

# An Investigation on Chinese Teachers' Realistic Problem Solving Abilities and Beliefs

**Limin Chen**

*Shenyang Normal University, China  
Center for Instructional Psychology and Technology  
Katholieke Universiteit Leuven, Belgium*

**Wim Van Dooren**

**Lieven Verschaffel**

*Centre for Instructional Psychology and Technology  
Katholieke Universiteit Leuven, Belgium*

*Some years ago, Verschaffel, De Corte, and Borghart (1997) investigated Flemish pre-service teachers' own abilities to solve word problems realistically as well as their evaluations of pupils' reactions to these problems, and found that their participants behaved quite unrealistically and rather systematically on both tasks. We replicated the study with Chinese pre-service teachers. When compared to the participants from the study of Verschaffel et al., results first revealed that Chinese pre-service teachers behaved much more realistically not only when solving the seven problematic word problems themselves, but also when evaluating elementary school pupils' problem solving performance. Second, a strong and straightforward relationship was found between the teachers' realistic reactions in the problem solving test, and their evaluations of the pupils' responses in the problem solving questionnaire.*

**Key word:** problem solving, comparative study, Belgium and China

## Theoretical and Empirical Background

Since the nineties, several researchers have looked at the (non-)realistic nature of children's representations and solutions of school arithmetic word problems (see Verschaffel, Greer, & De Corte, 2000, and Verschaffel, Greer, Van Dooren, & Mukhopadhyah, 2009 for overviews of

this research). In a pioneering study, Verschaffel, De Corte, and Lasure (1994) confronted a group of 10-11 years old Flemish pupils with a set of 20 word problems, half of which were standard problems (S-problems) that could be straightforwardly solved by applying the most obvious arithmetic operation(s) with the given numbers (e.g., “A boat sails at a mean speed of 45 kilometres per hour. How long does it take this boat to sail 180 kilometres?”), whereas the other half were contextually problematic problems (P-problems) wherein the appropriate mathematical model or solution is neither obvious nor indisputable (e.g., “John’s best time to run 100 metres is 17 seconds. How long will it take him to run 1 kilometre?”). Pupils’ reactions to the P-problems yielded an alarmingly small number of realistic reactions (RRs), i.e. either a realistic answer (e.g., “Certainly more than 170 seconds” for the above P-problem) or other types of answer which was accompanied with a realistic comment (e.g., “ $10 \times 17 = 170$  seconds. If I neglect that John may get tired after a while and thus may not be able to continue to run at his record speed” for the above P-problem). Replications, using translations of essentially the same set of ten P-problems, have been carried out in many countries, basically with the same results (see Verschaffel et al., 2000, 2009). A recent Chinese replication study by Xin (2009) revealed that Chinese pupils were also performing rather poorly on the P-problems from Verschaffel et al.’s (1994) study, even though the overall percentage of RR’s was somewhat higher than that in most other studies.

The disappointing results concerning children’s non-realistic approach to P-problems led to the question: Where do pupils’ difficulties in solving these P-problems come from? According to Reusser and Stebler (1997), Schoenfeld (1991), and Verschaffel et al. (1994, 2000), they result from the type of mathematics teaching the children received. More specifically, these authors claim that pupils develop such a tendency implicitly, gradually, and tacitly through being immersed in the culture and practice of the mathematics classrooms in which they engage. This enculturation process is claimed to be caused by two aspects of pupils’ instructional practice and culture, namely (1) the nature of the word problems given to the pupils and (2) the way in which teachers conceive and actually treat the problems in their mathematics classroom (Reusser & Stebler, 1997; Verschaffel et al., 1994, 2000).

