International Journal of Instruction *July 2022 ● Vol.15, No.3*
e-ISSN: 1308-1470 ● www.e-iji.net *p-ISSN: 1694-609X*
pp. 1079-1098

Article submission code: 20210902063208
Received: 02/09/2021
Revision: 27/03/2022
Accepted: 19/04/2022
OnlineFirst: 24/06/2022

Incorporating Humor into Mathematical Word Problems: Is there a Negative Effect on Students' Performance?

Csaba Csíkos
University of Szeged, Hungary, *csikoscs@edpsy.u-szeged.hu*

Fanni Biró
University of Szeged, Hungary, *fanni.biro@gmail.com*

Judit Szitányi
ELTE Eötvös Loránd University, Hungary, *szitanyi.judit@tok.elte.hu*

In this research we aim to reproduce and extend the positive results about the effect of humor in solving mathematical word problems. We examined whether solving humorous word problems has some effect on either word problems demanding realistic consideration or word problems of the well-known routine type. Fifth- and sixth-grade students (N=1,153) solved different types of tasks including humorous, routine and "problematic" word problems, answered further questions concerning their attitude towards mathematics. Our results suggest that while sixth-graders outperformed their fifth-grade peers, the positive effect of humorous word problems proved to be even more relevant than the school grade. However, there has been some possible negative effect of humor when solving routine word problems. Furthermore, students expressed a positive attitude toward humorous word problems by selecting these as their favorite tasks compared to other types of word problems. The results obtained provide further evidence on the desirability and feasibility of introducing (more) humor in mathematics education.

Keywords: mathematical word problems, humor, attitude, lower secondary school, upper primary school

INTRODUCTION

In this research we aim to reveal the possible positive effects of introducing humor in mathematics education by means of analyzing students' answers to both routine and non-routine word problems which were administered to them accompanied by humorous tasks looking also mathematical. In general, we aim to reproduce and extend the positive results published by our Leuven colleagues (Van Dooren et al., 2019), and examine whether there may be a possible negative effect of humor on routine word problems. First, the issue of 'problematic' word problems is acknowledged, then we make a short

Citation: Csíkos, C., Biró, F., & Szitányi, J. (2022). Incorporating humor into mathematical word problems: Is there a negative effect on students' performance? *International Journal of Instruction*, *15*(3), 1079-1098. https://doi.org/10.29333/iji.2022.15357a

cultural-historical outline of the role humor has or may have played in the school, and the currently available evidence on the effects of humor will be addressed.

Review of Literature

"Problematic" Realistic Word Problems

Ample evidence has been gathered documenting the phenomenon that students are inclined to use superficial algorithms when solving mathematical word problems. In the case of simple arithmetic (or arithmetic-like) word problems, one such superficial strategy that is frequently used and has additionally been widely observed and documented leads to students focusing merely on figures in the text and connecting them to an arithmetical operation to be executed (Verschaffel et al., 2000). It cannot be denied that this strategy works well in many cases, especially when students easily recognize that the arithmetic operations currently being studied had been presented as a word problem.

Several different approaches have been used to meet the challenge of students' resistance to changing their already well-functioning superficial word problem solving strategy. Among the intervention programs aimed to change and improve students' word problem-solving strategies, those that imply the complex transformation of the classroom atmosphere, learning materials and instructional methods (e.g., Verschaffel et al., 1999) have not yet explicitly involved the phenomenon of integrating humor into mathematics classroom activities.

Humor and (mathematics) learning

There is extensive and long-standing literature on different aspects of humor. For the purposes of the current study, some elements of the philosophical and psychological theories are to be reviewed. More than a century ago, Kline (1909), and Hall and Allin (1897) called forth for scientific empirical data collection pondering towards a sound theory of humor, and they labelled the various earlier definitions misleading or inadequate verbal descriptions grabbing only partial aspects of humor. As is the case with many-many scientific concepts, the roots go back to Aristotle.

Aristotle's thoughts on humor and comedy (and the possible lack of it) became particularly well-known due to Eco's fictitious novel ("The name of the rose"), and why the existence and disappearance of a book from Aristotle on humor could become the main conflict source in the novel is brightly justified by Verene (2017, p. 156): "Truth is the grim face that forbids laughter". Furthermore, Verene cites Eco himself about the plausibility that Aristotle intentionally lost his text on humor (the second part of Poetics), and the reason may have been Aristotle's own incapability to be lucid enough to write that essay. Eco's famous novel preserved the dual spirit in which Christianity, Europe's most widespread religion, "has been famously unable … to speak in one voice on the topic" of laughter (Burde 2010, p. 216). According to Burde, Christianity's ambivalence can be rooted in Aristotle's ("The Philosopher's") different treatments of humor in his three different works; one of which, Poetics, considered laughter as low and even uncouth, assuming that comedy represents men as worse than in actual life.

During the last centuries and also nowadays (see Laineste 2002), humor has often been thought of as anti-authoritarian. From topical jokes (reflecting actual, often political news) to jokes focusing on various social groups or classes, there is a contrast between serious and dramatic issues and the way people are able to handle those contents cheerfully. Consequently, any institutions or social systems that have either incontrovertibly or not necessarily a status based on some level of authority might be inclined to refuse the introduction of humor at all levels and interactions within the system.

Avoiding a deeper analysis of how educational institutions inherited actions, values and even space arrangements from the churches, the ambiguity concerning the status of humor in the church does explain why humor had generally a bad reputation in education (Morreall, 2014) However, within the educational system there is surely plenty of potential for incorporating humor into classroom activities; besides mathematics, other subjects also present opportunities for a more humorous approach to problem solving activities (Zsolnai and Zsolnai, 1987). Zsolnai and Zsolnai argue that utilizing humor demands the establishment of a tolerant classroom atmosphere, and incorporating funny and light-hearted content is possible within different school subjects, such as in Ziyaeemer and Kumar's (2014) research on English lessons.

