

# Computational requirement and the misunderstanding of language inconsistent word problems

Yves Bestgen (yves.bestgen@psp.ucl.ac.be)

Department of Psychology, 10 Place du Cardinal Mercier  
Louvain-la-Neuve, B-1348 Belgium

## Abstract

Undergraduates meet with difficulties when they have to solve comparison problems in which key words in the problem prime inappropriate operations (e.g., the word "less" priming subtraction when addition is required to solve the problem). This is a comprehension error. In this paper, I examine one factor suspected of making the comprehension of some problems more difficult than others: the computational requirement of the problem. In two experiments, the size of the operand involved in the computation and the number of mathematical operations required to solve a problem were manipulated. Both factors affected the number of comprehension errors leading to the conclusion that increasing the computational requirement of a problem drove some of the solvers to use a superficial mode of understanding the text.

**Keywords:** Arithmetical word problem; language comprehension; computational requirement.

## Introduction

Over the last twenty-five years, arithmetical word problems have been the subject of a large amount of research (e.g., Andersson, 2007; De Corte, Verschaffel, & DeWinn, 1985; Hegarty, Mayer, & Monk, 1995; Kintsch & Greeno, 1985; Lewis & Mayer, 1987; Pape, 2003; Reed, 1999; Riley, Grenno, & Heller, 1983; Reusser & Stebler, 1997; Verschaffel, De Corte, & Pauwels, 1992). These problems are interesting with respect to their role in the teaching of arithmetics, but they are also fascinating materials for cognitive scientists because solving them requires the integration of several competencies: language understanding, problem solving strategies and arithmetical abilities (Andersson, 2007; Cummins, Kintsch, Reusser, & Weimer, 1988; LeBlanc & Weber-Russell, 1996; Thevenot & Oakhill, 2008). This is specifically the case for problems like the following one (Lewis & Mayer, 1987): *At Shell, gas costs \$1.12 per gallon. This is 5 cents less than gas at Texaco. How much do 3 gallons of gas cost at Texaco?*

Children find such problems particularly difficult to solve. According to Riley, Grenno and Heller's classification (1983), this problem is one of the most difficult "two step" arithmetical problems. It is a *comparison* problem because it includes a relational statement that compares the value of two variables. Such problems are more difficult than change or combine problems (Riley et al., 1983). However, there is a more important factor that explains why the example problem is particularly difficult. It is a *language inconsistent* problem because the relational term "less" primes the use of an inappropriate arithmetical operation. To

solve the problem, it is necessary to add 5 to 1.12 rather than subtract it. Use of the inappropriate arithmetical operation, called a reversal error, is often observed with inconsistent problems, while this error is rare for problems with consistent relational statements such as '*Gas at Texaco is 5 cents more than at Shell*' (Lewis, 1989; Verschaffel et al., 1992). Stern (1993), summarizing six studies in which elementary school children solved comparison problems, reported far more errors with inconsistent problems than with consistent problems. Astonishingly, this is also true for undergraduates. Lewis and Mayer (1987) observed that 13% of undergraduates made reversal errors with inconsistent comparison problems versus less than 1% for consistent problems. Furthermore, one third of the 32 undergraduates participating in the study by Hegarty, Mayer and Green (1992) made at least two reversal errors with four comparison problems, of which two were inconsistent and two consistent. The present study aims at understanding why undergraduates meet with difficulties when they have to solve inconsistent comparison problems. Resting on studies that showed that errors originate in the use of a superficial comprehension strategy by the solvers, I evaluated one factor suspected of directing the strategy selection: the computational requirement of the problem.

## Why Do People Err in Solving Inconsistent Problems?

A major change in the study of arithmetical word problems emerged during the eighties when more and more researchers stressed the central role of language comprehension in explaining the difficulties solvers meet (Cummins et al., 1988; De Corte, Verschaffel, & DeWinn, 1985; Kintsch & Greeno, 1985; Lewis & Mayer, 1987). Their conviction mainly arose from two observations. First, it appeared that children often solve arithmetical problems more easily when presented numerically rather than as word problems, even though both require exactly identical computational steps (Cummins et al., 1988). Second, researchers were able to improve dramatically the solution performance by introducing a slight change in the wording of the problem (De Corte et al., 1985; Hudson, 1983; Staub & Reusser, 1995).

