, and Western Australia. The teachers were asked to administer the assessment tasks and return the student work. Some teachers used the rubrics to mark the responses. All the results were further marked by two markers and validated by a team of researchers to ascertain the usefulness of the scoring rubric and the accuracy of the data entry. The second set of data – the project data, was taken from 377 students from seven high schools situated in lower socioeconomic regions with diverse populations across New South Wales, South Australia, Victoria, and Western Australia. The project schoolteachers were asked to mark and return the raw score instead of individual forms to the researchers. The project schools received two 3-day face-to-face professional learning sessions on developing mathematical reasoning. They also had access to a bank of teaching resources and four on-site visits to support their teaching effort.

Findings

Table 1 shows the overall percentage breakdown of student responses for GSODA. Similar to the data obtained for other tasks (e.g., see Seah & Horne, 2020), there were many no responses especially for GSODA2 and GSODA3. GSODA2 appears to be slightly more difficult to solve than the other two items. Students in Year 8 and 10 project schools performed slightly better than those in trial schools. The Year 7 project students' performance was weaker for all three items in comparison to their counterparts. Unlike the project school data, which show a gradual improvement for each item across the year levels, the trial school cohort's performance was erratic, with the Year 7 outperforming the other year levels in GSODA1 and GSODA2. Note that the Year 9 and 10 trial data were collected from three different States. Small samples may have influenced these results, but other factors may be at play. Furthermore, only four students (2%) from the trial schools, one in Year 7 and 8, and two in Year 9, answered all items correctly, while in the project schools this was 5.6% (4.2% Year 8 and 1.3% Year 10).

Table 1
Breakdown of Student Responses for GSODA

Score	Trial Data					Project Data			
	Year 7	Year 8	Year 9	Year 10	Total	Year 7	Year 8	Year 10	Total
GSODA1	*n = 82*	*n = 69*	*n = 50*	*n = 20*	*n = 221*	*n = 171*	*n = 204*	*n = 37*	*n = 377*
0	40.2	37.7	34	45	38.6	28.6	37.4	47.1	36.3
1	4.9	21.7	22	15	15	27.7	10.5	0	14.1
2	9.8	1.5	8	0	5.9	8.4	7.9	4.4	7.4
3	4.9	5.8	6	0	5	5.9	5.3	4.4	5.3
4	40.2	33.3	30	40	35.9	29.4	39	44.1	36.9
GSODA2									
0	62.2	84.1	56	75	69.1	83.2	63.7	54.4	68.1
1	9.8	4.4	10	10	8.2	11.8	9	13.2	10.6
2	14.6	7.3	24	5	13.6	1.7	6.3	4.4	4.5
3	4.9	2.9	6	10	5	3.4	5.3	10.3	5.6
4	8.5	1.5	4	0	4.6	0	15.8	17.7	11.1
GSODA3					*n=161*				
0	46.2	87	65	60	49.6	77.3	63.2	63.2	67.6
1	30.8	5.8	5	15	10.9	13.5	12.6	4.4	11.4
2	17.3	0	20	20	7.7	7.6	11.1	7.4	9.3
3	5.8	7.3	10	5	5	1.7	13.2	25	11.7

Choice of Strategy for GSODA1

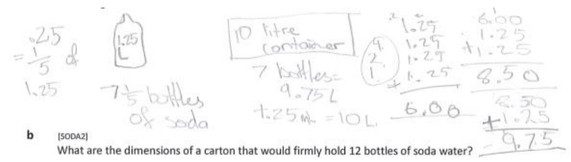
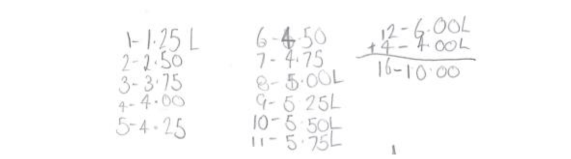
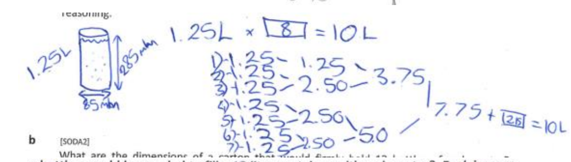
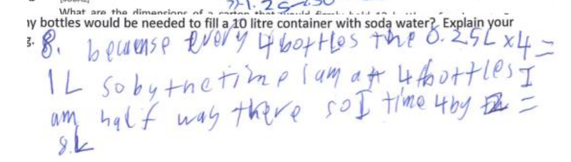
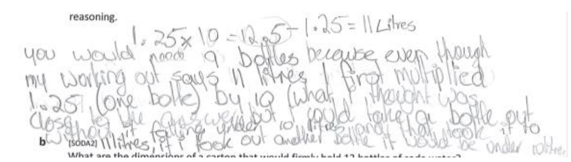
In the trial schools, 143 out of 221 students wrote a response. The most frequent solution was multiplying 1.25 by 8, followed by repeated addition (23.8%) and writing ‘8’ with no reason given (22.4%) (Table 2). Seven of the eight students who divided 10 by 1.25 answered correctly.

