Exploring mathematical discussion in word problem-solving

Authors:
Percy Sepeng[1,2]
Paul Webb[2]

Affiliations:
[1]Marang Centre for Mathematics and Science Education, University of the Witwatersrand, South Africa

[2]Centre for Educational Research, Technology and Innovation, Nelson Mandela Metropolitan University, South Africa

Correspondence to:
Percy Sepeng

Email:
percy.sepeng@wits.ac.za

Postal address:
Private Bag X3, WITS 2050, South Africa

Dates:
Received: 10 Oct. 2011
Accepted: 22 June 2012
Published: 15 Aug. 2012

How to cite this article:
Sepeng, P., & Webb, P. (2012). Exploring mathematical discussion in word problem-solving. *Pythagoras, 33*(1), Art. #60, 8 pages. http://dx.doi.org/10.4102/pythagoras.v33i1.60

This study explored whether discussion as a teaching strategy in mathematics classrooms could have positive gains in improving learners' problem-solving performance, as well as their ability to make sense of real wor(l)d problems. This article discusses the partial findings of a bigger study that used a pre-test intervention or post-test mixed-method design and utilised both quantitative and qualitative data. Analysis of the data generated from both pre-tests and post-tests suggests that the intervention strategy significantly improved the experimental group's problem-solving skills and sense-making performance. The statistical results illustrate that the experimental group performed significantly better than the comparison group in the post-test. The main finding of this study is that in classrooms of experimental schools in which discussion technique was successfully implemented, there was a statistically significant improvement in the learners' competence in solving word problems.

Introduction

Problem solving and integrated assessment are seen as the cornerstones of school mathematics and the *Principles and Standards for School Mathematics* (National Council of Teachers of Mathematics [NCTM], 2000) called for mathematics instruction and assessment to focus more on conceptual understanding than on procedural knowledge or rule-driven computation (Hamilton, 2004; Kilpatrick, Swafford & Findell, 2001). Major arguments for including word problems in the school mathematics curriculum have always been their potential ability to promote realistic mathematical modelling and problem-solving. Solving word problems also help learners to develop the skill of knowing when and how to apply classroom mathematical knowledge as well as everyday life-knowledge when solving problems. In this article we therefore argue that there are multiple benefits and good outcomes when learners participate in mathematical discussion in the classroom.

This article begins with a discussion of the current debates on using discussion as a strategy for the teaching and learning of mathematics in general. The study draws largely on the results of a research project conducted in six schools in the Eastern Cape Province of South Africa (Sepeng, 2010). Much of the work in this study relies both theoretically and methodologically on notions of classroom mathematics discourse and mathematical modelling. The main argument in this article is that careful use of discussion as a teaching strategy in mathematics classrooms appears to have positive gains in improving learners' problem solving performance in word problems.

The hypothesis for the study was that the introduction of *mathematical discussion* in the teaching and learning of word problems would improve and enhance learners' *problem solving* performance and the ability to *make sense* of real wor(l)d problem-solving. The purpose of this article is therefore to demonstrate how the results of the study were found to support this hypothesis.

Solving word problems

Word problems have been defined differently in different studies. For the purpose of this study, the definition provided by Verschaffel, Greer and De Corte (2000) is used. These researchers define word problems as 'textual descriptions of situations assumed to be comprehensive to the reader, within which mathematical questions can be contextualised' (p. ix). They also stress that word problems 'provide, in convenient form, a possible link between the abstractions of pure mathematics and its applications to the real-world phenomena' (p. ix). According to Palm (2009) mathematical word problems include pure mathematical tasks 'dressed up' in a real-world situation that require students to 'undress' these tasks and solve them (p. 60).

A further methodological issue, which socio-cultural approaches have yet to address satisfactorily, arises from the increasingly multicultural nature of mathematics classrooms. Students' interpretations of mathematics classroom interaction relate in part to their different social, cultural and linguistic backgrounds. When classroom interaction is analysed, therefore, some way of taking account of this diversity needs to be found, otherwise there is the risk that a single cultural perspective, that of the researcher, will be imposed. Discursive psychology has the potential to address some of the above-mentioned issues.

