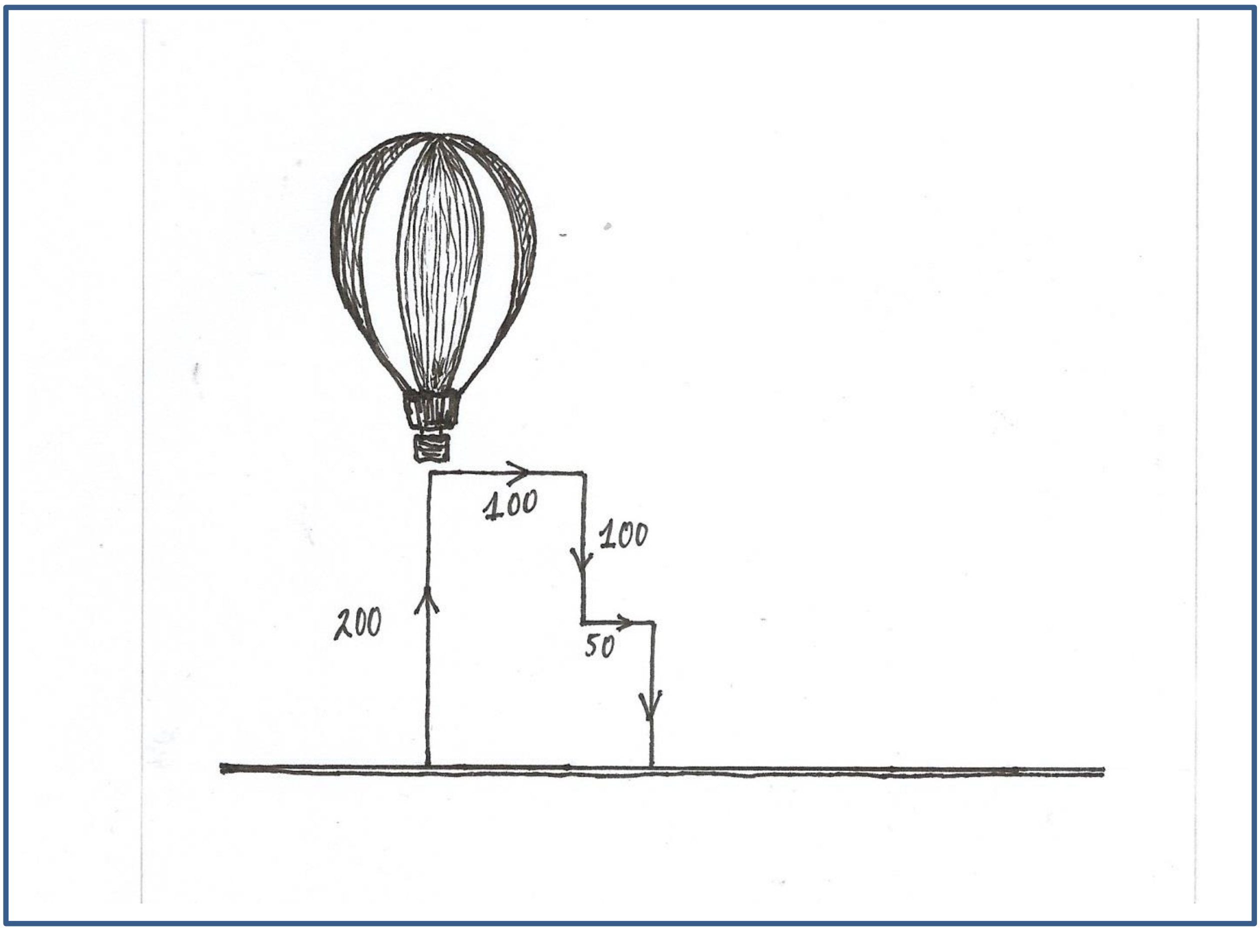
CHAPTER 3

Word problem solving in contemporary math education: A plea for semantic-linguistic skills training

Anton J. H. Boonen, Björn B. de Koning, Jelle Jolles, & Menno van der Schoot

(under review)



A balloon first rose 200 meters from the ground, then moved 100 meters to the east, then dropped 100 meters. It then traveled 50 meters to the east, and finally dropped straight to the ground. How far was the balloon from its original starting point?

Abstract

This study pointed out that word problem solving instruction in Realistic Math Education (RME) may have insufficient attention for teaching semantic-linguistic skills to handle semantic complexities in word problems. We investigated the performances of 80 sixth grade students, classified as successful and less successful word problem solvers based on a standardized mathematics test from the RME curriculum, on word problems that ask for both sophisticated representation skills and semantic-linguistic skills. The results showed that even successful word problem solvers had a low performance on semantically complex word problems, despite adequate performance on semantically less complex word problems. Less successful word problem solvers had low scores on both semantically simple and complex word problems. Results showed that reading comprehension was only related to the successful word problem solvers' performance on semantically complex word problems. On the basis of this study, we concluded that semantic-linguistic skills should be given a (more) prominent role during word problem solving instruction in RME.