It seems that there are no really serious objections against the introduction of (more) humor into the complex system of interactions in education. However, as it could be inferred from the historical-cultural considerations, humor is by nature a complex phenomenon, and several aspects should be taken into account when considering its introduction into the educational system. This complexity of research might explain why educational research on the effects of humor is rare (see Egmir et al., 2017).

For sake of simplicity, we chose a definition provided by a recent handbook on the psychology of humor (Martin and Ford 2018, p. 3). Humor is a broad, multifaceted term that represents anything that people say or do what others perceive as funny and tends to make them laugh, as well as the mental processes that go into both creating and perceiving such an amusing stimulus, and also the emotional response of mirth involved in the enjoyment of it. According to this definition, humor is of inherently societal nature. Those many aspects which should be taken into account are the type of humor, the age of students, the timing and proportion of humor in the system, and – above all – the socio-cultural environment. As for the types of humor, scientists went much further than the branches of theories proposed by Hall and Allin (1897), and Kline (1909). Martin and Ford (2018) summarize the main aspects of contemporary humor theories as follows. First, the person must be in an appropriate mood and arousal state in order to experience humor. Second, the person must possess the cognitive prerequisites necessary to detect and resolve incongruities both in verbal and non-verbal communication. Third, the person must perceive the humorous situation as both violation (in the sense of threatening beliefs about the world) and benign.

The psychological models that describe the possible mechanisms for incorporating humor in education are summarized in studies by Van Dooren et al. (2019) and Weber (2016). For the most part, false beliefs or incongruence may be viewed as supplying a rich source of humor in mathematical word problems, such as when something (usually unexpectedly) violates our regular mental patterns or conclusions. Indeed, when solving insight problems or realistic word problems, it is often useful to let our thoughts construct and reconstruct different problem fields and mental representations.

In mathematics education, it is the genre of word problems that gives space for humor for as long as at least several thousands of years. The very first tasks that went beyond the pure practice of arithmetic or the immediate usefulness of a problem (e.g., tasks that aimed to provide entertainment for the learners, see Pickover 2009) may have evoked some sense of humor in the learners. The first well-known example of applying humor blended with a mathematical word problem is Flaubert's famous "how old is the captain" task in which there are lots of seemingly meaningful quantities provided about a steamship carrying grain, and at the end the unexpected question comes about the age of the captain (Verschaffel et al., 2000).

The cultural-historical overview provided in this section pictured a not so favorable situation for introducing humor into mathematics education. Nonetheless, the ice is already broken, and possibly new empirical studies can contribute to the growing evidence on the positive effects of introducing humor in mathematics education.

The Possible Effects of Incorporating Humor into Word Problem Solving

The positive relationship between humor and academic performance in general has been documented in several investigations in different age-groups throughout the educational system. Studies with university students (Berk and Nanda, 2006; Ford et al., 2012) examined the effect of humor on mathematics performance through anxiety, i.e., lowering the anxiety level by means of humorous instructions or tasks increased mathematical performance. According to Bieg et al. (2017, p. 31) "emotions do mediate between humor and performance". They found that among upper secondary students, teachers' humor related to mathematics content dispels boredom and increases the enjoyment of the lesson. As for lower secondary students (aged 10-14), Masten (1986) used different measures of humor, and found positive correlations with both IQ and academic competence. According to Batchelor et al. (2019) there are still many unresolved or controversial issues in the complex relationships of mathematics performance and the affective constructs of young children, therefore the place for fine-grained investigations is open. In Van Dooren et al.'s (2019) pioneering empirical study, besides the emotional aspects and benefits of incorporating humor, they conveyed the idea o possible cognitive effects from humor. By means of encouraging students to see problem situations from a different perspective, their metacognitive components needed for word problems solving may be activated and adjusted.

A Pioneering Study

Focusing on the relationship between humor and mathematics word problem solving, Van Dooren and his collaborators (Van Dooren et al., 2017; Van Dooren et al., 2019) revealed the positive effect humorous word problems have on sixth-grade students' solutions to realistic word problems. They found that when inserting humorous cartoons containing word problem-like jokes, the rate of students' realistic responses increased.

In their study, they administered both “problematic” word problems (tasks that require realistic consideration to be solved) and standard, routine word problems as well. This inspirational investigation opened the possibility to study three variables in a new research design; namely, short-term changes in students’ attitudes, possible negative effects of humor, and the issue of age limit for introducing humor in arithmetic word problems.

Humor and Students’ Attitudes

Short-term change in students’ attitudes towards mathematics in general or word problems specifically is an option worth exploring. According to Gresham and Burleigh (2019), it is not mathematics from where mathematics anxiety comes but the way it is presented to children. Eliminating anxiety by means of introducing humor may and should have a positive impact on students’ attitudes. Changes in the characteristics of the affective constructs of human personality occur usually slowly, and the affective characteristics are in general modelled as mainly preconditions or antecedents for changes in the cognitive sphere and not vice versa (see, for example, Reyes 1984). Nevertheless, we aimed to gather data on different aspects of attitude towards mathematics and mathematics learning, since the introduction of humor into mathematics education certainly require further considerations about the possible changes in students’ attitudes towards mathematics in general, and towards word problems in special. In an educational experiment lasting for eight weeks, there were only slight changes in 3rd grade students’ affect towards mathematics (Csíkos et al., 2012).

