Why does placing the question before an arithmetic word problem improve performance? A situation model account

Thevenot, Catherine; Devidal, Michel; Barrouillet, Pierre Noël; Fayol, Michel

How to cite

THEVENOT, Catherine et al. Why does placing the question before an arithmetic word problem improve performance? A situation model account. In: The quarterly journal of experimental psychology, 2007, vol. 60, n° 1, p. 43–56. doi: 10.1080/17470210600587927

This publication URL: https://archive-ouverte.unige.ch/unige:88300

Publication DOI: 10.1080/17470210600587927

The Quarterly Journal of Experimental Psychology

ISSN: 1747-0218 (Print) 1747-0226 (Online) Journal homepage: http://www.tandfonline.com/loi/pqje20

Why does placing the question before an arithmetic word problem improve performance? A situation model account

Catherine Thevenot , Michel Devidal , Pierre Barrouillet & Michel Fayol

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2007, 60 (1), 43–56

Psychology Press
Taylor & Francis Group

Why does placing the question before an arithmetic word problem improve performance? A situation model account

Catherine Thevenot
University of Sussex, Brighton, UK

Michel Devidal
Université Blaise Pascal, Clermont-Ferrand, France

Pierre Barrouillet
Université de Bourgogne, Dijon, France

Michel Fayol
Université Blaise Pascal, Clermont-Ferrand II, France

The aim of this paper is to investigate the controversial issue of the nature of the representation constructed by individuals to solve arithmetic word problems. More precisely, we consider the relevance of two different theories: the situation or mental model theory (Johnson-Laird, 1983; Reusser, 1989) and the schema theory (Kintsch & Greeno, 1985; Riley, Greeno, & Heller, 1983). Fourth-graders who differed in their mathematical skills were presented with problems that varied in difficulty and with the question either before or after the text. We obtained the classic effect of the position of the question, with better performance when the question was presented prior to the text. In addition, this effect was more marked in the case of children who had poorer mathematical skills and in the case of more difficult problems. We argue that this pattern of results is compatible only with the situation or mental model theory, and not with the schema theory.

In arithmetic word problems the instructions tell a story or, in other words, describe verbally a situation (Verschaffel, Greer, & De Corte, 2000). In order to solve this kind of problem, adults and children have to mentally represent the relations between the different elements described in the text (Carpenter & Moser, 1982; Coquin-Viennot, 2000; Coquin-Viennot & Moreau, 2003; De Corte & Verschaffel, 1981; De Corte, Verschaffel, & De Win, 1985; Hudson, 1983; Kintsch & Greeno, 1985; Riley, Greeno, & Heller, 1983). The nature of this representation remains a controversial issue. The aim of this paper is to clarify the level and the nature of the representation constructed by children to solve arithmetic word problems.

For Kintsch and Greeno (1985) a single high-level representation is constructed from the text

Correspondence should be addressed to Catherine Thevenot, FPSE, Université de Geneve, 40, bd du Pont d'Arve, CH-1205, Geneve, Switzerland. E-mail: Catherine.Thevenot@pse.unige.ch
We would like to thank Jane Oakhill for her useful comments and her corrections on the last version of this paper.

http://www.psypress.com/qjep
DOI:10.1080/17470210600587927

base. This "problem model" contains only the mathematical information relevant and necessary to solve the problem. Propositions are taken from the text base and are assigned to appropriate slots of a set schema in the problem model. A set schema is an abstract structure that is stored in long-term memory and that is designed to represent the different states of the problem. It contains four attributes that correspond to the object, the quantity, the specification, and the role slots. For example, let us consider the combine problem "Joe has 3 marbles. Tom has 5 marbles. How many marbles do they have altogether?" (see Riley et al., 1983, for a semantic classification of word problems). The sentence "Joe has 3 marbles" contains three propositions—that is, Joe; have; 3 (marbles). The proposition "3 (marbles)" cues representation of a set. Then, the propositions are sorted into the object (marbles), quantity (3), and specification (Joe) slots in the set schema. A second set is represented when the proposition "Tom has 5 marbles" is encountered. A third set schema is added to the problem model from the question "How many marbles do they have altogether?". The have-altogether proposition contained in this question cues assignment of the role of this third set as a superset. When a superset is created, the model posts a request to assign subset roles to other sets in its representation. Note that the slots corresponding to the roles of the first two set schemata were still empty until this point. The different set schemata are then brought together by subordinate problem schemata, which are triggered by specific linguistic expressions in the text of the problem, such as "altogether", "less . . . than" or "more . . . than". In our example, the proposition "altogether" triggers the so-called part–whole schema as well as a superset strategy, which will allow the resolution of the problem (see Riley et al., 1983, for a description of the different problem schemata, and Kintsch & Greeno, 1985, for a more detailed description of their model). The difficulty for problem solving is thus explained by mismatches between the linguistic expression contained in the text and the schema chosen for the resolution (Cummins, 1991; Cummins, Kintsch, Reusser, & Weimer, 1988; Kintsch, 1987; Lewis & Mayer, 1987). The improvement in performance with age is then explained by practice, which allows the constitution and the consolidation of the problem schema in long-term memory and, consequently, leads to a decrease in the mismatches.