Notwithstanding the existence of a vast amount of (theoretical) analyses of and reflections upon the characteristics of children’s mathematics lessons that may be responsible for their tendency to approach word problems in a non-realistic way, empirical research on these instructional factors is

rather scarce. One study that sheds some partial light on this issue is an investigation by Verschaffel, De Corte, and Borghart (1997), which analyzed future elementary school teachers' views on the role of real-world knowledge concerning the problem context in the modeling of school arithmetic word problems. Participants came from the first and last year of three teacher-training institutes in Flanders. They were administered a problem solving test wherein participants were asked to solve seven P-problems and seven S-problems, and a problem solving questionnaire wherein participants were asked to evaluate different pupils reactions to the same 14 problems as those in the problem solving test. First, the study revealed that Flemish pre-service teachers themselves produced many non-realistic answers (NAs) to the P-items in the problem-solving test. Second, they gave higher scores for the NAs than for the realistic answers (RAs) to the P-items in the problem solving questionnaire. Third, a strong and straightforward relationship was found between the *non-realistic* reactions (NRs) in the problem solving test and the evaluations of the pupils' RAs and NAs in the problem solving questionnaire, while the congruence between the RRs in the problem solving test and the scorings of the RAs and NAs in the problem solving questionnaire was less straightforward. Finally, there was a significant difference in the overall number of RRs and the scores of the RAs and NAs between the first-year and the third-year pre-service teachers in favor of the latter group.

However, Verschaffel et al. (1997) warn that their results were obtained in the context of the Flemish teacher training system and that it remains a question whether their results can be generalized to other cultural c.q. educational contexts. The goal of the present study was to replicate that study in a Chinese context. More particularly, we wanted to investigate Chinese pre-service teachers' P-problem solving abilities and their evaluations of pupils' reactions to the same P-problems, and compare their results to those obtained by the Flemish pre-service teachers in Verschaffel et al.'s (1997) study. Such a comparison with Chinese pre-service teachers seemed particularly interesting, because Chinese elementary and secondary school students have repeatedly shown to be better performers on mathematical tasks measuring basic mathematical knowledge and routine mathematical skills than their Western peers (Cai & Nie, 2007). Moreover, Ma's (1999) comparative study of the nature and the development of elementary school teachers' elementary mathematics content and pedagogical content knowledge in China and the United States has revealed big differences in favor of the Chinese teachers. So these studies may raise the expectation that Chinese pre-

service teachers will perform relatively well on Verschaffel et al.'s (1997) test. However, there is also some empirical evidence which causes us to be cautious, as some other studies have shown that Chinese students do not outperform their Western counterparts on complex, open-ended tasks measuring non-routine, creative problem solving (Cai & Nie, 2007; Chen, Van Dooren, Chen, & Verschaffel, 2007).

## Method

### Participants

Participants were 208 pre-service elementary school teachers from three different grade levels selected from Shenyang Normal University in China. The number of first-, second-, and third-year pre-service teachers was 72, 66, and 70, respectively. The training program for pre-service teachers in Shenyang Normal University consists of four years, after which these pre-service teachers are awarded a bachelor degree.

### Instruments

The instruments consisted of a problem solving test and a corresponding problem solving questionnaire. In the problem solving test participants were asked to solve seven P-problems and seven S-problems, which were literally translated from those used by Verschaffel et al. (1997). Only the names of the protagonists were replaced by typical names of Chinese children. Table 1 lists the seven P-problems.

Table 1

### Seven P-Problems Involved in the Study

Name	Problem
Buses	450 soldiers must be bused to their training site. Each army bus can hold 36 soldiers. How many buses are needed?
School	Mingming and Dongdong go to the same school. Mingming lives at a distance of 17 kilometres from the school and Dongdong at 8 kilometres. How far do Mingming and Dongdong live from each other?
Runner	Zhanghong's best time to run 100 metres is 17 seconds. How long will it take him to travel 1 kilometre?
Flask	This flask is being filled from a tap at a constant rate. If the depth

of the water is 3.5 centimetres after 10 seconds, how deep will it be after 30 seconds? (This problem was accompanied by a picture of a partly-filled cone-shaped flask)

- Rope A man wants to have a rope long enough to stretch between two poles 12 metres apart. But he has only pieces of rope which are 1.5 metres long. How many of these pieces would he need to tie together to stretch between the poles?
- Planks Lifang has bought 4 planks of 2.5 metres each. How many planks of 1 metre can he get out of these planks?
- Friends Zhangdong has 5 friends and Mingming has 6 friends. Zhangdong and Mingming decide to give a party together. They invite all their friends. All friends are present. How many friends are there at the party?