Possible Negative Effects of Humor

Several studies pointed out the possible negative effects of using humor in school settings. We now dispense the analysis of the effects of negative humor, i.e., when humor is offensive or malign. According to Banas et al. (2011), while the forms of negative humor have been intensively studied in educational settings, and, of course, they should be avoided, the possible negative effects of using the otherwise positive humor are understudied. Such negative consequence might be the overuse of humor. According to McMorris et al. (1985), there are individual differences in how incorporating humor in academic achievement tests is perceived.

The question of whether students’ performance during standard routine tasks is affected by humor in either a positive or negative way is debatable. When citing Freudenthal’s thoughts regarding a routine-like word problem, Greer (1997) emphasized that a seemingly standard word problem can be modeled in different ways. It cannot be denied that another possible effect of using humor in conditions when students are expected to apply a straightforward algorithm to solving a word problem might even be counterproductive. This debate leads to the examination of whether there is a place for using humor in a given classroom or culture within the didactic contract (see Brousseau et al. 2014) for solving mathematical word problems. The aim of the current investigation is to replicate results regarding the positive effect humor may have on word problem-solving, while further studying the advantages or disadvantages of using humor in mathematics classrooms. Our study is not an exact but a so-called conceptual replication of the study by Van Dooren et al. (2019) in the sense that it aims at “probing the phenomena in question to understand those results more deeply” (Cai et al. 2018, p. 4).

Possible Age Limitations

Finally, the age groups in which the connection between humor and mathematics performance can or should be studied in an age range from university students to elementary pupils. Since there seems to be no upper age limit in investigating the connection, the issue of the earliest possible incorporation arose. According to various studies on the development of the sense of humor, there is a remarkable milestone in understanding pictorial incongruity in cartoons, i.e., around the age of 6 to 8, children become capable not only perceiving but resolving incongruity in cartoon humor. Many researchers relied on the Piagetian sense of developmental stages in the development of the sense of humor thus focusing and emphasizing the cognitive structures necessary for understanding incongruities (see Semrud-Clikeman and Glass 2008). Dowling (2014, p. 122) found the sixth grade of schooling as the age when “abstract thinking and incongruities are better understood”.

Hypotheses

In compliance with our review of the literature and the research design we have developed, the following hypotheses are to be tested.

- Humorous word problems have a positive effect on the rate of realistic answers, while they have no negative effect on solving routine word problems.

- Sixth-grade students’ behavior will be similar to that of fifth-graders with the possible exception that more appropriate arithmetical answers will be yielded for routine word problems.

- Students tend to express their positive attitude towards humorous word problems, while they generally have a negative attitude towards word problem solving.

METHOD

Participants

A total of 1,153 students (547 fifth-graders and 606 sixth-graders) from nine schools located in the capital city of Hungary were involved in this investigation. Since the participating schools are all located in different parts of the city, it can be stated that the sample represents various SES-background populations. All fifth- and sixth-grade students at the participating schools were asked to complete the test during one of their mathematics classes.

Measures and Procedure

Three versions of a paper-and-pencil test were developed, all of which contained two of the three subtests. The three subsets are (1) five problematic word problems, (2) five standard word problems, and (3) five humorous word problems.

Five Problematic Word Problems

The first subtest comprised of five tasks based on a seminal work by Verschaffel et al. (1994). In that work, ten pairs of simple arithmetic word problems were collected and referred to as either standard (S) or problematic (P). In each pair of the S- and P-items the same arithmetical operation was expected to be used by the students, and whereas in the S-items the mere execution of the arithmetic operation yielded the right and mathematical correct answer, in the P-items some realistic considerations were to be taken into account. An example for a pair of such items:

Chris made a walking tour. In the morning he walked 8 kilometers and in the afternoon 15 kilometers. How many kilometers did Chris walk?

S-item where the answer comes merely from executing an arithmetical operation, i.e., 8 + 15 = 23.

Bruce and Alice go to the same school. Bruce lives at a distance of 17 kilometers from the school and Alice at 8 kilometers. How far do Bruce and Alice live from each other?

P-item where mere execution of either an addition or a subtraction indicates only partial insight of the problem.

Five out of these P-items were reconstructed as closed-question format items by Csíkos et al. (2011). While these items were labelled as P5, P9, P6, P3 and P1 items in Verschaffel et al.'s work (1994), in subsequent publications within the profession, these problems became designated as the *Runner*, *Rope*, *School*, *Water* and *Friends* items. The closed-question-format versions of these five items (slightly modified as compared to the ones in Csíkos et al. 2011) formed the first subset of word problems used in the current investigation. The item stem of these five items are as follows:

Runner: John's best time to run 100 meters is 17 seconds. How long will it take him to run 1 kilometer?

Rope: A man wants to have a rope long enough to stretch between two poles 12 meters apart, but he has only pieces of ropes 1.5 meters long. How many of these pieces would he need to tie together to stretch between the poles?

School: Bruce and Alice go to the same school. Bruce lives at a distance of 17 kilometers from the school and Alice at 8 kilometers. How far do Bruce and Alice live from each other?

Water: What will be the temperature of water in a container if you pour 1 liter of water at 80°C and 1 liter of water at 40°C into it?

Friends: Carl has 5 friends and Georges has 6 friends. Carl and Georges decide to give a party together. They invite all their friends. All friends are present. How many friends are there at the party?

The exact look of the tasks was as follows in the example of the *Runner* item:

John's best time to run 100 meters is 17 seconds. How long will it take him to run 1 kilometer?

a) It is unambiguous.

1 kilometer = 1000 metres; 1000:100=10

10 times 17 = 170 seconds.

It takes 170 seconds.

b) He needs more than 10x17 seconds because he would have run in 170 seconds, if he had not got tired.

It takes more than 170 seconds.

c) We do not know how the runner has got tired.