Keywords: sophisticated representation strategies; semantic-linguistic skills; word problem solving

Introduction

In the last decades, mathematical word problem solving has gained much attention from both researchers and educational practitioners (Boonen, Van der Schoot, Van Wesel, De Vries, & Jolles, 2013; Campbell, 1992; Depaepe, De Corte, & Verschaffel, 2010; Hegarty, Mayer, & Monk, 1995; Hajer, 1996; Hickendorff, 2011, 2013; Moreno, Ozogul, & Reisslein, 2011; Swanson, Lussler & Orosco, 2013). Mathematical word problems refer to mathematical exercises that present relevant information on a problem as text, rather than in the form of mathematical notation (Rasmussen & King, 2000; Timmermans, Van Lieshout, & Verhoeven, 2007). Hence, effectively solving a mathematical word problem is assumed to depend not only on students' ability to perform the required mathematical operations, but also on the extent to which they are able to accurately understand the text of the word problem (Hegarty et al., 1995; Jitendra & Star, 2012; Lewis & Mayer, 1987; Van der Schoot, Baker-Arkema, Horsley, & Van Lieshout, 2009). Both of these aspects are related in such a way that developing a deeper understanding of the text of the word problem serves as a crucial step before the correct mathematical computations can be performed. Hence, a key challenge for word problem solvers is to get an adequate understanding of the problem statement (Boonen et al., 2013; Lee, Ng, & Ng, 2009; Thevenot, 2010).

Two individual skills are relevant in this regard. First, an important factor contributing to a deeper understanding of the text of the word problem is the ability to construct a rich and coherent mental representation containing all (the relations between the) solution-relevant elements that are derived from the text base of the word problem (De Corte, Verschaffel, & De Win, 1985; Hegarty et al., 1995; Pape, 2003). That is, word problem solvers have to use a problem-model strategy in which they translate the problem statement into a qualitative mental representation of the problem situation hidden in the text (Pape, 2003; Van der Schoot et al., 2009). This mental representation subsequently allows them to make a solution plan and execute the required mathematical operations. Although successful word problem solvers appear to employ such a representational strategy, less successful problem solvers often adopt an impulsive, superficial direct translation strategy, in which they only focus on selecting the presented numbers that, in turn, form the basis for their mathematical calculations (Hegarty et al., 1995; Verschaffel, De Corte, & Pauwels, 1992).

The second important individual skill in word problem solving success substantiated by research evidence is the influence of a student's reading comprehension abilities (Boonen et al., 2013; Pape, 2004; van der Schoot et al., 2009). It has been suggested that reading comprehension abilities are especially helpful in dealing with semantic-linguistic word problem characteristics such as the sequence of the known elements in the text of the word problem, the degree to which the semantic relations between the given and unknown quantities of the problem are made explicit, and the relevance of the information in the text of the word problem (De Corte et al., 1985; De Corte, Verschaffel, & Pauwels, 1990; Marzocchi, Lucangeli, De Meo, Fini, & Cornoldi, 2002; Verschaffel et al., 1992).

Moreover, semantic-linguistic skills appear to be more important in overcoming such textual complexities than being able to apply sophisticated representation strategies (De Corte et al., 1985; 1990). This might explain why the use of a sophisticated mental representation strategy is not sufficient in all circumstances. That is, word problems containing semantically complex features require both accurate representation skills and reading comprehension skills, whereas for word problems with a lower semantic-linguistic complexity, sophisticated representational skills might be sufficient.

Teaching word problem solving in Realistic Mathematics Education

These findings suggest that, to teach students how to effectively solve mathematical word problems, sophisticated representational skills and semantic-linguistic skills should both be part of the mathematics education program. Particularly, paying attention to semantic-linguistic skills is relevant to help students improve their word problem solving success, as word problems become semantically more complex as students progress in their educational career, for example, when they make the transition to secondary education. Word problems offered in secondary school subjects like geometry, physics and biology, include more verbal information and generally contain more complex semantic-linguistic text features (Helwig, Rozek-Tedesco, Tindal, Heath & Almond, 1999; Silver & Cai, 1996).

