INFLUENCE OF SITUATIONAL AND MATHEMATICAL INFORMATION ON SITUATIONALLY DIFFICULT WORD PROBLEMS

Santiago VICENTE[1], Josetxu ORRANTIA[1], Lieven VERSCHAFFEL[2]

[1] University of Salamanca, Faculty of Education
Pso. Canalejas 169, 37008 Salamanca, Spain
E-mail: sanvicente@usal.es; orrantia@usal.es

[2] Catholic University of Leuven, Centre for Instructional Psychology and Technology
Vesaliusstraat, 2, B-3000 Leuven, Belgium
E-mail: Lieven.Verschaffel@ped.kuleuven.be

Abstract: Arithmetical word problem solving has a double nature, textual and mathematical, that is, several cognitive processes related to the comprehension of a word problem as a text and as a mathematical structure to be grasped and manipulated are involved in the word problem solving process and influence children's task performance. Theoretical models and empirical studies have accounted for the implications of these different processes for word problem solving. Based on the rewording methodology a new study was set up to assess the influence of two kinds of rewording - mathematical and situational - on two-step change problems of two different levels of mathematical and situational difficulty. Results showed that while mathematical rewording improved children's performance on mathematically as well as situationally difficult problems, situational rewording was not useful for any of them.

Key words: word problem solving, conceptual knowledge, rewording, mental model

The cognitive processes underlying arithmetical word problem solving have been intensively studied by cognitive and instructional psychologists over the past 25 years. As far as addition and subtraction word problems involving only one arithmetical operation are concerned, there are three major types of problems: change, combine, and compare problems (Riley, Greeno, Heller, 1983; see also Fuson, 1992; Reed, 1999; Verschaffel, De Corte, 1993, 1997). Some authors have suggested a fourth category, namely "equalize problems", which involve a change of a quantity within a compare situation and, thus, combine features of a change and a compare problem. Within each of these four basic semantic categories, further distinctions have been made based on the direction of the action or the relation (e.g., an increase or a decrease of the initial set in a change structure) and on the identity of the unknown set (e.g., the start set, the change set, or the result set in a change structure), resulting in 20 different types of one-step addition and subtraction problems.

Notwithstanding the vast amount of research that has been done on this type of word problem over the past 25 years, there is still much unclearness and discussion about which cognitive processes underlie

Acknowledgments: This research was supported by Grant SEJ2006/11249 to J. Orrantia from the Ministerio de Educación y Ciencia in Spain and by Grant GOA 2006/01 from the Research Fund K.U. Leuven in Belgium.

This article was accepted before May 1, 2008, in accordance with previous publishing rules.

children's understanding and solution of these problems. Especially the nature of the cognitive schemes and processes - other than the (purely) mathematical ones - that are involved in the initial phases of the word problem solving process is still a matter of serious debate (Verschaffel, Greer, De Corte, in press). The main goal of this article is to assess the importance of a good and rich understanding of the situation evoked by the problem text, especially when the problem difficulty is mainly or also caused by situational (rather than mathematical) complexities.

THEORETICAL FRAME

Two different kind of theoretical models have been proposed for explaining how people solve arithmetical problems. Most classical ones (Briars, Larkin, 1984; Riley et al., 1983; also Kintsch, Greeno, 1985) are based on the idea that word problems can be solved by applying mathematical knowledge only, that is, by using mathematical knowledge to select the numerical data and the arithmetical operation needed for solving them. This mathematical knowledge allows children to detect and operate with the part-whole structure underlying the three kind of arithmetical problems, and to build a mathematical mental model in which the sets involved in the problems and their mutual relations are represented. More recent models (Reusser, 1988; Kintsch, 1988, 1998) are based on the assumption that word problem solving involves not only mathematical but also situational thinking. For example, Reusser's model claims that, before the mathematical (quantitative) model is generated, a necessary step in the problem solving process is the generation of a qualitative representation (= the Episodic Situation Model) of the situation described by the problem's text. Only when this qualitative mental model has been generated can the mathematical model be generated by abstracting the sets involved in the situation and the part-whole relations between them.

Empirical studies have found that small changes in the wording of problem texts, aimed at modifying (see e.g., Hudson, 1983) or clarifying (see e.g., Cummins, Kintsch, Reusser, Weimer, 1988; Cummins, 1991; Davis-Dorsey, Ross, Morrison, 1991; De Corte, Verschaffel, De Win, 1985; Staub, Reusser, 1992; Stern, Lehrndorfer, 1992) the mathematical and/or the situational problem model, may have a dramatic (positive) impact on children's solution processes and skills. Considering the two kind of theoretical models (mathematical and situational), these studies on rewording can be grouped according to the kind of extra information they regard as helpful for improving children's understanding of word problems: studies that focused on the explicitation of the mathematical structure of the problem (Cummins, 1991; Davis-Dorsey et al., 1991; De Corte et al., 1985) and those in which some aspects of the situation in which the problem is embedded was modified (Hudson, 1983; Cummins et al., 1988; Davis-Dorsey et al., 1991; Staub, Reusser, 1992; Stern, Lehrndorfer, 1992). While studies on mathematical rewording were quite uniform in their rewording of the problems[1], studies on situational rewording were much less homogeneous in their definition and operationalization of the

[1] These studies did so first, by including a sentence identifying the initial state set that was added at the beginning of change-final set unknown problems; second, by including the expressions "X of these ____" and "the rest" for combine-part unknown problems; and finally, by modifying the usual question in compare problems: "How many more birds than worms are there?" is changed to another question: "How many birds won't get a worm?"

rewording factor (see Table 1), and suffered from some important methodological limitations. First, when modifying the situational context of the problem, some of these studies also modified the problem's mathematical features. For example, by introducing in the question an action that children could easily imagine, Hudson (1983) actually transformed the original compare problem into an equalize problem, which is known to be easier to solve (see Table 1). Second, a detailed and formal description and account of the (different kinds of) modified texts and their relation to the original standard versions is missing. For instance, Cummins et al. (1988) did not clarify the nature of the relation between the information added and the information given in the problem. A second example can be found in Staub and Reusser's (1992) study: they included very different kinds of modifications in the same reworded problem - syntactical as well as situational - and, consequently, only an approximative speculation could be made about the relative influence of each type of modification distinguished by these authors.