However, according to Reusser (1989) and Staub and Reusser (1995), as well as Nathan, Kintsch, and Young (1992), Kintsch and Greeno's (1985) model relies excessively on the schema theory. Therefore, these authors suggest the existence of a "nonmathematical" representation: the situation model (as first described by Van Dijk & Kintsch, 1983), which corresponds to a level of representation that specifies the agents, the actions, and the relationships between the events in everyday contexts. In this conceptualization, the problem model, as described by Kintsch and Greeno (1985), is constructed not only by the mobilization of specific knowledge stored in long-term memory—that is, the problem schema—but also on the basis of information contained in the situation model. Within that framework, the difficulty of the problem is determined by the difficulty in understanding the situation described. The same type of representation—namely, a mental model—is described by Johnson-Laird (1983) in the domain of text comprehension and reasoning. For this author as well, solving a problem requires the construction of a mental representation of the situation described by the problem and not a mental representation of the problem itself (see Zwaan & Radvansky, 1998).

To sum up, a schema and a situation model differ in their structure and content as well as in the way that they are maintained in memory. A schema is a problem frame stored in long-term memory that is activated by specific textual clues. The schema contains invariant characteristics related to a category of problem and empty slots, which are filled by the pieces of information that are specific to the problem to be solved (i.e., numbers and objects). Moreover, it contains the procedures that are necessary to solve the problem, and these procedures are triggered automatically as soon as the empty slots have been filled (Kintsch, 1988; Kintsch & Greeno, 1985; Rumelhart, 1980; Schank, 1975; Schank & Abelson, 1977). On the contrary, a situation or mental model is a temporary structure stored in working memory that contains, in addition to the mathematical information necessary to solve the problem, nonmathematical information that is related to the context in which the situation described by the problem takes place. A situation model is therefore more qualitative and less formal than a schema (Nathan et al., 1992; Reusser, 1989; Staub & Reusser, 1995; Stern & Lehrndorfer, 1992). Indeed, functional, temporal, and structuring elements described in the text of the problem can be integrated in the situation model, and those elements could influence individuals' performance and strategies. Note that such influences are difficult to interpret within the schema theory because contextual information is not represented in a "problem model".

The relevance of these two frameworks, namely the schema and the situation or mental model frameworks, is still put to the test nowadays (Moreau & Coquin-Viennot, 2003; Stetic, 1999). A specific finding that needs to be explained by any theory is that there is a substantial advantage to be presented with the question before rather than after the text of the problem. This finding was first accounted for by the schema theory (Devidal, Fayol, & Barrouillet, 1997; Fayol, Abdi, & Gombert, 1987); however we have recently proposed an alternative explanation (Thevenot, Barrouillet, & Fayol, 2004). In a standard word problem, the question often contains all the information needed to describe the relations between the different elements of the text as well as specific linguistic expressions. Therefore, according to Devidal et al. (1997), the question placed at the beginning of the text, via the specific information it contains, directly triggers the appropriate schema (e.g., the expression "less than" will trigger a comparison schema). This early activation allows the numerical data to be integrated into the schema as soon as they are presented, and the calculations can be achieved during reading. The subsequent release of working-memory resources would explain the improvement in performance. However, by using an original paradigm that allows us to determine precisely when the calculations are achieved (see Thevenot & Oakhill, 2005b, for a description of the paradigm), we confirmed the facilitatory effect of the question at the beginning but showed that this effect is observable whether or not the calculations are achieved during reading (Thevenot et al., 2004). Consequently, the increase in the number of successful resolutions when the question is placed before the text rather than after cannot be attributable only to the online integration of the numerical data into the activated schema. We proposed that the aid provided by the position of the question before the text would only be due to its facilitatory effect on the construction of the representation required to solve the problem. Just like a title before a narrative text (Hunt & McDaniel, 1993; Mayer, 1983; Rawson & Kintsch, 2002), the question before an arithmetic problem would facilitate the subsequent encoding and integration of the information in the representation. This interpretation is alternative to the one formulated within the schema framework since it is also compatible with an approach that considers the representation as a specific and temporary mental structure constructed in working memory, such as a situation model (Reusser, 1989) or a mental model (Johnson-Laird, 1983).

The aim of this paper is to validate one of those interpretations and to invalidate the other by presenting children who differ in their mathematical skills with problems of varying difficulty. If the facilitatory effect of the position of the question before the problem is higher for high-skilled participants and for easy problems, then the effect can be attributed to the early activation of the problem schema. Conversely, if the facilitatory effect is higher for lower skilled children and for more difficult problems, then the effect can be attributed to the better integration of a temporary and specific mental representation. Indeed, we know that schemata are stored in long-term memory as a result of repetitive contact with a specific situation (Kintsch & Greeno, 1985; Rumelhart, 1980; Schank, 1975; Schank & Abelson, 1977). Therefore, the higher the frequency of correct resolutions of the problem, the higher the probability of the construction of its associated schema. By definition the frequency of correct resolutions for a specific type of problem is lower for low- than for high-skilled children. The latter are thus more likely to have stored the appropriate schema than the former. Consequently, if the facilitatory effect of the position of the question before the text is higher for high-skilled participants, then we can conclude that the better performance is due to the early activation of the appropriate schema, in which the subsequent information is integrated. Following the same reasoning, difficult problems, which by definition are often failed, are less likely to be associated with the appropriate schema than are easier problems, which are often correctly solved. As a consequence, if the facilitatory effect of the position of the question before the text is more important for easier than for more difficult problems, the schema-based interpretation is to be maintained.