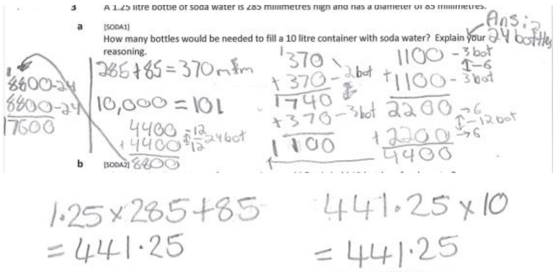
Table 2
Types and Percentages of Responses for GSODA1

1.25 × 8	Repeated addition	No reason	Grouping	Use all numbers	10 ÷ 1.25	1.25 × 10	Combination
24.5	23.8	22.4	9.1	6.3	5.6	4.2	3.5

Depending on the reasonableness of the answer, different scores were assigned to students who wrote a response without reason. A “0” was scored for 600 ml, 12.5 L or 637, and “1” if students wrote: 4, 6, 9 or 11 bottles. Seven students said 8 bottles with no reasoning and were scored 3. One student altered the capacity to a whole number and stated that “each bottle of coke [sic] is 500 ml, 500 + 500 = 1 L so 20 × 500 ml bottles are needed.” As the response did not address the question concerned, it is scored as 0.

Of the 34 students who used a repeated addition strategy, ten were not successful. Their errors included incorrect addition (see student A in Figure 3), change of increment (student B), or correct answer but incorrect calculation (student C). Twelve students were able to use a grouping strategy (combine addition and multiplication) in their solution (student D). Conversely, seven students chose to multiply 1.25 by 10 (Student E) and produced an incorrect answer. Nine students failed to ignore the irrelevant information. This combined with a lack of understanding of unit, led student F to generate an answer that could match 10,000 and student G to use all the numbers in some sort of procedures.

Student solutions	Interpretation of responses
	Student A: *7ML29112003 (score 2) Added four 1.25 incorrectly, combined this with two lots of 1.25 and then added another 1.25 to get 9.75. Since 0.25 is 1fifth of 1.25L, concluded that $7^{1}/_{5}$ bottles of soda must be needed.
	Student B: 7KK1922004 (score 2) Use repeated addition strategy. However, lost track of the increment and instead of increasing by 1.25, increased by 0.25 from the third lot instead.
	Student C: 8BM2112003 (score 3) Listed 7 lots of 1.25. worked out that two lots of 1.25 is 2.5, two lots of 2.5 is 5. Incorrectly added 3.75 and 5 to produce 7.75. Since an extra 2.15 is needed to make 10L, it must need 8 bottles.
	Student D: 7KR852003 (score 4) 8. because every 4 bottles the 0.25Lx4=1L so by the time I am 4 bottles I am halfway there so I time 4 by 2=8L.
	Student E: 7RD1972003 (score 2) You would need 9 bottles because even though my working out says 11 litres. I first multiplied 1.25 (one bottle) by 10 (what I thought was close to the answer but I could take a bottle out without it falling under. 10 could take a bottle and that took it to 11 times, if I took out another bottle it would be under 10 litres.