Table 1. Rewordings used in previous studies

Note: CH - change; CB - combine; CP - compare. Words in italics indicate the extra information added compared with the standard formulation

Study	Problem type	Example(s) of standard problems	Example(s) of reworded problem(s)
Hudson (1984)	CP1	"There are five birds and three worms. How many more birds than worms are there?"	"Here are some birds and here are some worms. *Suppose the birds all race over and each one tries to get a worm.* How many birds *won't* get a worm?"
De Corte et al. (1985)	CH5	Joe won 3 marbles. Now he has 5 marbles How many marbles did Joe have in the beginning?"	*"Joe had some marbles.* He won 3 *more* marbles. Now he has 5 marbles. How many marbles did Joe have in the beginning?"
	CB2	"Tom and Ann have 9 nuts. Three nuts belong to Tom. How many nuts does Ann have?"	"Tom and Ann have 9 nuts *altogether*. Three *of these* nuts belong to Tom. *The rest belong to* Ann. How many nuts does Ann have?"
	CP1	"There are six children and three chairs. How many more children than chairs are there?"	"There are 6 children *but* there are *only* 3 chairs. How many children *won't get* a chair?"

Table continues

Table 1 (continued)

Study	Problem type	Example(s) of standard problems	Example(s) of reworded problem(s)
Davis-Dorsey et al. (1991)	CH5 CB2 CP1	"John walked 3/5 of a mile to see a movie. Later he walked to Mike's house. John walked 4/5 of a mile altogether. How far did John walk from the movie to Mike's house?"	The wording was done the same as in De Corte et al. (1985), and additionally, there was personalization with children's favorite movie, household pets' names, favorite food, and friends' names: "*(Best friend)* walked 3/5 of a mile to see *(favorite movie)*. Later he walked to *(other friend's)* house. *(Best friend)* walked 4/5 of a mile altogether. How far did *(best friend)* walk from the *(favorite movie)* to *(other friend's)* house?
Cummins (1991)	CB2	"Joe won 3 marbles. Now he has 5 marbles How many marbles did Joe have in the beginning?"	Same as De Corte et al. (1985), except for the first sentence, which is: "There are 9 nuts" instead of *"Joe had some marbles"*.
Cummins et al. (1988)	CB2 CH5 CH6 CP5	"Jane paid 13 dollars for her racquet. She paid 5 dollars more than Mimi. How many dollars has Mimi paid?"	Example of a CP 5 problem (a similar rewording was applied to the other three kinds of problem): *Jane and Mimi play tennis together twice a week. They both always try hard to beat each other. Both of them decided to buy new tennis racquets. So far* Jane has paid 13 dollars for her racquet. She paid 5 dollars more than Mimi. How many dollars has Mimi paid?

Table continues

Table 1 (continued)

Study	Problem type	Example(s) of standard problems	Example(s) of reworded problem(s)
Stern & Lehndorfer (1992)	CP1 CP2 CP3 CP4 CP5 CP6	"Peter has 6 crayons. Laura has 4 crayons. How many fewer crayons does Laura have than Peter?"	Example of a CP 2 problem (a similar rewording was applied to the other five kinds of compare problem): *Peter is Laura's older brother. Because he is older, his bedroom is larger and his toys are more expensive than Laura's. Peter also gets more pocket money than Laura and he has a new bike whereas Laura has Peter's old bike. When Peter does his homework, Laura doodles a little bit*, Peter has 6 crayons. Laura has 4 crayons. How many fewer crayons does Laura have than Peter?
Staub & Reusser (1992)	CH1 CH2 CH5 CH6	"Peter gave Mary 7 apples. Now he has 4 apples. How many apples did Peter have?"	Example for a CH 6 problem (a similar rewording was applied to the other kinds of problem): Peter has 4 *apples now*. *Today* Peter gave Mary 7 apples. How many apples did Peter *pick yesterday?*)

More recently, some new rewording studies have been realized, which all took Reusser's (1988) model as their explicit theoretical framework (Moreau, Coquin-Viennot, 2003; Coquin-Viennot, Moreau, 2007; Vicente, Orrantia, Verschaffel, in press), but the results have not shown that including situational information has any (positive) influence on children's problem-solving performance. Moreau and Coquin-Viennot (2003) found that fifth-grade children were able to distinguish the information that is essential for *solving* two-step change problems (= mathematical information) from information that is necessary for *understanding* the problem (= situational information). In addition, Coquin-Viennot and Moreau (2007) found that children's achievement was better when solving compare (but not change) problems that included compatible situational descriptions of the situation (i.e., "A bicycle slightly damaged is 120 euro less expensive than a perfect one, which costs 750 euro. How much does the damaged bicycle cost?" than those in which an incompatible situation was described (i.e., "A fully-equipped bicycle is 120 euro less expensive than a normal one that costs 750 euro. How much is the fully-equipped bicycle?").

However, in Moreau and Coquin-Viennot's (2003) first study nothing was said about the influence of situational information on children's actual problem-solving performance, and because the second one (Coquin-Viennot, Moreau, 2007) did not include any standard versions of the word problems as a control variable, it was impossible to assess the relative positive effect of compatible contexts or the relative negative impact of incompatible ones on children's understanding of standard word problems.

Finally, Vicente et al. (in press) developed a fine-grained and formal description and account of the textual adaptations, and applied this descriptive system to generate three different reworded versions of change, two-step problems. Standard versions did not contain any additional information; mathematically reworded versions highlighted the mathematical (part-whole) structure of the problem, and situationally reworded versions described the intentional, temporal and causal structure of the situation, following Reusser's (1988) definitions of Episodic Situation Model. The results showed that while mathematical rewording significantly improved children's achievement, situational rewording did not. After having experimentally shown that these differential results for mathematical and situational rewording could not be explained by the length of the resulting problem text, these authors presented two alternative explanations for the ineffectiveness of situational rewording: first, their original word problems were situationally so transparent and easy for the children that there was no need for additional information about that situation; second, word problem solving is basically, as the classical models imply, a mathematical activity, in which the construction of an appropriate Episodic Situation Model (as claimed by Reusser, 1988) is of minimal importance.