The crucial role of language comprehension is particularly evident in the case of language inconsistent compare problems. It is the way the problem is worded that makes it more difficult to solve. For any inconsistent problem, it is possible to design a consistent problem that requests exactly the same computation steps to solve it.

Solvers are more successful with the language consistent problems while they often err when facing the language inconsistent versions. Typically, they make reversal errors (Cummins et al., 1986; Lewis & Mayer, 1987), caused by the use of a deficient strategy to understand the problem. For Hegarty et al. (1992, 1995), unsuccessful problem solvers, i.e. solvers who make a large number of reversal errors, do not build a problem model representing the situation described in the text, while successful problem solvers do. Using a direct translation strategy, unsuccessful solvers try to construct a solution plan by devoting their attention only to keywords such as relational terms and numbers. In the case of an inconsistent problem, they are misled by the relational term that primes the corresponding but incorrect operation ("*less*" primes subtraction and "*more*" addition), leading to a reversal error.

Analysis of the eye fixations collected during the reading of word problems by successful and unsuccessful solvers brought empirical data to support this theory. Unsuccessful solvers spent more time than successful solvers fixating on keywords and numbers, but less time on background information useful to construct a situation model of the problem. Moreover, Hegarty et al. (1995, see also Cummins et al., 1986) observed that successful and unsuccessful problem solvers do not remember the relational sentence in the same way. Unsuccessful solvers reported more often than successful solvers a relational statement that was the reverse of the relation really expressed in the problem, but preserved the original relational keyword they had seen. Such an error is expected if they take a superficial view of the problem and use a direct translation strategy based on the keywords to construct the solution plan.

In summary, research on individual differences suggests that successful solvers tend to construct a meaningful and rich representation of the problem while unsuccessful solvers tend to use a "short-coming" strategy. However, it remains to be explained why successful solvers do not always manage to answer an inconsistent problem correctly and why unsuccessful solvers do not always err (Hegarty et al., 1995, p. 29). Are there factors that make some inconsistent compare problems more prone to reversal errors? Discovering such factors would have several implications. It would point out in which direction one should look to find individual differences that cause the frequent use of an inefficient strategy by some solvers. Furthermore, the distinction between a deep and powerful strategy that leads to the construction of a mental model and a deficient strategy based on a superficial representation of the problem is pervasive in text comprehension research (Graesser, Millis, & Zwaan, 1997; Kintsch, 1998; Noordman & Vonk, 1998; Rinck, 2000). Studying factors that determine the comprehension strategy used should be beneficial for this area too.

In the present study, I evaluated one factor suspected of directing the strategy selection in understanding inconsistent problems: the computational requirement of the problem. Several studies showed that the comprehension and the

computational processes necessary to solve a problem are not serial, but are at least partially overlapping (Hegarty et al., 1992, 1995; Kintsch, 1998). For instance, Hegarty et al. (1992) observed that the additional time necessary for interpreting the inconsistent information does not occur during the first reading of the problem, but later. Solvers wait until they have read the entire problem before going back to the relational statement and start to determine the arithmetical operation requested. It is during this second reading that they build the situational model necessary to solve the inconsistent problem. Because attention resources are limited and simultaneously required for computation and comprehension, increasing the complexity of the computation should capture resources that otherwise would be devoted to the understanding of the relational statement. This should direct some solvers to shift from a deep and elaborate strategy to a more superficial and short-cut one. To test this hypothesis, participants were asked to solve two inconsistent problems that were very different in computational requirement. Obviously, more arithmetical errors for the computationally more difficult problems were expected. The most interesting analysis focused on the comprehension errors. If computational requirement affects comprehension, there should have been more comprehension errors for the computationally harder problems.

## Experiment 1

### Method

#### Participants

Fifty-three participants, all second year psychology students who were native speakers of French, took part in the experiment for course credit.

