Running head: Humor and realistic word problem solving

Improving realistic word problem solving by using humor

Wim Van Dooren, Stephanie Lem, Hannelore De Wortelaer, and Lieven Verschaffel

Centre for Instructional Psychology and Technology, KU Leuven, Belgium

Mailing address:

Wim Van Dooren

Centre for Instructional Psychology and Technology

Dekenstraat 2 – PO Box 3773

3000 Leuven – Belgium

E-mail: wim.vandooren@kuleuven.be

Fax: +32 16 326274

Tel: +32 16 325755

Abstract

Research has shown that children have a strong tendency to exclude real world considerations when solving word problems. In this study, we investigated a novel way to try to change this tendency. We tested whether children would adapt their behavior when solving word problems in which realistic considerations are required (P-items) when these problems are embedded in a humoristic context as compared to when they are offered in a typical word problem solving context. 148 sixth graders solved four P-items in a humor condition versus a word problem condition. It was found that overall significantly more realistic reactions were given in the humor condition, and this was the case for three of the four problems.

Implications of these findings for further research and for classroom instruction are discussed.

Key words: mathematics education; word problems; humor; realistic problem solving; intervention

1. Introduction

An important goal of mathematics education in general and mathematical word problem solving in particular is that pupils learn to use mathematics in real life situations (Verschaffel, De Corte, & Lasure, 1994). However, research has shown that this is a very difficult goal to reach as pupils often approach mathematical word problems in a superficial way, without making the connection between real life and mathematics (Van Dooren, Verschaffel, Greer, & De Bock, 2006). Many researchers have tried to improve pupils' realistic word problem solving, with varying success. We tried to increase the number of realistic reactions to word problems in a group of 148 Flemish pupils in a novel way, by inserting humoristic elements in a test aimed at assessing pupils' realistic word problem solving. Humor has several assets that can help in improving learning in general (Ivy, 2013), and in stimulating realistic considerations in word problem solving in particular, as we will explain in the next sections.

1.1 Word problems in the mathematics classroom

Word problems are defined by Verschaffel, Greer, and De Corte (2000) as “verbal descriptions of problem situations wherein one or more questions are raised the answer to which can be obtained by the application of mathematical operations to numerical data available in the problem statement” (p. ix). They are typically used as a way to build a bridge between mathematics and the real world. Schoenfeld (1991) stresses that word problems can act as a vehicle to make the step from formal to applied mathematics. Verschaffel et al. (2000) call this the application function of word problems. Besides this application function, they describe several other functions of word problems in the mathematics classroom, such as: to motivate pupils, to evaluate their intelligence and mathematical abilities, to train pupils' creative and heuristic reasoning, and to develop new mathematical concepts and skills.

The application function hence is a major reason for including word problems in the elementary school mathematics curriculum. Pupils should develop the skill to know when and how to apply mathematics in everyday life. For a long time, word problems have played this function without much critical concern, but during the last two decades, it has been shown that the current school practices do not at all reach this goal. Apparently, as a result of their year-long participation in traditional mathematical word problem solving lessons, many pupils approach word problems in a superficial and artificial way. They just search for the mathematical operation(s) to perform with the given numbers, with little or no attention to the meaningfulness of their solution, and they merely report the numerical result of their calculation(s) as the answer to the word problem, without proper interpretation of the meaning of that result in relation to the problem situation (Lave, 1992; Reusser & Stebler, 1997; Schoenfeld, 1991; Verschaffel, Greer, Van Dooren, & Mukhopadhyay, 2009).

Word problems can be categorized in different ways. Galbraith and Stillman (2001) distinguish four types of word problems based on the way they integrate the real world into the word problem. First, “injudicious problems” do not include the real world into the word problem in any way. On the contrary: they violate real-world aspects. For instance, “One person can dig a hole in the ground in 6 hours. 10 persons working at the same rate can dig the hole in how much time?” (Bahmaei, 2011, p. 5). For this word problem it is possible to get to a numerical answer, but it is not very realistic to be digging one hole with ten persons at the same time and at the same speed. Second, in “context-separable problems”, the word problem is embedded in a real-world context, but it is not necessary to use this context to solve the problem. For instance, “There are 6 apples in one bag, how many apples do you get when you buy 4 bags?” Third, in “standard application problems” some simple realistic considerations need to be made due to the peculiarities of the real-world situation at stake and the outcome of the arithmetical operations cannot be communicated in a straightforward way. For instance, “An

army bus holds 36 soldiers. If 1128 soldiers are being bussed to their training site, how many buses are needed?” (Carpenter, Lindquist, Mathhews, & Silver, 1983, p.656). When only doing the calculation $1128/36$ one arrives at the solution “31.33 buses”, but by applying some real life knowledge, one knows that correct solution should be “32 buses”. Finally, in “genuine modelling problems” the solver has to use his/her own real-world knowledge about the problem at hand to build up the appropriate model of the problem situation and to make proper decisions about the mathematical operations to be performed. For instance, “How much does a family that eats on average 1.5 bread a day, spend on bread on a yearly basis?” One needs to know the number of days in a year and come up with an estimate of the average price of bread, to come up with a realistic solution.