Discussion in mathematics classrooms

The use of discussion as a tool to increase reasoning has gained emphasis in classrooms worldwide, as earlier reports had predicted (Yore, Bisanz & Hand, 2003). Discussion, however, requires scaffolding and structure in order to support learning (Norris & Phillips, 2003). Wood, Williams and McNeal (2006) found variations in students' ways of seeing and reasoning, and these were attributed to the particular differences established in classrooms early in the year pertaining to *when* and *how to contribute* to mathematical discussions and *what to do as a listener*. Their conclusions are consistent with the findings reported by several other researchers (e.g. Dekker & Elshout-Mohr, 2004; Ding, Li, Piccolo & Kulm, 2007; Gillies & Boyle, 2006; Webb, Nemer & Ing, 2006), who also suggested that participation obligations put boundaries around the opportunities for students to share their ideas and to engage in mathematical practices. When they make a difference through classroom discourse, teachers shift students' cognitive attention toward making sense of their mathematical experiences, rather than limiting their focus to procedural rules. In doing so, students become less engaged in solutions to problems than in the reasoning and thinking that lead to those solutions (Yackel & Cobb, 1996).

Students develop a mathematical disposition through the patterns of interaction and discourse created in the classroom and the process of ascribing meaningfulness to one another's attempts to make sense of the world. Learning about other ways to think about ideas, to reflect, and to clarify and modify thinking is fundamental to moving learning forward. Carpenter, Franke and Levi (2003) maintain that the very nature of mathematics presupposes that students cannot learn mathematics with understanding and without engaging in discussion. However, more talk in the classrooms does not necessarily enhance student understanding. Better understanding is dependent on particular pedagogical approaches, purposefully focused on developing a discourse culture that elicits clarification and produces consensus within the classroom community.

A variety of situations may arise in which the outcomes are not fully realised. For example, a number of studies have reported that some students appear to thrive more than others in whole-class discussions. In their respective research, Baxter, Woodward and Olson (2001) and Ball (1993) found that highly articulate students tend to dominate classroom discussions. Low academic achievers usually remain passive, and when they do participate visibly, their contributions are comparatively weaker, and their ideas sometimes muddled. Nevertheless, pedagogical practices that create opportunities for students to explain their thinking and to engage fully in dialogue have been reported in research undertaken by Steinberg, Empson and Carpenter (2004). In a study which was part of their Cognitively Guided Instruction Project, they found that classroom discussion was central to a sustained change in students' conceptual understanding.

Method

Sample in this study

The sample consisted of Grade 9 learners (n = 176) and their teachers in six township secondary schools, four of which were experimental schools (n = 107) and two of which were comparison schools (n = 69) (where no intervention took place). The six schools chosen were a convenient sample of a cluster of similar schools in Port Elizabeth. All the schools were functional (as opposed to dysfunctional – which is the case in many instances in South Africa), had similar characteristics in their approach to teaching and learning contexts and were public and previously marginalised schools. The schools drew their learners from communities of low economic status. The two comparison schools were chosen randomly within the group identified.

Research design

In this study we used a pre-test – intervention – post-test design. We investigated the situation in terms of the problem-solving performance of Grade 9 learners with English as their second language, using a pre-test to establish how the learners solved mathematics word problems and what problems they might have mathematically. Then we wanted to find out if introducing *discussion* into the classroom as part of the learning experience had any effect on learners' problem-solving performances and/or strategies. A post-test, which was exactly the same as the pre-tests, investigated if there were any changes. We also used this test to find explanations or reasons for any changes in learners' sense-making and solutions of real wor(l)d problems. In the next section we present the intervention strategy used to promote *discussion* in the teaching and learning of word problems.

Design type and the intervention

In this study an intervention strategy was used to investigate the effect of discussion on the problem-solving performance of Grade 9 second language learners. The intervention took place over a period of six weeks and focused on the use of writing and mathematical discussion to solve word problems. The intervention was implemented in the experimental group and its aim was to assist learners not only to identify problem situations that were problematic from a realistic point of view, but also to consider the (in)appropriateness of applying mathematical operations directly as their solutions. Moreover, the teachers of the experimental schools were introduced to a typical collaborative learning context(s) where the teacher posed a task (or a problem) and the learners, after sufficient time to complete the task, engaged in a discussion of the solution methods and/or strategies that they had developed in small groups. The whole class discussion continued for, at most, 15 minutes before another short segment (three to five minutes) of group task. This cycle was repeated five to six times in a double period of about 90 minutes. During the small group interaction learners had to *develop reasons to support their thinking* and/or *think about* some relevant issue(s) or question(s) instead of solving a specific mathematics question. Because of the continual emphasis on both justification and reasoning, whole classroom discussions resulted in the emergence of key concepts in wor(l)d *problem solving* within the contexts of realistic considerations in particular, and *sense-making* in general. As a consequence, the teaching approach appeared to have considerable potential for in-depth conceptual development growing out of the learners' discursive activity.

