

A Cross-sectional Analysis of Students' Answers to a Realistic Word Problem from Grade 2 to 10

Gabriella Ambrus ¹, Eszter Herendiné Kónya ², Zoltán Kovács ³, Judit Szitányi ⁴, Csaba Csíkos ^{4*}

¹ *Mathematics Teaching and Education Centre, Institute of Mathematics, Faculty of Science, ELTE Eötvös Loránd University, HUNGARY*

² *Institute of Mathematics, University of Debrecen, HUNGARY*

³ *Institute of Mathematics and Science, University of Nyíregyháza, HUNGARY*

⁴ *Department of Mathematics, Faculty of Primary and Pre-School Education, ELTE Eötvös Loránd University, HUNGARY*

* CORRESPONDENCE: ✉ csikos.csaba@tok.elte.hu

ABSTRACT

Several investigations have revealed that students tend to exclude their real-world knowledge when solving simple, routine-like mathematical word problems. The current research is a cross-sectional developmental analysis with students from grade 2 to 10 (N=1346). Other than describing the development (or lack thereof) in students' realistic answers, connections with math-related background variables and possible class-level effects have been investigated.

Keywords: word problems, realistic mathematics education, cross-sectional development

INTRODUCTION

Mathematical word problems are an effective tool in encouraging students to apply their mathematical knowledge and skills. As Pollak (1969) argues, it is the realm of mathematics word problems through which students in schools may become involved in the application of mathematics. Moreover, word problems are not only manifestations of mathematical structures dressed up in text, but the opposite is also true: word problems can be used to facilitate the learning of mathematical concepts (English & Sriraman, 2010). Mathematical word problems must to some extent necessarily refer to everyday concepts and relations. The extent to which everyday concepts and relations have a role in the solution process may help in categorizing word problems. For example, Galbraith and Stillman (2001) proposed a taxonomy in which four ideal types of word problems are listed: injudicious, context-separable, standard applications and modeling problems. From the modelling perspective of mathematics education, word problems play a leading role in the process of contextual modelling through which the very human nature of constructing word problems as they relate to everyday experiences is emphasized (Kaiser and Sriraman, 2006).

A branch of mathematical word problems can be labelled as "realistic". Searching for an appropriate definition, both characteristics of the word problem itself and the level of students' engagement should be taken into account (Hiebert et al., 1996). For the purpose of the current study we have used the classical Treffersian (1993) approach of horizontal mathematization, i.e., in the solution process students are expected to use mathematical tools to solve everyday life problems. Recently, Jupri and Drijvers (2016) suggested a refinement in the description of horizontal mathematization: reflection as the process of checking the solution can be crucial when encountering realistic word problems. Realistic tasks require translations between reality and mathematics and the process of translation is often labelled mathematical modelling. By "reality," we

Article History: Received 18 December 2018 ♦ Revised 18 February 2019 ♦ Accepted 18 February 2019

© 2019 by the authors; licensee Modestum Ltd., UK. Open Access terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>) apply. The license permits unrestricted use, distribution, and reproduction in any medium, on the condition that users give exact credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if they made any changes.

refer to meaning of the word as defined by Pollak (1979), i.e., the “rest of the world” outside mathematics including nature, society, everyday life and other scientific disciplines.” (Blum & Borromeo Ferri, 2009, p. 45).

Since the 1980s and 1990s, a number of investigations have revealed that students tend to exclude their real-world knowledge when solving mathematical word problems that seem to be solvable by following a series of routine steps. A set of ten such word problems (labeled as P-items, i.e. problematic items) were developed by Verschaffel, De Corte and Lasure (1994), and several replication studies have been conducted throughout the world. Results have revealed that throughout all educational systems children around the ages of 10 and 11 face the same difficulties when they encounter mathematical tasks that are seemingly easy-to-solve by means of a usual superficial (term borrowed from Verschaffel and De Corte, 1997) strategy: after searching for numerical data in the text of the word problem, students select and execute one or more of the basic arithmetic operations and finally provide a numerical answer which is usually the result of the arithmetic calculations.