1.2 Word problem solving with or without realistic considerations

Even though the ability to use real-life knowledge when solving word problems is part of the goals of mathematics education in many countries, many studies have shown that this goal is very difficult to reach. In one line of research where this is shown very convincingly. Researchers made a distinction between standard word problems (or S-items) and problematic word problems (or P-items). An example of an S-item is: “If there are 6 pizzas for 4 children at a party, how should they be shared out equally?”, which can easily be solved by dividing 6 by 4, coming to the perfectly possible answer of 1.5 pizzas per child. This item can be transformed into a P-item as follows: “If there are 6 balloons for 4 children at a party, how should they be shared out?” As it is not realistically possible to give 1.5 balloon to each child, the problem solver needs to take real life knowledge into consideration to come up with a meaningful solution. Greer (1993) constructed a set of eight pairs of mathematically parallel S- and P-items and looked at how 100 13-14-year olds solved them. An example P-item from this study is: “An athlete’s best time to run a mile is 4 minutes and 7 seconds. About how long would it take him

to run 3 miles?”, in which one needs to take into account that the athlete will get tired and cannot keep up the same pace for the complete three miles. In such a problem, it is impossible to give a precise numerical answer, but it would be acceptable to say that one cannot know how long it would take, or say that it will almost certainly take him longer than three times as long as for one mile, or give a numerical estimate in the form of an interval (e.g., “between 12 and 15 minutes”).

The general conclusion of Greer’s (1993) study was that the participants solved the word problems in a stereotypical way without taking into account realistic considerations, resulting in a very good performance on S-items and a very low performance on P-items. Verschaffel et al. (1994) conducted a similar study with 75 10-11-year-olds. A similar set of ten pairs of S- and P-items was constructed, and the scholars looked at the extent to which the participants took realistic considerations into account in their reactions to the P-problems. An example P-item in this study is: “Steve has bought 4 planks of 2.5m each. How many planks of 1m can he get out of these planks?” In this case one needs to be aware that pieces of 0.5m will be obtained in the process of cutting, so an adequate response would either be that 8 planks can be made, or that 10 planks can be made if one would manage to glue the pieces together. Like Greer (1993), Verschaffel et al. (1994) concluded that participants did not include real life knowledge when solving the P-problems. Studies using the same or similar items have been conducted in many countries, including Belgium, China, Germany, Hungary, Japan, Switzerland, The Netherlands, Turkey, and Venezuela (for an overview, see Greer et al., 2009). In all these studies, the overall percentage of realistic reactions was very low (between 5 and 20%).

1.3 Causes of unrealistic word problem solving

Verschaffel, De Corte, Van Vaerenbergh, Lasure, Bogaerts, and Ratinckx (1998) discuss several difficulties in pupils’ word problem solving processes, of which we will discuss

two that can specifically explain why pupils do not solve word problems in a realistic way. First, it is possible that pupils lack metacognitive competences such as planning the problem solving process, monitoring its progress, and evaluating its outcome. This leads to a superficial approach of solving word problems. Especially the phases of understanding the problem situation and of interpreting and checking the outcome of the computational work are often omitted by pupils. They quickly jump to calculations based on a superficial reading of the word problem, dedicate most attention to the actual calculations, and once the outcome of the calculations is found they immediately report it as their answer to the problem. A second difficulty can be found in the beliefs pupils have about mathematics and mathematics education in general and about word problems in particular. Beliefs that can hinder the realistic solving of word problems are for example the idea that every word problem has exactly one correct solution which moreover needs to be numerical in nature, and that all numeric information in a word problem should be used to find the solution. Also, pupils do not see that there is a connection between what they see in their mathematics classes and what happens in the real world.