Recently, Vicente et al. (in press) argued that Cognitive Load Theory (Sweller, 1999; Sweller, Chandler, 1991; Paas, Renk, Sweller, 2003) may help to interpret the findings of Vicente et al.'s study. This theory claims that all cognitive tasks imply three kinds of working memory load: first, *intrinsic load*, which comes from the difficulty of the task itself, and which cannot be modified without changing the essentials of the task, second, *germane load,* which is related to the processes that allow a deeper understanding of the information and lead to a better achievement, and, third, *extraneous load*, referring to all processing that is not necessary or useful for attaining the goal. Intrinsic load cannot be modified, but germane and extraneous load can be. So, taking into account that these three kinds of processing share the resources of the working memory, the goal of instructional design is, according to this theory, to try to increase germane processing (without overloading working memory's resources) and decrease extraneous processing (by removing everything unnecessary or irrelevant or misleading). So, in terms of this theory, by rewording word problems Vicente et al. tried to add relevant information with the goal of enhancing germane processing without (simultaneously) increasing extraneous processing. According to their results and interpretation, extra mathematical information can be considered to be relevant and to imply germane processing while extra situational information can be considered to be irrelevant and to imply extraneous load. So, the results obtained by Vicente et al. fit nicely into Cognitive Load Theory.

The goal of the present study was to know whether extra situational information is useful for problem solving when the problematicity of the task has not only a mathematical but also a situational source. We defined and operationalized situational difficulty and situational rewording as in

Staub and Reusser's (1992) study: we created situationally difficult problems by designing problems in which the order in which the events are described in the text does not follow the actual sequence of the events and then we generated versions of these problems that did or did not contain extra situational information that was expected to be helpful in disentangling the actual temporal structure of the events. By doing this, we wanted to investigate whether additional situational c.q. temporal information is useful for solving problems whose difficulty comes (also) from situational complexities, just as additional mathematical information had already proved to be an effective scaffold for representing and solving mathematically complex problems (Vicente et al., in press).

METHOD

Sample

The task was administered to a sample of 131 children - 38 3rd-graders, mean age 8 years 4 months; 44 4th-graders, mean age 9 years 3 months; and 49 5th-graders, mean age 10 years 5 months - from a school in Salamanca, Spain. 62 of them were male and 69 were female.

All students' mathematical ability was measured by means of the mathematical scale of a Spanish test called BADYG (Yusté, 1985). Based on children's score on this test, the sample was divided into three different groups: low-achievement (score below C 30, n = 20), average-achievement (score between C 30 and C 70, n = 43) and high-achievement children (score higher than C 70, n = 68).

Tasks

All problems used for the studies were two-step change problems like the one shown in Figure 1.

Variables included in the problems were mathematical difficulty (easy vs. difficult), situational difficulty (regular vs. irregular order of events), and type of extra information (standard, mathematical, or situational).

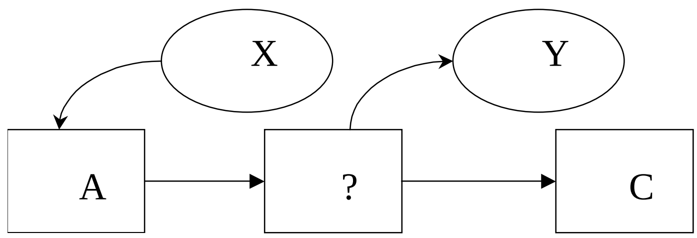


Figure 1. Schematic representation of the experimental problems

Note: the schematic representation corresponds to the following prototypical problem: *Peter had A meters of cable. He bought X meters more cable. He used Y meters of cable and he ended up with C meters of cable. How many meters of cable did he buy/ use?*

Mathematical Difficulty Level

Problems could be easy or difficult depending on the nature of the unknown set in the problem (the meters of cable bought (= set A, see Figure 1) or the meters of cable used (= set B), respectively). Mathematically easy problems (MATH-EASY) implied a combination of two consistent[2] change situations. Mathematically difficult (MATH-DIFF) problems implied a combination of two inconsistent change situations (Riley et al., 1983) (see Figure 1).

Situational Difficulty Level

This variable was defined by the order in which the events of the situation were described. Following Staub and Reusser (1992), two different versions of the problems were generated: "situationally easy" and "situationally difficult" problems (SIT-EASY, and SIT-DIFF, respectively). The former describe the events by following the natural order in which they would appear in real life. SIT-DIFF problems give the information about the initial moment of the action string only at the very end of the problem text.

Rewording

Three different versions of reworded problems were included: standard, mathematical and situational. Standard problems did not contain any additional information. The mathematical reformulation makes the mathematical role of the total set shared by the two one-step change situations[3] explicit, just as Vicente et al. (in press) did in their study. Problems with situational information contained information added to the standard problem in order to highlight the temporal structure of the problem situation. Taking into account Vicente et al.'s (in press) results, we decided to enrich the standard problem text with additional temporal markers, but with neither causal nor intentional ones. The extra situational information underlined the position of each moment in the temporal sequence: initial state, first change, second change and final state. We also decided to include a new moment into this sequence ("after buying those meters of cable...), by means of an action (...he began a renovation") that highlighted the "transition" from the first change situation to the second one. In sum, the total number of experimental problems was 2 (MATH-EASY vs. MATH-DIFF problems) X 2 (SIT-EASY vs. SIT-DIFF) X 3 (standard, situational, or mathematical) = 12. Examples for each one of the experimental problems are shown in Appendix 1.

A propositional analysis was carried out to determine the differences between the different kinds of problem at a textual level, following Graesser and Goodman's (1985; see Vicente et al., in press) propositional analysis system. This analysis allowed us to check that the different reworded versions contained the same

[2] Consistent problems are problems in which superficial structure is in agreement with the operation needed for solving the problem. Inconsistent problems are problems in which superficial structure and the operation needed are opposite. That is, while for consistent problems the operation elicited by the key word (in our case: "win") is the same as the actual operation to be done (i.e., an addition), in inconsistent problems the operation to be performed (namely: subtraction) is different from the one suggested by the key word (i.e., "win). For this reason, consistent problems are easier to solve than inconsistent ones.

[3] As we have already described, the two steps in a two-step problem comprised a simple change situation. Each simple situation consisted of a start set, a change set and a result set, and the set that played the role of result set in the first simple change problem was at the same time the start set of the second simple change problem.

statement nodes[4] across mathematical and situational difficulty levels, meaning that textual features of the problems were similar in all conditions. Hence this variable should not influence the results.