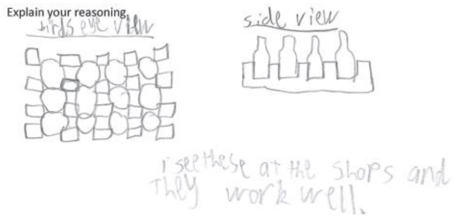
Student F: 7XK17112003 (score 2)
Added 285 and 85 to get 370. Assumed that 10L = 10,000. Proceeded to double 370 and one more gave 1100, 6 lots of 370 gave 2200. Double that gave 4400, which is 12 lots of 370. Double that gave 8800. Attempted to repeat doubling but realised the answer exceeded 10,000, so settle for 24 bottles.
Student G: 10ST632001 (score 0)
Multiplied 1.25 by the height and added the diameter. Multiplied the total by 10 to obtain 441.25.

Figure 3. Analysis of student solutions for GSODA1 (* The first number indicates the year level).

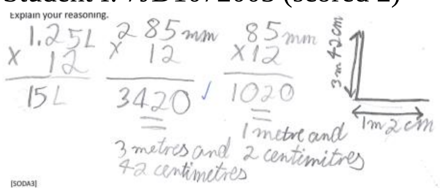
Solving Problem Requiring Multiplication

As shown in Table 1, GSODA2 received the least correct responses when compared with GSODA1and GSODA3. Of the 96 students who wrote a response, 28 gave an irrelevant response such as ‘a big rectangle’, ‘12.5L’ or ‘15L’. On average, 38.5% of the trial school cohort produced some form of drawing (43.6% Year 7, 37.5% Year 9, and 33.3% Year 8 and 10). Some drew an array with no explanation while others drew what they would see from the top and side view but did not give the dimensions (Student H in Figure 4). Around 14.6% of the students multiplied each dimension by 12 (Student I). Often, their efforts were hindered by poor computational skills (Student J) and lack of checking the reasonableness of their answer.

Student H: 7JL13102003 (scored 0)



Student I: 7JD1072003 (scored 2)



Student J: 7SM3172003 (scored 3)

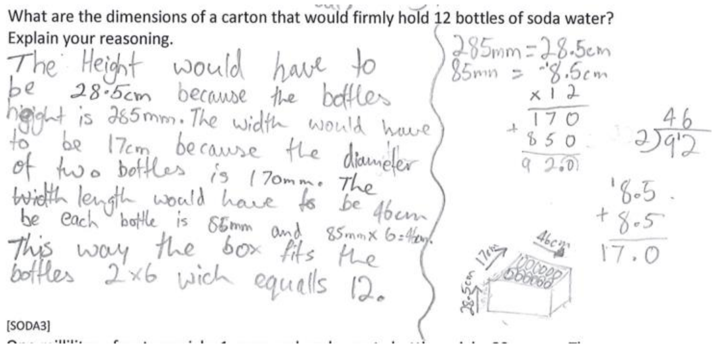


Figure 4. Analysis of student solutions for GSODA2.

Solving GSODA3 involves coordinating three components: 1) recognising the unit conversion, 2) working out the number of litres and hence the number of grams of liquid in 12 bottles, plus adding the mass of the bottles themselves, and 3) adding the weight of the box in the final calculation. Of the 59 students who wrote a response, seven gave an irrelevant response. Of the 52 students who scored at least a “1”, a quarter assumed that 1.25 L is 125 grams, 125 kg, or 1025 ml (see Student K and L in Figure 5). Another 9.6% multiplied 750 grams by 12 or used the height dimension as part of the calculation respectively. The most common error was neglecting to multiply either the soda water or empty bottles by 12 (Student M). These students appeared to have difficulty coordinating all the different aspects of the tasks in a clear plan to find the solution.