Although learners had some time to explore their reasoning with one another other during the limited time allocated to small group discussion, the *interruptions* brought about by the teachers in the classroom discussion did not give them much time to pursue their own ideas. However, learners were expected to accept the obligation and engage in *thinking about the issues at hand* and in *sharing their thinking* within their smaller groups. As such and since the discussion inevitably focused on their reasons, learners were in a good frame of mind to compare and contrast their reasons with those of others. Consequently, their thoughts and discussions formed a basis for engaging meaningfully in the subsequent classroom discussion. In some instances, concept cartoons in mathematics were used as a stimulus or trigger for discussion while learners were solving problems in their small groups. The purpose of introducing discussion was to help learners seek, share and construct knowledge when engaging in solving word problems. The discussions took the form of dialogue and talk (formal and informal) in both English and the learners' home language (isiXhosa). In promoting discussion, learners were expected to engage critically with problems and build positively on what others had said.

The observations during and after the intervention were done with the aim of measuring teachers' implementation of the strategies that they had learnt during teacher workshops. These workshops gave teachers from experimental schools the opportunity to be trained on how to get learners discussing, arguing and writing about their views and experiences when they solved mathematics word problems. The aim of promoting these strategies in their teaching was to develop and improve their approaches to the teaching and learning of word problem solving in their classrooms. These teachers were introduced to and trained in strategies to improve their pedagogical content knowledge and their ability to promote the teaching and learning of mathematics when solving word problems.

The intervention also focused on the language of mathematics embedded within word problems. Simple translations were provided for phrases that are often used in mathematical word problems to simplify the meaning of these problem statements. Table 1 shows some examples of the translations provided.

The problem-solving tasks

The pre-test and post-test consisted of the following three modelling problems adapted from Verschaffel, Greer and Van Dooren (2009. Learners were expected to complete these *problem-solving* (*PS*) *tasks* and give a written explanation of how they arrived at their answers:

PS1: 100 children are transported by minibuses to a summer camp at the seaside. Each minibus can hold a maximum of 8 children. How many minibuses are needed?

PS2: Two boys, Sibusiso and Vukile, are going to help So nwabo rake leaves on his plot of land. The plot is 1200 square meters. Sibusiso rakes 700 square meters in four hours and Vukile does 500 square meters in two hours. They get 180 rand (R) for their work. How are the boys going to divide the money so that it is fair?

PS3: John's best time to run 100 meters is 17 seconds. How long will it take him to run 1 kilometre?

Coding of the tests

The three PS tasks were coded using a schema that was an elaboration of the classification schema developed by Verschaffel, De Corte and Lasure (1994). The classification schema comprised fourteen categories, which were reduced to three general categories:

- *Realistic reaction* (*RR*): All cases where a learner either gave the (most) correct numerical solution that also took into account the real-world aspects of the problem context, as well as cases where there was a clear indication that the learner tried to take into account those real-world aspects, without giving the mathematically and situationally (most) accurate numerical answer.
- *No reaction* (*NR*): All those cases where there is no indication that the learner was aware of the realistic modelling difficulty, for example, mathematically correct but situationally inaccurate and/or incorrect or inappropriate responses, computational errors, etc. This category also provides a measure of the problem-solving performance of the learner.

TABLE 1: Translations of word problem phrases.

In a mathematics word problem this phrase:	Usually means you will need to:
How many more?	subtract the smaller number from the larger number
How many altogether?	add
What is the difference?	subtract the smaller number from the larger number
How many are left?	subtract the smaller number from the larger number
'each' in a problem with the phrase 'How many altogether or in total'	multiply
'each' as in 'How many do they each have?'	divide
least	select the lowest value (number)
most	select the highest value (number)
Find the sum or the total	add
Find the product	multiply
Find the difference	subtract
Find the quotient	divide

Source: Zhang, L.J., & Anual, S. (2008). The role of vocabulary in reading comprehension: The case of secondary students learning English in Singapore. *RELC Journal*, *39*(1), 52–77

- *Other reaction* (*OR*): All cases where a learner did not provide a numerical response and did not give any written comment that indicated that the learner was aware of the realistic modelling difficulty that prevented him or her from answering the problem, as well as instances where the learner generated incorrect responses with mathematical (or computational) errors.

Overview of reliability and validity in this study

The primary strategy utilised in this study to ensure external validity was the provision of thick, rich and detailed descriptions so that anyone interested in transferability would have a solid framework for comparison (Merriam, 1988). Nixon and Power (2007) point out that warranting of claims must fulfill the criteria of trustworthiness, soundness, coherence, plausibility and fruitfulness. Trustworthiness refers to the quality of qualitative data collected (Anastas, 2004); and in the sensof neutrality in the findings or decisions of the study (Guba & Lincoln, 2005).