The causes of these difficulties lie in several experiences that pupils have in the mathematics classroom, namely the nature of the word problems that are used, the way these problems are handled by the teacher, and the classroom culture that emerges as a consequence of it (Verschaffel, De Corte, Van Vaerenbergh, Lasure, Bogaerts, & Ratinckx, 1999). Word problems that are used in the classroom are often very stereotypical and unrealistic. As the goal of word problem solving is that pupils learn to see the connection between the real world and mathematics, this is an important problem. It makes pupils think that word problems have nothing to do with the real world (Verschaffel et al., 2000). On top of that, Cooper and Harries (2003) discuss the strong assessment culture, in which stereotypical tests are used in which word problems can be solved in a superficial and straightforward way by the simple application

of one or more calculations on the numbers given in the problem statement. Of course, it is easier to allow a wide range of possible answers and the inclusion of a variety of realistic considerations in interactions in a mathematics class than in a large-scale written assessment. But also in classroom interactions, teachers often do not use instruction methods that are powerful enough to teach pupils to use and the aforementioned metacognitive competences that are necessary for solving word problems correctly. Often, pupils are required to solve word problems by imitation, using a fixed scheme, and with little attention for the variety of approaches and of considerations that could be made (Verschaffel et al., 1999). It would be better to take the time to think about the word problem and use a problem solving scheme that includes reflection, as the routine application of a superficial solution scheme will not help pupils see the connection between word problems and reality (Brunner, Pauli, & Reusser, 2010). The classroom culture, finally, can also have an important impact on the extent to which pupils learn to take real life considerations into account when solving word problems. When pupils are always confronted with stereotypical word problems that can be solved in a superficial way, without relating them to real-life experience, pupils' beliefs will develop accordingly.

1.4 Improving realistic word problem solving

Researchers have made various attempts to increase the number of realistic considerations pupils make when solving word problems. A major distinction in these attempts can be made between classroom interventions of a longer duration (the effect of which is typically tested by a pretest posttest design) and experimental manipulations of the test setting itself. Various studies (e.g., Bonotto, 2003, 2006; English & Lesh, 2003; Verschaffel et al., 1998) have shown that classroom interventions that are spread out over a longer period can be successful in increasing the number of realistic reactions given by pupils. For example, Verschaffel et al.

(1998) constructed a learning environment that aimed at creating more active, strategic and motivated problem solvers. This intervention ran about 12 weeks and took 15 hours in total. There were positive effects on the development of strategies and problem solving skills of the participants. Also the beliefs and attitudes of the pupils were more positive after the intervention than before.

However, other lines of research (e.g., Defranco & Curcio, 2011; Dewolf, Van Dooren, Ev Cimen, & Verschaffel, 2014; Dewolf, Van Dooren, & Verschaffel, 2011; Krawitz, Schukajlow, & Van Dooren, 2016; Mellone, Verschaffel, & Van Dooren, 2017; Yoshida, Verschaffel, & De Corte, 1997) attempt to counter the unrealistic word problem solving tendency in the short term, by means of experimental manipulations in the test setting itself. For instance, Yoshida et al. (1997) added a warning to a word problem test, pointing out to pupils that complexities in the problem statement would make the word problems less straightforward or even impossible to solve. Despite this warning, these participants did not use more realistic considerations when solving the word problems than participants who did not receive this warning. Another example is that of Dewolf et al. (2014) who added representative illustrations to P-items in an attempt to increase the number of realistic reactions. The illustrations did not have the expected effect. Interviews with the participants showed that the participants focused on the text of the word problem and thought that the illustrations were not related to the word problem. These studies all involved an experimental manipulation of the test setting, but were still done within the context of the mathematics test. Other studies have explicitly manipulated the tendency to give unrealistic reactions by comparing performance on word problems given in a mathematical test versus similar problems given in another setting. Defranco and Curcio (1997), for example, first gave pupils the above-mentioned bus problem in which the correct solution had to be rounded up as one cannot order half buses. In the traditional test setting, only two out of twenty pupils were able to find the correct solution.

When the same twenty pupils were given a more authentic version of the same task, and were asked to actually call a bus company to order buses for a school party, 16 of them did order the correct number of buses. Another example is the study by Dewolf et al. (2011) where it was shown that when the same word problem, that contained various ethical dilemmas, was embedded in a test that was taken in a religion class, pupils involved more realistic considerations and acknowledged that a word problem could get a variety of answers, as compared to the same word problem embedded in a traditional mathematics test. These findings show that when word problems are offered in a setting outside the (traditional) mathematics classroom, pupils are more inclined to reason realistically.

Whereas the long-term interventions that were mentioned above are primarily set up with the goal to directly improve educational practice (by delivering effective teaching materials), the short-term studies wherein experimental manipulations are tested have primarily a theoretical goal: By testing if a specific manipulation leads to an increase in the number of realistic reactions to P-items, researchers gain a better understanding of pupils' thinking processes and of the reasons for the absence of realistic considerations in their word problem solving behavior. Hereafter, we describe a study that investigates the impact of a new experimental manipulation of the test setting itself, namely whether inserting humor in a word problem test increases the number of realistic answers.

1.5 Humor and learning

The literature on humor (e.g., Ivy, 2013; Nicewonder, 1994) points to many possible advantages of using humor in educational settings. It can reduce stress and tension, attract attention, motivate and change attitudes, increase self-confidence, improve the relation between pupils and teacher, and stimulate creativity. These effects can be helpful in solving word problems in a more realistic way. For instance, Gadanidis, Gadanidis, and Huang (2004) concluded that

humor can help to develop a positive attitude towards mathematics and mathematics education, which may further help pupils in seeing the connection between everyday life and word problems (Teslow, 1995).