Several replicative studies of this Flemish investigation have been conducted worldwide (for an extensive review (see Verschaffel, Van Dooren, Greer, & Mukhopadhyay, 2010) with the same general conclusion. Students by the age of roughly ten do indeed tend to exclude their real-world knowledge when solving routine-like, simple arithmetic word problems. According to the actual content and the arithmetic operations to be computed, in some cases none of the students could provide a realistic answer or at least show some hesitancy indicating that e.g., some information is lacking or ambiguous. Changing the context of the task, i.e. providing a hint about the non-solvability of the problem or placing the simple word problem among puzzle-like tasks caused no relevant changes in the solution strategies used (Yoshida, Verschaffel, & De Corte, 1997). Research by Csíkos, Kelemen and Verschaffel (2011) has revealed that when changing the task format from open-ended to multiple-choice, the increase in fifth-grade students’ performance can be explained by the guessing chance generated by the multiple-choice format.

The reasons why students usually struggle with word problems that require them to activate their real-world knowledge have been intensively examined in recent decades (see e.g., Baruk, 1989; Ramharter, 2002; Peled & Balacheff, 2011; English & Shriraman, 2010; Prediger, 2010).

Varga (1988, p. 295) nicely describes his nephew’s attempts to solve such kinds of word problems.

Here is the problem:

Somebody tells a joke on Monday to five people. The next day, on Tuesday, each of the five tells the joke to six other people. Each of the latter tells it to seven people on Wednesday. How many will have heard it on Wednesday?

Tamás Varga’s nephew had it as homework, and found three different solutions, depending on the interpretation of the text:

- Five people heard the joke on Monday, five times six or 30 on Tuesday, five times six times seven or 210 on Wednesday. The answer is 210.
- In another interpretation the answer is $5+30+210$ or 245. Those who heard the joke on Monday or on Tuesday will have heard it on Wednesday together with the 210 who heard it precisely that day.
- The one who told the joke to the first five people must have heard it previously (unless he or she invented it) so the answer is 246.

But the nephew was desperate and had a great dilemma. “If I come up with any of these solutions,” said to his uncle, “the teacher may have another solution in mind and she will make fun of me in front of the class because I could not find the real solution. The whole class will laugh at me!” (Varga, 1988, p. 295).

As of the 1990s, research on math-related beliefs has proved a powerful tool in elucidating students’ difficulties in solving realistic word problems. Reusser and Stebler’s (1997) seminal work on the role of individual beliefs, Yackel and Cobb’s (1996) study on classroom-level norms along with Brousseau’s didactical contract idea have all provided new impetus toward designing intervention studies that foster students’ mathematical thinking. There are promising results regarding how to change the classroom norms, the teaching methods and the tasks themselves to improve the chances of students’ being engaged in word problem-solving (Verschaffel, De Corte, Lasure, Van Vaerenbergh, Bogaerts, & Ratinckx, 1999) It appears possible that mathematics education impairs students’ capability to apply their mathematical knowledge in real-world situations from the very beginning of formal elementary schooling. The aim of the current research project is to reveal developmental trends in students’ realistic answers using simple routine-like word

problems. We hypothesized that older students will attain better results in terms of performing more realistic reactions. We additionally hypothesized that a word problem containing larger figures and more demanding arithmetic operations would have an impact on the rate of the students' realistic reactions. We subsequently aimed to characterize those students who provide realistic answers, i.e. whether they possess strikingly different beliefs or have reached a different level of academic achievement compared to their peers. Furthermore, our sampling procedure enabled us to investigate whether there are "crystallization points," i.e. classes where the majority of the students provide realistic solutions.

METHODS

Sample

The current two studies were conducted in Hungary. According to recent PISA studies, the mathematics achievement of Hungarian students is below the OECD average. More information and some characteristics of Hungarian mathematics education can be found in Szendrei (2007) and in Csíkos, András, Rausch and Shvarts' (2019).