Consequently, this problem cannot be solved.

Five Standard Word Problems

A second set of five items consisted of the standard item version of the five problematic word problems (Runner, Rope, School, Water, Friends) with their original, open-ended version.

'*Runner*': A boat sails at a speed of 45 kilometers per hour. How long does it take this boat to sail 180 kilometers?

'*Rope*': A man cuts a clothesline of 12 meters into pieces of 1.5 meters each. How many pieces does he get?

'*School*': Chris made a walking tour. In the morning he walked 8 kilometers, and, in the afternoon, he walked 15 kilometers. How many kilometers did Chris walk?

'*Water*': A shopkeeper has two containers for apples. The first container contains 60 apples and the other 90 apples. He puts all the apples into a new, bigger container. How many apples are there in that new container?

'*Friends*': Pete organized a birthday party for his tenth birthday. He invited 8 boy friends and 4 girl friends. How many friends did Pete invite for his birthday party?

Five Humorous Word Problems

A third set of five items consisted of humorous word problems (see Table 6 for more details). Two jokes ("Birds" and "Melons") were borrowed from the Flemish colleagues while three others were developed based on three ideas from a humor newspaper published for teenagers in the 1980s and 1990s.

Unlike the study by Van Dooren et al. (2017) in which humorous word problems were not to be solved by the students but they were inserted as cartoons in between mathematical tasks, in the current investigation we handled the humorous word problems in exactly the same manner as the problematic word problems. All humorous items were provided in closed-question format followed by the answer pattern similar to that of provided with the problematic word problems. Option a) always expressed a straightforward, one-operation arithmetic solution; option b) presented the point of the joke (while it presented an estimation or an interval for the problematic word problems); while option c) always stated the same option, i.e., that “not all the necessary data had been given, therefore the task could not be solved”.

The structure of the three test versions was the following:

Version A: Five standard items + five problematic items

Version B: Five humorous items + five standard items

Version C: Five humorous items + five problematic items

With the aim of testing the hypotheses listed above, this research design enabled us to investigate the effect of humorous items on both standard and problematic items. Our methods simultaneously allowed us to gather unbiased base solution rates for Version A. The three versions were randomly assigned to the students in each of the participating classes.

Questionnaire Items

In all three test versions, additionally, seven items were to be judged on five-point Likert-scale (see Table 5 for the item wording). The first two items asked students’ attitudes towards mathematics in general, and specifically while in the classroom. The third and fourth items were borrowed from Berze and Csíkos (2017), and addressed the content of parental supervision, and thus also the presumed importance of mathematics. The last three items focused on students’ attitudes towards three complementary fields of school mathematical experiences: word problem solving, mental calculation and geometrical constructions. At the very end of the test, an open-ended question asked the students to write down which of the ten previously solved tasks was their most favorite one.

Coding

The coding procedure of the closed-question format items were straightforward. During the coding process of the open-ended versions of the standard items two kinds of data were derived: the numerical answer provided by the student, and a dichotomous score about whether he or she showed some hesitancy in connection to the solvability of the word problem. Since standard word problems are routine tasks and the majority of upper elementary school students can usually solve them via the straightforward application of an arithmetic operation, by means of scoring the possible hesitancy about the solvability, it became possible in Version B to reveal whether those who answered the humorous items first were more inclined to express their hesitancy about the solvability of the standard word problem or point to some possible alternative solutions.

Qualitative analyses

The open-ended format of the routine word problems provided the opportunity to conduct qualitative analysis on students' written answers. Based on the considerations given by Johnson and Onwuegbuzie (2004), our investigation proved to be of the mixed-method type. Having formulated quantitative research objectives, and having collected both qualitative and quantitative data, the analysis of our qualitative data emphasized the so-called "seeing plausibility" strategy (see Cohen et al., 2002).

FINDINGS

Comparing the Results to a Previous Study

The closed-question format enabled comparisons between the results of the current investigation and results yielded by Csíkos et al. (2011). Since there were three options from which student could choose one, the mere guessing chance is 33.3%. Table 1 presents the b) and c) options jointly, since both can be considered realistic reactions.

Table 1
Rate of realistic reactions to the problematic word problems (%)

Problematic word problem	Selection rate according to Csíkos et al. (2011) – Fifth-grade students (N = 4037)	Selection rate in the current investigation	
		Fifth-grade students (N = 352)	Sixth-grade students (N = 409)
Runner	55.4	53.1	58.0
Rope	62.4	60.5	69.6
School	50.7	69.1	78.7
Water	59.5	66.5	80.4
Friends	51.7	74.1	77.8

Having compared the three dichotomized datasets (correct vs incorrect items), one-way ANOVA showed significant between-age-group differences in all items with the exception of the Runner item. On the other four items, according to the inhomogeneity of variances in the samples (Levene-test p values < 0.001), Dunnett's T3 post-hoc analysis could have been used, and it revealed significant differences among the following groups. On the Rope item: {6th grade, 2020} > {5th grade, 2011}, p = 0.04. On the School item: {6th grade, 2020} > {5th grade, 2020} > {5th grade, 2011}, p < 0.01 in all cases. On the Water item: 6th grade, 2020} > {5th grade, 2020} > {5th grade, 2011}, p < 0.03 in all cases. On the Friends item: {6th grade, 2020} > {5th grade, 2011, p < 0.001, and {5th grade, 2020} > {5th grade, 2011}, p < 0.001.

The results suggest that the selection rate for the realistic options is tendentiously similar or higher than those in the previous study. Furthermore, sixth-grade students have higher percentage values than fifth graders in the current study. The differences between the results of the two surveys can be due to the fact that previously we had a nationwide representative sample (students from each types of settlements) whereas currently the participants were recruited only from the capital city in which students perform better in general according to the tendencies in nationwide educational surveys.