\*Contrary to all other P-problems, the school problem was presented with two non-realistic response alternatives, because the typical non-realistic response to this P-problem is *either*  $17 - 8 = 9$  or  $17 + 8 = 25$  (Verschaffel et al., 1997).

In the problem solving questionnaire, participants were asked to score four different answers from pupils to the same 14 word problems as in the problem solving test with either 1 point if they considered the pupil's response as totally appropriate,  $\frac{1}{2}$  point for a partly appropriate response, or 0 point for a completely inappropriate response, as in the study of Verschaffel et al. (1997).

The four response alternatives to the seven P-problems in the problem solving questionnaire belonged to four different categories:

- Non-realistic answer (NA), which results from the straightforward and uncritical application of the arithmetic operation elicited by the problem statement (e.g., for the above-mentioned flask item, the NA was 10.5, which is the product of  $3 \times 3.5$ ).
- Realistic answer (RA), which follows from the effective and appropriate use of real world knowledge about the context elicited by the problem statement (the RA for the flask item was "it is impossible to give a precise answer.")
- Technical error (TE), which results from the straightforward and uncritical application of the arithmetic operation elicited by the problem statement, but which differs from the NA because of a purely technical mistake in the execution of the arithmetic operation(s) (e.g.,

responding to the flask item, the answer  $3 \times 3.5 \text{ cm} = 11.5 \text{ cm}$ ).

- Other answer (OA), involving an answer that could not be classified into one of the former categories; for instance, solving the flask item with the result of a wrong operation, such as an addition ( $3.5 \text{ cm} + 20 \text{ cm}$ ) instead of a multiplication.

Again, these NAs and RAs were the same as those in Verschaffel et al.'s (1997) study.

In contrast with the seven P-problems, the seven corresponding S-problems had, of course, no realistic and non-realistic response alternative. The four response alternatives to the S-problems belonged to the following categories: (1) the correct answer, (2) a technical error, (3) a wrong-operation error, and (4) a response that simply stated that the problem was unsolvable.

#### *Procedure and task administration*

Strictly following the procedure for task administration of Verschaffel et al. (1997), the problem solving test and questionnaire were administered on the same day. All pre-service teachers received the problem solving test first. Immediately after they had finished and handed in it, they were given the problem solving questionnaire. To prevent order effects between the different P-problems from each test, two different versions of the problem solving test and questionnaire with inverted sequences of presentation of the P-problems were used in a counterbalanced design. The administration of the test and questionnaire was always done by the teacher educator who was responsible for the mathematics education program in which the pre-service teachers participate.

#### **Data Coding**

*Problem solving test.* Following the data coding system of Verschaffel et al. (1997), pupils' reactions to the P-problems were classified into RRs, i.e. either a realistic answer or other kind answer accompanied with a realistic comment (More details see Verschaffel et al. (1997)), and NRs, whereby neither the answer nor the additional comments showed any evidence of activation of real-world knowledge.

*Problem solving questionnaire.* In accordance with the study of Verschaffel et al. (1997), the analysis of the pre-service teachers' reactions to the seven P-problems in the problem solving questionnaire focused on the score (1,  $\frac{1}{2}$ , or 0) given to the RAs and NAs.

## Results

### Results for the Problem Solving Test

In the problem solving test, the Chinese pre-service teachers demonstrated highly realistic considerations when confronted with the P-problems. Indeed, 76% of all reactions to the seven P-problems could be considered as RRs, which is considerably higher than the 37% that Xin (2009) found for Chinese pupils. Moreover, the Chinese pre-service teachers gave significantly more RRs than their Flemish counterparts, both for all P-problems together ( $z = 6.43, p = .00$ ), and for all individual P-problems (all  $p = .00$ ). Table 2 lists the percentage of RRs for each of the seven P-problems for the Chinese pre-service teachers, as well as the results for the Flemish pre-service teachers from the study of Verschaffel et al. (1997). Interestingly, the differences in the number of RRs between the various P-problems are very similar to those observed in the Flemish pre-service teachers.