Reliability is the degree to which the instrument consistently measures whatever it is measuring (Ary, Jacobs & Razavier, 1990; Best & Kahn, 2003). According to Silverman (1999), reliability refers to the degree of consistency with which instances are assigned to the same category by different observers or by the same observer on different occasions. Neuman (2003) suggests reliability has to do with the issue of dependability. Dependability of the data in this study was established by capturing the observations on a tape and video recorder, and transcribing them both manually in writing and with computer software. Attempts were made to reproduce the interview scripts as accurately as possible to eliminate possible threats to the reliability of the instruments used in this study. Creswell (2005) defines threats as the problems that threaten our ability to draw correct cause-and-effect inferences that arise because of the experimental procedures or the experiences of participants.

Ethical considerations

The Education, Research Technology and Innovation Committee gave prior permission to conduct this research as part of the Integrated School Development and Improvement project offered by the Centre for Educational Research, Technology and Innovation at the Nelson Mandela Metropolitan University. After obtaining ethics clearance, the first author approached the principals and teachers of the participating schools, where their roles as participants and their rights to choose to be participants and to participate or not in the study were explained to them. Both the teachers and parents (on behalf of the learners) gave their informed consent. They were assured of confidentiality and also that participation was voluntary. They were given a guarantee that they could withdraw from the study at any time and that no personal details would be disclosed. They were promised confidentiality of information collected in the schools and were assured that no portion of the data collection would be used for any purpose other than this research. Learners were also assured that the test results would not influence their school marks.

Statistical analysis of data

The quantitative statistical data generated from the pre-tests and post-tests were captured in a Microsoft Office Excel spreadsheet and subjected to repeated measure ANOVA techniques (using the STATISTICA Software package) which simultaneously accounted for pre-test and post-test data of the experimental and comparison groups in order to provide both descriptive and inferential statistics. ANOVA techniques were used to determine the statistical (non)significance of the results, based on mean differences between experimental and comparison groups before and after the intervention. Where necessary, the statistical technique of *matched-pairs t-tests* was computed to compare the mean scores of the comparison and experimental groups.

Cohen's d statistics were calculated to determine whether statistically significant ($p < 0.0005$) pair-wise differences were practically significant. A *small* practical significance is noted where $0.2 < d < 0.5$; a *moderate* practical significance is noted if $0.5 < d < 0.8$; and a *large* practical difference is recorded if $d > 0.8$. Expressed differently, an effect size of less than 0.2 is considered to be insignificant, an effect size between 0.2 and 0.5 is considered to be of *small* significance; an effect size between 0.5 and 0.8 is considered as being *moderately* significant, while an effect size of 0.8 and greater is considered to be *highly* significant. Effect size as expressed by the Cohen's d statistics is defined as the difference in means divided by the pooled standard deviation and is a measure of magnitude (or significance) of the differences between the pre-test and post-test scores (Gravetter & Walnau, 2008).

Findings

The word problems in this study, which are examples of a central part of mathematics learning, can be seen as attempts to connect mathematical reasoning to learners' everyday life experiences and/or knowledge (Sepeng, 2011). In other words, these problem-solving tasks can be viewed as a manifestation of the notion that mathematics is or should be part of mundane practices in everyday life (Verschaffel et al., 2009).

Effect of intervention on problem solving

Table 2 depicts a summary of the statistical analysis of learners' word problem solving achievements as well as their realistic reactions. Analysis of pre-test results indicate a statistically significant difference between the experimental and comparison groups, with the experimental group's performance being worse than that of the comparison group ($p < 0.0005$). After the intervention, the experimental group's performance was statistically significantly better ($p < 0.0005$) than the comparison group's, with a mean difference ($\Delta\bar{x}$) of 29.14. In Table 2 a positive mean difference implies that the mean score of the experimental group was more than that of the comparison group in the post-test.

Practical significance of the differences

As $p < 0.0005$ in all cases, Cohen's d was calculated in order to gauge the effect size of the practical significance of the differences in the experimental and the comparison groups. A large practical significance (*d* = 1.56) was noted. When the effect sizes on the RR difference between the experimental and the comparison groups were compared, a moderate practical significance (*d* = 0.57) was calculated (see Table 2).