Predictions

The study aimed at testing the following five predictions, the first one being a replication of what was already found in our previous study (Vicente et al., in press).

First, children will perform better on the reworded versions that enhance the creation of the mathematical model than on the standard versions, but only for MATH-DIFF + SIT-EASY problems (and not for the three other problem types, namely MATH-DIFF + SIT-DIFF, MATH-EASY + SIT-EASY and MATH-EASY + SIT-DIFF). We don't expect any influence of extra mathematical information on these problems because intrinsic load of MATH-DIFF + SIT-DIFF will be too high, leaving no room for the germane processing induced by extra information; MATH-EASY + SIT-EASY are so easy that no extra information is needed to solve them; and finally the difficulty of MATH-EASY + SIT-DIFF does not have a mathematical source, so extra mathematical information will not be helpful.

[4] Following Graesser and Goodman, "a statement node is very similar but not equivalent to a proposition. Like a proposition, a statement node contains a predicate and one or more arguments (...) but may contain more than one proposition" (1985, p. 115). They argued for example, that arguments needed for some verb predicates (those which denote cognition, perception, intentions and communication) must be included in the same node. This way, the sentence "The Egyptian knew that wagons have wheels" contains 5 propositions: "Knew (A1, Proposition 2)"; "Possess (X1, X2)"; "Egyptian (A1); "Wagons (X1)" and "Wheels (X2") although it can be considered as one statement node.

Second, children will perform better on the reworded versions that enhance the creation of the situational model than on the standard versions, but only on MATH-EASY + SIT-DIFF problems (and not on the three other problem types, namely MATH-DIFF + SIT-DIFF, MATH-EASY + SIT-EASY and MATH- DIFF, + SIT-EASY for the same reasons as our first prediction for MATH-DIFF + SIT-DIFF and MATH-EASY + SIT-EASY, and because the difficulty of MATH-DIFF + SIT-EASY has not a situational source, so extra situational information will not be helpful).

Third, by including children's mathematical ability level as a variable we expect to assess the cognitive cost of processing mathematical and situational information. That is, if the processing of the extra information that is provided in the reworded versions implies a low cognitive load, children of all ability levels should improve their performance on mathematically reworded versions of MATH-DIFF + SIT-EASY problems and on situationally reworded versions of MATH-EASY + SIT-DIFF problems compared with standard versions. However, if the cognitive cost of processing these textual aids is high, only high ability children will show such improvement on their achievement.

Fourth, regarding the type of errors expected for each problem type and version, two different predictions can be formulated. First, we expect children to commit more errors implying some mathematical comprehension on mathematically reworded problems compared with the standard ones, but only on MATH-DIFF + SIT-EASY problems (and not on MATH-DIFF + SIT-DIFF, MATH-EASY + SIT-EASY and MATH-EASY + SIT-DIFF) (= prediction 4a); second, children will demonstrate higher levels of mathematical understanding (via a better situational understanding) when making an error than on standard versions, but only on SIT-DIFF + MATH-EASY problems (and not on MATH-DIFF + SIT-DIFF, MATH-EASY + SIT-EASY and MATH-DIFF, + SIT-EASY) (= prediction 4b).

Our fifth and last prediction concerns the relative impact of the mathematical and situational difficulty levels on children's performance. If the mathematical and situational complications have a similar influence on the task difficulty, effect size of the differences between the MATH-EASY and the MATH-DIFF problems and between the SIT-EASY and the SIT-DIFF problems should be similar (= prediction 5a). If, however, increasing the mathematical difficulty of the problems has a greater impact on children's performance than increasing their situational difficulty, effect size of the differences between MATH-EASY and MATH-DIFFICULT problems should be bigger than between SIT-EASY and SIT-DIFF problems (= prediction 5b).

Procedure

The 12 experimental problems were distributed over three different administrations, each one with 4 experimental problems. During each administration we added two control items: two one-step problems (a compare and a combine) for third grades, and two two-step problems (one involving two simple compare problems and one comprising two simple combine problems) for fourth and fifth graders. Moreover, four different versions of the tests were made, in which the order of the items was counterbalanced in order to avoid disturbing order effects.

Data Coding

Children's responses were coded as correct (1 point) or incorrect (0 points). We considered as correct all solutions in which the correct arithmetical operations with the appropriate numbers were chosen, without taking into account the computational exactness of the final result, because we wanted to assess children's comprehension of the problems and not their calculation ability.

Children's erroneous answers were analyzed too. To categorize these erroneous answers we relied on categorizations developed by Cummins et al. (1988), Vergnaud (1985) and Verschaffel (1994). Based on these classifications we decided to distinguish two major error types, depending on the level of understanding denoted. First, answers that reflect an incomplete, but not a totally incorrect, understanding of the situation. And, second, a category with all the mistakes that reflect no comprehension at all. Illustrations are provided in Table 2.

RESULTS

Main results on every problem type and its reworded version obtained by children of all grades and ability levels are shown by Tables 3 and 4.

Main results were analyzed by using non-parametric tests (Kruskal-Wallis test for analyzing general results by grade and ability level; Friedman test for analyzing rewording as a main effect; Mann-Whitney U test for comparisons between grades and between ability levels; and Wilcoxon two paired samples for analyzing results by mathematical difficulty, situational difficulty and rewording).