Table 2

#### Percentage of Realistic Reactions (RRs) of the Chinese and Flemish Pre-Service Teachers to the Seven P-Problems From the Problem Solving Test

Problem Name	Chinese Pre-Service Teachers ( $n = 208$ )	Flemish Pre-Service Teachers ( $n = 332$ )
Buses	98	90
School	84	48
Runner	58	31
Flask	76	39
Rope	64	37
Planks	94	64
Friends	59	29
Total	76	48

Similar to the results in Verschaffel et al. (1997), there was a significant difference in the overall number of RRs between the three grade levels in favor of the third-year group ( $F(2, 1453) = 9.64, MSE = 1.74, p = .00$ ). More specifically, the percentage of RRs to the P-problems for the three grade levels was 74.8% for the first-year, 70.6% for the second-year,

and 82.4% for the third-year pre-service teachers. Pairwise comparisons showed that the difference between the first- and second-year pre-service teachers was not significant ( $p = .37$ ), but that there was a significant difference between first- and third-year ( $p = .01$ ), and between second- and third-year pre-service teachers ( $p = .00$ ).

### Results for the Problem Solving Questionnaire

The pre-service teachers' strong disposition toward realistic problem solving was also revealed by their evaluations of the RAs and the NAs on the seven P-problems of the problem solving questionnaire. Specifically, in 70% of cases did the RAs receive a score of 1 (versus only 47% in the Flemish study); 11% of the RAs received a  $\frac{1}{2}$ -score and in only 19% of the cases the RA was scored with a 0. On the other hand, the NA was scored with a 1 in only 23% of the cases (versus 56% in the Flemish study), whereas 55% and 22% of the NAs received a  $\frac{1}{2}$ - and 0-score, respectively (see the total percentages given at the bottom of Table 3). Table 3 gives the percentages of 1-,  $\frac{1}{2}$ -, and 0-scores for the RA and for the NA for each of the seven P-problems.

Table 3

#### Percentages of 1-, $\frac{1}{2}$ - and 0-Scores for the RA\* and the NA\*\* by the Chinese and Flemish Pre Service Teachers to the Seven P-Problems From Problem Solving Test

Problem Name	Chinese pre-service teachers						Flemish pre-service teachers					
	RA			NA			RA			NA		
	1	$\frac{1}{2}$	0	1	$\frac{1}{2}$	0	1	$\frac{1}{2}$	0	1	$\frac{1}{2}$	0
Buses	91	8	1	1	92	7	95	3	2	6	73	21
School	68	12	20	9	72	19	46	5	49	66	19	15
Runner	75	10	15	27	44	29	24	5	72	81	12	7
Flask	78	3	19	20	40	40	42	3	55	59	22	19
Rope	80	15	5	20	73	7	24	7	69	72	20	8
Planks	54	9	37	35	25	40	68	13	19	33	18	49
Friends	43	21	36	49	41	10	28	8	64	75	15	10
Total	70	11	19	23	55	22	47	6	47	56	26	18



\* RA: realistic answer; \*\* NA: non-realistic answer

Summarizing these findings for the questionnaire, they reveal, first, that the Chinese pre-service teachers' overall evaluation of the RA to the P-problems was considerably more positive than for the NA,  $\chi^2(2) = 53.31$ ,  $p = .00$ . Compared to their Flemish counterparts from the study of Verschaffel et al. (1997), the Chinese pre-service teachers held a much more positive disposition towards realistic problem solving, since they evaluated the RAs much more with 1-scores than the Flemish pre-service teachers,  $\chi^2(2) = 17.82$ ,  $p = .00$ , and the NAs much less with 1-scores,  $\chi^2(2) = 24.57$ ,  $p = .00$ .

Table 4 reveals that the scores given to the RAs and NAs between the first-, second-, and third-year pre-service teachers in the problem solving questionnaire were very similar, and, thus, did not reveal a significant effect of grade,  $\chi^2(4) = 0.75$ ,  $p = .95$  for the RAs and  $\chi^2(4) = 0.11$ ,  $p = 1.00$  for the NAs.

