IMPACT OF ORDER OF DATA IN WORD PROBLEMS ON DIVISION OF A WHOLE INTO UNEQUAL PARTS

Jarmila Novotná[1]✉, Martin Chvál[2]

✉[1]Faculty of Education, Charles University, Prague, Czech Republic, jarmila.novotna@pedf.cuni.cz
[2]Faculty of Education, Charles University, Prague, Czech Republic

Highlights
- *The presented results are of interest for mathematics education as well as for construction of tests*
- *Item response theory was used for evaluating differences in the difficulty of problems*
- *Analyses of pupils' solutions allow to get deeper insight into pupils' difficulties when solving word problems*

Abstract

The paper investigates how the order of numerical data in word problems on division of a whole into unequal parts affects achievement and reasoning of 14-16-year old pupils. The variable was altered in two word problems, in one of which also the context was changed (psychological variable) and in the latter "if-clause" is or is not used (linguistic variable). 182 pupils were involved in the experiment. The solutions were analysed quantitatively using Item Response Theory as well as qualitatively. The data suggest that pupils' success is affected by the order of numerical data in the statement in an unfamiliar context. The presence of "if-clause" in the statement was studied in a two-level problem. The order of numerical data played its role in case of formulations without "if". The results of the experiment are of interest for mathematics education as well as for construction of tests. The paper is an extended version of the paper by Novotná (2018).

Keywords

Context, "if-clause" formulation, mathematics, order of numerical data, word problems

Introduction

The area of word problems has been identified as an area of concern in many countries because of pupils' difficulties while solving them. Teachers and researchers report that word problems are one of the areas in which pupils show the poorest performance, see e.g. (Hembree,1992; Zohar and Gershikov, 2008). Wijaya, van den Heuvel-Panhuizen and Doorman (2015) give recommendations on how to improve the opportunities-to-learn while solving word problems as well as on how to conduct further research on this topic.

Novotná and Vondrová (2017) investigated the impact of context on Grade 8 pupils' choice of strategies when solving two missing value word problems with the same mathematical model but different contexts. Vondrová, Novotná and Havlíčková (2019) studied the impact of the order of numerical data, context, position of the unknown transformation and the length of the text in an additive word problem on the performance and reasoning of primary pupils from grades 4 and 5.

The study described here analyses only one type of word problems (problems on division of a whole into unequal parts) in terms of the impact of the variable "the order of numerical data in the word problem statement". The study is a part of a wider research within the project of the Grant Agency of the Czech Republic aimed at investigating variables influencing the difficulty of word problems. The project is unique at least in the Czech context because the research is conducted in cooperation of researchers in linguistic, mathematical education and psychology. This enables the project team to investigate the problem from different standpoints.

In this paper, the result of coding of word problem statement is called the *legend* (Novotná, 2010). The order of numerical data in the problem statement is called *proper* if the data appear in the sequence needed to solve the problem, otherwise it is called *mixed* (Vondrová, Novotná and Havlíčková, 2019). For the purposes of this paper, the *order of information recorded in a legend* is called *direct* if the data appear in the order used in the problem statement, otherwise it is called *indirect*. It is obvious that a proper order of numerical data does not imply that the solver records it in the proper legend and vice versa.

Searle, Lorton and Suppes (1974) reported that the order of data was a major predictive variable for pupils from Grades 4, 5 and 6 in problem solving. Similarly, Hembree's (1992) pointed better performance in case of proper order of data. Neither reference specifies how exactly the order of data was changed, though. On the other hand, Nesher, Hershkovitz and Novotná (2003) found that the effect of the order of presenting comparison relations was negligible in simple comparison problems.

Vicente, Orrantia and Verschaffel (2008) altered both mathematical and situational variables in word problems. The mathematically difficult problems included a combination of two inconsistent change situations; in the situationally difficult problems, the information about the initial moment of the action string was given at the end of the text. They found statistically significant differences in the performance of pupils in Grades 3 to 5 for both variables with a larger effect size coming from the mathematical difficulty level.

Soneira, González-Calero and Arnau (2018) analyse the solvers' translation of problem statements from natural language into language of algebra. They report that "the student performs a translation from natural language into algebraic language keeping the same order of the symbols in the equation as the key words in the statement, regardless of the meaning of the expressions" (2018: 44).

The paper answers two research questions: How does the order of numerical data in the problem statement in problems on division of a whole into three unequal parts influence the achievement

and reasoning of Grade 9 (age 14-16) pupils? Do the solving strategies used by these pupils vary according to the numerical order of data in the statement? Attention in the analyses is also be paid to the impact of change in context or wording from the linguistic perspective on the choice of solving strategies.