The Effect of Humor on Problematic Word Problems

The results presented in Table 1 took into account all data from students who solved the closed-question format of the problematic word problems either in test version A or C. To analyze the possible effect of humorous items, in Table 2 the rate of realistic responses yielded in the two versions are compared to one another. The samples of the fifth- and sixth-grade students are examined separately.

Table 2
The rate of realistic reactions in test versions A and C (version C reflects the effect of humorous word problems) (%).

Problematic word problem	Fifth-grade students		Sixth-grade students	
	Version A	Version C	Version A	Version C
Runner	43.4	65.2	48.5	67.7
Rope	48.4	75.2	57.6	81.9
School	68.7	69.7	73.5	84.2
Water	58.7	76.3	76.6	84.3
Friends	68.4	81.3	69.5	86.1

An overview of descriptive statistical nature tells that students performed better in tasks of Version C as compared to Version A. In Version C, students first solved five humorous tasks, whereas in Version A, they solved five routine word problems first. According to a series of Welch-tests, among the groups of fifth-graders the difference between the results in the two versions proved to be significant in the Runner, Rope and Water tasks ($t = 3.61$, 4.58, and 3.02, respectively, with $p < 0.01$ in each case. There were no significant differences in the School and Friends tasks ($t = 0.28$, $p = 0.78$; $t = 1.44$, $p = 0.15$). In the group of sixth-graders, significant differences were found in the Runner, Rope, and Friends tasks ($t = 3.56$, 4.71, and 3.62, respectively, with $p < 0.001$ in each case, however, there were no significant differences in the School and Water tasks ($t = 1.83$, $p = 0.07$; $t = 1.52$, $p = 0.13$), nevertheless, the tendencies in favor of the Version C results are clear.

The general linear model (two-way ANOVA; Grade X Test version) computed separately for each realistic task showed the relative effect size of the test version and the school grade. Table 3 presents the partial eta-squared values for the two factors.

Table 3
ANOVA-statistics and effect sizes of the test version and the school grade for the performance on the five problematic word problems

Problematic word problem	F	p	Partial eta-squared (%)	
			Test version (A vs C)	Grade (5^{th} vs 6^{th})
Runner	9.13	< 0.001	3.0	0.2
Rope	15.96	< 0.001	4.9	0.5
School	2.51	0.06	0.2	0.5
Water	7.75	< 0.001	1.4	1.1
Friends	4.90	.002	1.4	< 0.01

The usual milestones in the interpretation of the effect sizes are 1% (small), 5% (medium) and 15% (large). According to this, the test version, i.e., whether students encountered humorous tasks before solving problematic word problems had a medium-sized effect for the Runner and Rope tasks, a small effect for the Water and Friends tasks, and a negligible effect for the School problem.

The Effect of Humor on Routine Word Problems

In order to investigate the effect of humorous word problems on the solving of standard-type (routine) word problems, versions A and B were used. The humorous word problems in Version B may have triggered more hesitancy, feeling of non-solvability or similar obstacles for students while solving the routine word problems (Table 4). The number of students showing hesitancy or expressing non-solvability tended to be higher among sixth-grade students.

Table 4
The number of students' reactions demonstrating hesitancy or non-solvability in the course of standard word problems

Problematic word problem	Fifth-grade students		Sixth-grade students	
	Version A	Version B	Version A	Version B
Runner –standard	1	1	2	4
Rope – standard	0	0	0	0
School – standard	1	1	2	8
Water –standard	1	6	4	11
Friends – standard	1	4	3	5

A careful analysis of students' responses illustrates the phenomenon how the presence of humorous items inclined some students to bluntly express their opinions about the solvability of these tasks.

The following examples (Figure 1 and 2) illustrate students' hesitancy about the solvability of different tasks that were intended to be of the routine and unambiguous task type.

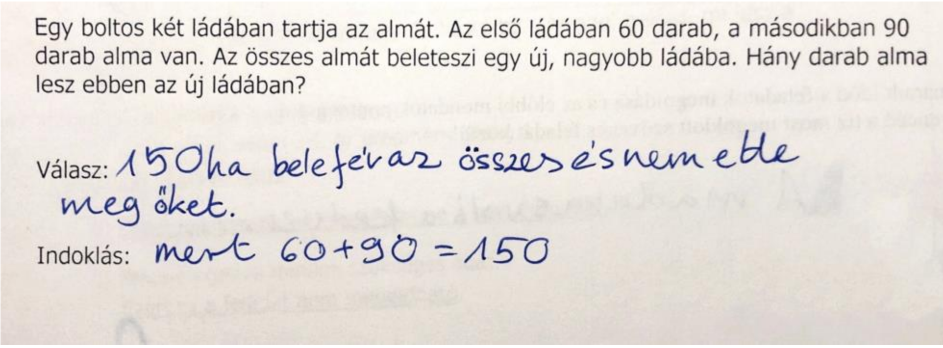

Figure 1
Example of a student expressing his hesitancy about the solvability of a standard word problem

The word problem (standard version of the 'Water' problem) is presented as follows: „A shopkeeper has two containers for apples. The first container contains 60 apples and the other 90 apples. He puts all the apples into a new, bigger container. How many apples are there in that new container?" The student's answer is the following: *Answer: „150 if all the apples could fit in the container, and he didn't eat them."* Comments: "Because 60+90=150".

In another example yielded from the same task the student pondered towards whether any of the apples should be thrown away because of rottenness.