Sense-making (or realistic considerations) of word problem solving (PS tasks)

Table 3 shows the number and percentage of learners who succeeded in producing three, two, one, and zero RRs to the PS tasks. It illustrates that learners performed rather poorly in these tasks, which required not only computational skills, but realistic sense-making as well. Only 1% of learners in the experimental schools produced three situationally accurate answers or reacted three times in a way that showed attention to the realistic modelling complexity of the problems. A closer look at the post-test results for the experimental schools shows a 10% improvement in the production of 2 RRs compared to a drop of 6% in the comparison schools.

The mean difference (difference between the mean scores) shows [$\Delta\bar{\chi}$ = 0.47] a statistically significant ($p < 0.0005$) difference between the experimental and the comparison groups for the RR after the intervention. The positive mean score shows that, although the comparison group had a tendency to consider reality and sense-making when solving the word problems before the intervention ($\Delta\bar{\chi}$ = -0.02), its performance was well below the experimental group's after the intervention.

Matched-pairs *t*-tests

Table 4 shows the results of a *matched-pairs t-test* that was used to test whether there was a significant mean difference between experimental and comparison groups before and after the intervention (or pre-test and post-test). In addition to this, data in Table 4 depict mean scores of the experimental and comparison groups for word problem solving and sense-making (RRs) in this study.

At the $p < 0.0005$ significance level, the study gives overwhelming evidence that the problem-solving scores of the experimental group improved by 17.08 after the intervention, whilst the practical significance calculated for the experimental group is moderate. Although the comparison group's mean score was higher than that of the experimental group before the intervention (pre-test), a negative mean difference ($\Delta\bar{\chi}$ = -26.67) suggests that the comparison group not only scored well below the experimental group after the intervention, but that their scores were significantly lower than the experimental group's.

TABLE 2: Experimental minus comparison groups' scores (mean difference).

Category	Statistic	a	b	RRa	RRb	*d*	RR*d*
Mean and SD	Mean	-14.61	29.14	-0.02	0.44	43.75	0.47
	SD	5.53	5.56	-0.01	0.43	10.98	0.20
Statistical significance based on mean difference	*t*-statistic	-4.89	8.24	-0.24	4.39	10.12	3.70
	df	174	174	174	174	174	174
	p-value	< 0.0005	< 0.0005	0.812	< 0.0005	< 0.0005	< 0.0005
Practical significance based on mean difference	Cohen's *d*	0.75	1.27	n.a.	0.68	1.56	0.57
	df	2	2	2	2	2	2
Statistical significance based on frequency distribution	Chi2-stat	25.11	0.41	9.28	13.58	54.89	14.76
	p-value	< 0.0005	0.813	0.010	0.001	< 0.0005	0.001
Practical significance based on frequency distribution	*df*	1	1	1	1	1	1
	Cramér's V	0.38	n.a.	0.23	0.28	0.56	0.29

a, pre-test; b, post-test; RR, realistic reactions; *d*, difference; n.a., not applicable.

TABLE 3: Number of learners producing realistic reactions.

Group	Pre-test										Post-test									
	3 RR		2 RR		1 RR		0 RR		Total		3 RR		2 RR		1 RR		0 RR		Total	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%	*n*	%
Experimental	1	1	3	3	29	27	74	69	107	100	2	2	14	13	29	27	62	58	107	100
Comparison	0	0	4	6	18	26	47	68	69	100	0	0	0	0	10	14	59	86	69	100

RR, realistic reactions; *n*, number of learners.

TABLE 4: A test of a significant mean difference (pre-tests and post-tests in the experimental and comparison groups using a matched-pairs *t*-test).

Group	Category	Mean	SD	Difference	SD	Matched-pairs *t*-test			Cohen's *d*
						t	*df*	*p*	
Experimental	a	40.64	21.33	17.08	31.76	-5.56	106	0.000*	0.54
	b	57.72	24.92						
Comparison	a	55.25	15.80	-26.67	20.79	10.65	68	0.000	-1.28
	b	28.58	19.36						
Experimental	RRa	0.36	0.59	0.23	0.89	-2.73	106	0.007	0.26
	RRb	0.59	0.79						
Comparison	RRa	0.38	0.60	-0.23	0.69	2.80	68	0.007	-0.34
	RRb	0.14	0.35						

a, pre-test; b, post-test; RR, Realistic reactions; *p*, probability value.
*, Note that reported p = 0.000 implies $p < 0.0005$

TABLE 5: Results summary.