The Netherlands, like many other countries, currently places great emphasis on the teaching of word problem solving in contemporary mathematics education (Elia, Van den Heuvel-Panhuizen, & Kovolou, 2009; Ruijssenaars, Van Luit, & Van Lieshout, 2004;). The teaching of mathematics in the Netherlands takes place within the context of a domain-specific instructional approach, called Realistic Mathematics Education (RME, Van den Heuvel-Panhuizen, 2003), where the process of mathematical word problem solving plays an important role (Barnes, 2005; Hickendorff, 2011; Prenger, 2005; Van de Boer, 2003; Van den Heuvel-Panhuizen, 2005). Studies investigating the educational practice of RME show that the teaching of sophisticated representation strategies receives a lot of attention in word problem solving instruction (Elia et al., 2009; Van Dijk, Van Oers, & Terwel, 2003; Van den Heuvel-Panhuizen, 2003). However, the training of semantic-linguistic skills appears to be less explicitly trained in the instructional practice of RME, in spite of its proven importance in previous studies (e.g., De Corte et al., 1985, 1990; Hegarty et al., 1992). This is presumably because teachers may underestimate or are not aware of the importance of semantic-linguistic skills for solving word problems (Hajer, 1996; Van Eerde, 2009). Thus, the current approach to teaching word problem solving appears to emphasize the development of representation skills, but seems to pay less attention to the role of semantic-linguistic skills.

In this respect, educational practice regarding teaching word problem solving does not seem to be aligned with what is currently known from research about the factors involved in effective word problem solving. This study aims to provide evidence for the claim that semantic-linguistic skills receive little attention in word problem solving instruction in RME, thereby identifying an important area of concern with respect to the way word problem solving is currently taught in the Netherlands. To test this claim, we compared students’ performance on word problems obtained while following the RME curriculum to their performances on an independent word problem solving task. First, we classified students as successful or less successful word problem solvers with the help of a mathematics test that is part of the RME curriculum, viz., the CITO (Institute for Educational Measurement) Mathematics test. This test can be considered a method-specific (i.e., RME-specific) mathematics test of students' word problem solving performance, as it builds upon the currently used instructional method for word problem solving. Hence, this test reflects the skills that students learn in the RME classroom, in order to solve word problems (Doorman, Drijvers, Dekker, Van den Heuvel-Panhuizen, De Lange, & Wijers, 2007; Hickendorff, 2011). Second, we examined students' performance on an independent word problem solving test, which contained either word problems that could be solved by only using a sophisticated mental representation strategy, or word problems that required them to also use their semantic-linguistic skills.[1]

Based on the assumption that word problem solving instruction in RME pays little attention to handling the semantic-linguistic features of the problem text, we hypothesized that it is likely that a key aspect that differentiates successful from less successful word problem solvers concerns their ability to construct a sophisticated mental representation of the problem text. Previous studies have shown that asking students to solve compare problems, especially inconsistent compare problems (see Example 1), is a suitable method for investigating whether or not they use a sophisticated representation strategy (e.g., Pape, 2003; Van der Schoot et al., 2009).

[Example 1]

At the grocery store, a bottle of olive oil costs 7 euro.

That is 2 euro *more than* at the supermarket.

---

[1] This procedure provides an advantage over prior studies of, among others, Hegarty et al. (1995), Pape (2003), and van der Schoot et al. (2009), which typically use the main dependent variable of the study (i.e., problem solving success) as an outcome measure as well as a means to classify students into successful and less successful word problem solvers. The classification used in the present study, on the other hand, is based on an external, well-established measure of mathematical word problem solving, which is independent of the main dependent variable of the study (i.e., word problem solving success). This allows us to make more meaningful group comparisons.

If you need to buy 7 bottles of olive oil, how much will it cost at the supermarket?

In this example, the translation process requires the identification of the pronominal reference 'that is' as the indicator of the relation between the value of the first variable ('the price of a bottle of olive oil at the grocery store') to the second ('the price of a bottle of olive oil at the supermarket'). This identification is necessary to become cognizant of the fact that, in an inconsistent compare problem, the relational term 'more than' refers to a subtraction operation rather than to an addition operation. So, inconsistent word problems create greater cognitive complexity than consistent word problems, requiring students to ignore the well-established association between *more* with increases, and addition and *less* with decreases and subtraction (Schumacher & Fuchs, 2012). Empirical evidence corroborates this interpretation by showing that word problem solvers make more (reversal) errors on inconsistent than on consistent word problems (i.e., consistency effect, Lewis & Mayer, 1987; Pape, 2003; Van der Schoot et al., 2009). Especially students who fail to build a high-quality mental representation of the problem statement, and thus immediately start calculating with the given numbers and relations, seem to be less successful on inconsistent word problems (Hegarty et al., 1995).

In the present study, we expected neither successful nor less successful problem solvers to experience difficulties with solving consistent compare word problems. However, we did assume that successful word problem solvers in the RME curriculum would experience less difficulties with correctly solving inconsistent compare problems as a result of their use of a sophisticated representation strategy (acquired during word problem solving instruction in RME), than less successful problem solvers who employ a more superficial problem solving approach (Van der Schoot et al., 2009; Verschaffel et al., 1992).

It is important to keep in mind that this only holds for consistent and inconsistent compare problems with a low semantic complexity; that is, problems that only tap into students' ability to construct a sophisticated mental representation. If the semantic complexity of compare problems increases, even students classified as successful word problem solvers (according to our classification based on the RME instruction) may come to experience difficulties with correctly solving inconsistent compare problems. In this case, correctly solving a word problem requires students to use both mental representational skills and semantic-linguistic skills, while word problem solving instruction in RME has provided students only with considerable training in the first of these two skills.