Table 2. Categorization of wrong answers

	MATH-EASY PROBLEMS	MATH-DIFF PROBLEMS
Examples	Peter had 37 meters of cable. He bought 100 meters more cable. He used some more meters of cable and he ended up with 11 meters of cable. How many meters of cable did he use?	Peter had 37 meters of cable. He bought some more meters of cable. He used 126 meters of cable and he ended up with 11 meters of cable. How many meters of cable did he buy?
MISTAKES		
COMPREHENSION	100 + 37; 100 - 11	126 +11; 126 - 37
	(100 + 11) - 37	(126 - 11) + 37
NO COMPREHENSION	Other mistakes, i.e.: 100 + 37 + 11; 100 - 37 - 11; 100 - 11; 100 X 11 No answer	

Note: MATH-EASY - mathematically easy problems; MATH-DIFF - mathematically difficult problems; SIT-EASY - situationally easy problems; SIT-DIFF - situationally difficult problems

Table 3. Results obtained per grade

	MATH-EASY						MATH-DIFF						Tot. (grade)
	SIT-EASY			SIT-DIFF			SIT-EASY			SIT-DIFF			
	St.	Mat	Sit	St.	Mat	Sit	St.	Mat	Sit	St.	Mat	Sit	
Grade 3	0.71	0.66	0.61	0.42	0.61	0.29	0.13	0.21	0.08	0.16	0.16	0.18	0.35
Grade 4	0.84	0.84	0.70	0.64	0.61	0.52	0.23	0.50	0.30	0.23	0.20	0.16	0.48
Grade 5	0.88	0.90	0.86	0.70	0.72	0.74	0.44	0.56	0.50	0.36	0.42	0.40	0.62
Tot. (rew.)	0.81	0.80	0.72	0.59	0.65	0.52	0.27	0.42	0.29	0.25	0.26	0.25	
Tot. (ord.)	0.78			0.58			0.33			0.25			
Tot. (Diff.)	0.68						0.29						

Note: MATH-EASY - mathematically easy problems; MATH-DIFF - mathematically difficult problems; SIT-EASY - situationally easy problems; SIT-DIFF - situationally difficult problems

Table 4. Whole results obtained per ability level

	MATH-EASY						MATH-DIFF						Tot. (ability)
	SIT-EASY			SIT-DIFF			SIT-EASY			SIT-DIFF			
	St.	Mat	Sit	St.	Mat	Sit	St.	Mat.	Sit	St.	Mat.	Sit	
Low	0.50	0.55	0.55	0.55	0.50	0.25	0.15	0.25	0.00	0.10	0.10	0.10	0.30
Average	0.81	0.79	0.67	0.44	0.60	0.58	0.23	0.37	0.23	0.19	0.19	0.23	0.45
High	0.91	0.90	0.83	0.71	0.72	0.59	0.35	0.54	0.45	0.35	0.38	0.32	0.59
Tot. (rew)	0.74	0.75	0.68	0.57	0.61	0.48	0.24	0.39	0.23	0.21	0.22	0.22	
Tot (ord.)	0.72			0.55			0.29			0.22			
Tot (Diff.)	0.64						0.25						

Note: MATH-EASY: mathematically easy problems; MATH-DIFF: mathematically difficult problems; SIT-EASY: situationally easy problems; SIT-DIFF: situationally difficult problems

An overall analysis of the results by grade (3, 4 and 5) and ability level (low, average and high) of the pupils, and by problem difficulty (MATH-DIFF + SIT-DIFF, MATH-EASY + SIT-EASY and MATH-EASY + SIT-DIFF, and MATH-DIFF + SIT-EASY) and rewording (standard, mathematical and situational) was performed. First, we analyzed the influence of grade and ability level on the results. Regarding grade, differences on success rates were significant, $\chi^2 = (2, 132) = 30.26$, $p < .001$. Significant differences were found between third and fourth graders ($z = 2.74$; $p < .007$), fourth and fifth graders ($z = 3.19$; $p < .002$), and third and fifth graders ($z = 5.25$; $p < .001$) Differences of children's success rates per ability level was also significant, $\chi^2 = (2, 132) = 22.56$, $p < .001$. Differences were significant between low and average ($z = 2.26$; $p < .03$), average and high ($z = 3.04$; $p < .003$) and low and high ($z = 4.25$; $p < .001$) ability children.

Secondly, the effects of the three task variables - mathematical difficulty level, situational difficulty level, and rewording - all, proved to be significant: mathematical difficulty ($z = 8.85$; $p < .001$); situational difficulty ($z = 5.51$; $p < .001$); and rewording $\chi^2 = (2, 132) = 13.34$, $p < .002$). However, although the rewording effect was significant, only the difference between standard and mathematically reworded versions was significant ($z = 2.40$; $p < .02$).

A more detailed analysis of the results allowed us to get information about the different predictions that were stated before. Results were in line with our first prediction, that is, mathematically reworded versions increased children success rates only for MATH-DIFF + SIT-EASY problems ($z = 3.20$; $p < .03$) and not for MATH-EASY + SIT-EASY, MATH-DIFF + SIT-DIFF or MATH-EASY + SIT-DIFF problems.

The results did not support our second prediction, because children did not perform better on reworded problems that enhance the creation of the situational model of the problem, as previously reported. In fact, for situationally reworded versions the only significant difference from standard versions was found for MATH-EASY problems, on which standard versions allowed children a better achievement ($z = -2.32$; $p < .03$). However, this was the case for MATH-EASY + SIT-EASY problems ($z = -2.20$; $p < .03$) only.

As regards the influence of mathematical ability level on the results (= prediction 3), on the one hand, mathematically reworded problems increased success rates on MATH-DIFF + SIT-EASY problems, especially for fourth graders ($z = 3.00$; $p < .004$) and high ability children ($z = 2.71$; $p < .008$). Furthermore, although for SIT-DIFF problems no significant differences were found between standard and mathematically reworded versions, mathematically reworded MATH-EASY + SIT-DIFF problems increased third graders' ($z = 2.11$; $p < .04$) and average-ability children's ($z = 1.19$; $p < .004$) success rates. According to our third prediction, these results mean that the germane cognitive load linked to extra mathematical information is high, since only high achievers performed better on the mathematically reworded versions of MATH-DIFF + SIT-EASY problems than on the standard ones, and only average achievers solved the mathematically reworded versions of MATH-EASY + SIT-DIFF problems more successfully than the standard ones. Extra situational information, on the other hand, hindered low achievers' performance on MATH-EASY + SIT-DIFF problems ($z = -2.44$; $p < .02$). So, with respect to our third prediction, we can conclude that extra mathematical information implies a high level of germane processing, resulting in an advantage for average and high achievers (but not for low achievers), whereas extra situational information implied extraneous processing that led children from all ability levels, but especially the low ability pupils, to perform more weakly on situationally reworded versions of the MATH-EASY + SIT-DIFF problems than on the standard versions.