If the opposite pattern of results is obtained, then this facilitatory effect cannot be attributable to the activation of a schema-based representation stored in long-term memory, and, consequently, a specific and temporary representation-based interpretation will be supported. Moreover, the opposite pattern of results would specifically be explained by the mental model or situation model theory. We know that, in children, one of the best predictors of arithmetic word problem performance is text comprehension abilities (Swanson, Cooney, & Brock, 1993). It turns out that poor comprehenders are precisely those children who fail to construct an adequate mental model of the situation described in a text (Oakhill, 1996; Oakhill, Cain, & Yuill, 1998; Singer & Ritchot, 1996). More precisely, poor comprehenders have difficulty in integrating new information within the current mental model: They build partial, but not complete, models of text (Cain & Oakhill, 1999; Yuill, Oakhill, & Parkin, 1989). Therefore, lower skilled children in the domain of word problem solving would naturally be those children who benefit more from an aid for the construction of the mental model. The question at the beginning of the text constitutes an ideal aid for the integration of the information in the mental model since, as a title prior to a text, it constitutes a guide for the organization of the information within the mental representation. A study by Wilberschied and Berman (2004) in the domain of foreign language comprehension shows that such advance organizers are particularly helpful for younger and less proficient students. Moreover, the more difficult the problem, the higher the aid provided by the question prior to the text should be. Indeed, we know that the difficulty of a problem is not determined by the nature of the operation to be performed but, rather, by the complexity of the situation described in the text (De Corte & Verschaffel, 1991; Riley et al., 1983). The more complex the situation, the more difficult the construction of the mental representation of the relations described in the text (e.g., De Corte & Verschaffel, 1985). Therefore, the aid provided by the question prior to the text for the organization of the representation should be higher for difficult problems.

As in the previous experiments that showed a facilitatory effect of the question prior to the text, the problems were presented segment by segment on a computer screen (Devidal et al., 1997; Thevenot et al., 2004; Thevenot & Oakhill, 2004, 2005a, 2005b, in press). Indeed, the question has to be presented separately from the text of the problem to ensure that it is not read before the text even when it is at the end of the problem. Moreover, as in the previous experiments, the text itself was segmented, and the self-presentation times of the different segments were measured. If a schema-type representation is constructed by children to solve the problem, then the segments that contains information that is not necessary for a mathematical representation (i.e., contextual segments) will be self-presented for a shorter time when the question is before rather than after the text. Indeed, if, as assumed by Devidal et al. (1997), the schema is activated as soon as the question is read, the solver would go directly to the segment that contains the information that can be assigned to the free slots of the schema. As seen previously, contextual information is not integrated in the set and subordinate schemata (Kintsch & Greeno, 1985). This effect should be even more pronounced for those children whose schemata are well consolidated in long-term memory—namely, high-skilled children. On the contrary, the mean self-presentation times of the segments that contain the names of the protagonists and the quantities of objects should be exactly the same whatever the position of the question because those pieces of information are integrated in both schemata and situation models.

In the present experiment, two groups of fourth-graders who differed in their skills in mathematics were presented with three different types of problem described by Riley et al. (1983): Combine 1 (Joe has 5 marbles. Tom has 3 marbles. How many marbles do Joe and Tom have altogether?); Compare 1 (Joe has 5 marbles. Tom has 3 marbles. How many marbles does Joe have more than Tom?); and Compare 2 problems (Joe has 5 marbles. Tom has 3 marbles. How many marbles does Tom have less than Joe). These specific problems were chosen because we needed problems that only differ in their question and not in the text associated with the question. Indeed, it was imperative that children could not infer the question from the text of the problem. If that were the case, the manipulation of the position of the question would not have made any sense as children could have guessed it when it was at the end of the problem. Out of the 14 problems studied by Riley et al., only the 3 problems presented above fit this constraint. Moreover, precise data on their relative difficulty are available in the literature: We know that compare problems are more difficult than combine problems (Briars & Larkin, 1984; Morales, Shute, & Pellegrino, 1985; Riley & Greeno, 1988) and that Compare 1 problems are easier than Compare 2 problems, in which the expression "less than" is inconsistent with the operation required to solve them (Briars & Larkin, 1984; Hegarty, Mayer, & Monk, 1995; Lewis & Mayer, 1987; Pape, 2003; Verschaffel, 1994; Verschaffel, De Corte, & Pauwels, 1992). The choice of the population on which this experiment was conducted was motivated by the fact that fourth-graders are supposed to be completely familiar with the type of problems that they were presented with, which should allow the abstraction of the appropriate schema.

Method

Participants

Two groups of fourth-graders were selected on the basis of their score on a mathematical test containing 15 different problems involving numeration, arithmetic, arithmetic word problems, and problems about measurement.

A total of 244 French-speaking children were tested individually. In order to divide the population into two heterogeneous groups, only 72 children were selected to take part in the experiment. The 36 children who comprised the group of high-skilled participants obtained a mean score of 12.2 out of 15 on the mathematical test (from 11 to 13) and were 10;6 years old in average. The 36 children who comprised the lower skilled group scored 8 on the mathematical test on average (from 5 to 10) and their mean age was 10;6 years.

Material

All the problems were constructed on the same model. They all contained a segment describing the context (e.g., from the garage), a segment presenting the first protagonist (i.e., always Tom) and a verb (e.g., Tom took), a segment corresponding to the first quantity of objects (e.g., 6 bottles), a segment presenting the second protagonist (i.e., always Tim) and the same verb as previously (e.g., Tim took), a segment corresponding to the second quantity of objects (e.g., 3 bottles), and finally a segment corresponding to the question (e.g., How many bottles did Tom and Tim take from the garage altogether?).

The question could take three different forms as a function of the type of problem: "How many bottles did Tom and Tim take from the garage altogether?", for example, for Combine 1 problems; "How many bottles did Tom take more than Tim?", for example, for Compare 1 problems; and "How many marbles did Tim take fewer than Tom?", for example, for Compare 2 problems.[1] Each problem was presented twice, once with the question at the beginning and once with the question at the end of the text. Two problems were presented in each of the six experimental conditions: 3 (type of problem) × 2 (question position). Children were thus presented with 12 problems in a random order. In order to make the task less boring for children, all the problems differed in the "contextual phrase", the nature of the objects, and the verbs. Finally, the numbers used in the problems were all one-digit numbers, and to ensure the coherence of the problem, the first number was always superior to the second one.