A relatively well-studied and accepted way to increase the semantic complexity of (inconsistent) compare problems is to manipulate the relational term (Lewis & Mayer, 1987; Van der Schoot et al., 2009). That is, we can increase the semantic complexity of a word problem by making a distinction between an unmarked ('more than'), and a marked ('less than'), relation term. Research has shown that students find it easier to convert the unmarked relational term 'more than' into a subtraction operation than the marked relational term 'less than' into an addition operation (Clark, 1969; Kintsch, 1998; Lewis & Mayer, 1987; Pape, 2003; Van der Schoot et al., 2009). The difficulties experienced with solving marked inconsistent word problems lie in the fact that these problems draw on students' use of a sophisticated representation strategy as well as on their semantic-linguistic skills. As the effect of semantic-linguistic complexity only starts to play a role when the problem statement has been mentally represented accurately, the influence of semantic-linguistic skills is restricted to the group of successful problem solvers. So, although our group of successful word problem solvers may use a sophisticated representation strategy, the lack of attention to semantic-linguistic skills in the educational practice of RME is likely to cause them to experience difficulties with correctly solving (semantically complex) marked inconsistent word problems.

According to several researchers, the extent to which successful word problem solvers might be able to overcome difficulties with correctly solving marked inconsistent word problems is related to their semantic-linguistic skills (e.g., Lee, Ng, Ng, & Lim, 2004; Van der Schoot et al., 2009). Translating a marked relational term like 'less than' into an addition operation is found to be closely associated with a general measure of semantic-linguistic skills, and with reading comprehension in particular (Lee et al., 2004; Van der Schoot et al., 2009). This suggests that reading comprehension skills, together with sophisticated representation skills, might be necessary to deal with semantically complex word problems. The present study therefore also takes into account students' general reading comprehension ability.

In sum, the present study aimed to test the claim that the current Dutch instructional approach used in RME pays limited attention to the semantic-linguistic skills that allow students to handle linguistic complexities in a word problem. To this end, we tested the following hypotheses:

1. We hypothesized that, as a result of difficulties with constructing a coherent mental representation of word problems, less successful word problem solvers in the RME curriculum would make more errors on both unmarked and marked inconsistent word problems than on unmarked and marked consistent word problems
2. We hypothesized that, as a result of paying insufficient attention to semantic-linguistic skills in the teaching of word problem solving, successful word problem solvers in the RME curriculum would experience difficulties with solving semantically complex, marked inconsistent word problems, but not with solving semantically less complex, unmarked, inconsistent word problems.
3. We hypothesized that, as a result of the alleged relation between reading comprehension ability and the ability to overcome the semantic-linguistic complexities of a word problem, a positive relation for successful problem solvers exists between reading comprehension ability and the number of correctly solved marked inconsistent word problems.

Methods

Selection of participants

Data from 80 Dutch sixth-grade students (42 boys, M_{age} = 11.72 years, SD_{age} = 0.39 years and 38 girls, M_{age} = 11.71 years, SD_{age} = 0.41 years) from eight elementary schools in the Netherlands were collected. These students were almost equally divided in two groups (by means of the median split method) on the basis of their score on the CITO Mathematics test (2008). This selection procedure resulted in a group of less successful word problem solvers (*N* = 41) and a group of successful word problems solvers (*N* = 39). The CITO Mathematics test is a nationwide standardized test that reflects the way in which word problem solving is instructed in Realistic Mathematics Education. The test contains elements like *mental arithmetic* (addition, subtraction, multiplication and division), *complex applications* (problems involving multiple operations) and *measurement and geometry* (knowledge of measurement situations), all of which are offered as mathematical word problems. The internal consistency of this test was high (Cronbach's alpha = .95, Janssen, Verhelst, Engelen & Scheltens, 2010).

Parents provided written informed consent based on printed information about the purpose of the study.

Instruments and procedure

The two measurement instruments that were used in this study were administrated to the students by three trained independent research assistants in a session of approximately 45 minutes.

Inconsistency task

The inconsistency task contained eight two-step compare problems that were selected from the study of Hegarty et al. (1992) and translated into Dutch. All of the word problems consisted of three sentences. The first sentence of each compare problem was an assignment statement expressing the value of the first variable, namely the price of a product at a well-known Dutch store or supermarket (e.g., At Albert Heijn a bottle of olive oil costs 4 euro). The second sentence contained a relational statement, expressing the value of the second variable (i.e., the price of this product at another store or supermarket) in relation to the first (e.g., At Spar, a bottle of olive oil costs 3 euro more than at Albert Heijn). In the third sentence, the problem solver was asked to find a multiple of the value of the second variable (e.g., If you need to buy three bottles of olive oil, how much will you pay at Spar?). The answer to these compare problems always involved first computing the value of the second variable (e.g., 4 + 3 = 7), and then multiplying this solution by the quantity given in the third sentence (e.g., 7 times 3 = 21).