In order to measure developmental tendencies in students' possible realistic considerations to open word problems, the research design included a cross-sectional survey applied from grades two to ten. According to the results of a previous exploratory study with a structurally analogous task (Ambrus, & Szűcs, 2016), the school grades involved were selected with the aim of revealing fine-tuned developmental stages in lower grades. In both in Study 1 and Study 2, students worked individually, but sampling was done on a classroom basis and the tasks were solved during regular math classes in the middle of the school year.

Sample for study 1

In our study two open structurally analogous word problems were used (see Measures). For the easier task (computation), students from grades 2 to 6 were selected. First-graders in the middle of their first school year cannot be expected to work individually on word problems.

In Hungary, the great majority of children attend what is known as a "general school" from first to eighth grade, and can then apply to different types of secondary schools. The participating general schools were asked to involve their classes in grades 2, 3, 4 and 6. Sixth-grade classes were divided into two parts in order to provide a comparative matching sample for Study 2. The number of students involved from the four school grades were 171, 200, 226 and 129, respectively.

Sample for study 2

In the schools involved in Study 1, half of the sixth- and eighth-grade classes participated. Additionally, tenth-grade classes were selected from different (upper secondary) secondary schools. The tenth-grade branch of our sample represents a population that is not merely two years older than eighth-grade students but were also selected from a higher-achieving stratum of the total student population as a result of the selection process for secondary school. The number of students involved in Study 2 were 159, 119 and 279, respectively. The reason why the number of students in Study 2 are fairly different is the following: in order to have a sample of sixth-grade students we involved additional schools whose sixth-grade classes were recruited.

Measures

Each survey participant was asked to solve a mathematical word problem; after its completion, students were also asked to fill in a questionnaire.

Two versions of a word problem were developed. Each differed in the magnitude of the numbers involved and the story (wording) built around the otherwise identical mathematical structure. Both versions described a situation in which the main actor receives some money on a weekly basis. Using the weekly base and the overall amount the students had to calculate the number of days that had passed. The reason of having two versions of a word problem in a cross-sectional study can be justified by not merely considering the number of digits present in the text, but by previous exploratory studies (Ambrus, & Szűcs 2016; Ambrus 2016).

Word problem in Study 1

The king's new valet gets one gold coin for a week's service, which the valet has saved. How many days has the valet been in the king's service if he has already saved 6 gold coins?

Table 1. The coding system with examples

Code	Description	Example
1	Expected answer which results from the anticipated straightforward and unproblematic application of the arithmetic operations elicited by the problem statement.	$35 \times 7 = 245$; or $6 \times 7 = 42$.
2	Technical error which accompanies the straightforward application of the expected arithmetic operations elicited by the problem statement but differs from Code #1 in that there is a technical error or inaccuracy in the execution of the arithmetic operation(s).	There is a mistake in the multiplication.
3	Realistic answer which follows from the effective use of real-world knowledge about the context elicited by the problem statement in one or more stages of the solution process.	The real situation at least partly considered; e.g., the answer: from 245 to 251 days
0	This category was used when the pupil did not answer the problem or when (s)he wrote down that (s)he could not answer it.	
4	All answers that could not be classified into one of the former categories.	$5000 : 7 = 714$ days

Word problem in Study 2

Since Pisti moved to a new house with his family, he has received his pocket money of 1,000 Hungarian forints, weekly. He has saved all of his pocket money since they moved. How many days have they spent in their new home if Pisti has already saved 35,000 Hungarian forints?

Both problems may seem to require straightforward application of a usual word problem-solving strategy (which is called “superficial” by Verschaffel and De Corte, 1997), i.e. by using the figures explicitly given in the text, finding one or two appropriate arithmetic operations, and executing them. Thus the word problem in Study 1 is expected to be solved predominantly by calculating $6 \times 7 = 42$, while the task in Study 2 is foreseen to be solved as $(35\,000 : 1000) \times 7 = 245$.