Procedure

As in the previous experiments that manipulated the position of the question (Devidal et al., 1997; Fayol et al., 1987), we used a segmented self-presentation procedure (Pynte, 1974) in which information is displayed segment by segment on a computer screen. For example, the successive segments took the following form: "In a box // Tom has // 8 marbles // Tim has // 2 marbles // How many marbles does Tom have more than Tim? //" for a Compare 1 problem with the question after the text. This specific segmentation was chosen to replicate Devidal et al. (1997). When the participants had read a segment, they pressed the space bar to replace it with the next segment. The different segments' self-presentation times were recorded by the computer. At the end of the problem, the child had to give his answer out loud, and this answer was written down by the experimenter.

Results

Proportions of successfully solved problems

A 2 (children's mathematical skill: high or low) × 2 (position of the question: before or after) × 3 (type of problem: Combine 1, Compare 1, Or Compare 2) analysis of variance (ANOVA) with the two last factors as repeated measures and the first as a between-subjects measure was performed on the proportions of successfully solved problems (Table 1).

Evidence for the reliability and appropriateness of the test that we used to assess the children's abilities was provided by the proportion of successfully solved problems, which was higher for high-skilled children (.81) than for lower skilled children (.66), $F(1, 70) = 10.62$, $MSE = 0.23$, $p < .002$.

The classical effect of the position of the question was significant: The performance was better when the question was placed before (.83) rather than after the text (.64), $F(1, 70) = 47.62$, $MSE = 0.08$, $p < .001$. Contrary to our expectations, the main effect of the type of problem was not significant, $F(2, 140) = 1.97$, $MSE = 0.13$. However, by definition, this main effect does not take the position of the question into account. As all the previous studies determined the relative difficulty of the problems with the question after the text, we conducted a subsidiary analysis, taking into account only the resolution performance when the question was after the text. This analysis revealed an effect of the type of problem on the proportion of successfully solved problems, $F(2, 140) = 3.54$, $MSE = 0.11$, $p < .03$. As expected, Combine 1 problems were easier (.72) than Compare 1 (.61) and Compare 2 problems (0.58), $F(1, 70) = 3.13$, $MSE = 0.12$, $p < .08$, and $F(1, 70) = 6.49$, $MSE = 0.11$, $p < .01$, respectively. However, contrary to the results in the literature, Compare 1 problems did not lead to higher proportions of correct answers than did Compare 2 problems, $F < 1$.

The interaction between the children's mathematical skill and the position of the question was significant, $F(1, 70) = 4.32$, $MSE = 0.08$, $p < .04$. Lower skilled children benefited more from the position of the question before the text

[1] In French, *"Combien Tom et Tim ont-ils sorti de voitures ensemble?"*, *"Combien Tom a-t-il sorti de bouteilles de plus que Tim?"*, and *"Combien Tim a-t-il sorti de voitures de moins que Tom?"* for Combine 1, Compare 1, and Compare 2 problems, respectively.

Table 1. *Proportions of successfully solved problems as a function of the children's mathematical skill, the position of the question, and the type of problem*

	High-skilled children			*Lower skilled children*		
Question position	*Combine1*	*Compare1*	*Compare2*	*Combine1*	*Compare1*	*Compare2*
Before	.92 (.19)	.85 (.33)	.87 (.25)	.78 (.35)	.74 (.35)	.85 (.33)
After	.80 (.30)	.72 (.39)	.71 (.36)	.64 (.39)	.51 (.40)	.46 (.37)

Note: Standard deviations are in parentheses.

(+.25 compared to the condition in which the question is after the text) than did high-skilled children (only + .14). Moreover, there was an interaction between the type of problem and the position of the question, $F(2, 140) = 3.81$, $MSE = 0.06$, $p < .02$. The benefit gained from the position of the question before the text was higher for Compare 2 (+.28) than for Combine 1 problems (+.13), $F(1, 70) = 6.86$, $MSE = 0.06$, $p < .01$. However, the difference in terms of benefit between Compare 1 (+.18) and Combine 1 problems was not significant, $F < 1$. Finally, the benefit was higher for Compare 2 than for Compare 1 problems, $F(1, 70) = 4.31$, $MSE = 0.04$, $p < .04$.

The interaction between the children's mathematical skill, the type of problem, and the position of the question was not significant, $F(2, 140) = 1.52$, $MSE = 0.06$.

Mean resolution times

We considered the self-presentation times of the last segment of the problem as the resolution time, since it was when this segment was presented on the screen that the child had to give an answer to the problem. When the question was presented before the problem, the last segment corresponded to s5 (second amount of objects) whereas the question (s6) corresponded to the last segment when presented at the end of the problem. As the question was longer than the fifth segment, only significant interactions between the different variables are of interest in this analysis. Only problems that were correctly solved were taken into account. A 2 (children's mathematical skill: high or low) × 2 (position of the question: before or after) × 3 (type of problem: Combine 1, Compare 1, or Compare 2) ANOVA with the two last factors as repeated measures and the first as a between-subjects measure was performed on those mean resolution times (Table 2).

Resolution times were a less sensitive measure of the children's mathematical skills than the proportion of correctly solved problems as the interaction between this variable and the position of the question was not significant, $F(1, 70) = 3.77$, $MSE = 89{,}914$, $p = .06$.