The eight compare problems were separated in four different word problem types by the crossing of two within-subject factors: *consistency* (consistent vs. inconsistent) and *markedness* (unmarked vs. marked). Consistency referred to whether the relational term in the second sentence was consistent or inconsistent with the required arithmetic operation. A consistent sentence explicitly expressed the value of the second variable (At Spar a bottle of olive oil costs 3 euro [more/less] than at Albert Heijn) introduced in the prior sentence (At Albert Heijn a bottle of olive oil costs 4 euro). An inconsistent sentence related the value of the second variable to the first by using a pronominal reference (That is 3 euro [more/ less] than at Albert Heijn). Consequently, the relational term in a consistent compare problem primed the appropriate arithmetic operation ('more than' when the required operation is addition, and 'less than' when the required operation is subtraction). The relational term in an inconsistent compare problem primed the inappropriate arithmetic operation ('more than' when the required operation is subtraction, and 'less than' when the required operation is addition). Markedness expressed the semantic complexity of the relational term. A marked relational term (i.e., less than) is semantically more complex than an unmarked relational term (i.e., more than).

The stimuli were arranged in four material sets. Each participant was presented with eight word problems, two from each word problem type. The order in which the word problems were presented in each set was pseudorandomized. Each set was presented to 20 participants. Across sets and across participants, each word problem occurred equally often in the unmarked/consistent, marked/consistent, unmarked/inconsistent and marked/ inconsistent version to ensure full combination of conditions and materials. Across word problems, we controlled for the difficulty of the required calculations, and for the number of letters in the names of the variables (i.e., stores) and products. To ensure that the execution of the required arithmetic operations would not be a determining factor in students' word problem solving performance, the operations were selected on the basis of the following rules: (1) the answers to the first step of the operation were below 10; (2) the final answers were between 14 and 40; (3) none of the first steps or final answers contained a fraction of a number or negative number; (4) no numerical value occurred twice in the same problem; and (5) none of the (possible) answers were 1. The numerical values used in consistent and inconsistent problems of each word problem type were matched for magnitude (see Van der Schoot et al., 2009).

For the analyses, we looked at students' accuracy (i.e., the amount of correct answers) on each of the four word problem types: (1) unmarked/consistent; (2) marked/consistent; (3) unmarked/inconsistent; and (4) marked/inconsistent. The internal consistency of this measure in the present study was high (Cronbach's alpha = .90).

Reading comprehension

The (Grade 6 version of the) normed standardized CITO (Institute for Educational Measurement) Test for Reading Comprehension (2010) of the Dutch National Institute for Educational Measurement was used to assess children's reading comprehension level. This test is part of the standard Dutch CITO pupil monitoring system and is designed to determine general reading comprehension level in elementary school children. This test consists of two modules, each involving a text and 25 multiple choice questions. The questions pertained to the word, sentence or text level, and tapped both the text base and situational representation that the reader constructed from the text (Kintsch, 1998). On this test, children's reading comprehension level is expressed by a reading proficiency score, which, in this study, ranged from 15 to 95 (M = 40.51, SD = 13.94). The internal consistency of this test was high with a Cronbach's alpha of .89 (Weekers, Groenen, Kleintjes & Feenstra, 2011).

Data analysis

A 2 x 2 x 2 analysis of variance (ANOVA) was conducted with Consistency (consistent vs. inconsistent) and Markedness (unmarked vs. marked) as within-subject factors and Group (less successful vs.

successful word problem solvers) as the between-subject factor. Follow-up tests were performed using paired sample t-tests.

In the present study, the role of reading comprehension in the four word problem types was examined by calculating the correlations (Pearson's *r*) between reading comprehension and the difference score between the unmarked inconsistent and consistent word problem types, and the correlation between reading comprehension and the difference score between the marked inconsistent and consistent word problem types. These difference scores reflect the differences in performance between the consistent and inconsistent word problem types, and can be taken as a measure of the extent to which students are able to construct a mental representation of the described problem situation. The lower the difference score, the less word problem solvers suffer from the inconsistency. The correlations were calculated for less successful and successful word problem solvers separately.

This approach deviates from, but provides an important advantage over, the study by Van der Schoot et al. (2009), who added reading comprehension as a covariate in the repeated measures ANOVA. That is, the results obtained by Van der Schoot et al. (2009) could provide only limited insight into the exact locus of the covariate's effect, as it was not known which group (less successful or successful word problem solvers) or in which word problem type (consistent unmarked/marked or inconsistent unmarked/marked) reading comprehension played a role. Moreover, it turns out that the repeated measures ANCOVA does change the main effects of the repeated measures compared to assessing the main effects via a simple repeated measures ANOVA (see Thomas, Annaz, Ansari, Scerif, Jarrold, & Karmiloff-Smith, 2009). So, the approach used in the present study enabled us to obtain more specific insight into the precise role of reading comprehension in word problem solving. All the analyses had an alpha of .05.