#### Materials

The experimental material consisted of two inconsistent two-step problems: one was a comparison problem and the other was an equalize problem. In both the comparison and equalize problems, the first sentence assigned a value to a variable. The second sentence was the relational statement that expressed the value of a new variable in terms of the first variable, and the third sentence asked for some kind of computation on the new variable. In both problems, the relational term was the marked term «*less*», which is known to produce many more reversal errors than the unmarked term «*more*» (Lewis & Mayer, 1987). The only difference between the comparison and equalize problems was in the way the relational statement was worded. In the equalize problem, the relational statement told how to change one variable to make it equal to the other.

To manipulate the computational requirements, I wrote two versions of each experimental problem by altering the first two sentences. In both versions, numbers of the same magnitude were used, but computation in the low-requirement version was simplified by using round numbers (See Table 1).

Table 1: Problems for Experiment 1 (numbers in parenthesis refer to the high requirement version).

Comparison Problem
Peter is paid EUR 1900 (1930) a month. That's 300 (355) less than Fabrice. How much does Fabrice earn in 2 months?
Equalize Problem
A Shell stock costs EUR 460 (465) If an Esso stock was EUR 100 (128) less, it would be at the same price as a Shell stock. How much do 3 Esso stocks cost?

There were also nine filler problems, mostly of the combine and change type requiring three or four steps to be solved. None of them were comparison or equalize problems, and none included an inconsistent relational term.

Two counter-balanced sets of problems were constructed. Each set contained the two experimental problems and the nine filler problems. The comparison problem was presented in position 3 and the equalize problem in position 10. In one set, the first experimental problem was the low requirement version and the second experimental problem the high requirement version. The reverse was true for the second set. Therefore, each participant saw both experimental problems, one in its low-requirement version and the other in the high-requirement version. Twenty-seven participants answered the first set and twenty-six the second.

### Procedure

The arithmetical problem task was interleaved with a text comprehension task in which participants had to read 250 word texts and answer four comprehension questions. The arithmetic problems were to be solved between the reading and question answering tasks. To maximize the cognitive resource required by the task, participants were instructed to solve each arithmetical problem mentally without writing any intermediate computations and to do so in 30 seconds. Texts and problems were assembled in booklets. Texts, problems and comprehension questions were printed on separate pages.

Participants were tested in groups of ten to fifteen in a classroom. To keep subjects synchronized over the different tasks, the time to process each task was fixed according to a pilot study. At each transition point, the experimenter gave the signal to start.

### Results

First, an analysis was conducted to determine whether the high requirement problems were, as expected, more difficult than the low requirement problems. Results were clear-cut. None of the 53 participants provided the exact numerical answer for the high requirement problems, while 29 did for the low-requirement problems. High requirement problems were clearly harder.

However, this observation is not sufficient to demonstrate that participants made more comprehension errors in the high-requirement condition, because not all of the erroneous answers were the result of an error in the reversal of the relational term. There were a large number of numerical errors, due to the requirement of computing the solution mentally. The scoring was thus focused on the way participants understood the inconsistent relational term. An answer was scored as *correct* when it was close or equal to the answer the participants should have produced if they correctly reversed the relational term. An answer was scored as a *comprehension error* when it was close or equal to the answer participants would have produced if they did not reverse the relational term. *Undetermined* answers corresponded mainly to cases where participants provided no answer, but also when they did not take into account the relational term, producing their answer by using the first and the third statements only. Table 2 shows that 62% of the comprehension errors were observed in the high requirement condition (see Table 2).

Table 2: Number of Correct Answers, Comprehension Errors and Undetermined Answers for the Low and High Requirement Conditions in Experiment 1

	Low	High
Correct answer	34	10
Comprehension error	13	21
Undetermined answer	6	22

Since each participant answered a low and a high requirement problem, it was possible to apply a nonparametric sign test to the participants who produced no undetermined answers. Twenty-eight participants were in this category. Nine of them produced two correct answers, five produced two comprehension errors, and the remaining 14 produced only one comprehension error. Thirteen of these 14 made their comprehension errors in the high requirement problem while they solved the low requirement problem correctly. Only one participant had the reverse profile, making an error only in the low requirement problem. The sign test on these values confirmed that the high requirement condition led to significantly more comprehension errors than the low requirement condition (*Sign-test*<sup>1</sup> ( $N = 14$ ;  $X = 1$ ),  $p = .0018$ ). In summary, increasing the computational requirement of an inconsistent problem interfered with the comprehension of the problem and led to more reversal errors.