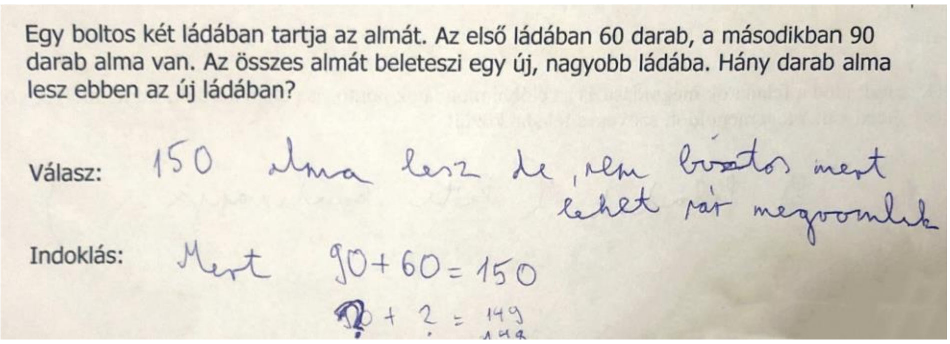

Figure 2
The student's answer is the following: Answer: „*There will be 150 apples, but it is not sure because it is possible that some of the apples go rotten.*" Comments: "Because 90+60=150, ?+?=149, 148 …"

According to the frequency values presented in Table 2, it was this "Apple" task (standard version of the Water problems) that encouraged students the most to express their hesitancy about the solution. Possibly as an effect of the humorous items just solved before, the following answer itself contains a bit of humor (Figure 3).

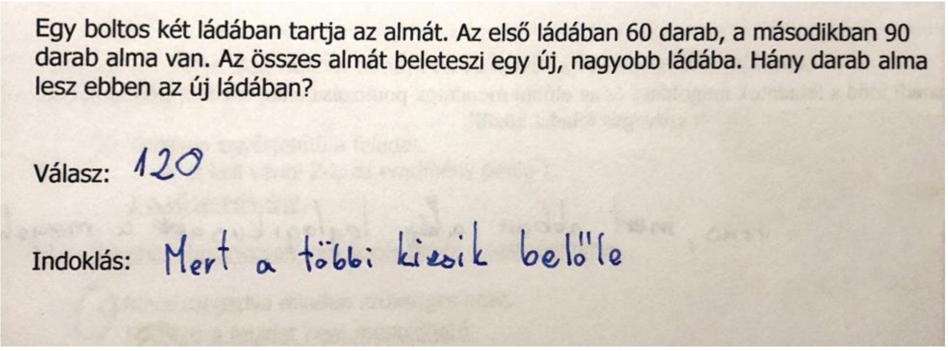

Figure 3
The student's answer is the following: Answer: „120" Comments: *"Because the others fall out"*

Figure 4 illustrates the possible negative role humor may have played in possibly distracting the students in the process of solving routine word problems. The example comes from the standard version of the School problem.

All these four illustrations come from Version B where humorous tasks may have triggered the problem-solving strategies in students' minds which went well beyond the usual superficial word problem solving procedures.

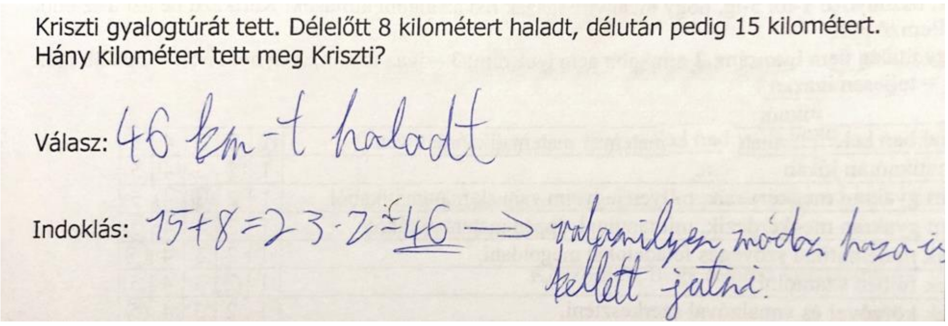

Figure 4
„Chris made a walking tour. In the morning she walked 8 kilometers, and, in the afternoon, she walked 15 kilometers. How many kilometers did Chris walk?" The student's answer is the following: *Answer: „She walked 46 kilometers." Comments: "15+8=23×2=46→ she had to get home somehow"*

Students' Attitude and its Connections to Word Problem Solving

Table 5 presents the basic descriptive statistical values for the seven items concerning attitude towards mathematics.

Table 5
Students' reactions to the mathematics attitude-related questions, Mean values (SD in parentheses)

Item	5th grade	6th grade
I wake up happier on those days when having mathematics class. *	2.67 (1.22)	2.40 (1.12)
I am in a good mood during mathematics classes. *	3.36 (1.15)	3.20 (1.04)
My parents often ask about my school math marks.	3.82 (1.30)	3.73 (1.30)
My parents often ask what we are currently learning in mathematics.	3.67 (1.30)	3.55 (1.33)
I like to solve mathematics word problems.	2.56 (1.34)	2.67 (1.36)
I like mental calculation. *	3.07 (1.43)	3.27 (1.41)
I like geometric constructions with compasses and ruler. *	3.95 (1.28)	3.73 (1.34)

Note: N ranged from 533 to 540 in Grade 5, and from 688 to 695 in Grade 6; significant differences between the mean values of the two groups are marked by *.

The first two items asked about a general attitude towards mathematics as a school subject. Grade 6 students were significantly less positive in this respect. The third and fourth items revealed that parents were indeed interested in the student' mathematics results, as well as in the topics they are learning about. The last three items gave information about three different fields of school mathematics learning, and word problem solving proved to be much less favorable than the other two; paired-wised t-test comparison showed significant differences in both age-groups between the attitude towards word problem solving and the other two fields.