Results

In Figures 1 and 2, word problem solving performance is presented as a function of consistency (consistent vs. inconsistent) and markedness (marked vs. unmarked) for less successful problem solvers (Figure 1), and for successful problem solvers (Figure 2), respectively.

Inspection of both figures shows that, as expected, the effects of consistency and markedness differed for less successful and successful word problem solvers. As shown in Figure 1, for less successful word problem solvers there was a consistency effect for both marked and unmarked word problems (Consistency: F (1,40) = 10.94, $p < .01$, $\eta_p^2 = .22$; Consistency x Markedness interaction: F (1,40) = 0.25, $p = .62$, $\eta_p^2 = .01$, indicating a large and small effect size respectively, according to Pierce, Block & Auguinis [2004]). So, less successful word problem solvers performed significantly lower on both the unmarked and marked inconsistent word problem types, compared to the consistent unmarked and marked word problem types (t (40) = 2.22, $p < .05$; t (40) = 3.02, $p < .01$ respectively).

However, as displayed in Figure 2, different findings were obtained in the group of successful problem solvers. In this group, the consistency effect was present for marked but absent for unmarked word problems (Consistency: F (1, 38) = 13.00, $p < .001$, $\eta_p^2 = .26$; Consistency x Markedness interaction: F (1, 38) = 16.03, $p < .001$, $\eta_p^2 = .30$, which can be considered to be large effects according to Pierce et al., [2004]). This indicates that successful word problem solvers performed significantly lower on marked inconsistent compared to marked consistent word problems (t (38) = 4.67, $p < .001$); whereas performance on unmarked consistent and unmarked inconsistent word problem types did not differ significantly (t (38) = 1.07, $p = .29$). This pattern of findings regarding the successful and less successful problem solvers was evidenced by a significant three-way interaction between consistency, markedness, and group (F (1,78) = 4.32, $p < .05$, $\eta_p^2 = .05$, indicating a medium-sized effect).

In sum, these findings show that less successful word problem solvers performed lower on both semantic-linguistically simple and complex word problems, whereas successful word problem solvers only performed lower when the word problem text contained complex semantic-linguistic features.

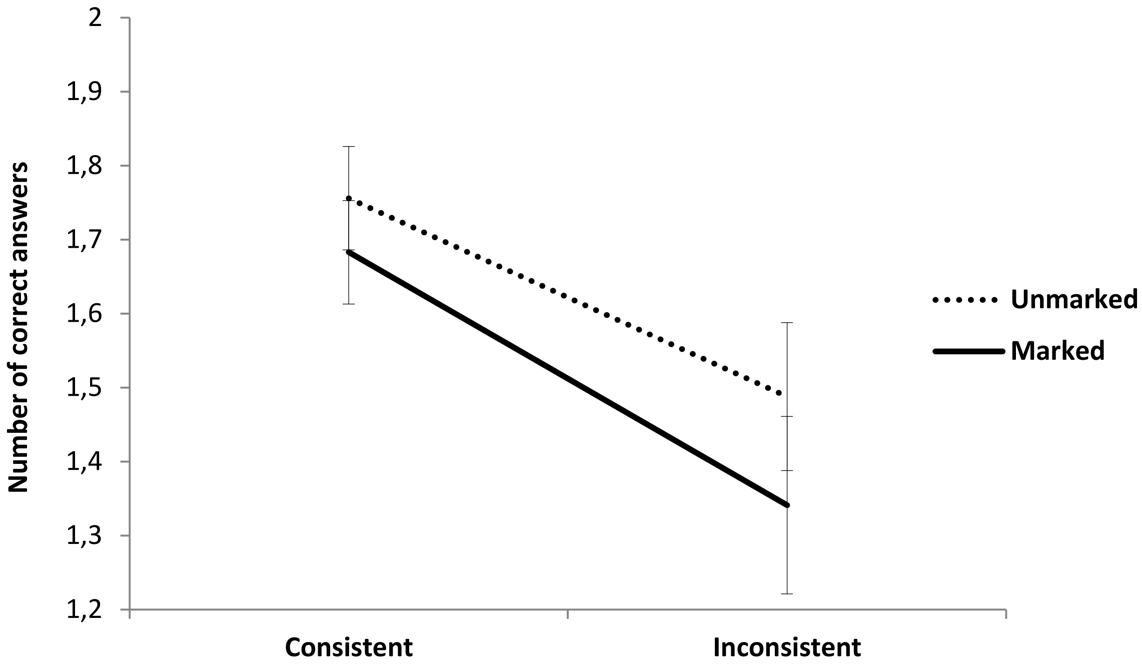


Figure 1. Less successful word problem solvers: Interaction effect Consistency x Markedness x Group