## Experiment 2

The second experiment tests another manipulation of the computational component of a problem: the number of steps required to solve a problem. One-step problems require only

<sup>1</sup> Based on the classical sign test, called *nonrandomized conditional sign test* in Coakley and Heise (1996), which is recommended by many textbooks even if it is quite conservative.

one arithmetic operation to be solved, the operation stated in the relational sentence. The last sentence does not require any supplementary operation. Two-step problems resemble the one used in our first experiment. They end with a question that requires a computational operation on a variable. The more steps a problem requires, the more the demand on the solver's attention resources.

In this second experiment, I contrasted very easy one step problems with far more difficult several step problems. The most important difference between this "number of step factor" and the "magnitude factor" already studied is the position where the manipulation occurs in the problem. The magnitude manipulation affects the relational statement while the number of steps manipulation affects the last sentence of the problem, the sentence that follows the relational statement. If the computational complexity of this last sentence affects the number of errors in understanding the previous sentence, it would increase confidence that comprehension and computational processes interact during the solving of an arithmetical problem.

## Method

### Participants

One hundred and one participants, all second year psychology students who were native speakers of French, took part in the experiment for course credit.

### Materials

The experimental material consisted of two inconsistent comparison problems. To manipulate the computational requirements of the third sentence, I wrote two versions of each problem (see Table 3). In the OSP version, the last sentence included no supplementary operation, while in the 3SP version, it included three supplementary operations: two multiplications and one addition.

Table 3: Comparison Problem for Experiment 2.

OSP	The chemical formulary is 32 pages long. It's 9 pages less than the mathematical one.
	How many pages are in a mathematical formulary?
3SP	The chemical formulary is 32 pages long. It's 9 pages less than the mathematical one.
	How many pages are in 3 chemical and 2 mathematical formularies?

Two counter-balanced sets of materials were constructed. Each set contained the two experimental problems. In one set, the first experimental problem appeared in the OSP condition and the second in the 3SP condition. The reverse was true for the second set. Fifty-one participants answered the first set and fifty answered the second set.

There were also two filler problems that participants answered between the two experimental problems. These were very difficult logical problems adapted from Casey (1993) that were expected to prevent participants from

perceiving the structural similarity between the two experimental problems.

### Procedure

Problems were presented in booklets. The first and last were always the arithmetical word problems and the two intermediates were the filler problems. Participants were tested in groups of five to ten in a classroom. They were instructed to solve the problems mentally without writing any intermediate computations. The time allowed to solve each problem was not limited because, as observed in the first experiment, this restriction prevented some participants from producing an interpretable answer in the available time. Participants typically took between five and twelve minutes to solve the four problems.

### Results

As expected, the OSP problems were far easier to solve than the 3SP problems. Ninety-six percent of the participants provided the exact numerical answer for the OSP, while only 66% did for the 3SP problems.

To test our prediction that participants made more comprehension errors in the 3SP condition than in the OSP condition, I scored participants' answers based on how they understood the inconsistent relational term. The scoring procedure was identical to the procedure used in Experiment 1, except that it was not necessary to use an undetermined category as all the participants produced interpretable numerical answers.

Table 4: Number of Correct Answers and Comprehension Errors for the OSP and 3SP Conditions in Experiment 2

	OSP	3SP
Correct answer	97	84
Comprehension error	4	17

As summarized in Table 4, 80% of the comprehension errors were observed in the 3SP condition. Overall, 81 participants made no comprehension errors and one made an error both in the OSP and in the 3SP problem. The remaining 19 participants made only one comprehension error. Sixteen of them made their comprehension error in the 3SP problem while they solved the OSP problem correctly. Only three participants had the reverse profile, making an error in the OSP problem only. The sign test on these values confirmed that the 3SP condition led to significantly more comprehension errors than the OSP condition (*Sign-test* ( $N = 19$ ,  $X = 3$ ),  $p = .0044$ ).