Humor may also stimulate children to see the problem situation from a different perspective. This seems most relevant for the current study. If we want pupils to make realistic considerations when solving word problems, they need to see the situation described in the word problem from a different perspective than the one they are used to. This can help pupils to reason less superficially when solving a word problem, trying to see different sides of the word problem and use alternative solution strategies.

The literature proposes various theories of humor, including for example relief theory, superiority theory and incongruity theory (Buijzen & Valkenburg, 2002). These theories often are considered as complementary. Each theory tends to focus on specific kinds of situations, but in many cases more than one theory can explain the instantiation of humor. In the present study, we focus on incongruity theory (e.g., Attardo, 1997; Feyaerts, 2008), or more generally on the incongruity-resolution mechanism that is central in it. Incongruity resolution has also been incorporated in other humor theories such as the script-based semantic theory of humor (Raskin, 1985) and in the general theory of verbal humor (Raskin & Attardo, 1991). The incongruity relief mechanism is specifically relevant for our study as it focuses on the perception of the same situation from two different, seemingly incongruent, perspectives. Only one of these two interpretations of the situation is considered to be the more plausible one, and this one will occur in the listener/reader when the situation is presented. Humor originates when it becomes clear that the alternative, less plausible interpretation – the one the listener/reader did not think of initially – ultimately turns out to be true. The reaction to this unexpected experience of incongruity is one of laughter. For example, consider the following joke: The teacher asks: “When there are five flies on the table and I kill two of them, how many are still

on the table?” Calvin responds: “Only the two dead ones, the others were scared and flew away”. The answer that many readers of this joke expected is ‘three flies’, as $5-2 = 3$. The interpretation Calvin gave of the problem makes the reader look again at the word problem from a different, more realistic perspective. Reading such jokes could help pupils to look at word problems from a different perspective. We want to remark that what we have called a “joke” here is not necessarily experienced as humoristic by the reader. This is a highly subjective experience. In order to experience humor, it is critical that the reader shares the initial interpretation intended by the author of the joke, and not immediately thinks of the alternative interpretation yet. Otherwise, the joke will not evoke humor but merely involve a description of two possible interpretations of the same situation.

1.6 Rationale and research goals

The mechanism of incongruity leading to humorous experience as elaborated in the above theoretical framework seemed particularly useful for addressing the need for making realistic considerations when solving word problems. Pupils interpret the word problem solving situation in line with their previous experiences, in which typically there is no need to extensively think of the real world situation that is described in the word problem and whether peculiarities in that real world situation urge for a reconsideration of the outcomes of the arithmetical situations. In P-items, however, there is a need to take these peculiarities into account. This experience of an incongruity between two interpretations of what is required in solving word problems might be stimulated by using humor. For this study, we specifically looked for jokes presented in the form of word problems involving two incongruent interpretations with respect to the acceptability and desirability to include real-world considerations into their solution. This type of joke can be found in many children’s books, calendars, magazines, etcetera. Once more, we emphasize that these kinds of jokes will not

necessarily evoke laughter in every reader. The reader needs to adhere the initial interpretation that is intended to be evoked, and only later come to see that an alternative interpretation of the same situation was also possible. In this specific case, the selected jokes would only be considered funny by learners who initially interpret the word problem situation as one that is requiring no real-world considerations at all, and only later come to realize that real-world knowledge could be applied, leading to a very different solution which is seemingly incompatible with the classroom word problem solving setting.

Our expectation was that by surrounding the P-items by jokes that specifically addressed the incongruity between sticking to the rules and norms of the mathematics classroom (the more plausible interpretation) and using real-world considerations (the less plausible interpretation) children would be more inclined to reason realistically when solving the P-items.

2. Method

2.1 Participants

Participants were 148 6th grade pupils from four Flemish elementary schools which were all situated in small to medium-sized villages in the center of the country. Eighty-two of the pupils were boys. Pupils in 6th grade typically are between 11 and 12 years old.

2.2 Design and material

Participants were randomly assigned to one of two conditions: the word problem condition ($n = 75$), or the humor condition ($n = 73$). Assignment to conditions happened within classrooms in order to obtain two groups which were comparable in their word problem solving ability. In both conditions, pupils solved four P-items adapted from Verschaffel et al. (1994). These four P-items, which can be found in Table 1, were offered in exactly the same way in both conditions. Figure 1 illustrates how this was done for one P-item, i.e. the swimming item. The

adaptation we made to the P-items was that we embedded them in the context of Calvin and Hobbes, two cartoon characters, as these were also used for the jokes in the humor condition, as explained below.