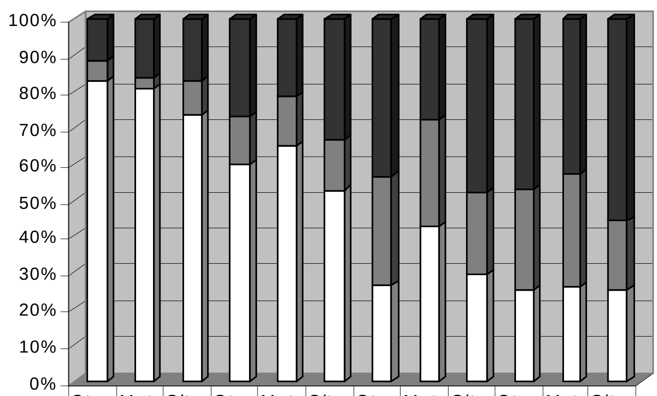
Regarding the qualitative analysis of the answers (= prediction 4, see Figure 2), a non-parametric Chi-square test was applied to analyze the nature of children's wrong answers. The results of the analysis of children's wrong answers were in line with our prediction 4a but not with 4b. In line with prediction 4a, for MATH-DIFF + SIT-EASY problems the standard version showed a significantly higher level of no-comprehension mistakes than mistakes denoting (some) mathematical comprehension, χ^2 (1, N = 52) = 6.23, $p < .02$, while this difference was not significant for the mathematically reworded version, χ^2 (1, N = 46) = 2.17, $p = .14$. The same results were observed for MATH-DIFF + SIT-DIFF problems, χ^2 (1, N = 99) = 6.313, $p = .01$, for standard versions; χ^2 (1, N = 97) = 2.320, $p = .12$ for mathematically reworded version. Taken together, these results support the idea that, although there were no significant differences between the mathematically reworded and standard versions in children's success rate, the extra mathematical information allowed children to understand better the mathematically reworded versions of MATH-

DIFF + SIT-DIFF than the standard versions.

However, contrary to our prediction 4b, for MATH-EASY + SIT-DIFF problems, standard and situationally reworded versions showed a higher level of no-comprehension mistakes than errors denoting some mathematical comprehension, χ^2 (1, N = 52) = 10.24, p < .002; and χ^2 (1, N = 61) = 6.23, p < .02.

Finally, and related to our fifth prediction, we compared the influence of both sources of difficulty of our study, namely, mathematical and situational. In order to do so, mean success rates were analyzed in a 2 (mathematical difficulty: easy vs. difficult) X 2 (situational difficulty: easy vs. difficult) ANOVA with repeated measures. Both main effects proved to be significant: mathematical difficulty, F(1, 132) = 193.74, p <.0001, η^2 = .59; and situational difficulty, F(1, 132) = 39.22, p < .0001, η^2 = .23 respectively. However, and in line with our prediction 5b, the bigger effect size coming from the mathematical difficulty level (.59 vs. .23) suggests that although the situational complications affected the difficulty level of the task, the influence of mathematical difficulty was considerably bigger.

To summarize, the results obtained from this study supported our predictions 1, 4a and 5b. In line with our first prediction, mathematical rewording led to better performance compared with the standard version for MATH-DIFF + SIT-EASY problems but not for MATH-EASY + SIT-DIFF + ones, MATH-EASY + SIT-



	MATH-EASY						MATH-DIFF					
	SIT-EASY			SIT-DIFF			SIT-EASY			SIT-DIFF		
	St.	Mat.	Sit	St.	Mat.	Sit	St.	Mat.	Sit	St.	Mat.	Sit
No comprehension errors	11,3	15,9	16,6	26,5	21,2	32,5	43,1	27,2	46,9	46,9	42,4	55,3
Comprehension errors	5,30	3,03	9,09	12,8	13,6	13,6	30,3	29,5	22,7	28,0	31,0	18,9
Right answers	81,0	80,0	72,0	59,0	65,1	52,0	26,5	42,0	29,0	25,0	26,0	25,0

Figure 2. Qualitative analysis of the errors committed by children when solving each version of each problem type

EASY problems, or for MATH-DIFF + SIT-DIFF. According to our prediction 4a children committed less no-mathematical comprehension errors when solving mathematically reworded MATH-DIFF + SIT-EASY problems than the standard versions. Furthermore, mathematically reworded MATH-DIFF + SIT-DIFF problems showed a higher level of errors indicating some comprehension, which points to the idea that mathematical information had a slight influence on the results on these problems. Finally, in line with prediction 5a, the influence of the mathematical difficulty complication was bigger than the effect of the situational difficulty complication.

Our results did not support prediction 2, 4b, and 5b. Against prediction 2, extra situational information did not seem useful for any kind of problem, even though we did our best to create a well-established temporal structure for the situation described by the problem, which in turn is one of the central dimensions of the Episodic Situation Model defined by Reusser (1988). On the other hand, and against prediction 4b, no differences were found between the kind of errors committed by children on situationally reworded versions and standard ones of MATH-EASY + SIT-DIFF problems, and, against prediction 5b, effect size of the situational difficulty complication was much lower than the mathematical one.

Finally, regarding prediction 3, results by ability level showed that the processing of mathematical information implied a high cognitive cost, because though all children showed a better performance on mathematically reworded MATH-DIFF + SIT-EASY problems than on standard versions, only for high ability children was this difference significant. In addition, low ability children were most harmed by situational information and showed the worst achievement on situationally reworded versions both on MATH-DIFF + SIT-EASY and MATH-EASY + SIT-DIFF problems.

DISCUSSION

To compare the relative effectiveness of the different kinds of rewording on elementary school children's understanding and solution of word problems, Vicente et al. (in press) carried out a set of studies in which they found that, while mathematical rewording leads to better achievement, situational rewording does not. These differential results could not be explained in terms of the fact the resulting problem text was much longer in the latter than in the former case. The goal of the present study was to test two alternative explanations for these results. A first possible alternative explanation was that the problems contained no serious situational complexities, so no extra situational information was required to solve them properly. According to the second explanation, providing extra situational information was useless because word problem solving is after all a mathematical task, and because of this, only extra mathematical information will help problem solvers towards the successful completion of the task whereas providing extra situational details will not bring the problem solver more quickly or easily to his goal and may even hinder him in reaching that goal. To investigate these two hypothetical explanations, we set up a study in which not only the mathematical problem difficulty was manipulated, but also the situational difficulty. Mathematical difficulty was manipulated by including two consistent and two inconsistent two-step change problems. As in the study of Staub and Reusser (1992), situational difficulty was manipulated by working with regular problems, in which the order wherein the sets were introduced followed nicely the order of actual events in the situational model, as well as problems in which the problem text did not follow the regular order of events. Afterwards we compared the effectiveness on performance and on level of understanding (as evidenced by the kinds of error of the provision of additional mathematical and situational information in these mathematically and situationally easy and/or difficult problems.