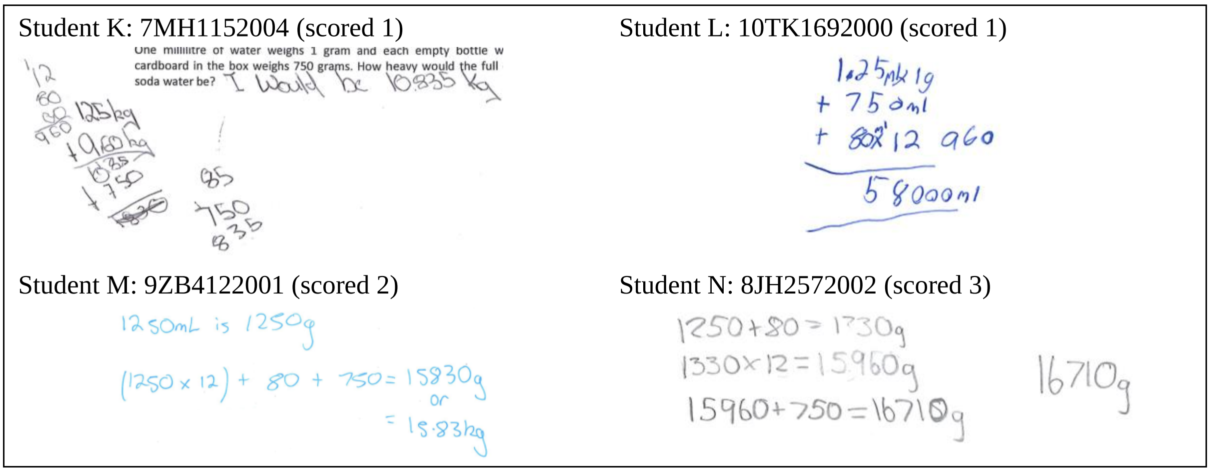


Figure 5. Analysis of student solutions for GSODA3.

Discussion

A problem is only a problem when the solution is not straightforward, as in the GSODA task. Most students did not already know the solution or how to obtain a solution. To successfully solve this problem, the students needed to coordinate several information and communicate their solution. We can see evidence of Sfard's (2021) four components (*endorsed narratives, keywords, visual mediators*, and *routines*) or the lack of them at play. To begin, the students' solutions reported here show a narrative that, while understood by the researchers, were in many instances not endorsed. For example, we found student G in Figure 3 who did not use any *keywords* in their working but added all the numbers, thus showing a lack of understanding of measurement concepts and dimensionality. Conversely, student I in Figure 4 used the measurement terms but multiplied 85 mm by 12 obtaining an answer of 1 m and 2 cm with the drawing of the final solution showing over 3 m height, thus demonstrating a lack of understanding of the relationship between these terms.

In an endorsed narrative, linguistic, symbolic/algebraic, and diagrammatic tools serve as *visual mediators* to facilitate and support effective communication. We observed a convoluted explanation of why 9 bottles are needed by student E and a succinct linguistic explanation of why it should be 8 bottles instead by student D in Figure 3. Others such as student B, who relied on a symbolic tool were often unable to multiply decimal numbers. Some students, such as student J in Figure 4 used a diagram to effectively show the final solution. While student H's diagrams clearly demonstrated how the bottles were visualised in the real situation, they did not help to obtain a correct solution.

The large numbers of no responses and low numbers of correct responses for all three items clearly show that many students did not have an established *routines* upon which they could call when trying to solve such problems. The drawing of diagrams, representing the information from the question on those diagrams and identifying the actual nature of the question are routines in problem solving that would have assisted many. Even routines of simple calculations were lacking as seen in many of the solutions in Figure 3. These combined with a lack of understanding of measurement concepts and dimensionality resulted in students unable to determine the dimensions of a desired carton or to assume that 1.25 L equal 125 kg or to multiply 1.25ml by 1 g (see student K and student L in Figure 5).

The implications of this for curriculum, assessment and pedagogy are fourfold. First, the teaching of measurement concepts needs to emphasise conceptual understanding and the connection between the concepts. Rather than memorising the conversion of units and arithmetic processes, greater emphasis must be placed on the concept of volume (which requires understanding and visualisation of three-dimensional objects) and its relationship to capacity and mass. Second, these concepts need to be presented in problems set in the real contexts where students are encouraged to visualise the problem mentally and on paper, and to manipulate and test the mental model of the problem to generate a solution. Third, routines need to be established, which encouraged the use of a range of communication tools such as diagram, symbols and language associated with the context. The language of explanation, argument and justification needs to be taught specifically and used regularly in classrooms to help students learn to explain their reasoning and justify their solutions (Seah & Horne, 2021). In this, assessment tasks should focus on the reasoning process rather than finding the right answer as shown in many multiple-choice questions. Finally, there should be frequent opportunities for students to solve multi-step problems. Students should be encouraged to discuss in pairs and in groups to fully comprehend the problems, develop all types of communication tools and deciding which type of communication tools work best for different situations. Only by the inculturation of problem solving as part of a socio-mathematical norm and mathematical classroom practices (Cobb & Yackel, 1996) can true change be realised.