Name	Formulation	Example of NR and RR
Item 1 Rope	Calvin and Hobbes play badminton in the garden. As a net, they use a rope that is stretched between two poles that are 12 m apart. They only have pieces of rope that are 1.5 m long. How many pieces to they have to tie together to stretch a rope between the poles?	NR: <i>"$12 / 1.5 = 8$ pieces of rope"</i> RR: <i>"More than 8 pieces because of the knots" / "You cannot know but more than 8"</i>
Item 2 School	Calvin and Inge attend the same school. Calvin lives at a distance of 17 km from school, and Inge lives at a distance of 21 km. How far do they live from each other?	NR: <i>"$21 - 17 = 4$ km" / "$21 + 17 = 38$ km"</i> RR: <i>"It is not clear where they live" / "Between 4 and 38 km"</i>
Item 3 Swimming	When Calvin goes to school, Hobbes sometimes takes a swim. His best time to swim 25 m is 20 seconds. How long does it take Hobbes to swim 500m?	NR: <i>"$20 * 20 = 400$ seconds"</i> RR: <i>"Probably more than 400 seconds" / "You cannot know as he will get tired" / "He can't keep that pace"</i>
Item 4 Christmas	It is almost Christmas. Each year, Calvin's mom buys a large Christmas tree. The man who is selling the Christmas trees told Calvin that he sold 243 Christmas trees in December. How many do you think he will sell in January, February and March altogether?	NR: <i>"$243 * 3 = 729$ trees"</i> RR: <i>"Probably less than 729" / "Impossible to tell" / "Not many!"</i>

Table 1: The P-items with examples of non-realistic (NR) and realistic (RR) reactions (items adapted from Verschaffel et al., 1994).

Figure 1. Illustration of the presentation of P-items (c.q. the swimming item)

The four sheets containing the P-items were embedded in a booklet. However, two different versions of the booklet were created, one for each condition. In the humor condition, the P-items were alternated with pages containing jokes; in the word problem condition the P-items were alternated with S-items. We will now discuss both in more detail.

In the humor condition, each joke page contained two jokes addressing the central incongruity, i.e. whether word problems need to be interpreted by sticking to the rules and norms of the mathematics classroom or whether common sense reasoning is adequate. Pupils were asked to indicate which joke they found most funny and why. As such, the answer given by pupils was irrelevant to our study. The assignment for pupils to select the best joke was simply given to ensure that they would properly process the jokes instead of skipping these pages. An example joke page is presented in Figure 2.

Figure 2. Exemplary joke page included in the Humor Condition test booklet

In the word problem condition, the joke pages were replaced by pages containing two S-items. Figure 3 presents an example. On these pages pupils were asked to solve two typical word problems that were found in a fifth grade mathematics textbook, in order to put these pupils more into a traditional mathematics classroom setting. To ensure comparability with the Humor condition and with the P-items in the booklet, these word problems were also adapted to contain the Calvin and Hobbes characters. These S-items required different calculations than in the P-items, and in none of them the obtained solution needed to be modified or interpreted differently taking into account the real world situation that was described in the word problem. We chose word problems from a fifth grade textbook as they would not be too difficult for the participants, making them focus more on the P-items.

Each booklet version was made in four different orders of the items, in order to avoid order effects. However, each time, the booklet either started with a page with two jokes or a page with two S-items, after which the first P-item was provided.

Figure 3. Exemplary worksheet included in the Word Problem Condition test booklet

2.3 Procedure

At the start of the experiment, each class group was randomly divided in two groups, by alphabetically ordering the pupils by their name and assigning the first pupil to the Word Problem Condition, the second to the Humor Condition, and so on. They then went to two separate rooms to take the test. In one room, the word problem booklet was handed out; in the other room the humor booklet was handed out. The same instructions were given in both conditions. Pupils were told that the researchers came to test some new teaching materials, so pupils were asked to fill out the booklets. In the instructions, no specific reference to mathematics was made. Neutral answers were given when pupils asked questions during the test. Although 30 minutes were reserved to fill out the booklets, participants were not informed about this as we did not want them to work under time pressure. The available time was sufficient for all pupils. A calculator was allowed during the test.

The experiment took place during school time but not during the time usually spent to a mathematics lesson. After the experiment, the pupils collectively got a debriefing in which the aim of the study and the hypotheses were clarified, and pupils received an explanation about the kind of realistic considerations that could be made for each of the P-items.