However, as in the previous analysis, the interaction between the position of the question and the type of problem was significant, $F(2, 140) = 7.11$, $MSE = 98{,}859$, $p < .001$. When the question was before the text, the resolution times were always shorter than when the question was after the text, and, more interestingly, this difference was higher for Compare 2 (−362 cs) and Compare 1 problems (−209 cs) than for Combine 1 problems (−83 cs), $F(1, 70) = 19.07$, $MSE = 73{,}443$, $p < .001$; $F(1, 70) = 3.73$, $MSE = 75{,}947$, $p < .05$, respectively. However, the difference was not significantly higher for Compare 2 than for Compare 1 problems, $F(1, 70) = 2.99$, $MSE = 147{,}187$, $p = .09$.

Self-presentation times

A 2 (children's mathematical skill: high or low) × 2 (position of the question: before or after) × 3 (type of problem: Combine 1, Compare 1, or Compare 2) × 6 (type of segment: context, s1; first protagonist's action, s2; first quantity of objects, s3; second protagonist's action, s4; second quantity of objects, s5; question, s6) ANOVA with the three last factors as repeated

Table 2. *Auto-presentation times*[a] *of the different segments of the problem as a function of the type of problem, the position of the question, and the children's mathematical abilities*

		High-skilled children		*Lower skilled children*	
Type of problem	*Segment*	*Q1*	*Q2*	*Q1*	*Q2*
Combine 1	Q	484	—	490	—
	S1	159	192	187	212
	S2	155	158	146	176
	S3	148	173	168	188
	S4	135	138	129	153
	S5	433	164	516	200
	Q	—	568	—	547
Compare 1	Q	455	—	428	—
	S1	159	179	177	210
	S2	157	157	148	168
	S3	158	163	179	183
	S4	137	144	130	146
	S5	400	162	638	182
	Q	—	715	—	742
Compare 2	Q	414	—	441	—
	S1	152	177	174	201
	S2	143	158	151	169
	S3	140	164	178	196
	S4	127	137	126	149
	S5	328	152	420	194
	Q	—	702	—	771

Note: Question position: Q1 before, Q2 after the text.
[a]In cs.

measures and the first as a between-subjects measure was performed on the mean self-presentation times (Table 2).

There was a main effect of the type of problem, $F(2, 140) = 3.16$, $MSE = 17{,}380$, $p < .04$. The mean self-presentation times were higher for Compare 1 problems (263 cs) than for Combine 1 (251 cs) or Compare 2 problems (248 cs). There was also an effect of the type of segment, $F(5, 350) = 352.07$, $MSE = 33{,}046$, $p < .001$. While the mean self-presentation times were quite similar for the first four segments (182 cs, 157 cs, 170 cs, 138 cs for s1, s2, s3, and s4, respectively), they were higher for s5 (316 cs) and even higher for s6 (563 ms).

More interestingly, the effect of the type of segment interacted with the position of the question, $F(5, 350) = 120.30$, $MSE = 23{,}223$, $p < .001$. When the question was before the text of the problem, all the mean self-presentation times were shorter than when the question was after the text, except in the case of s5 for which the reverse pattern was obtained (s1, 167 vs. 195 cs for the question before and after respectively; s2, 150 vs. 164 cs; s3, 162 vs. 178 cs; s4, 131 vs. 145 cs; s5, 456 vs. 175 cs; s6, 452 vs. 674 cs), all *F*s $(1, 70) > 16.30$, all ps $< .001$.

Moreover, there was an interaction between the position of the question, the type of segment, and the children's mathematical skill, $F(5, 350) = 2.56$, $MSE = 23{,}556$, $p < .02$. While the differences in the mean self-presentation times as a function of the position of the question were the same for high- and low-skilled children for s1, s3, and s6 ($F < 1$ for the three type of segment), those differences were more pronounced for low-skilled participants in the case of s2 (i.e., first protagonist's action), $F(1, 70) = 6.99$, $MSE = 1{,}151$,

$p < .01$, s4 (i.e., second protagonist's action), $F(1, 70) = 4.93$, $MSE = 1{,}035$, $p < .03$, and s5 (i.e., second quantity of objects), $F(1, 70) = 6.31$, $MSE = 47{,}026$, $p < .01$.

Discussion

The results of this experiment show that children with low skills in mathematics are those who benefit most from the position of the question before the text of an arithmetic word problem. Indeed, although this result was not obtained for resolution times, the proportion of problems correctly solved by children followed this pattern. This result is difficult to interpret within the schema framework, which postulates that arithmetic problems are solved via the mobilization of mental structures stored in long-term memory (Devidal et al., 1997; Kintsch & Greeno, 1985; Riley & Greeno, 1988; Riley et al., 1983). According to this theory, the facilitatory effect of the question before the text would be explained by the early activation of the appropriate schema. The information would therefore be integrated into the schema during reading, and this would result in a better encoding of the problem and, thus, better performance (Devidal et al., 1997). Following this reasoning, children who are more likely to benefit from this facilitarory effect would be those children who are more likely to have extracted a schema for a specific type of problem—that is, high-skilled children. As stated above, the reverse result was obtained here: It was the low-skilled children who benefited most from the position of the question before the problem.

Therefore, an interpretation of the facilitatory effect of the position of the question before the text has to be drawn from a framework that considers the representation constructed to solve the problem as temporary and specific. As explained in the Introduction, the mental model (Johnson-Laird, 1983) or situation model (Reusser, 1989) theories provide such a framework. While the simple fact that high-skilled children are not those who benefit more from the position of the question before the text is sufficient to rule out the schema interpretation, the fact that the inverse result is obtained supports the mental model theory. Indeed, we know in the domain of text comprehension that low-skilled children are those who experience the highest difficulty in the construction of the mental model of the situation (Cain & Oakhill, 1999; Oakhill, 1996; Oakhill et al., 1998; Singer & Ritchot, 1996; Yuill et al., 1989). It therefore seems natural that they are the same children who benefit most from an aid to the construction of the representation—that is, from having the question before the text.