### Conclusion

Language inconsistent arithmetical problems have been the subject of a large amount of research over the last twenty-five years. It has shown that difficulties occur during the comprehension of the problem. It has also pointed out important individual differences in the ability to solve these kinds of problems due to the strategy used to understand the

problem. Successful solvers construct a meaningful and rich representation of the problem while unsuccessful solvers use a short-cut strategy that leads to reversal errors. The present paper addresses a related question: why are some language inconsistent problems more difficult than others or, alternatively stated, why do solvers sometimes use an inefficient strategy and sometimes an efficient one? More specifically, I focused on one factor suspected of directing the strategy selection: the computational requirement of the problem. In two experiments I manipulated the size of the operand involved in the computation and the number of mathematical operations required to solve the problem. Both factors affected the number of reversal errors produced by the solvers. The harder the problem was at the computational level, the larger the number of comprehension errors. These results can be interpreted in the framework of Hegarty et al.'s (1992, 1995) explanations of individual differences in solving inconsistent problems. Increasing the computational requirements of a problem directs some solvers to shift from a deep and elaborate strategy to a more superficial and short-cut one.

The limited scope of this paper must be stressed. This study specifically addresses the difficulties met by undergraduates who are able to correctly answer inconsistent problem, but who sometimes make comprehension errors. Explaining the difficulties children have to master inconsistent problems is another question studied, for instance, by Stern (1993). She observed that first graders did not seem to understand that statements like "Jane has 2 marbles more than Peter" and "Peter has two marbles fewer than Jane" express the same situation (Stern, 1993). They were consequently unable to reverse the relational term. This was not the case for the undergraduates who participated to our experiments because almost all of them solved correctly at least one of the two inconsistent problems.

However, by focusing on a factor that directs the comprehension strategy used by the readers, this study has implications that go further than the processing of arithmetical problems to reach text comprehension in general. The strategy selected to understand a text and the nature of the representation build by the reader is central to current models of text comprehension (Graesser, Millis, & Zwaan, 1997; Kintsch, 1998; Rinck, 2000; Tzeng, van den Broek, Kendeou, & Lee, 2005; van den Broek, Rapp, & Kendeou, 2005). A text can be understood at a superficial level leaving readers with a propositional representation of its content or at a deeper level if readers build a mental model of what the text is about. Only the latter allows the reader to fully understand the text (Kintsch, 1994). Research on text comprehension showed that several factors determine the depth of processing. Readers who have relevant prior knowledge about a text tend to build a mental model while low-knowledge readers build a representation closer to the surface of the text (McNamara & Kintsch, 1996; Noordman & Vonk, 1992). Characteristics of the text are also important. Rewriting a poor text to make it more

coherent favours the building of a mental model (Britton & Gülgöz, 1991; Roebben & Bestgen, 2006)). McNamara et al. (1996) and O'Reilly and McNamara (2007) showed furthermore that these readers' and textual factors interact in determining the comprehension strategy. Even if our study is focused on arithmetical word problems, it sheds a complementary light on this question by identifying another factor that directs the selection of the comprehension strategy. Increasing the computational requirement of a problem was enough to conduct some of the solvers to use a superficial mode of understanding the text. This was observed even if solvers had the prior knowledge necessary to understand the problem, that is knowing that *if gas at Shell is five cents less than at Texaco, gas at Texaco is five cents more than at Shell*. They just skipped to use it. Moreover, the use of a deficient comprehension strategy was forced by manipulating not the comprehension difficulty of the text, but its computational difficulties. Further research is now needed to determine if such effects can be observed with other genre of texts like expository texts and to test other manipulations that could direct the selection of a superficial strategy.

## Acknowledgments

Yves Bestgen is a Research Associate of the Belgian Fund for Scientific Research (F.R.S-FNRS).