2.4 Coding and analysis

In line with earlier studies using P-items (e.g., Dewolf et al., 2014; Verschaffel et al., 1994), pupils' reactions were coded as realistic (RR) whenever they gave some indication of making realistic considerations in their solution, either in their calculations or in their comments (e.g., by taking the realistic aspects into account in their calculations, by indicating that there is no correct numerical answer, or that the word problem is somewhat strange). When none of these elements were present in the answer and only straightforward arithmetical calculations on the given numbers were conducted, reactions were coded as non-realistic (NR). Examples of RR and NR for each P-item are given in Table 1. One author coded all responses and to check the inter-rater reliability, another author coded 20% of all responses, which resulted in a 100% agreement.

The dependent variable of interest in all analyses was dichotomous: realistic reaction (1) or non-realistic reaction (0). As all participants answered multiple items, we chose to use logistic regression using the generalized estimating equations approach in SPSS to account for the correlation between responses of the same participants (Liang & Zeger, 1986).

3. Results

Table 3 shows the percentages of realistic reactions per item and condition. We found a main effect of condition on the number of realistic reactions, $\chi^2(1, N = 148) = 6.903, p = .009$, with more realistic reactions in the humor condition (28.08%) than in the word problem condition

(16.00%). On top of that, we found a main effect of item, $\chi^2(3, N = 148) = 83.989, p < .001$, with percentages of realistic reactions varying from 7.44% to 54.60%. Pair-wise comparisons showed that all items had statistically significant different rates of realistic reactions, except from the rope and the school item that had the exact same rate of realistic reactions.

Condition	Rope item	School item	Swimming item	Christmas item	Total
Humor ($n = 73$)	6.85	13.70	28.77	63.02	28.08
Word problem ($n = 75$)	7.99	1.33	9.33	45.33	16.00
Total	7.44	7.44	18.93	54.60	21.97

Table 3. Percentage of realistic reactions to the four P-items, by condition

We also found an interaction effect of item and condition, $\chi^2(3, N = 148) = 11.504, p = .009$. Pair-wise comparisons were performed to see what this interaction effect means. For the rope item, there was no significant difference between both conditions. For the other three items, on the other hand, the humor condition yielded significantly more realistic reaction than the word problem condition. Odds ratios of this effect ranged between 1.954 (Christmas item) and 11.587 (School item).

A closer look at the realistic reactions that were given to the P-items helped to reveal the nature of the impact of the humor condition. As explained above, a pupils' reaction could be coded as realistic for several reasons: The realistic nature of the reaction could lie in the kind of calculations that were conducted or in the refusal to give a precise, numerical response, and the provision of an explanation why it was impossible to do so. A qualitative exploration of the answers given by the pupils indicated that the higher number of realistic reactions in the Humor Condition was mainly due to the latter reason: Pupils typically refused to give a numerical

answer to a P-item and briefly stated, for instance, that it was impossible or meaningless to perform the calculation suggested in the problem and/or to give an answer, often accompanied by an explanation that referred to the specific realistic complication involved in the word problem. Examples of such answers are: “Additional rope is needed to make knots” (Rope item), “A tricky question” (School item), “Depends how fast he swims, this was a record!” (Swimming item), and “Impossible to know. If he is lucky he may still sell a few extra trees” (Christmas item).

4. Conclusion and discussion

Many researchers have attempted to enhance pupils’ tendency to make real-life considerations when solving word problems (P-items; e.g., Dewolf et al., 2014; Yoshida et al., 1997). Typically, short-term interventions that were restricted to manipulating the test materials or to changing the test setting had only small or even no effects, while long-term interventions that taught about problem solving during mathematics classes, were more successful. In the current study, we found that the use of humor by alternating the P-items with jokes was a successful strategy in eliciting considerably more realistic reactions in pupils. This finding indicates that pupils’ tendency to provide unrealistic reactions to word problems indeed can be attributed at least partly to their expectations regarding what needs to be done when solving word problems at school. Pupils are able to make the required realistic considerations, and they can be made aware of the tension between the (implicit) expectations that hold in a traditional classroom word problem solving setting versus the real world. In a typical scholastic setting most often they do not show this ability while they do seem to possess it. These results were obtained by a rather simple manipulation of the test material itself, while previous studies doing manipulations of the test materials typically found no or only small effects. So, the finding that humor may lead to a substantial increase in realistic reactions is not only theoretically

significant, as it helps us to understand the phenomenon as such. It also has significance for educational practice. It may be valuable for teachers to try to incorporate humor in their mathematics classes in order to make it easier for their pupils to look at word problems from an alternative perspective, and change pupils' beliefs and attitudes in the longer run. However, there are still several issues that should be addressed in further research in order to gain a deeper understanding of the role of humor and to better exploit its impact on educational practice.