Further analyses showed that although the attitude towards mathematics is not expected to change suddenly, still the answers varied according to the test versions the students were just solving. Two-way ANOVA (Grade x Version) revealed that the test version had a marginal effect on the attitude items, with the exception of the fifth item on word problem solving which had an eta-square value of 0.4%, and in both Grade 5 and 6 the students who solved Version B or C (the ones which contained the humorous items) had a higher mean than students with Version A; however, the difference was not significant ($F(2,534) = 1.36$, $p = 0.26$ in Grade 5; $F(2,692) = 1.32$, $p = 0.27$ in Grade 6).

Students' Favorite Word Problems

At the end of the test, students were asked which of the previously solved tasks had been their favorite. This question was posed without asking for any reason or explanation. Restricting the analysis to those who solved either the B or C version of the test, the results show that humorous word problems were received rather favorably among the ten tasks they had solved. Table 6 presents the selection rate for humorous word problems. According to the results, 81.8% of the students chose one of the humorous word problems as their favorite, while the B and C versions did not display relevant differences in this respect.

Table 6
The popularity of humorous word problems

Humorous word problem	Selection rate (%)
If your sibling has three apples and you take away two of them, what will be the results? (a fight)	19.7
If there are five birds sitting on the branch of a tree and two of them are shot, how many are left? (none)	34.3
If I have one melon in my hand, and two in the other, what do I have? (strong arms)	17.6
How would you distribute 98 peaches among 7 children? (by making jam first)	3.0
Starting on Monday, throughout an entire week we do 20 push-ups each day. What will be the result by Sunday evening? (terrible muscle strain)	7.2
Five humorous word problems together	81.8

DISCUSSION

Novelty of the Research

The novelty of our investigation lies in extending previous results regarding the possible effect of humor may have on word problem-solving in three ways. First, our research design enabled for conducting between-age comparison analyses. Second, we aimed to reveal any possible negative effect humor might have when solving traditional, routine word problems. Third, the short-term changes in students' attitudes towards mathematics were investigated by a brief questionnaire.

The between-age comparison here is limited to two school grades; 5th and 6th grade students in Hungary are 11 and 12 years old. According to the results of two-way ANOVA, it was not the school grade but the positive effect of humor that had a greater impact on students' performance on problematic word problems. An open question remains whether there would be similar effects of humor at younger ages. According to the Piagetian sense of cognitive development, the appearance of the necessary cognitive prerequisites for understanding humor is usually expected at the fourth stage of cognitive development (from the age of 11-12 years). In further research, the role and effects of humor can be investigated among students from a larger span of years.

Positive Versus Negative Effects of Humor

Our results reaffirmed the potential positive effect of humor on the solution of realistic word problems that had otherwise been labelled as problematic in previous research. The results confirm findings by the Leuven colleagues in that an increase in realistic response rates could be revealed. One major difference from their research design was the use of a closed question format; we therefore first compared our results to a previous study containing items that possessed a closed-question format with three options. The differences summarized in Table 1 are not large and can be attributed either to changes in the learning materials during the past ten years or to a difference in the sampling. The positive effect of humor was observed both in fifth and sixth grades.

Despite this finding, a possible marginal effect related to humor was revealed in the sense that after reading and "solving" humorous word problems students showed a stronger tendency to misinterpret or overcomplicate word problems that had been labelled and intended for use as standard word problems. Since this marginal effect was revealed only among 1 to 2% of the students, this aspect of the potential disadvantage to using humor should be acknowledged yet can practically be ignored. Items on different aspects of attitude towards mathematics revealed that students have generally negative attitude towards word problem solving as compared to other fields of mathematics. Although the students who had just previously solved the test version containing humorous items gave higher scores to that item, the difference was not significant. It suggests that while we did observe a short-term increase in students' realistic reactions on problematic word problems, there was hardly any change in students' affective sphere even in the short run.

Educational Implications

According to our results, using humor can have more advantages than possible disadvantages as regards mathematical word problem-solving. Understanding jokes while reflecting upon the process of incongruity in our thinking can be a powerful tool for stimulating children to see word problems from other perspectives (Van Dooren et al., 2017). Having addressed the phenomenon of the potential effect humor may have on word problem-solving in mathematics, the serious question of how to integrate humor into the mathematics classroom remains open to discussion. In our current investigation, students themselves expressed their positive attitude toward humorous word problems by selecting them as their favorite among other types of word problems. Gazit (2013) has provided further empirical evidence that pre-service elementary school teachers have an overall positive attitude toward integrating humor in the classroom. Gresham and Burleigh (2019) advocated the incorporation of humor into teacher education courses. Positive classroom atmosphere (humor was mentioned as part of it) belongs to the crucial factors preservice teachers participating in a mathematics methods course mentioned as being evidently effective. Teacher educators themselves may be examples of how to incorporate humor in mathematics, and the sense of humor as a psychological characteristic of a good teacher is constantly and unquestionably revealed in empirical studies for long decades throughout the world (in Hungary, e.g., Suplicz 2012). Further research may and should address the issue of parents' views and beliefs on incorporating humor into mathematics (and other subjects), since they are key actors in shaping their children's views and attitudes towards both mathematics and humor.

ACKNOWLEDGMENTS

The current research was supported by the MTA-ELTE Complex Mathematics Education Research Group and the MTA-SZTE Metacognition Research Group.

REFERENCES

Banas, J. A., Dunbar, N., Rodriguez, D., & Liu, S-J. (2011). A review of humor in educational settings: Four decades of research. *Communication Education, 60*, 115-144.