## References

- Andersson, U. (2007). The contribution of working memory to children's mathematical word problem solving. *Applied Cognitive Psychology*, 21, 1201–1216.
- Britton, B. K., & Gülgöz, S. (1991). Using Kintsch's computational model to improve instructional text: effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psychology*, 83, 329–345.
- Casey, P. J. (1993). "That man's father is my father's son": The role of structure, strategy, and working memory in solving convoluted verbal problems. *Memory and Cognition*, 21, 506–518.
- Coakley, C. W., & Heise, M. A. (1996). Versions of the sign test in the presence of ties. *Biometrics*, 52, 1242–1251.
- Cummins, D. D., Kintsch, W., Reusser, K., & Weimer, R. (1988). The role of understanding in solving word problems. *Cognitive Psychology*, 20, 405–438.
- DeCorte, E., Verschaffel, L., & DeWinn, L. (1985). The influence of rewording verbal problems on children's problem representation and solutions. *Journal of Educational Psychology*, 77, 460–470.
- Graesser, A.C., Millis, K.K., Zwaan, R.A. (1997). Discourse comprehension. *Annual Review of Psychology*, 48, 163–189.
- Hegarty, M., Mayer, R. E., & Green, C. E. (1992). Comprehension of arithmetic word problems: evidence

- from students' eye fixations. *Journal of Educational Psychology*, 84, 76-84.
- 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, 18-32.
- Hudson, T. (1983). Correspondences and numerical differences between disjoint sets. *Child Development*, 54, 84-90.
- Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist* 49, 294-303.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.
- Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review*, 92, 109-129.
- LeBlanc, M. D., & Weber-Russell, S. (1996). Test integration and mathematical connections: A computer model of arithmetic word problem solving. *Cognitive Science*, 20, 357-407.
- Lewis, A. B. (1989). Training students to represent arithmetic word problems. *Journal of Educational Psychology*, 81, 521-531.
- Lewis A. B., & Mayer R. (1987). Students' miscomprehension of relational statements in arithmetic word problem, *Journal of Educational Psychology*, 79, 363-371.
- McNamara, D. S., Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence, *Discourse Processes*, 22, 247-288.
- Noordman, L., & Vonk, W. (1992). Reader's knowledge and inferences in reading. *Language and Cognitive Processes*, 7, 373-591.
- Noordman, L. & Vonk, W. (1998). Discourse comprehension. In A. D. Friederici (Ed.), *Language comprehension: A biological perspective* (229-262). Berlin: Springer.
- O'Reilly, T. & McNamara, D.S. (2007). Reversing the reverse cohesion effect: Good texts can be better for strategic, high-knowledge readers. *Discourse Processes*, 43, 121-152.
- Pape, S.J. (2003). Compare word problems: Consistency hypothesis revisited. *Contemporary Educational Psychology*, 28, 396-421.
- Reed, S. K. (1999). *Word problems: Research and curriculum reform*. Hillsdale: Erlbaum.
- Reusser, K., & Stebler, R. (1997). Every word problem has a solution — The social rationality of mathematical modeling in schools. *Learning and Instruction*, 7, 309-327.
- Riley, M. S., Greeno, J. O, & Heller, J. L. (1983). Development of children's problem-solving ability in arithmetic; In H. P. Ginsberg (Ed.), *The development of mathematical thinking* (pp.153-196). New-York: Academic Press.
- Rinck M., (2000). Situation models and text comprehension - Findings and problems. *Psychologische Rundschau*, 51, 115-122.
- Roebben, N., & Bestgen, Y. (2006). Reading and expertise: The impact of connectives on text comprehension in the financial field. In S. Carliner, J. P. Verckens and C. de Waele (Eds.), *Information and Document Design: Varieties in recent research* (pp.149-165), Benjamins.
- Staub, F., & 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* (pp. 285-305). Hillsdale: Erlbaum.
- Stern, E. (1993). What makes certain arithmetic word problems involving the comparison of sets so difficult for children? *Journal of Educational Psychology*, 85, 7-23.
- Thevenot, C., & Oakhill, J. (2008). A generalization of the representational change theory from insight to non-insight problems: The case of arithmetic word problems. *Acta Psychologica*, 129, 315-324.
- Tzeng, Y., van den Broek, P., Kendeou, P., & Lee, C (2005). The computational implementation of the Landscape Model: Modeling inferential processes and memory representation of text comprehension. *Behavioral Research Methods, Instruments & Computers*, 37, 277-286.
- Van den Broek, P., Rapp, D. N., Kendeou, P., 2005. Integrating memory-based and constructionist processes in accounts of reading comprehension. *Discourse processes*, 39, 299-316.
- 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.