One precaution to bear in mind while discussing these results is that there is more possibility for improvement in the low- than in the high-skilled children. Our results could follow from this artifact. Yet, a strong argument to rule out this possibility is that no ceiling effect is observed on the participants' performance as a function of the position of the question. Indeed, low-skilled children show an improvement of .25 whereas higher skilled children show an improvement of only .14. However, the position of the question before the text of the problem could have allowed the latter to make the same improvement as the former since higher skilled children showed a proportion of .74 of correct answers when the question was positioned after the text.

Just as two different predictions could be drawn from the schema and mental model theory as a function of the children's skill level, so two different predictions could be drawn as a function of the difficulty of the problem. Following the same reasoning as before, the schema theory would have been supported if the facilitatory effect of the position of the question before the text had been greater for easy rather than difficult problems. Again, the general reverse pattern was obtained on the proportions of problems correctly solved, as well as on resolution times. Among the three problems we studied, the facilitatory effect was always more substantial for the most difficult problem—namely, Compare 2—than for the easiest one—namely, Combine 1 problem. The results obtained on Compare 1 problems, which were the intermediate problems in terms of difficulty in our experiment, are less clear. Indeed, while the facilitatory effect is the same for Compare 1 and Compare 2 problems if we consider the resolution times, it is higher for Compare 2 than for Compare 1 problems if we consider the proportions of problems correctly solved. Whereas the results obtained on the proportions of problem correctly solved are coherent with the literature, the results obtained on resolution times are coherent with the data of our experiment. Compare 2 problems are known to be more difficult than Compare 1 problems (Briars & Larkin, 1984; Lewis & Mayer, 1987; Hegarty et al., 1995; Verschaffel, 1994; Verschaffel et al., 1992). Accordingly to the situation model theory, we should therefore obtain higher facilitatory effects for Compare 2 than for Compare 1 problem, which is what we observed for the proportions of problems correctly solved. However, in our experiment, Compare 1 and 2 problems were of equivalent difficulty, and therefore we should not obtain any differences for those two types of problem, which we obtained for the resolution times. It is difficult to find any plausible interpretation for those surprising results. But, nevertheless, they cannot support the schema theory because, descriptively, the benefit of seeing the question prior to those problems (both for the resolution times and for the proportions of problems correctly solved) was always greater than for the easier—namely Combine 1—problems.

Another interesting result in this study is the fact that when the question is placed before the problems, the differences in their relative difficulty almost completely disappear (.85, .80, and .86 for Combine 1, Compare 1, and Compare 2 problems, respectively, for the proportions of problems correctly solved). How could the schema theory account for this result? As mentioned in our Introduction, this theory states that practice, which allows the construction and the consolidation of the problem schema, is responsible for the improvement in performance. In other words, compare problems are seen as more often failed (i.e., more difficult) than combine problems because they are learnt later by children and because they are less often solved. While the placement of the question before the text could have facilitated the early activation of the appropriate schema, it cannot possibly consolidate the compare problem schema in long-term memory. Consequently, the schema theory cannot explain why placing the question before the text improves the performance in such a way that difficult problems become as easy as simpler problems. However, the mental model theory allows an interpretation of this fact. Within this framework, the difficulty of the problem depends on the difficulty for the solver to figure out the situation described by the text. In this sense, compare problems are difficult because, whereas the protagonists' quantities of marbles are presented as two distinct sets (e.g., John has 8 marbles, Tom has 7 marbles), John's ownership has to be represented later as the sum of Tom's marbles with an extra set that represents the difference. In other words, to solve the problem, one of the two quantities has to be conceived later as a part of the other. Combine problems are easier because, conversely, the initial representation induced by the text is the appropriate one to solve the problem (i.e., "two distinct quantities have to be considered in combination", Riley et al., 1983, p. 161). However, when the question is before the text, the solver is directly informed of the relation between the two quantities, and, in the case of compare problems, he or she no longer engages erroneously in the construction of a representation in which the quantities are independent. In other words, the position of the question before the problems could allow the direct construction of the appropriate representation for both combine and compare problems, which could explain why they are equal in difficulty in this specific condition.

In addition to the children's performance, the mean self-presentation times of the different segments that comprised the problem were analysed. As explained earlier, contrary to situation models, schemata do not contain information that is not necessary for the resolution of the problem. Therefore, if a schema is activated as soon as the question is read, the contextual segments should be of no interest for the solver and, therefore, should be read faster when the question is before rather than after the text. This is indeed the result that we obtained. However, this result was not limited to contextual segments. In fact, almost all the segments, even those that are necessarily integrated both in a schema and in a situation model (i.e., information about protagonists and quantities), were read faster when the question was before the text of the problem. As a consequence, it is impossible to conclude that the faster self-presentation times of the contextual segments when the question is before the problem are due to the fact that they are not integrated to the representation constructed by the solver. A more probable interpretation is that, just like a title before a narrative text, the question before an arithmetic problem facilitates the subsequent encoding and integration of the information in the representation. This interpretation is compatible both with the schema and with the situation model theories. The only segment that followed an inverse pattern was s5—namely, the segment that describes the second quantity of objects. This result is not surprising since when the question is before the text, this segment is the last in the problem. Thus, contrary to the condition in which the question is after the text, the resolution time is comprised in its self-presentation time, which necessarily results in higher self-presentation times when the question is before rather than after the text. Note that the very same interpretation is put forward to explain the higher difference of self-presentation time for s6 (i.e., the question) than for s1, s2, s3, or s4: When the question is after the text, it constitutes the last segment of the text. Its self-presentation time therefore contains the resolution time in addition to the reading time, which necessarily results in higher self-presentation times when it is after the text rather than before.