Both versions of the task can be considered realistic for the students according to Gravemeijer and Terwel (2000), since the situations may be experientially real for them. The day of arrival at the new home or the king’s castle, and the day within a week where payment is done, and the actual day in the text of the word problem define a wide range of possibilities for the solution (see Ambrus, 2016). Therefore, the number of days the king’s valet has been in the king’s service might span from 30 to 48 days, and the days of living in a new place might span from 233 to 251 days. The “pocket money” problem can be considered as a modeling task as well, since modeling tasks are problem-oriented, authentic word tasks whose solution demands the construction, application and validation of models. Validation of the models means checking if the answer to the question is appropriate in the given context (see Schukajlow, 2011).

Additionally, in both studies a short questionnaire (Table 4) was administered to the students; the questionnaire contained seven items employing a five-point Likert-scale and inquiring into their beliefs and opinions about mathematics and the task they had just solved. Furthermore, students were requested to name the school subjects which they liked the best and another one which they preferred the least. The questionnaire was used from Grade 4.

Procedures and Analysis

The teachers who supervised the survey in a class received a written information sheet before the survey was conducted. This sheet detailed for them how to react if a question arose while the students were working. A general rule was to avoid any help and give only neutral answers such as to read the text again. The teachers were asked not to discuss the solutions with the students before the next day. Before the task began, the teachers briefed the students about the survey based on the information provided in the aforementioned sheet.

The students first received the word problem sheet with the task and could work on it for ten minutes, then they handed it in and received and filled out the questionnaire.

Students’ answers to the word problem were coded according to the coding system developed by our Flemish colleagues (Verschaffel, De Corte, & Lasure, 1994, see Table 1). The coding process was done by two graduate students who were previously informed and trained in the coding process by the first author of this paper. The code of the answers was in most of the cases obvious and unambiguous for the coding persons; relatively few problems were discussed with the first author.

Amióta Pistiék új lakásba költöztek, hetente kapja a zsebpénzét, 1000 Forintot, amit azóta mindig félre is tesz.

Hány napja laknak ott, ha már 35 000 Forintot gyűjtött így össze?

Indokold válaszodat!

$$\frac{35000}{1000} = 35 \Rightarrow 35 \text{ hét} \Rightarrow 245 \text{ nap}$$

Visszant a feladat nem említi, hogy ez a hetente történő fizetés pontosan hogyan jörszódik le. Pl.: tegyük fel hogy minden hét 1. napján kapja meg a pénzt (Hétfő) de mivel nincs megadva a költözés pontos dátuma ezért lehetséges hogy pl (kedd) * érkeztek és csak az ezt (hétfő) követő hét hétfőjén kapta meg először a "fizetését" ilyen vagy ehhez hasonló esetekben az érték változhat. 1 dolog csak a biztos misztig az, hogy Pistibet 35-öszer fizették ki.

* minden

Figure 1. Illustration of a realistic answer showing hesitancy about the uniqueness of the solution

Table 2. Relative frequencies (%) of students' answers to the first word problem (king's valet version) (please consult **Table 2** for a detailed description of the categories)

Answer code	Grade			
	2	3	4	6
0	9.4	1.5	0.9	0
1 (expected)	30.4	62.0	78.8	80.5
2 (expected, but with calculation error)	2.9	1.5	0.9	0.8
3 (realistic)	0	0	0.4	3.1
4 (miscellaneous)	57.3	35.0	19.0	15.6

In several cases, there appeared to be some hesitancy about the unambiguity or solvability of the tasks, despite having provided a computationally correct answer being labelled as a Code #1 answer. **Figure 1** presents the answer from a tenth grade student who first gave a technically correct answer and then expressed his hesitancy about the uniqueness of the solution.

Consequently, in line with Verschaffel, De Corte and Lasure's (1994) coding strategy, a second code was given to each answer according to whether the answer showed some hesitancy in performing the straightforward operation due to the activation of real-world knowledge. An answer is called "realistic reaction" if either it has Code #3 as the main code or the second code indicates realistic considerations. The term "realistic answer" is reserved for Code #3 type answers in the forthcoming analyses.

RESULTS

Study 1

Our first research question concerned whether there is a developmental trend in the main answer code. **Table 2** shows the relative frequencies of the main answer codes according to school grades.