First, the impact of humor was different for the various P-items, and for the rope item there was even no effect at all. The present study does not allow to hypothesize why embedding the P-items in a humoristic setting seems to have a differential effect on different types of P-items. Further research could shed a light on this, for instance by asking children to think aloud while solving the problems and by probing the possible realistic considerations they make during the different stages of the solution process. On top of that, the addition of humor could be tested on a variety of items, taking into account different item characteristics. For instance, items could be categorized using a categorisation scheme for word problems as the one used by Galbraith and Stillman (2001).

Second, this study investigated the momentaneous effect of humor on the number of realistic reactions on these specific word problems, and did not look at effects in the long run when pupils solve other word problems. Further research could investigate whether pupils would continue to give more realistic reactions in a regular school setting after having been confronted with mathematical word problems involving realistic considerations. Further, it is worth investigating whether children's attitudes towards solving word problems would change as a consequence. It also seems valuable to investigate whether the inclusion of more elaborated materials involving humor in the classroom practice over a longer time period, and involving activities such as classroom discussions could lead to larger and longer lasting effects than observed in the current study.

Third, it would be interesting to look at the possible differential effects of humor in different schools and classes. Class culture, didactics and characteristics of the items used in the classroom can have an impact on pupils' beliefs about word problems, influencing the amount of realistic considerations they use when solving them. It is hence possible that humor has a larger effect in some schools or classes than in others.

Finally, the current study focused on only one specific mechanism of humor that may be beneficial for stimulating children's word problem solving, namely the incongruity of two interpretations of a situation. Various other potentially beneficial effects are suggested in the literature, such as motivational effects. These effects could be studied in further research. And as suggested by Weber (2016), humor can also be used to better understand mathematical reasoning as such.

References

- Attardo, S. (1997). The semantic foundations of cognitive theories of humor. *International Journal of Humor Research*, 10, 395-420.
- Attardo, S., & Raskin, V. (1991). Script theory revis(it)ed: Joke similarity and joke representation model. *Humor*, 4, 293–347.
- Bahmaei, F. (2011) Mathematical modeling in primary school, advantages and challenges. *Journal of Mathematical and Application*, 1, 3-13.
- Bonotto, C. (2003). Investigating the mathematics incorporated in the real world as a starting point for mathematics classroom activities. In N. A. Pateman, B. J. Dougherty, & J. Zilliox (Eds.), *Proceedings of the 27th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp.129-136). Honolulu, HI: Hawaii University.

- Bonotto, C. (2006). Extending students' understanding of decimal numbers via realistic mathematical modeling and problem posing. In J. Novotná, H. Moraová, M. Krátká, & N. Stehlíková (Eds.), *Proceedings of the 30th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 193-200). Prague, Czech Republic: Prague Charles University.
- Brunner, E., Pauli, C., & Reusser, K. (2010). Understanding-oriented mathematics instruction using the example of solving a word problem. *Journal für Mathematik und Didaktik*, 31, 31-50.
- Buijzen, M., & Valkenburg, P. M. (2004). Developing a typology of humor in audiovisual media. *Media Psychology*, 6, 147-167.
- Carpenter, T. P., Lindquist, M. M., Matthews, W., & Silver, E. A. (1983). Results of the third NAEP mathematics assessment: Secondary school. *Mathematics Teacher*, 76, 652-659.
- Cooper, B., & Harries, T. (2003). Children's use of realistic considerations in problem solving: some English evidence. *Journal of Mathematical Behavior*, 22, 451-465.
- DeFranco, T. C., & Curcio, F. R. (1997). A division problem with a remainder embedded across two contexts: Children's solutions in restrictive vs. real-world setting. *Focus on Learning Problems in Mathematics*, 19, 58-72.
- Dewolf, T., Van Dooren, W., Ev Cimen, E., & Verschaffel, L. (2014). The impact of illustrations and warnings on solving mathematical word problems realistically. *Journal of Experimental Education*, 82, 103-120.
- Dewolf, T., Van Dooren, W., & Verschaffel, L. (2011). Upper elementary school children's understanding and solution of a quantitative problem inside and outside the mathematics class. *Learning and Instruction*, 21, 770-780.