The results are, again, very clear: generally speaking adding mathematical information led to better performance whereas extra situational information did not. While this general result already points towards the second above-mentioned explanation, two additional results point further in that direction. First, the proportion of variance explained by the mathematical difficulty of the task. Second, extra mathematical information improved children's achievement on MATH-DIFF + SIT-EASY problems (and also on MATH-EASY + SIT-DIFF problems for third-graders and average-ability children), whereas situational rewording was not helpful in any of these cases. This brings us to the question: why was extra mathematical information only helpful for the two types of problems with only one difficulty (either MATH-DIFF *or* SIT-DIFF problems) and why it did not improve children's achievement on MATH-EASY + SIT-EASY problems or on MATH-DIFF + SIT-DIFF problems? Based on the Cognitive Load Theory (Paas, Renk, Sweller, 2003; Sweller, 1999; Sweller, Chandler, 1991) germane processing of cognitive tasks improves performance when the sum of germane load and intrinsic load does not overload available working memory resources. In our study, MATH-EASY + SIT-EASY problems may have been too simple for children to demand additional information (so there was no need for extra information), while the intrinsic load of MATH-DIFF + SIT-DIFF problems may have been too high, leaving no room for germane processing induced by extra information. The other two types of problems, on the other hand, which both were of an intermediate level of complexity (either mathematically *or* situationally difficult), profited from adding functional information.

In addition, the results obtained per ability level also fit well with Cognitive Load Theory: as only high-ability children showed a significant improvement in their performance, we can assume that germane load linked to extra mathematical information was high. Arguably, high achievers had more cognitive resources available for germane processing and, because of this, mathematical information was especially useful to them on MATH-DIFF + SIT-EASY problems. Furthermore, when intrinsic load decreased, namely, for MATH-EASY + SIT-DIFF problems, average-level achievers showed such an improvement on mathematically reworded versions. Furthermore, extraneous processing linked to situational information harmed to low achievers especially, because precisely they had fewer cognitive resources to face the task and because some of them had to be devoted to processing situational information.

In sum, our results again force us to conclude that we overestimated the power of situational rewordings, at least for this population of students and for these kinds of task. From a theoretical perspective, we can conclude that our results - as did those from previous studies (Vicente et al., in press) - failed to demonstrate the importance and impact of the episodic situational model in the solution process of word problems. In addition our results can be nicely interpreted in terms of Cognitive Load Theory: while the processing of (task relevant) extra mathematical information produced germane load, at least for tasks of intermediate complexity, and, consequently, led to better problem-solving performance, extra situational information, which seemed to be task-irrelevant, led to processes that produced extraneous load on the solvers' working memory and, therefore, to a decrease in success rate.

Nevertheless, our results as well as their theoretical implications must be taken very carefully. At a theoretical level, and as we have pointed out, rewording word problems and analyzing its effects on children's performance is only an indirect way of accounting for the cognitive processes implied in the task. Therefore, further research is needed in which the results of these studies based on indirect measures are complemented by and confronted with those from investigations that try to assess more directly to what extent children try to build up a (rich) situational model when solving word problems, what difficulties they encounter when doing so, how this situational model is related to their further problem-solving steps, and how the existence and richness of this model is influenced by various task, subject and context variables.

Received August 13, 2007

REFERENCES

BRIARS, D.J., LARKIN, J.H., 1984, An integrated model of skill in solving elementary word problems. *Cognition and Instruction*, 1, 245-296.

COQUIN-VIENNOT, D., MOREAU, S., 2007, Arithmetic problems at School: When there is an apparent contradiction between the situation model and the problem model. *British Journal of Educational Psychology*, 77, 69-80.

CUMMINS, D.D., KINTSCH, W., REUSSER, K., WEIMER, R., 1988, The role of understanding in solving word problems. *Cognitive Psychology*, 20, 405-438.

CUMMINS, D.D., 1991, Children's interpretations of arithmetic word-problems. *Cognition and Instruction*, 8, 261-289.

DAVIS-DORSEY, J., ROSS, S.M., MORRISON, G.R., 1991, The role of rewording and context personalization in the solving of mathematical word-problems. *Journal of Educational Psychology*, 83, 61-68.

DE CORTE, E., VERSCHAFFEL, L., DE WIN L., 1985, Influence of rewording verbal problems on children's problem representations and solutions. *Journal of Educational Psychology,* 77, 460-470.

FUSON, K.C., 1992, Research on learning and teaching addition and subtraction of whole numbers. In: G. Leinhardt, R. Putnam, R.A. Hattrup (Eds.), *Analysis of arithmetic for mathematics teaching* (pp. 53-187). Hillsdale, NJ: Erlbaum.

GRAESSER, A.C., GOODMAN, S.M., 1985, Implicit knowledge, question answering, and the representation of expository text. In: B.K. Britton, J.B. Black (Eds.), *Understanding expository texts* (pp. 109-171). Hillsdale, NJ: Erlbaum.

HUDSON, T., 1983, Correspondences and numerical differences between disjoint sets. *Child Development*, 54, 84-90.

KINTSCH, W., 1988, The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95, 163-182.

KINTSCH, W., 1998, *Comprehension: A paradigm for cognition.* Cambridge: Cambridge University Press.

KINTSCH, W., GREENO, J., 1985, Understanding and solving word arithmetic problems. *Psychological Review*, 92, 109-129.

MOREAU, S., COQUIN-VIENNOT, D., 2003, Comprehension of arithmetic word problems by fifth-grade pupils: Representations and selection of information. *British Journal of Educational Psychology*, 73, 109-121.

PAAS, F., RENKL, A., SWELLER, J., 2003, Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38, 1-4.

REED, S.K., 1999, *Word Problems: Research and curriculum reform.* Mahwah, NJ: Lawrence Erlbaum Associates.

REUSSER, K., 1988, Problem solving beyond the logic of things: Contextual effects on understanding and solving word problems. *Instructional Science*, 17, 309-338.