*Table 4*  
**Percentage of 1-, ½-, and 0-Scores for the RA\* and the NA\*\* on the Seven P-Problems From Problem Solving Test for the Chinese Pre-Service Teachers on the Three Grade Levels**

Grade Level	RA			NA		
	1	½	0	1	½	0
1	23	5	7	9	18	8
2	22	3	6	7	18	7
3	25	3	6	7	19	7
Total	70	11	19	23	55	22

\* RA: realistic answer; \*\* NA: non-realistic answer

### **Relationship Between Problem Solving Test and Problem Solving Questionnaire**

Similar to Verschaffel et al. (1997), we also explored to what extent the pre-service teachers' evaluations for the NAs and the RAs in the problem solving questionnaire matched their own performance in the problem solving test, by separately analyzing the scores for the RAs and the NAs following the NRs (348 out of 1456, i.e. 24%) and the RRs (1108 out of 1456, i.e. 76%) in the problem solving test.

Scores for the RAs and NAs in the problem solving questionnaire for the RRs in the problem solving test. Table 5 presents the distribution of the different combinations of RA and NA scorings over the seven P-problems of the problem solving questionnaire for the 76% of the cases wherein a RR was produced in the problem solving test.

Table 5

**Percentage of Combinations of Scores (1, ½, or 0) for the RA\* and for the NA\*\* Over the Seven P-Problems of the Problem Solving Questionnaire for the Chinese Pre-Service Teachers who Themselves Produced a Realistic Reaction in the Problem Solving Test (76% of all Cases)**

NA-Scores	RA-Scores			Total
	1	½	0	
1	3	2	8	13
½	52	5	5	62
0	23	1	1	25
Total	78	8	14	100

\* RA: realistic answer; \*\* NA: non-realistic answer

From Table 5, we can see that the relationship between the RRs in the test and the evaluations of the RAs and NAs in the questionnaire was rather strong and in the expected direction. Indeed, 78% of all the cases where a participant reacted on a P-problem in realistic way in the problem solving test, was followed by scoring the RA on that same item with a 1-score in the questionnaire. Accordingly, almost all participants who gave a RR to a P-problem in the problem solving test were reluctant to give a 1-score to the NA to that item in the questionnaire, but preferred to give it either gave it a ½ (62%) or 0 (25%).

So, the two most common combinations – which together accounted for 75% of all cases in which participants had produced a RR on the same P-problem themselves in the problem solving test – were a 1-score for the RA combined with a ½-score (52%) or a 0-score (23%) for the NA (see Table 5). Participants who combined a 1 for the RA with a 0 for the NA based their scorings on whether or not attention was paid to the realistic modelling complication. With their 0-score for the NA, they expressed that seeing this complication – as they had done themselves in the problem solving test and as

the pupil with the RA had done in the questionnaire – was the crucial aspect of the problem. Hereafter we give one typical example of participants' explanations for their 1-score for the RA in combination with a 0-score for the NA.

Participant III-11/Flask item<sup>1</sup> (RA: It is impossible to give a precise answer; NA:  $3 \times 3.5 = 10.5$ . After 30 seconds, the level of the water will be 10.5 cm)

Motivation for the 1-score for the RA: "The shape of the flask is not regular, so it is indeed impossible to give a precise answer."

Motivation for the 0-score for the NA: "This pupil did not read the problem very carefully, and mistook the coned-shaped flask as a cylinder-shaped flask with which (s)he probably is more familiar."

However, the very high percentage of combinations of a 1-score for RA and a  $\frac{1}{2}$ -score for NA (namely 52%) indicated that in many instances where participants had reacted themselves to a P-problem in a realistic manner, they were nevertheless rather understanding and tolerant towards pupils who interpreted and solved these P-problems without taking into account the relevant real world knowledge. According to their written explanations, they thought that even though the pupils did not consider the reality involved in the problem, they selected the correct mathematical operation to solve the problems, which they considered to be an important appropriate step in the problem solving process, and so they gave a  $\frac{1}{2}$ -score. This is illustrated in the following example.

Participant III-50/Buses item (RA: 450 divided by 36 is 12.5. So 13 buses are needed; NA: 450 divided by 36 is 12.5. So 12.5 buses are needed)

Motivation for the 1-score for the RA: "He/she considered the reality that the remaining 18 soldiers also need one bus."

Motivation for the  $\frac{1}{2}$ -score for the NA: "The mathematical operation is correct, but the answer is wrong because the number of buses can't be a decimal; so I scored it with  $\frac{1}{2}$  point."