Category	Pre-test and post-tests
Problem-solving (mathematically correct answers)	• The comparison group scored statistically significantly (d < 0.0005) higher in problem-solving in the pre-test but performed well below the experimental group after the intervention. • The experimental group's mean score for problem-solving improved statistically significantly more than the comparison group's did after the intervention (post-test). A large practical significance (d = 1.56) was also noted.
Sense-making (realistic consideration of problem context)	• The comparison group scored slightly higher than the experimental group in sense-making of word problems in the pre-tests, but the difference was not statistically significant. • The experimental group scored statistically significantly (p < 0.0005) higher than the comparison group in sense-making of word problems in the post-tests. A moderate practical significance (d = 0.57) was noted.

A practical non-significance was also calculated for the comparison group. Despite a small practical significance found for the experimental group's RRs, the experimental group did significantly better than the comparison group with a marginally significant improvement in sense-making scores ($\Delta\bar{x} = 0.23$) in the experimental group after the intervention.

Table 5 provides a brief summary of findings based on learners' solving and sense-making of word problems.

Discussion

Problem solving

Analysis of the data obtained from pre-tests and post-tests revealed that learners' problem-solving performance in the post-tests improved over time after the intervention. Statistical results illustrate that there was a statistically significant difference between the experimental and comparison groups before the intervention (pre-tests), with the comparison group performing significantly better than the experimental group. However, the experimental group performed statistically significantly better than the comparison group after the intervention. In other words, it appeared that the intervention strategy in this study (see introduction of the discussion) had a positive influence on learners' skills in solving word problems. As a result, a large practical significance was also noted. The experimental group improved significantly from the pre-test to the post-test.

Sense-making

The overall results of this study illustrate that learners' performance changes dramatically when discussion is introduced into the mathematics classroom as a teaching strategy. When word problems are taught, teachers should consider carefully how to model the situation (or context), and also whether the information provided is relevant and sufficient for solving the problem (Säljö, Riesbeck & Wyndham, 2009). The results of this study showed that before the intervention (pre-test) learners had a tendency to respond to the problems even if the information given was irrelevant to the information needed to answer the given question. It is interesting that intercultural comparison studies show similar findings (Säljö et al., 2009; Verschaffel et al., 2000; Xin, 2009; Xin, Lin, Zhang & Yan, 2007).

The statistical results revealed a significant difference between the experimental group and comparison group. The experimental group appeared to show a tendency to consider reality marginally better than the comparison group. A large significant practical difference between the experimental group and the comparison group was also noted after the intervention.

The results of the study demonstrate that the introduction of discussion in the teaching and learning of word problems in mathematics not only had a positive effect on learners' problem-solving performance, but also on their ability to consider reality when they had to solve word problems. The data generated in this study also suggest that whole-class discussion and problem-based approaches to the teaching of word problems can be applied appropriately and successfully (to certain degree) in second language teaching and learning settings, and can assist both mathematics teachers and learners to improve their knowledge of the real wor(l)d effects of mathematics problems.

Conclusion

In this study discussion as a strategy to improve second language learners' word problem solving and sense-making skills was explored. The literature suggests, and initial observations appeared to confirm, that the difficulties in solving word-problems are related to effective pedagogical strategies that advance problem-based and whole-class discussion approaches to the teaching and learning of word problems in mathematics (Verschaffel et al., 2000, 2009). The results of this study seem to illustrate and substantiate that *what teachers do* serves as a fundamental component to raising learning outcomes (Douglas, 2009). In the study teachers' pedagogical practices during the intervention resulted in lessons which involved group interaction and communication. The lessons were organised in such a way that there was some form of taking turns, where each member of a small group had to make their talk comprehensible to all (Heap, 1990). In this way the lessons in the mathematics classrooms came to be dominated by cooperative learning discussions where talk within teacher-learner interactions and/or learner-to-leaner(s) interactions were of high quality.

Statistical analysis of variance showed statistically significant and large practically significant evidence that the introduction of discussion in the teaching and learning of word problems in mathematics increases the problem-solving and sense-making performance of second language learners. The level at which the threshold of p was set in the study was 0.0005, which meant that there was a 0.05% chance that the results were accidental. The large practical significance noted in the study implies a research result that should be viewed as important for teaching practice in mathematics classrooms. The findings of this study suggest that when discussion was introduced into the mathematics classroom, better connections between classroom mathematics and out-of-school mathematics were made and that there was better integration between the learners' formal written mathematical language and their informal spoken mathematical language. In fact, learners not only generated more computationally correct responses, but also produced more situationally accurate and appropriate solutions to real-wor(l)d problems in the post-test (or after the intervention).