In parallel with the growing rate of "expected answers," there is a slight increment in the realistic answers, while the rate of computational errors and other types of answers is demonstrating a decrease. Since too few students could provide a realistic answer to the word problem used in Study 1, this subsample is too small to analyze from the point of view of whether there are connections between math-related beliefs and realistic answers.

Study 2

Table 3 shows the distribution of students' solutions to the second word problem used in Grades 6, 8 and 10.

Table 3. Relative frequencies (%) of students' answers to the first (?) word problem (pocket money version (please consult **Table 1** for a detailed description of the categories)

Answer code	Grade		
	6	8	10
0	0	0	1.4
1 (expected)	66.0	80.7	90.0
2 (expected, but with calculation error)	3.8	1.7	0.7
3 (realistic)	0.6	4.2	3.6
4 (miscellaneous)	29.6	13.4	4.3

Table 4. Comparing Likert-scale mean (SD in parentheses) values from the RR group and the other students

	RR group (N = 35)	Others (N = 509)
Mathematics is important in life.	4.17 (0.86)	4.09 (1.00)
I like mathematics.	3.60 (1.22)	3.35 (1.22)
For scientists, mathematics is very important.	4.63 (0.60)	4.58 (0.72)
Mathematics is one of the most important school subjects.	3.91 (0.89)	3.92 (1.04)
I only like those tasks which can be solved quickly.	3.00 (1.19)	3.11 (1.20)
In the task we have just solved, everything was given for an unambiguous solution.	2.40* (1.50)	4.02 (1.22)
I liked the task we have just solved and found it interesting.	2.86* (1.24)	3.35 (1.30)

Note. Significant differences are marked with an *

The same tendencies have been shown as regards the second word problem. Please note that it contained much larger figures and instead of a trivial division by 1, here a division by 1,000 may and should have occurred. Notwithstanding, almost the same relative frequencies are observable. For answer code #3, the absolute frequency values are 1, 5, and 10, respectively.

In their original article, Verschaffel, De Corte and Lasure (1994) used the term "realistic reaction" to answer patterns that have either answer code #3 or which showed some hesitancy based on real-world knowledge when eventually executing the expected operation. The number of students whose answers could be categorized as realistic reaction were 5, 8 and 24 in Grades 6, 8 and 10, respectively. Therefore, in Study 2, a subsample of 37 students (RR group) can be compared to the others. Due to some missing questionnaires in both groups, **Table 4** provides descriptive statistics for two subsamples with 35 and 509 students.

The two groups of students generally have similar beliefs in connection to the importance of mathematics; their attitudes toward school subjects and quickly-solvable tasks are similar as well. However, strikingly marked differences were found in the opinion regarding whether the tasks they had just previously solved contained all the necessary data for an unambiguous solution. Of course, those who provided either a realistic answer or showed some hesitancy in executing the expected straightforward operations disagreed with this penultimate statement. More interestingly, even though they were able to provide a realistic reaction to a task that was thought to be interesting to them (at least on the part of the researchers), they did not show any enthusiasm towards the word problem. The non-RR group of students found the task significantly more interesting. As for the two types of RR reactions, students with a Code #3 RR answer could be compared to the students with other kinds of RR. The only significant difference between these two groups concerned the fourth statement on the importance of mathematics as a school subject. Students who showed hesitancy based on real-world knowledge when solving the task by means of executing straightforward operations consider mathematics much less important than the Code #3 RR group.

The connection between students' school marks in mathematics and their answer patterns was analyzed by means of ANOVA. Math marks are given on a five-point scale, 5 being the best and 1 is the worst. The mean (SD in parentheses) of the Code # 3 RR group (N = 16), the other RR group (N = 17), and the rest of the sample (N = 487) were as follows: 4.56 (0.51); 4.06 (0.90); 3.82 (1.06). ANOVA with Tukey post-test, $F(2, 517) = 4.26$, $p = .015$, revealed that there is a significant difference between the smallest and largest mean values, but not between these two and the middle ones.