References

Australian Curriculum, Assessment and Reporting Authority (ACARA). (n.d). *The Australian Curriculum: Mathematics. http://www.australiancurriculum.edu.au/*

Clarke, D., Goos, M., & Morony, W. (2007). Problem solving and working mathematically: An Australian perspective. *ZDM Mathematics Edcuation, 39*(5–6), 475–490. https://doi.org/10.1007/s11858-007-0045-0

Cobb, P., & Yackel, E. (1996). Constructivist, emergent, and sociocultural perspectives in the context of developmental research. *Educational Psychologist, 31*(3–4), 175–190. https://doi.org/10.1080/00461520.1996.9653265

Jonassen, D. H. (2011). *Learning to solve problems: A handbook for designing problem-solving learning environments*. Routledge. https://doi.org/10.4324/9780203847527

Lamb, S., Huo, S., Walstab, A., Wade, A., Maire, Q., Doecke, E., Jackson, J., & Endekov, Z. (2020). *Educational opportunity in Australia 2020: Who succeeds and who misses out.* Centre for International Research on Education Systems, Victoria University, for the Mitchell Institute.

Lortie-Forgues, H., Tian, J., & Siegler, R. S. (2015). Why is learning fraction and decimal arithmetic so difficult? *Developmental review, 38*, 201–221. https://doi.org/10.1016/j.dr.2015.07.008

Makar, K., Dole, S., Visnovka, J., Goos, M., Bennison, A., & Fry, K. (2020). Looking back and taking stock: Reflections on the MERGA research review 2012-2015. In J. Way, C. Attard, J. Anderson, J. Bobis, H. McMaster, & K. Cartwright (Eds.), *Research in Mathematics Education in Australasia 2016–2019* (pp. 7-26). Springer Nature Singapore. https://doi.org/https://doi.org/10.1007/978-981-15-4269-5_2

Polya, G. (1945). *How to solve it: A new aspect of mathematical model*. Princeton University.

Seah, R., & Horne, M. (2020). The influence of spatial reasoning on analysing about measurement situations. *Mathematics Education Research Journal, 32*(2), 365–386. https://doi.org/10.1007/s13394-020-00327-w

Seah, R., & Horne, M. (2021). Developing reasoning within a geometric learning progression: Implications for curriculum development and classroom practices. *Australian Journal of Education, 65*(3), 248–264. https://doi.org/https://doi.org/10.1177/00049441211036532

Sfard, A. (2021). The devil in details: Mathematics teaching and learning as managing inter-discursive gaps. In Y. H. Leong, B. Kaur, Y. H. Cho, J. B. W. Yeo, & S. L. Chin (Eds.), *Excellence in mathematics education: Foundations and pathways* (Proceedings of the 43rd annual conference of the Mathematics Education Research Group of Australasia), pp. 1-18. Singapore: MERGA.

Siemon, D., Callingham, R., Day, L., Horne, M., Seah, R., Stephens, M., & Watson, J. (2018). From research to practice: The case of mathematical reasoning. In J. Hunter, P. Perger, & L. Darragh (Eds.), *Making waves, opening spaces* (Proceedings of the 41st annual conference of the Mathematics Education Research Group of Australasia), pp. 40–49. MERGA.

Thomson, S., De Bortoli, L., & Underwood, C. (2016). *PISA 2015: A first look at Australia's results*. Australian Council for Educational Research.

Thomson, S., Wernert, N., O'Grady, E., & Rodrigues, S. (2017). *TIMSS 2015: Reporting Australia's results*. Australian Council for Educational Research.

Weber, K., & Leikin, R. (2016). Recent advances in research on problem solving and problem posing. In A. Guitierrez, G. C. Leder, & P. Boero (Eds.), *The second handbook of research on the psychology of mathematics education* (pp. 353–382). Sense Publishers.