RILEY, N.S., GREENO, J., HELLER, J.I., 1983, Development of children's problem solving ability in arithmetic. In: H.P. Ginsburg (Ed.), *The development of mathematical thinking* (pp. 153-196). New York: Academic Press.

STAUB, F., REUSSER, K., 1992 April, The role of presentational factors in understanding and solving mathematical word problems. Paper presented at the meeting of the American Educational Research Association, San Francisco.

STERN, E., LEHRNDORFER, A., 1992, The role of situational context in solving word-problems. *Cognitive Development*, 7, 259-268.

SWELLER, J., 1999, *Instructional design in technical areas.* Camberwell, Victoria, Australia: Australian Council for Educational Research.

SWELLER, J., CHANDLER, P., 1991, Evidence of cognitive load theory. *Cognition & Instruction*, 8, 351-362.

VERGNAUD, G., 1985, *L'enfant, la mathématique et la réalité.* Bern: Peter Lang.

VERSCHAFFEL, L., 1994, Using retelling data to study elementary school children's representations and solutions of compare problems. *Journal for Research in Mathematics Education,* 25, 141-165.

VERSCHAFFEL, L., GREER, B., DE CORTE, E., in press, Whole number concepts and operations. In: F.K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 557-628). Greenwich, CT: Information Age Publishing.

VERSCHAFFEL, L., DE CORTE, E., 1993, A decade of research on word-problem solving in Leuven: Theoretical, methodological and practical outcomes. *Educational Psychology Review*, 5, 239-256.

VERSCHAFFEL, L., DE CORTE, E., 1997, Word problems. A vehicle for promoting authentic mathematical understanding and problem solving in the primary school. In: T. Nunes, P. Bryant (Eds.), *Learning and teaching mathematics: An international perspective* (pp. 69-97). Hove, East Sussex, UK: Psychology Press.

VICENTE, S., ORRANTIA, J., VERSCHAFFEL, L., in press, Influence of situational and mathematical rewording on word problem solving. *British Journal of Educational Psychology.*

YUSTE, C., 1985, *BADYG (gráficos C, D y E).* Madrid: Ciencias de la Educación Preescolar y Especial.

APPENDIX 1

Examples of the experimental problems used in the study

EASY

Ordo-naturalis

Standard

Peter had 37 meters of cable. He bought 100 meters of cable. He used some meters of cable and he ended up with 11 meters of cable. How many meters of cable did he use?

Mathematical

Peter had 37 meters of cable. He bought 100 meters more and *joined them to those that he had. From the resulting total of meters of cable* he used some meters and he ended up with 11 meters of cable. How many meters of cable did he use?

Situational

Two days ago Peter had 37 meters of cable. *Yesterday* Peter bought 100 more meters of cable than those he already had. *After buying those meters of cable he began a renovation. While making the renovation* he has used some meters of cable, and *when he finishes* there are 11 meters of cable left. How many meters of cable has Peter used?

Non ordo-naturalis

Standard

Peter bought 100 meters of cable. He used some meters of cable and he ended up with 11 meters of cable. At the beginning Peter had 37 meters of cable. How many meters of cable did he use?

Mathematical

Peter bought 100 meters more and *joined them to those that he had. From the resulting total of meters of cable* he used some meters and he ended up with 11 meters of cable. At the beginning Peter had 37 meters of cable. How many meters of cable did he use?

Situational

Yesterday Peter bought 100 more meters of cable than those he already had. *After buying those meters of cable he began a renovation. While making the renovation* he has used some meters of cable, and *when he finishes* there are 11 meters of cable left. *Two days ago* Peter had 37 meters of cable. How many meters of cable has Peter used?

DIFFICULT

Ordo-naturalis

Standard

Peter had 37 meters of cable. He bought some more meters of cable. He used 126 meters of cable and he ended up with 11 meters of cable. How many meters of cable did he buy?

Mathematical

Peter had 37 meters of cable. He bought some more meters of cable and *joined them to those that he had. From the resulting total of meters of cable* he used 126 meters of cable and he ended up with 11 meters of cable. How many meters of cable did he buy?

Situational

Two days ago Peter had 37 meters of cable. *Yesterday* Peter bought some more meters of cable than those he already had. *After buying those meters of cable he began a renovation. While making the renovation* he has used 126 meters of cable, and *when he finishes* there are 11 meters of cable left. How many meters of cable did he buy?

Non ordo-naturalis

Standard

Peter bought some more meters of cable. He used 126 meters of cable and he ended up with 11 meters of cable. At the beginning Peter had 37 meters of cable. How many meters of cable did he buy?

Mathematical

Peter bought some more meters of cable and *joined them to those that he had. From the resulting total of meters of cable* he used 126 meters of cable and he ended up with 11 meters of cable. At the beginning Peter had 37 meters of cable. How many meters of cable did he buy?

Situational

Yesterday Peter bought some more meters of cable than those he already had. *After buying those meters of cable he began a renovation. While making the renovation* he has used 126 meters of cable, and *when he finishes* there are 11 meters of cable left. *Two days ago* Peter had 37 meters of cable. How many meters of cable did he buy?

VPLYV SITUAČNÝCH A MATEMATICKÝCH INFORMÁCIÍ NA SITUAČNE ZLOŽITÉ SLOVNÉ ÚLOHY

S. Vicente, J. Orrantia, L. Verschaffel

Súhrn: Riešenie aritmetických slovných úloh má dvojaký charakter: textový a matematický. Znamená to, že na ich riešení sa podieľa viacero kognitívnych procesov, ktoré umožňujú pochopenie slovnej úlohy ako textu a ako matematickej štruktúry - to všetko ovplyvňuje výkon detí v týchto úlohách. Teoretické modely a empirické štúdie objasňujú implikácie rozdielnych procesov na riešenie slovných úloh. Uskutočnili nový výskum, aby vyhodnotili vplyv matematického a situačného preštylizovania na riešenie lineárnych rovníc o jednej neznámej na dvoch úrovniach matematickej a situačnej obtiažnosti úlohy. Výsledky ukázali, že kým matematické preštylizovanie zvýšilo výkon detí pri riešení matematicky aj situačne zložitých úloh, situačné preštylizovanie nemalo žiaden vplyv na riešenie úloh.