Study 1 and 2 Databases Joint

There are two research questions for which the two studies are to be used together. As we have discussed in the Methods section of this paper, about half of the sixth-grade students solved the same version of the task as the younger children, whereas the others solved the second version used in upper school grades. As **Table**

2 and **3** indicated, the two sub-groups of sixth-grade students differed in how frequently the answer patterns appeared. These differences (if any of them is significant either in the statistical sense of the term or from a mere educational perspective) can be attributed to the difference in the magnitude of numbers in the task, or to the difference in the execution of arithmetic operations. Statistically speaking, the distributions differ significantly (Chi-square = 28.44). Obviously, this significant difference is due to the increase of technical errors in the calculation (Code #2) and the increase of Code #4 answers which indicate a diversity of unusual solutions that occurred in the second version of the task. Additionally, the rate of realistic answers dropped from 3.1% to 0.6% while the rate of the flawlessly calculated non-realistic answers dropped from 80.5% to 66.0%. All these changes in the answer pattern can be readily explained by the more demanding calculations due to the larger figures in the word problem and the one extra step of division by 1,000 instead of a literally non-executed division by 1.

Furthermore, we focused on whether there were some schools or classes which demonstrated a “biased” answer pattern, i.e. where there were more frequent realistic answers than elsewhere. This could either be due to some unfair collaboration between the students or might be attributed to a different local classroom culture (didactical contract à la Brousseau) where student dare to question the superficial word problem-solving strategy. Out of the 44 students who provided a realistic reaction, 12 was exceptional in his or her class with this kind of solution. There were only three classes out of which at least four students gave realistic reactions; two in Grade 10, and one on Grade 8. Nonetheless, in all these three classes they formed a minority compared to those whose answer could be categorized as non-realistic. This suggests that even if there might have been some unfair collaboration among students, this did not lead to a uniform solution strategy on the class-level.

As for the possibility of the influential role of the classroom culture, an interesting trend was revealed. In one of the Grade 10 classes with 4 realistic answers, students belonging to the group of “good mathematicians” (as indicated by good mathematics marks for all of them). In the other Grade 10 class where there were six realistic answers, these students had fairly low math marks that were slightly below the classroom average.

DISCUSSION

Our results show the development (or lack thereof) in students’ realistic answers. Albeit **Table 2** and **3** present the main answer codes (and not the realistic reaction as defined by Verschaffel, De Corte and Lasure, 1994), the main message of the frequency values is that providing a response to a routine-like word problem other than using the habitual superficial solution strategy is more likely to occur at a later stage in development. The drawback in sixth-grade students’ answers caused by the larger figures in the text of the second version of the word problem indicates the difficulties of the calculations. Our results may indicate that larger figures in the text of the word problem did not alert the students to the possibility of providing realistic answer to the word problems.

The main novel aspect of our research lies in its wide-span cross-sectional research design. The phenomenon of students’ bias towards using superficial word problem-solving strategies has been well documented in various age-groups. The ratio of students who followed a superficial solution strategy and executed one or two arithmetic operations increased from grade to grade in both samples.

The wide-span cross-sectional design suffers from one weakness that could not be eliminated: despite being structurally equivalent, the two tasks (from grade 2 to grade 6, and from grade 6 to grade 10) were not completely identical. Another type of limitation arises from the representativeness of the sample. Since students may change their school types at several points in the Hungarian educational system (after grade 4, 6 and 8) the populations from which the samples were selected are not necessarily equivalent. Consequently, there is a limitation in the generalizability of our results, i.e. the current study should be followed by large sample longitudinal investigations.

Given that several groups of stakeholders (policy makers, lay people, teachers) seem to be unhappy with the missed opportunity for developing flexible, genuine word problem modelling skills (as opposed to teaching the aforementioned superficial strategies) in schools, we suggest considering the following two-step process. First, stakeholders should agree upon what makes different word problem-solving strategies valuable or superficial. This issue further leads to philosophical questions on equity (Khisty & Chval, 2002) and power (Kolloosche, 2014) in mathematics education. Second, since primary and secondary teachers do not necessarily have the prerequisite knowledge and skills to apply and teach genuine modeling in word problem-solving, professional development trainings may and should address this second issue. Nevertheless, lacking the

agreement that genuine modeling (as opposed to superficial word problem-solving strategies) must be taught to and mastered by everyone, teachers themselves seem not to feel the need to change their current beliefs and practice.