Moreover, for an exclusive schema-based interpretation, the difference in reading times of the contextual segments as a function of the position of the question should have been more pronounced for high-skilled children, whose schemata would be more consolidated in long-term memory. This result was not obtained. In fact, the effects of the position of the question were more pronounced for lower skilled children and for the segments that described the protagonist's actions. Although those segments are integrated both in a mental model and a schema, the fact that the effect of the position of the question is higher for lower skilled than for high-skilled children is out of line with the schema theory.

Those last results show that the reading time collection of the different segments of the text is an informative method that can shed light on the way the problems are processed by individuals. Nevertheless, it has to be noted that this method of text presentation necessarily increases the demand of the task. Indeed, contrary to an ecological situation, the relevant information has to be maintained in working memory until the end of the problem for the resolution. This additional cost may influence the solvers' behaviour. However, because of the specific design of our experiment, this is unlikely to question the validity of our conclusions. Indeed, all our conclusions rely on the comparison of two conditions (i.e., question before or after the text) in which the additional cost due to the segmented presentation remains constant. Therefore, the differences that are observed as a function of the position of the question are improbably attributable to this additional cost.

Before concluding, it is worth noting that defending the relevance of the mental model theory in the construction of arithmetic word problem representations does not mean that the role of practice in the improvement of performance is ignored. The positive effects of practice are undeniable and lead necessarily to the conclusion that something, whatever it is, is stored in long-term memory at a certain point and therefore that the construction of an ad hoc representation in working memory cannot by itself account for this fact. But the literature provides us with little, if any, evidence for the fact that this knowledge corresponds to schemata associated with particular types of arithmetic problem. The improvement of linguistic or procedural knowledge, the increase in use of shortcuts or better strategies, and even better abilities to construct the representation all constitute factors that can account for the improvement in performance owing to practice without having to resort to the assumption that abstract structures and associated procedures are stored in long-term memory for each type of problem.

To conclude, we suggest that mental models for arithmetic problem solving result from a semantic analysis of the situation that the problem wording describes and that this semantic analysis does not differ from the one applied to other types of text. In other words, the construction of these models relies on general comprehension processes. We think that the correct representation of a problem only results from the understanding of natural language and terms such as "give", "lose", "have", or "together". The different sets are not given a definite mathematical status such as part, whole, subset, or superset. Instead, these statuses are implicit within the structure of the models. There is no need to predicate something such as "the set referred to as 'John 8' is a whole" in order to solve the problem "John had 8 marbles, he lost 3 marbles, how many marbles does he have now?". The meaning of the verb "to lose", which is not arithmetical in nature, should be sufficient to ensure the representation of a set of marbles that has been split up into two henceforth separate collections: the lost and the remaining marbles. Only a simple comprehension of the situation described by the text of the problem is therefore necessary to solve an arithmetic problem, and there is no need to invoke the existence of specific associations between mathematical expressions, frames, and procedures stored in long-term memory to account for this activity.

While our previous studies (Thevenot et al., 2004; Thevenot & Oakhill, 2004, 2005a, 2005b, in press) gathered evidence for the construction of mental models in arithmetic word problem solving, this paper constitutes for us a step forward in the sense that it confronts the two conceptions by drawing straightforward predictions from both theories. Our future research will follow up these findings.

Original manuscript received 3 February 2005
Accepted revision received 12 January 2006
First published online 30 March 2006

REFERENCES

Briars, D. J., & Larkin, J. H. (1984). An integrated model of skill on solving elementary word problems. *Cognition and Instruction, 1*, 245–296.

Cain, K., & Oakhill, J. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing, 11*, 489–503.

Carpenter, T. P., & Moser, J. M. (1982). The development of addition and subtraction problem-solving skills. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Coquin-Viennot, D. (2000). Lecture d'énoncés de problèmes arithmétiques: Effet d'une introduction thématique sur la construction de la représentation. *Archives de Psychologie, 68*, 41–58.

Coquin-Viennot, D., & Moreau, S. (2003). Highlighting the role of the episodic situation model in the solving of arithmetical problems. *European Journal of Psychology of Education, 3*, 267–279.

Cummins, D. (1991). Children's interpretations of arithmetic word problems. *Cognition & Instruction, 8*, 261–289.

Cummins, D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. *Cognitive Psychology, 20*, 405–438.

De Corte, E., & Verschaffel, L. (1981). Children's solution processes in elementary arithmetic problems: Analysis and improvement. *Journal of Educational Psychology, 73*, 765–779.

De Corte, E., & Verschaffel, L. (1985). Beginning first graders' initial representation of arithmetic word problems. *Journal of Mathematical Behavior, 4*, 3–21.

De Corte, E., & Verschaffel, L. (1991). Some factors influencing the solution of addition and subtraction word problems. In K. Durkin & B. Shire (Eds.), *Language and mathematical education.* Milton Keynes, UK: Open University Press.

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.

Devidal, M., Fayol, M., & Barrouillet, P. (1997). Stratégies de lecture et résolution de problèmes arithmétiques. *L'Année Psychologique, 97*, 9–31.

Fayol, M., Abdi, H., & Gombert, J. E. (1987). Arithmetic problem formulation and working memory load. *Cognition and Instruction, 4*, 183–202.