- English, L., & Lesh, R. (2003). Ends-in-view problems. In: R. Lesh & H. Doerr (Eds.), *Beyond constructivism. Models and modeling perspectives on mathematical problem solving, learning and teaching* (pp. 297-316). Mahwah, NJ: Erlbaum.
- Feyaerts, K. (2008). Taalkundig humoronderzoek: Tradities en tendensen [Linguistic research on humor: traditions and tendencies]. In B. Raymaekers (Ed.), *Denken en weten over de wereld* [Thinking and knowing about the world] (pp. 159-190). Leuven, Belgium: UP.
- Gadanidis, G., Gadanidis, J. M., & Huang, A. Y. (2004). Using humor to gain mathematical insight. *Mathematics Teaching in the Middle School*, 10, 244-250.
- Galbraith, P., & Stillman, G. (2001). Assumptions and context. Pursuing their role in modelling activity. In: J. F. Matos, W. Blum, S. K. Houston, & S. P. Carreira (Red.), *Modelling and mathematics education. ICTMA9: Applications in science and technology* (pp. 300-310). Chichester, U.K.: Horwood.
- Greer, B. (1993). The mathematical modeling perspective on wor(l)d problems. *The Journal of Mathematical Behavior*, 12, 239-250.
- Greer B., Verschaffel L., Van Dooren W., & Mukhopadhyay S. (2009). Making sense of word problems: past, present, and future. In: Verschaffel L., Greer B., Van Dooren W., & Mukhopadhyay S. (Eds.), *Words and Worlds. Modelling verbal descriptions of situations*, (pp. xi-xxvii). Rotterdam/Boston/Taipei: Sense Publishers.
- Ivy, L. L. (2013). Using humor in the classroom. *The Education Digest*, 79, 54-57.
- Krawitz J., Schukajlow S., & Van Dooren W. (2016). Effects of short-term practicing on realistic responses to missing data problems. In C. Csikos, A. Rausch, A., and J. Sztányi (Eds.), *Proceedings of the 40th Conference of the International Group for the Psychology of Mathematics Education* (Vol 3, pp. 131-138). Szeged, Hungary: PME.

- Lave, J. (1992). Word problems: A microcosm of theories of learning. P. Light & G. Butterworth (Eds.), *Context and cognition: Ways of learning and knowing* (pp. 74-92). New York: Harvester Wheatsheaf.
- Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73, 13-22.
- Mellone M., Verschaffel L., & Van Dooren W. (2017). The effect of rewording and dyadic interaction on realistic reasoning in solving word problems. *The Journal of Mathematical Behavior*, 46, 1-12.
- Nicewonder, C. (1994). Humor in the mathematics classroom? In J. L. Higbee, & P. L. Dwinell (Ed.), *Proceedings of the 18th Annual Conference of the National Association of Developmental Education* (pp. 31-32). Kansas City, Missouri: The University of Georgia.
- Raskin, V. (1985). *Semantic Mechanisms of Humor*. Dordrecht, The Netherlands: Reidel.
- 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.
[http://dx.doi.org/10.1016/S0959-4752\(97\)00014-5](http://dx.doi.org/10.1016/S0959-4752(97)00014-5)
- Schoenfeld, A. H. (1991). On mathematics as sense-making: An informal attack on the unfortunate divorce of formal and informal mathematics. In J. F. Voss, D. N. Perkins & J. W. Segal (Eds.), *Informal reasoning and education* (pp. 311-343). Hillsdale, NJ: Erlbaum.
- Teslow, J. L. (1995). Humor me: A call for research. *Educational Technology Research and Development*, 43, 6-28.
- Van Dooren, W., Verschaffel, L., Greer, B., & De Bock, D. (2006). Modelling for life: Developing adaptive expertise in mathematical modelling from an early age. In L. Verschaffel, F. Dochy, M. Boekaerts, & S. Vosniadou (Red.), *Instructional Psychology*:

Past, Present, and Future trends: Sixteen essays in honour of Erik De Corte (pp. 91-109).
Oxford, U.K.: Elsevier.

Verschaffel, L., De Corte, E., & Lasure, S. (1994). Realistic considerations in mathematical modeling of school arithmetic word problems. *Learning and Instruction*, 4, 273-294.

Verschaffel, L., De Corte, E., Van Vaerenbergh, G., Lasure, S., Bogaerts, H., & Ratinckx, E. (1998). *Leren oplossen van wiskundige contextproblemen in de bovenbouw van de basisschool* [Learning to solve mathematical context problems in elementary school]. Leuven, Belgium: Universitaire Pers Leuven.

Verschaffel, L., De Corte, E., Van Vaerenbergh, G., Lasure, S., Bogaerts, H., & Ratinckx, E. (1999). Learning to solve mathematical application problems: A design experiment with fifth graders. *Mathematical Thinking and Learning*, 1, 195-229.

Verschaffel, L., Greer, B., & De Corte, E. (2000). *Making sense of word problems*. Lisse, Nederland: Swets & Zeitlinger.

Verschaffel, L., Greer, B., Van Dooren, W., & Mukhopadhyay, S. (2009). *Words and worlds: Modelling verbal descriptions of situations*. Rotterdam, The Netherlands: Sense.

Weber, K. (2016). Mathematical humor: Jokes that Reveal how we think about mathematics and why we enjoy it. *The Mathematical Intelligencer*, 38(4), 56-61.

Yoshida, H., Verschaffel, L., & De Corte, E. (1997). Realistic considerations in solving problematic word problems: Do Japanese and Belgian children have the same difficulties? *Learning and Instruction*, 7, 329-338.