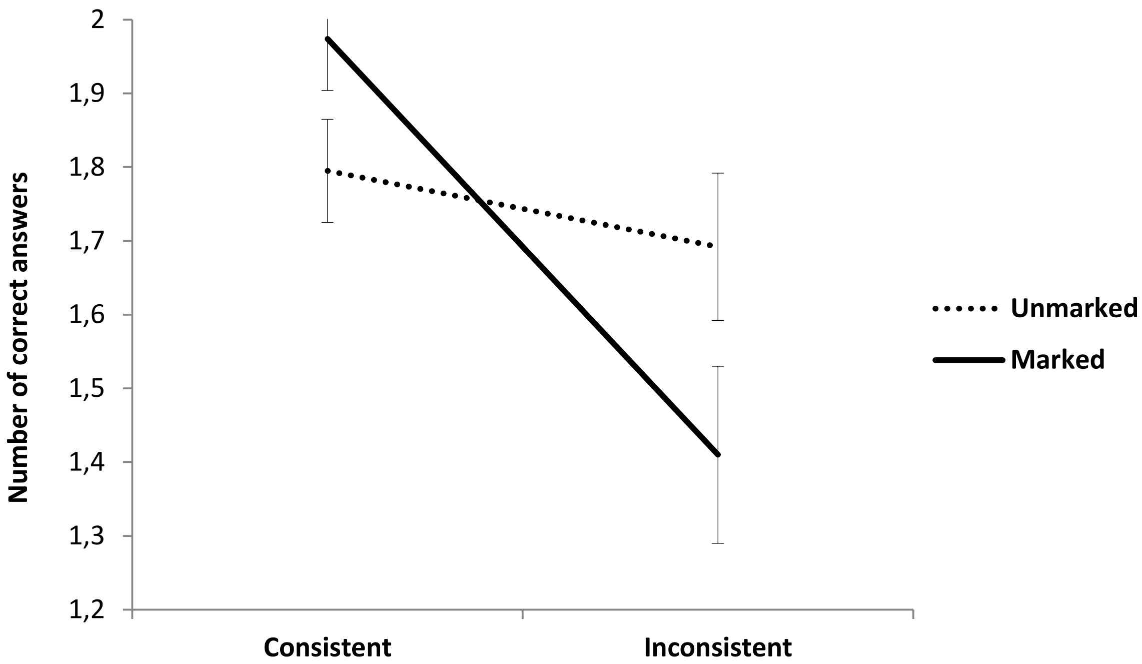


Figure 2. Successful word problem solvers: Interaction effect Consistency x Markedness x Group

Regarding the role of reading comprehension ability in word problem solving: overall successful word problem solvers (*M* = 46.42, *SD* = 2.66) scored significantly higher on the standardized reading comprehension test than less successful word problem solvers (*M* = 35.02, *SD* = 1.27), t (53.32) = 3.87, $p < .001$). To obtain more detailed insight into the role of reading comprehension skills in solving marked inconsistent word problems, reading comprehension ability was correlated with the difference scores (inconsistent - consistent) computed for the marked and unmarked word problem types.

In line with our expectations, the results of the correlational analyses show that only in the group of successful word problem solvers the difference score for the marked word problem type was significantly related to reading comprehension (Pearson's $R = -.40$, $p < .05$, $R^2 = .16$). That is, in the group of successful word problem solvers, a higher reading comprehension score was associated with a smaller difference score, which indicates that performance on marked word problems is higher for students who have higher reading comprehension abilities. This suggests that students with higher reading comprehension abilities appear to have a higher chance of overcoming problems with solving marked word problems.

Importantly, reading comprehension was not correlated with the successful word problem solvers' difference scores for unmarked word problems ($R = -.27$, $p = .10$). Furthermore, in the group of less successful word problem solvers, reading comprehension was also not correlated with the difference scores computed for either unmarked ($R = -.04$, $p = .76$) or marked word problems ($R = -.04$, $p = .83$).

Discussion

This study set out to investigate the claim that the contemporary RME approach pays limited attention to the teaching of semantic-linguistic skills during word problem solving instruction. We therefore designed a study in which we not only manipulated the extent to which a sophisticated representation strategy was required, but also varied the semantic complexity of the word problems by using a marked (i.e., high semantic complexity) or unmarked (i.e., low semantic complexity) relational term in the word problem text. Moreover, we classified students as successful and less successful word problem solvers on the basis of their performance on an independent and well-established RME-specific mathematics test.

Using this classification procedure, it was hypothesized that less successful word problem solvers would experience difficulties with correctly solving inconsistent word problems irrespective of their semantic complexity (Hypothesis 1). This hypothesis was confirmed by our analyses, which showed that less successful word problem solvers performed poorly on both marked and unmarked inconsistent word problems. Successful word problem solvers, on the other hand, were able to effectively solve inconsistent word problems that had a low semantic complexity. This finding suggests that the sophisticated representation skills required to solve non-obvious word problems are adequately learned in the RME curriculum, at least by successful word problem solvers.