Although the intervention of the study targeted only a limited number of teachers and schools in Port Elizabeth townships, and the conclusions drawn from the study cannot be generalised, the findings provide sufficient insights from which tentative recommendations for mathematics teacher development can be drawn. Analysis of the quantitative data suggests that promoting the introduction of discussion techniques in mathematics classrooms had benefits and in all probability promoted the participating learners' problem-solving performance and significantly increased the likelihood of realistic consideration of word problem solving. However, to successfully implement such a strategy, teachers need appropriate fundamental skills and the necessary knowledge of managing and maintaining classroom discourses that allow the development of formal written mathematical language and the skills necessary for the negotiation of meaning within informal spoken mathematical language.

Acknowledgements

This research is supported by the Centre for Educational Research, Technology and Innovation at the Nelson Mandela Metropolitan University. Any opinions, findings and conclusions or recommendations expressed in this study are those of the authors and do not necessarily reflect the views of the Centre for Educational Research, Technology and Innovation.

Competing interests

We declare that we have no institutional and/or personal or financial relationships which may have inappropriately influenced us in writing this article.

Authors' contributions

P.S. (University of the Witwatersrand) conducted all the research and wrote the manuscript. P.W. (Nelson Mandela Metropolitan University) was the promoter of the bigger study and made conceptual contributions to the bigger study and to the manuscript.

References

Anastas, J. (2004). Quality in qualitative evaluation: Issues and possible answers. *Research on Social Work Practice*, *14*(1), 57–65. http://dx.doi.org/10.1177/1049731503257870

Ary, D., Jacobs, L., & Razavier, A. (1990). *Introduction to research in education.* (4th edn.). Fort Worth, TX: Holt, Rinehart and Winston.

Ball, D.L. (1993). With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *Elementary School Journal*, *93*, 373–397.

Baxter, J., Woodward, J., & Olson, D. (2001). Effects of reform-based mathematics instruction on low achievers in five third-grade classrooms. *Elementary School Journal*, *101*, 529–549. http://dx.doi.org/10.1086/499686

Best, J., & Kahn, J. (2003). *Research in education.* Boston, MA: Pearson Education.

Carpenter, T., Franke, M., & Levi, L. (2003). *Thinking mathematically: Integrating arithmetic and algebra in elementary school*. Portsmouth, NH: Heinemann.

Creswell, J.W. (2005). *Educational research: Planning, conducting and evaluating quantitative and qualitative research*. (4th edn.). Boston, MA: Pearson Education.

Dekker, R., & Elshout-Mohr, M. (2004). Teacher interventions aimed at mathematical level raising during collaborative learning. *Educational Studies in Mathematics*, *56*, 39–65. http://dx.doi.org/10.1023/B:EDUC.0000028402.10122.ff

Ding, M., Li, X., Piccolo, D., & Kulm, G. (2007). Teacher interventions in cooperative-learning mathematics classes. *The Journal of Educational Research*, *100*(3), 162–175. http://dx.doi.org/10.3200/JOER.100.3.162-175

Douglas, K. (2009). Sharpening our focus in measuring classroom instruction. *Educational Researcher*, *38*(7), 518–521. http://dx.doi.org/10.3102/0013189X09350881

Gravetter, F., & Walnau, L. (2008). *Essentials of statistics for the Behavioral Sciences.* New York, NY: Wadsworth.

Gillies, R., & Boyle, M. (2006). Ten Australian elementary teachers' discourse and reported pedagogical practices during cooperative learning. *Elementary School Journal*, *106*, 429–451. http://dx.doi.org/10.1086/505439

Guba, E.G., & Lincoln, Y. (2005). Paradigmatic controversies, and emerging confluences. In N. Denzin, & Y. Lincoln (Eds.), *Handbook of qualitative research* (3rd edn.) (pp. 191–216). Thousand Oaks, CA: Sage.

Hamilton, R.J. (2004). Material appropriate processing and elaboration: The impact of balanced and complementary types of processing on learning concepts from text. *British Journal of Educational Psychology, 74*, 221–237. http://dx.doi.org/10.1348/000709904773839851, PMid:15130189

Heap, J. (1990). Applied ethnomethodology: Looking for the local rationality of reading activities. *Human Studies*, *13*, 39–72. http://dx.doi.org/10.1007/BF00143040

Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics.* Washington, DC: The National Academy Press.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: NCTM.

Merriam, S. (1998). *Qualitative research and case study applications in education.* San Francisco, CA: Jossey-Bass.

Neuman, W. (2003). *Social research methods: Qualitative and quantitative approaches.* Boston, MA: Pearson Education.