*Scores for the RAs and NAs in the problem solving questionnaire for the NRs in the problem solving test.* Table 6 presents the distribution of the

---

<sup>1</sup> For each example, we give the identification number of the participant who produced this scoring combination and the accompanying motivation. The first part of that number refers to the pre-service teacher's grade (I, II, or III) and the second part refers to his/her ID number within that grade. The complete text of the P-problems is provided in Table 1.

combinations of RA scorings (1,  $\frac{1}{2}$ , and 0) and NA scorings (1,  $\frac{1}{2}$ , or 0) over the seven P-problems of the problem solving questionnaire for the NRs (24%) in the problem solving test.

Table 6

**Percentage of Combinations of Scores (1,  $\frac{1}{2}$  or 0) for the RA\* and for the NA\*\* Over the Seven P-Problems of the Problem Solving Questionnaire for the Chinese Pre-service Teachers who Themselves Produced a Non-Realistic Reaction in the Problem Solving Test (24% of all Cases)**

NA-scores	RA-scores			Total
	1	$\frac{1}{2}$	0	
1	5	16	33	54
$\frac{1}{2}$	27	3	3	33
0	10	1	2	13
Total	42	20	38	100

\* RA: realistic answer; \*\* NA: non-realistic answer

Compared to the participants who had reacted realistically themselves on the P-problems of the problem solving test, for the participants who had themselves responded to these problems in a non-realistic way, the evaluation of the RAs and NAs in the questionnaire was somewhat less in line with their own solution behavior: Only in 49% of the cases wherein a participant had produced a non-realistic reaction to a given P-problem, did (s)he gave a 1-score to the NA in combination with a  $\frac{1}{2}$ - or a 0-score to the RA. Hereafter we provide a more detailed analysis of those scoring combinations.

Not surprisingly, the most frequent combination, which occurred in 33% of the cases in which a participant him/herself had responded in a non-realistic way, was a 1 for the NA in combination with a 0 for the RA. This scoring combination is perfectly in line with the participants' own non-realistic interpretation and solution of the P-problems in problem solving test. They scored the NA with 1 because they had given themselves a NA in the problem solving test, and they scored the RA with 0 because they could not understand or appreciate the context-based considerations underlying the RA alternative. Hereafter we give one example of such straightforward cases.

Participant II-09/Runner item (RA: It is impossible to answer precisely what Zhanghong's best time on 1 kilometre will be; NA:  $17 \times 10 = 170$ . Zhanghong's best time to run 1 kilometre is 170 seconds)

Motivation for the 0-score for the RA: "The pupil who selected this alternative misunderstood the problem. Maybe he/she thinks

Zhanghong took a rest for a while after he ran 100 metre, but the period for rest was not given in the problem statement.”

Motivation for the 1-score for the NA: “I think this alternative is completely correct; so I scored it with 1 point.”

As Table 6 further shows, 42% of the RAs in the problem solving questionnaire were given a 1-score either in combination with a 1 (5%), a  $\frac{1}{2}$  (27%), or a 0 (10%) for the NAs, even though the participants had provided a NR to a P-problem in the problem solving test themselves. According to their written explanations, in many cases the confrontation with the RA in the problem solving questionnaire seemed to have functioned as a “scaffold” toward a more realistic approach to the problem. So, these participants, who had reacted themselves in a non-realistic manner to a P-problem in the problem solving test, noticed the realistic modeling difficulty when reading the RA alternative in the questionnaire, and therefore awarded the RA with a 1 point. Remarkably, their 1-scores for the RAs were not always accompanied by 0-scores for the NAs. In only 10% of these 42%, the 1 for the RA was accompanied by a 0 for the NA; 27% gave a  $\frac{1}{2}$ -score for the NAs and 5% gave also a 1-score for the NA. The interpretation for that somewhat ambiguous score might be that in many cases the participants noticed the realistic modeling difficulty only at the time they saw the pupil’s RA in the questionnaire, and they remained quite understanding and tolerant towards elementary school pupils who interpreted and solved these P-problems without seriously taking into account the relevant real-world knowledge (as they as teachers initially had done themselves). Exemplary support for this interpretation is provided in the following example.