In a follow-up study we are now working on a pre-service teacher training program that might change both pre-service teachers' beliefs and future practices.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Gabriella Ambrus – Mathematics Teaching and Education Centre, Institute of Mathematics, Faculty of Science, ELTE Eötvös Loránd University, Hungary.

Eszter Herendiné Kónya – Institute of Mathematics, University of Debrecen, Hungary.

Zoltán Kovács – Institute of Mathematics and Science, University of Nyíregyháza, Hungary.

Judit Szitányi – Department of Mathematics, Faculty of Primary and Pre-School Education, ELTE Eötvös Loránd University, Hungary.

Csaba Csíkos – Department of Mathematics, Faculty of Primary and Pre-School Education, ELTE Eötvös Loránd University, Hungary.

REFERENCES

- Ambrus, G., & Szűcs, K. (2016). Fehleranalyse beim Lösen von offenen Aufgaben- Ergebnisse einer empirischen Studie in der Grundschule, *Teaching Mathematics and Computer Science*, 14, 83-113. <https://doi.org/10.5485%2Ftmcs.2016.0420>
- Ambrus, G. (2016). The Pocket Money Problem. In A. Kuzle, B. Rott, & T. H. Čadež (Eds.), *Problem Solving in the Mathematics Classroom – Perspectives and Practices from Different Countries* (pp. 49-59). Münster: WTM Verlag.
- Baruk, S. (1989). *Wie alt ist der Kapitän? Über den Irrtum in der Mathematik*. Basel: Birkhäuser. <https://doi.org/10.1007/978-3-0348-6354-4>
- Blum, W., & Borromeo Ferri, R. (2009). Mathematical modelling: Can it be taught and learnt? *Journal of Mathematical Modelling and Application*, 1, 45-58.
- Csíkos, C., Andras, S., Rausch, A., & Shvarts, A. (2019). Mathematical learning and its difficulties in Eastern European countries. In A. Fritz et al. (Eds.), *International handbook of mathematical learning difficulties* (pp. 145-163). Springer:Cham, Switzerland.
- 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. <https://doi.org/10.1007/s11858-011-0308-7>
- English, L., & Sriraman, B. (2010). Problem solving for the 21st century. In B. Sriraman & L. English (Eds.), *Theories of mathematics education. Seeking new frontiers* (pp. 263-290). Heidelberg-Dordrecht-London-New York: Springer. https://doi.org/10.1007/978-3-642-00742-2_27
- 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 (Eds.), *Modelling and mathematics education. ICTMA 9: Applications in science and technology* (pp. 300-310). Chichester, UK: Horwood. <https://doi.org/10.1533/9780857099655.5.300>
- Gravemeijer, K., & Terwel, J. (2000). Hans Freudenthal: a mathematician on didactics and curriculum theory. *Journal of Curriculum Studies*, 32, 777-796. <https://doi.org/10.1080/00220270050167170>
- Hiebert, J., Carpenter, T. P., Fennema, E., Fuson, K., Human, P., Murray, H., Olivier, A., & Wearne, D. (1996). Problem solving as a basis for reform in curriculum and instruction: The case of mathematics. *Educational Researcher*, 25, 12-21. <https://doi.org/10.3102/0013189x025004012>