Berk, R. A., & Nanda, J. (2006). A randomized trial of humor effects on test anxiety and test performance. *Humor, 19*, 425-454.

Bieg, S., Grassinger, R., & Dresel, M. (2017). Humor as a magic bullet? Associations of different teacher humor types with student emotions. *Learning and Individual Differences, 56*, 24-33.

Brousseau, G., Sarrazy, B., & Novotná, J. (2014). Didactic contract in mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 153-159). Dordrecht: Springer.

Burde, M. (2010). The Parodia sacra problem and medieval comic studies. In A. Classen (Ed.), *Laughter in the Middle Ages and early modern times: Epistemology of a fundamental human behavior, its meaning, and consequences* (pp. 215-242). Berlin – New York: de Gruyter.

Cai, J., Morris, A., Hohensee, C., Hwang, S., Robison, V., & Hiebert, J. (2018). The role of replication studies in educational research. *Journal for Research in Mathematics Education, 49,* 2-8.

Cohen, L., Manion, L., & Morrison, K. (2002). *Research methods in education*. Routledge.

Csíkos, C., Kelemen, R., & Verschaffel, L. (2011). Fifth-grade students' approaches to and beliefs of mathematics word problem solving: a large sample Hungarian study. *ZDM, 43*(4), 561-571.

Csíkos, C., Szitányi, J. & Kelemen, R. (2012). The effects of using drawings in developing young children's mathematical word problem solving: A design experiment with third-grade Hungarian students. *Educational Studies in Mathematics, 81,* 47-65.

Dowling, J. S. (2014). School-age children talking about humor: Data from focus groups. *Humor, 27*, 121-139.

Egmir, E., Erdem, C., & Kocyigit, M. (2017). Trends in Educational Research: A Content Analysis of the Studies Published in" *International Journal of Instruction". International Journal of Instruction, 10,* 277-294.

Ford, T. E., Ford, B. L., Boxer, C. F., & Armstrong, J. (2012). Effect of humor on state anxiety and math performance. *Humor*, *25*, 59-74.

Gazit, A. (2013). Pre service mathematics teachers' attitudes toward integrating humor in math lessons. *Israeli Journal of Humor Research*, *Issue No. 3*(June), 27-44.

Greer, B. (1997). Modelling reality in mathematics classrooms: The case of word problems. *Learning and instruction, 7*, 293-307.

Gresham, G., & Burleigh, C. (2019). Exploring early childhood preservice teachers' mathematics anxiety and mathematics efficacy beliefs. *Teaching Education, 30,* 217-241.

Hall, G. S., Allin, A. (1897). The psychology of tickling, laughing, and the comic. *The American Journal of Psychology, 9*, 1-41.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33,* 14-26.

Kline, L. W. (1909). The psychology of humor. *The American Journal of Psychology, 18*(4), 421-441.

Laineste, L. (2002). Take it with a grain of salt: The kernel of truth in topical jokes. *Folklore: Electronic Journal of Folklore, 20-22*, 7-25.

Masten, A. S. (1986). Humor and competence in school-aged children. *Child Development, 57*, 461-473.

McMorris, R. F., Urbach, S. L., & Connor, M. C. (1985). Effects of incorporating humor in test items. *Journal of Educational Measurement, 22,* 147-155.

Morreall, J. (2014). Humor, philosophy and education. *Educational Philosophy and Theory, 46*(2), 120-131.

Pickover, C. A. (2009). *The math book: from Pythagoras to the 57th dimension, 250 milestones in the history of mathematics*. New York: Sterling.

Reyes, L. H. (1984). Affective variables in mathematics education. *The Elementary School Journal, 84*, 558-581.

Semrud-Clikeman, M., & Glass, K. (2008). Comprehension of humor in children with nonverbal learning disabilities, reading disabilities, and without learning disabilities. *Annals of Dyslexia, 58,* 163-180.

Suplicz, S. (2012). Tanárok pszichológiai jellemzői diákszemmel. [Teachers' psychological characteristics in their students' eyes.] Ph.D. dissertation, University of Debrecen. Available on-line: https://dea.lib.unideb.hu/dea/bitstream/handle/2437/161946/Suplicz_Sandor_Ertekezes-t.pdf?sequence=5

Van Dooren, W., De Wortelaer, & Verschaffel, L. (2017). Are you joking? Or is this real? Increasing realistic responses to word problems via humor. In B. Kaur, W. K. Ho, T. L. Toh, & B. H. Choy (Eds.), *Proceedings of the 41st Conference of the International Group for the Psychology of Mathematics, Vol. 4.* (pp. 297-304). Singapore: PME.

Van Dooren, W., Lem, S., De Wortelaer, H., & Verschaffel, L. (2019). Improving realistic word problem solving by using humor. *The Journal of Mathematical Behavior*, 53*, 96-104.*

Verene, D. P. (2017). The pursuit of the pursuit of the truth. In S. G. Beardsworth & R. E. Auxier (Eds.), *The Philosophy of Umberto Eco* (pp. 153-177). Open Court Publishing Company.

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

Verschaffel, L., De Corte, E., Lasure, S., Van Vaerenbergh, G., Bogaerts, H, & Ratinckx, E. (1999). Design and evaluation of a learning environment for mathematical modeling and problem solving in upper elementary school children. *Mathematical Thinking and Learning, 1*, 95–229.

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

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.

Ziyaeemehr, A., & Kumar, V. (2014). The Relationship between Instructor Humor Orientation and Students' Report on Second Language Learning. *International Journal of Instruction, 7*, 91-106.

Zsolnai, J., Zsolnai, L. (1987). Mi a baj a pedagógiával? [What’s wrong with pedagogy?]. Budapest: Tankönyvkiadó.