Participant II-20/Rope item (RA: It is impossible to know how many pieces of rope you will need; NA:  $12 \div 1.5 = 8$ . Eight pieces of 1.5 metres are needed)

Motivation for the  $\frac{1}{2}$ -score for the NA: “From a purely mathematical perspective, we can compute like that, but taking into account the reality of the problem it is not right.”

## Discussion

The present study with Chinese pre-service teachers, which replicated Verschaffel et al.’s study (1997) with Flemish future teachers, investigated Chinese pre-service teachers’ own abilities to solve problems realistically and their evaluations of different pupil reactions to the same problems. Results

revealed, compared to the participants from the study of Verschaffel et al. (1997), Chinese pre-service teachers behaved much more realistically not only when solving the seven problematic word problems themselves, but also when evaluating elementary school pupils' problem solving performance. Furthermore, there was a strong and straightforward relationship between the RRs in the problem solving test, and the evaluations of the pupils' RAs and the NAs in the problem solving questionnaire, and a less straightforward congruence between the NRs in the problem solving test and the scorings of the RAs and NAs in the problem solving questionnaire. We end this article with a reflection on some theoretical, methodological, and educational issues that need to be addressed in further research.

A first intriguing issue is why the Chinese pre-service teachers reacted more realistically than their Flemish counterparts. Given that, first, neither the Flemish nor the Chinese sample was sufficiently large and representative for Flemish and Chinese pre-service teachers in general, and, second, that the Flemish data were collected already several years ago, it would be unwarranted to drive strong conclusions about the superiority of Chinese pre-service teachers as far as their disposition towards realistic word problem solving is concerned. But if the present findings concerning Chinese pre-service teachers' disposition towards realistic word problem solving would be confirmed in future studies, it would provide further evidence against the frequently heard argument that Chinese learners may be good at standard problems but weak on open-ended, non-routine problems (Cai & Nie, 2007), and would indicate the need for further research into what elements in the Chinese mathematics education culture and practice lead to this remarkably good performance of Chinese pre-service teachers, also on those P-problems requiring non-routine and realistic thinking.

Second, this study sheds some light on the complex relationship between pre-service teachers' knowledge and beliefs, on the one hand, and their teaching behaviour in the field of realistic mathematics on the other hand. It does so by showing strong links between how the pre-service teachers solve word problems themselves and how they evaluate pupils' answers (which can be considered as one very important aspect of their teaching). However, at the same time, it reveals that the link between a teacher's own word problem solving and his/her evaluation of a pupil's response is less simple and straightforward than one might expect at first sight. Participants' written explanations gave us some insight into how various kinds of pedagogical content knowledge and beliefs may interfere in this relationship. However,

several participants just gave scores without any further explanation or with only a very short and vague comment, which made it sometimes difficult to probe the considerations and arguments behind these scores. Therefore, to obtain richer data about these participants' beliefs on realistic problem solving, it would be interesting to complement, in future research, the administration of paper-and-pencil tests and questionnaires with individual interviews with a carefully selected sub-sample of participants.

Third, we emphasize that in the present study we did not assess (future) teachers' actual teaching c.q. evaluation behaviour in the real classroom, but only a kind of *simulation* of this important part of their complex and multi-dimensional teaching task. It would, therefore, be interesting to investigate how (future) teachers actually react to different kinds of student responses in the dynamic complexity of a real classroom situation. For some recent attempts into that direction, see Depaepe, De Corte, and Verschaffel, (2010).

Fourth, contrary to the Chinese pre-service teachers who participated in the present study, research by Xin (2009) has revealed that Chinese upper elementary school pupils perform rather poorly on P-problems. So we are confronted with the intriguing question of how to relate Chinese pupils' rather weak results on P-problems to the quite positive results for Chinese (future) teachers. If the available contrastive findings for Chinese pupils and (future) teachers would be replicated in future research (see next point), we would have to conclude that Chinese students' rather poor performances on these realistic problem solving items can certainly not be explained by teachers' weak realistic problem solving capacities and negative beliefs vis-à-vis realistic problems, as Verschaffel et al. (1997) did at the end of their study with Flemish (future) teachers. Alternative explanations referring to the broader (mathematics) educational contextual constraints wherein these teachers have to operate, might then help to explain what emerges from these contrastive data. For example, the pressure to complete the curriculum and to perform well on high-stake tests might force the teachers to focus their instructional attention on some more "classical" mathematical skills (that take the most dominant place in the curriculum, the textbook, and the assessment instruments), and, consequently spend less time and energy on less prominent skills, such as the ability to solve problems realistically, despite their own sufficiently sophisticated skills and beliefs (Cai & Nie, 2007).