- Jupri, A., & Drijvers, P. (2016). Students difficulties in mathematizing word problems in algebra. *Eurasia Journal of Mathematics, Science & Technology Education*, 12, 2481-2502. <https://doi.org/10.12973/eurasia.2016.1299a>
- Kaiser, G., & Sriraman, B. (2006). A global survey of international perspectives on modelling in mathematics education. *ZDM*, 38, 302-310. <https://doi.org/10.1007/bf02652813>
- Khisty, L. L., & Chval, K. B. (2002). Pedagogic discourse and equity in mathematics: When teachers' talk matters. *Mathematics Education Research Journal*, 14, 154-168. <https://doi.org/10.1007/bf03217360>
- Kollosche, D. (2014). Mathematics and power: An alliance in the foundations of mathematics and its teaching. *ZDM*, 46, 1061-1072. <https://doi.org/10.1007/s11858-014-0584-0>
- Peled, I., & Balacheff, N. (2011). Beyond realistic considerations: modeling conceptions and controls in task examples with simple word problems. *ZDM*, 43, 307-315. <https://doi.org/10.1007/s11858-011-0310-0>
- Pollak, H. O. (1969). How can we teach applications of mathematics? *Educational Studies in Mathematics*, 2, 393-404.
- Pollak, H. O. (1979). The Interaction between Mathematics and Other School Subjects. In UNESCO (Ed.), *New Trends in Mathematics Teaching IV*. (pp. 232-248). Paris: UNESCO.
- Prediger, S. (2010). Über das Verhältnis von Theorien und wissenschaftlichen Praktiken – am Beispiel von Schwierigkeiten mit Textaufgaben, *Journal für Didaktik der Mathematik* 31, 167–195. <https://doi.org/10.1007/s13138-010-0011-1>
- Ramharter, E. (2002). Wie alt ist der Kapitän wirklich? - Über die Bewährung der Mathematik an der Realität oder Wittgenstein, Regeln und der Mathematikunterricht. *Mathematica Didactica*, 25, 37-53.
- 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. [https://doi.org/10.1016/s0959-4752\(97\)00014-5](https://doi.org/10.1016/s0959-4752(97)00014-5)
- Schukajlow, S. (2011). *Mathematisches Modellieren. Schwierigkeiten und Strategien von Lernenden als Bausteine einer lernprozessorientierten Didaktik der neuen Aufgabenkultur*. Münster: Waxmann.
- Szendrei, J. (2007). When the going gets tough, the tough gets going problem solving in Hungary, 1970–2007: research and theory, practice and politics. *ZDM*, 39, 443-458. <https://doi.org/10.1007/s11858-007-0037-0>
- Treffers, A. (1993). Wiscobas and Freudenthal realistic mathematics education. *Educational Studies in Mathematics*, 25, 89-108. https://doi.org/10.1007/978-94-017-3377-9_6
- Varga, T. (1988): Mathematics Education in Hungary today. *Educational Studies in Mathematics*, 19, 291-298. <https://doi.org/10.1007/bf00312449>
- Verschaffel, L., & De Corte, E. (1997). Word problems: A vehicle for promoting authentic mathematical understanding and problem solving in the primary school? In T. Nunes & P. Bryant (Eds.), *Learning and teaching mathematics: An international perspective* (pp. 69–97). Hove, UK: Psychology Press.
- Verschaffel, L. De Corte, E., & Lasure, S. (1994). Realistic considerations in mathematical modeling of school arithmetic word problems. *Learning and Instruction*, 4, 273-294. [https://doi.org/10.1016/0959-4752\(94\)90002-7](https://doi.org/10.1016/0959-4752(94)90002-7)
- Verschaffel, L., De Corte, E., Lasure, S., Van Vaerenbergh, G., Bogaerts, H., & Ratinckx, E. (1999). Learning to solve mathematical application problems: A design experiment with fifth graders. *Mathematical Thinking and Learning*, 1, 195-229. https://doi.org/10.1207/s15327833mtl0103_2
- Verschaffel, L., Van Dooren, W., Greer, B., & Mukhopadhyay, S. (2010). Reconceptualising word problems as exercises in mathematical modelling. *Journal für Mathematik-Didaktik*, 31, 9-29. <https://doi.org/10.1007/s13138-010-0007-x>
- Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for research in mathematics education*, 458-477. <https://doi.org/10.2307/749877>
- 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. [https://doi.org/10.1016/s0959-4752\(97\)00007-8](https://doi.org/10.1016/s0959-4752(97)00007-8)