Hegarty, M., Mayer, R. E., & Monk, C. A. (1995). Comprehension of arithmetic word problems: A comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology, 87*(1), 18–32.

Hudson, T. (1983). Correspondences and numerical differences between disjoint set. *Child Development, 54*, 84–90.

Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language, 32*, 421–445.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge, MA: Harvard University Press.

Kintsch, W. (1987). Understanding word problems: Linguistic factors in problem solving. In N. Nagao (Ed.), *Language and artificial intelligence* (pp. 197–208). Amsterdam: North Holland.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction integration model. *Psychological Review, 95*, 163–182.

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review, 92*, 109–129.

Lewis, A., & Mayer, R. E. (1987). Students' miscomprehension of relational statements in arithmetic word problems. *Journal of Educational Psychology, 79*, 363–371.

Mayer, R. E. (1983). Can you repeat that? Qualitative effects of repetition and advance organizers from science prose. *Journal of Educational Psychology, 75*, 40–49.

Morales, R., Shute, V., & Pellegrino, J. (1985). Developmental differences in understanding and solving word problems. *Cognition & Instruction, 2*, 41–57.

Moreau, S., & Coquin-Viennot, D. (2003). Comprehension of arithmetic word problem by fifth-grade pupils: Representation and selection of information. *British Journal of Educational Psychology, 73*, 109–121.

Nathan, M. J., Kintsch, W., & Young, E. (1992). A theory of algebra word problem comprehension and its implication for the design of learning environments. *Cognition and Instruction, 4*, 329–390.

Oakhill, J. (1996). Mental models in children's text comprehension. In J. Oakhill & A. Garnham (Eds.), *Mental models in cognitive science* (pp. 77–94). Hove, UK: Psychology Press.

Oakhill, J., Cain, K., & Yuill, N. (1998). Individual differences in children's comprehension skill: Toward an integrated model. In C. Hulme & R. M. Joshi (Eds.), *Reading and spelling: Development and disorders* (pp. 343–367). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Pape, S. J. (2003). Compare word problems: Consistency hypothesis revisited. *Contemporary Educational Psychology, 28*, 396–421.

Pynte, J. (1974). Une expérience automatisée en psycholinguistique. *Informatique et Sciences Humaines, 22*, 45–46.

Rawson, K. A., & Kintsch, W. (2002). How does background information improve memory for text content? *Memory and Cognition, 30*(5), 768–778.

Reusser, K. (1989). *Textual and situational factors in solving mathematical word problems*. Bern, Switzerland: University of Bern.

Riley, M. S., & Greeno, J. G. (1988). Developmental analysis of understanding language about quantities and of solving problems. *Cognition and Instruction, 5*, 49–101.

Riley, M. S., Greeno, J. G., & Heller, J. I. (1983). Development of children's problem solving ability in arithmetic. In H. P. Ginsburg (Ed.), *The development of mathematical thinking.* New York: Academic Press.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theorical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Schank, R. C. (1975). *Conceptual information processing.* Amsterdam: North-Holland.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Singer, M., & Ritchot, K. (1996). The role of working memory capacity and knowledge access in text inference processing. *Memory & Cognition, 24*, 733–743.

Staub, F. C., & Reusser, K. (1995). The role of presentational structures in understanding and solving mathematical word problems. In C. A. Weaver, S. Mannes, & C. R. Fletcher (Eds.), *Discourse comprehension. Essays in honor of Walter Kintsch* (pp. 285–305). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Stern, E., & Lehrndorfer, A. (1992). The role of situational context in solving word problems. *Cognitive Development, 2*, 259–268.

Stetic, W. (1999). Word-problem solving as a function of problem type, situational context and drawing. *Studia Psychologica, 41*, 49–62.

Swanson, H. L., Cooney, J. B., & Brock, S. (1993). The influence of working memory and classification

ability on children's word problem solution. *Journal of Experimental Child Psychology, 55*, 374–395.
Thevenot, C., Barrouillet, P., & Fayol, M. (2004). Mental representation and procedures in arithmetic word problems: The effect of the position of the question. *L'Année Psychologique*, *104*, 683–699.
Thevenot, C., & Oakhill, J. (2004). Relationship between mental representations of arithmetic word problems and working memory capacities [Abstract]. *International Journal of Psychology, 39*(5/6), 558.
Thevenot, C., & Oakhill, J. (2005a). Constraint relaxation in arithmetic word problem solving. *Manuscript submitted for publication*.
Thevenot, C., & Oakhill, J. (2005b). The strategic use of alternative representations in arithmetic word problem solving. *Quarterly Journal of Experimental Psychology*, *58A*(7), 1311–1323.
Thevenot, C., & Oakhill, J. (in press). Representations and strategies for solving dynamic and static arithmetic word problems: The role of working memory capacities. *European Journal of Cognitive Psychology*.
Van Dijk, T., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.
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., De Corte, E., & Pauwels, A. (1992). Solving compare problems: An eye movement test of Lewis and Mayer's consistency hypothesis. *Journal of Educational Psychology, 84*, 85–94.
Verschaffel, L., Greer, B., & De Corte, E. (2000). *Making sense of word problems.* Heereweg, The Netherlands: Swets & Zeitlinger Publishers.
Wilberschied, L., & Berman, P. M. (2004). Effect of using photos from authentic video as advance organizers on listening comprehension in a FLES Chinese class. *Foreign Language Annals, 37*, 534–543.
Yuill, N., Oakhill, J., & Parkin, A. (1989). Working memory, comprehension ability and the resolution of text anomaly. *British Journal of Psychology, 80*, 351–361.
Zwaan, R., & Radvansky, G. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123*, 162–185.