However, on semantically complex word problems even the successful problem solvers experienced difficulties, as indicated by the large number of errors they made on marked inconsistent word problems (Hypothesis 2). More concretely, successful word problem solvers found it more difficult to translate a marked relational term ('less than') into an addition operation, than to translate an unmarked relational term ('more than') into a subtraction operation.

These findings once again support prior observations that (subtle) semantic-linguistic elements of a word problem, more specifically the marked relational term, influence word problem solving success (Clark, 1969; Kintsch, 1998; Lewis & Mayer, 1987; Pape, 2003; Van der Schoot et al., 2009). Moreover, they are in line with empirical work reporting processing problems with marked terms, which is suggested to be caused by the semantic representation of negative poles like 'less than' being more fixed and complex, and therefore less likely to be reversed, than that of positive poles like 'more than' (e.g., Lewis & Mayer, 1987; for a detailed explanation of the underlying mechanism, see e.g. Clark, 1969). For example, earlier studies have shown that students are less able to recall marked terms accurately in memory tasks (Clark & Card, 1969), have slower naming responses for marked terms in naming tasks (Schriefers, 1990), have slower solution times for problems with marked adjectives in reasoning problems (French, 1979), and experience problems with reversing a marked inconsistent word problem (e.g., Pape, 2003; Van der Schoot et al., 2009).

Importantly, our results showed that even successful students appear to be insufficiently equipped with the semantic-linguistic skills required to solve semantically complex word problems correctly. Given the current classification procedure, it is possible that students were simply not taught the necessary amount of semantic-linguistic skills during word problem solving instruction. This reinforces our premise that the development of semantic-linguistic skills receives little attention in contemporary RME instruction, thereby identifying an important aspect of current teaching practice of word problem solving in RME that could be reconsidered.

Building upon prior studies (e.g., Lee et al., 2004; van der Schoot et al., 2009), another aim of this study was to investigate whether reading comprehension ability could help (successful) word problem solvers to overcome the semantically complex marked relational term in an inconsistent word problem. In line with our expectations, reading comprehension was positively related to the performance on marked (but not unmarked) inconsistent word problems for the group of successful word problem solvers; whereas for the less successful group no significant relations were found between reading comprehension and word problem solving (Hypothesis 3).

These results provide corroborating evidence that general reading comprehension skills play an important role in students' ability to correctly solve semantically complex word problems. Moreover, our findings represent an advance over prior work by more specifically delineating which types of word problems and for which students reading comprehension ability might have an effect. This study shows that reading comprehension skills are especially helpful when it comes to improving the performance on semantically complex word problems by successful word problem solvers (as classified by the RME mathematics test). This suggests that despite having acquired limited semantic-linguistic skills during word problem solving instruction in the RME curriculum, (successful) students have the ability to rely on their reading comprehension skills to effectively solve semantically complex word problems.

In conclusion, the present study showed that students who performed well on word problems offered in RME, and therefore were characterized as successful word problem solvers, did not necessarily correctly solve word problems on an independent word problem test that contained problems that are semantically complex, and hence require both representational skills and semantic-linguistic skills. These findings suggest that word problem solving instruction in the RME curriculum is insufficient in the sense that little emphasis is placed on the explicit teaching of semantic-linguistic skills. This conclusion is particularly relevant for the educational practice of RME. The main implication is that word problem solving instruction should place greater emphasis on teaching the semantic-linguistic skills that enable students to process the semantic complexities that appear in the word problem statement adequately.

It is important to start developing such skills early in elementary school, as word problems get semantically more complex as students progress in their educational career, for example when making the transition from elementary to secondary education (Helwig et al., 1999; Silver & Cai, 1996). Making teachers in RME aware of the possible imbalance between the amount of instruction time being devoted to the teaching of strategic representation skills and semantic-linguistic skills, and encouraging them to pay more attention to semantic-linguistic skills, would provide a good starting point. Moreover, it is useful to make a distinction between learning to process more subtle semantic-linguistic text features (like a marked relation term) and dealing with more general semantic text complexities (like the relevance of the information in the word problem text, the explicitness of the described relations, and the sequence of the known elements in the word problem text).

These and other practical aspects of the results, such as finding the optimal balance between the amount of instruction in strategic representational and semantic-linguistic skills, remain to be addressed in future research. Presumably, currently effective intervention programs that focus on both strategic representational and semantic-linguistic skills, such as schema-based instruction (e.g., Jitendra, Star, Rodriguez, Lindell, & Someki, 2011; Jitendra, DiPipi, & Perron-Jones, 2002), and the *Solve It!* instruction method (Krawec, Huang, Montague, Kressler, & Melia de Alba, 2013; Montague, Warger, & Morgan, 2000), could provide a fruitful starting point in pursuing this challenge.