Nixon, A., & Power, C. (2007). Towards a framework for establishing rigour in a discourse analysis of midwifery professionalism. *Nursing Inquiry*, *149*(10), 71–79. http://dx.doi.org/10.1111/j.1440-1800.2007.00352.x, PMid:17298610

Norris, S.P., & Phillips, L.M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education, 87,* 224–240. http://dx.doi.org/10.1002/sce.10066

Palm, T. (2009). Theory of authentic task situations. In L. Verschaffel, B. Greer, W. Van Dooren, & S. Mukhopadhyay (Eds.), *Words and worlds: Modelling verbal descriptions of situations* (pp. 3–19). Rotterdam: Sense Publishers.

Säljö, R., Riesbeck, E., & Wyndham, J. (2009). Learning to model: Coordinating natural language and mathematical operations when solving word problems. In L. Verschaffel, B. Greer, W. Van Dooren, & S. Mukhopadhyay (Eds.), *Words and worlds: Modelling verbal descriptions of situations* (pp. 177–193). Rotterdam: Sense Publishers.

Sepeng, J.P. (2010). Grade 9 second-language learners in township schools: Issues of language and mathematics when solving word problems. Unpublished doctoral dissertation. Nelson Mandela Metropolitan University, Port Elizabeth, South Africa. Available from http://www.nmmu.ac.za/documents/theses/Johannes%20Percy%20Sepeng.pdf

Sepeng, P. (2011). Triadic dialogue: An analysis of interactions in multilingual mathematics primary classrooms. *US-China Education Review, 8*(4), 412–418. Available from www.eric.ed.gov/PDFS/ED520468.pdf

Silverman, D. (1999). Warriors or collaborators: Reworking methodological controversies in the study of institutional interaction. In S. Sarangi, & C. Roberts (Eds.), *Talk, work and institutional order: Discourse in medical, mediation and management setting* (pp. 401–425). Berlin: Mouton de Gruyter. http://dx.doi.org/10.1515/9783110208375.4.401

Steinberg, R., Empson, S., & Carpenter, T. (2004). Inquiry into children's mathematical thinking as a means to teacher change. *Journal of Mathematics Teacher Education*, *7*, 237–267. http://dx.doi.org/10.1023/B:JMTE.0000033083.04005.d3

Verschaffel, L., De Corte, E., & Lasure, S. (1994). Realistic considerations in mathematical modeling of school arithmetic word problems. *Learning and Instruction*, *4*, 273–294. http://dx.doi.org/10.1016/0959-4752(94)90002-7

Verschaffel, L., Greer, B., & De Corte, E. (2000). *Making sense of word problems.* Lisse, The Netherlands: Swets & Zeitlinger.

Verschaffel, L., Greer, B., Van Dooren, W., & Mukhopadhyay, S. (Eds.) (2009). *Words and worlds: Modelling verbal descriptions of situations.* Rotterdam: Sense Publishers.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education*, *27*(4), 458–477. Available from http://www.jstor.org/stable/749877

Yore, L.D., Bisanz, G.L., & Hand, B.M. (2003). Examining the literacy component of science literacy: 25 years of language arts and science research. *International Journal of Science Education, 25,* 689–725. http://dx.doi.org/10.1080/09500690305018

Webb, N.M., Nemer, K. M., & Ing, M. (2006). Small-group reflections: Parallels between teacher discourse and student behaviour in peer-directed groups. *Journal of the Learning Sciences, 15,* 63–119. http://dx.doi.org/10.1207/s15327809jls1501_8

Wood, T., Williams, G., & McNeal, B. (2006). Children's mathematical thinking in different classroom cultures. *Journal for Research in Mathematics Education, 37*(3), 222–252. Available from http://www.jstor.org/stable/30035059

Xin, Z. (2009). Realistic problem solving in China: Students' performances, interventions, and learning settings. In L. Verschaffel, B. Greer, W. Van Dooren, & S. Mukhopadhyay (Eds.), *Words and Worlds: Modelling verbal descriptions of situations* (pp. 161–176). Rotterdam: Sense Publishers.

Xin, Z., & Zhang, L. (2009). Cognitive holding power, fluid intelligence and mathematical achievement as predictors of children's realistic problem solving. *Learning and Individual Differences*, *19*, 124–129. http://dx.doi.org/10.1016/j.lindif.2008.05.006

Xin, Z., Lin, C., Zhang, L., & Yan, R. (2007). The performance of Chinese primary school students on realistic arithmetic word problems. *Educational Psychology in Practice, 23*(2), 145–159. http://dx.doi.org/10.1080/02667360701320853

Zhang, L.J., & Anual, S. (2008). The role of vocabulary in reading comprehension: The case of secondary students learning English in Singapore. *RELC Journal*, *39*(1), 52–77.