Finally, although the present study provided some insight into one of the instructional factors considered responsible for the tendency among children to neglect realistic considerations in their problem solving

endeavours, namely teachers' own capacity and willingness to take reality seriously in arithmetic word problem solving at school, it did not provide *direct* evidence of the causal relation between the teachers' realistic problem solving knowledge, beliefs and practices, on the one hand, and their students' abilities and beliefs on the other hand. This would require another kind of research – for instance a study with a longitudinal design with repeated measurements of pupils as well as their teachers – that allows researchers to explicitly and directly link data about the (un)realistic problem solving behaviour and accompanying beliefs of a group of pupils with data about the corresponding knowledge, beliefs and instructional practices of their teachers.

### References

- Cai, J., & Nie, B. (2007). Problem solving in Chinese mathematics education: Research and practice. *ZDM-The International Journal on Mathematics Education*, 39, 459–473.
- Chen, L., Van Dooren, W., Chen, Q., & Verschaffel, L. (2007). The relationship between posing and solving arithmetic word problems among Chinese elementary school children. *Journal of the Korea Society of Mathematical Education Series D: Research in Mathematical Education*, 11(1), 1-31.
- Depaepe, F., De Corte, E., & Verschaffel, L. (2010). Teachers' approaches towards word problem solving: Elaborating or restricting the problem context. *Teaching and Teacher Education*, 26, 152-160.
- Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. New York: Erlbaum.
- Reusser, K., & Stebler, R. (1997). Every word problem has a solution: the suspension of reality and sense-making in the culture of school mathematics. *Learning and Instruction*, 7, 309-328.
- Schoenfeld, A. H. (1991). On mathematics as sense-making: An informal attack on the unfortunate divorce of formal and informal mathematics. In J. F. Voss, D. N. Perkins, & J. W. Segal (Eds.), *Informal reasoning and education* (pp. 311-343). Hillsdale, NJ: Erlbaum.
- Verschaffel, L., De Corte, E., & Borghart, I. (1997). Pre-service teachers' conceptions and beliefs about the role of real-world knowledge in mathematical modelling of school word problems. *Learning and Instruction*, 4, 339-359.



- Verschaffel, L., De Corte, E., & Lasure, S. (1994). Realistic considerations in mathematical modelling of school arithmetic word problems. *Learning and Instruction, 4*, 273-294.
- 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: Modeling verbal descriptions of situations*. Rotterdam: Sense Publishers.
- Xin, Z. (2009). Realistic problem solving in China. In L. Verschaffel, B. Greer, W. Van Dooren, & S. Mukhopadhyay (Eds.), *Words and worlds: Modeling verbal descriptions of situations* (pp. 161–176). Rotterdam: Sense Publishers.

Acknowledgement: This research was supported by Grant Education Sciences Planning (JG10DB223) “Experimental research on the development of pupils’ problem posing ability in Shenyang City” from the Research Fund Shenyang Educational Committee and by Grant GOA 2006/01 “Developing adaptive expertise in mathematics education” from the Research Fund of the Katholieke Universiteit Leuven, Belgium.

**Authors:**

*Chen Limin*

*Shenyang Normal University*

*E-mail: Chenlimin\_leuven@126.com*

*Wim Van Dooren*

*Center for Instructional Psychology and Technology, Katholieke Universiteit Leuven, Belgium*

*E-mail: Wim.VanDooren@ppw.kuleuven.be*

*Lieven Verschaffel*

*Centre for Instructional Psychology and Technology, Katholieke Universiteit Leuven, E-mail: Lieven.Verschaffel@ppw.kuleuven.be*