Materials and Methods

Participants

As Novotná and Vondrová (2017: 280) report, pupils participating in the research are from four primary schools, which were purposefully sampled within the GAČR project (focusing on investigation of parameters influencing the difficulty of word problems). The schools were selected on the basis of Reports by the Czech School Inspection and of websites of the schools in with the aim of selecting schools with no specialisation, of medium size, attended by children from their immediate surroundings with a varied socio-economic background, with the percentage of children/foreigners not exceeding the average for the whole Czech Republic, not founded for children with special needs and placed in the area of outer Prague. An important criterion for inclusion of the school in the GAČR project was that the whole school would get involved. The *sample* used in the part of research presented here consists of 182 Grade 9 pupils. No selection of pupils was made; all the classes from the four schools of the same grade participated.

Materials

The pupils were given initial tests in mathematics and Czech language that enabled division of classes into equally abled groups. These groups were later given different versions of the same word problem differing in the order of numerical data in the problem statement or in the context or the presence of "if-clause". For the purpose of this text and based on our research questions, two word problems (labelled 9A and 9C in the following text) were selected. Both work with division of a whole into unequal parts where the whole and the relationships between the parts are given. In both problems, the variants are based on the same mathematical model. In Table 1 and 2, the statements are presented. In both, variants 1 and 3 have the same order of numerical data and variants 2 and 4 are their modifications in another context (9A1 versus 9A3 and 9A2 versus 9A4) or with another wording (9C1 versus 9C3 and 9C2 versus 9C4).

Familiar context	Unfamiliar context
9A1 Gymnastics, basketball and swimming clubs take place at the same time on Wednesday evening. Basketball club is attended by three times more children than gymnastics, swimming is attended by 114 more children than basketball. In total the three clubs are attended by 380 children. How many children are enrolled in each of the clubs?	**9A3** There are three political clubs in the parliament: Liberal, Conservative and Green. Each member of parliament can be member of only one of the clubs. The Liberal Club has three times more members than the Conservative Club, the Green Club has 114 more members than the Liberal Club. In total there are 380 members in all the three clubs. How many members are there in each of the clubs?
9A2 380 children are enrolled in three clubs: gymnastics, basketball and swimming. All these clubs take place at the same time on Wednesday evening. Basketball club is attended by three times more children than gymnastics, swimming is attended by 114 more children than basketball. How many children are enrolled in each of the clubs?	**9A4** 380 members of a parliament in a foreign country is divided into three political clubs: Liberal, Conservative and Green. Each member of parliament can be member of only one of the clubs. The Liberal Club has three times more members than the Conservative Club, the Green Club has 114 more members than the Liberal Club. How many members are there in each of the political clubs?

Table 1: Word problem 9A for Grade 9 in four variants (change of the order of numerical data in the statement and change of the context)

Formulated with "if" (linguistic variable)	Formulated without "if"
9C1 Students Kamil, Eva and David spent part of their holiday making money as tour guides at the Děčín palace. They earned 8 800 CZK, which they divided with respect to the number of tours. For one tour they earned 40 CZK. How did they divided the money if Kamil had 12 fewer tours than Eva and David had twice as many tours as Kamil?	**9C3** Students Kamil, Eva and David spent part of their holiday making money as tour guides at a palace. They earned 8 800 CZK. They divided this money with respect to the number of tours. For one tour they earned 40 CZK. Kamil had 12 fewer tours than Eva and David had twice se many tours as Kamil. How did they divide the money?
9C2 Students Kamil, Eva and David spent part of their holiday making money as tour guides at a palace. For one tour they earned 40 CZK. They divided the earned money with respect to the number of tours. How did they divide the money if Kamil had 12 fewer tours than Eva and David had twice as many tours as Kamil and in total they earned 8 800 CZK?	**9C4** Students Kamil, Eva and David spent part of their holiday making money as tour guides at a palace. For one tour they earned 40 CZK. They divided the earned money with respect to the number of tours. Kamil had 12 fewer tours than Eva and David had twice as many tours as Kamil. In total they earned 8 800 CZK. How did they divide the money?

Table 2: Word problem 9C for Grade 9 in four variants (change of the order of numerical data in the statement and change of formulation of the assigned conditions)

Note: The order of numerical data in the assignment, or context, or "if" formulation represents one of the mathematical, or psychological, or linguistic variables investigated in the GAČR project.

A brief analysis *a priori* in terms of solving strategies for both problems is presented below. Novotná et al. (2013) give a list of heuristic solving strategies used for this analysis. Both "school" algebraic and arithmetic solving strategies are appropriate. The following heuristic strategies are the most suitable: Guess – check – revise, Systematic experimentation, Solution drawing (see also Přibyl, Eisenmann, 2014).

There are commonalities as well as differences in Problem 9A and 9C statement structures. In both problems, the whole is known, there are three parts and one relationship between parts is multiplicative, the other additive. Part 2 is expressed in terms of Part 1 and Part 3 in terms of Part 2. There are two main differences between Problem 9A and 9C statements:

- In Problem 9A, both relationships are expressed by "more than". In Problem 9C, the additive relationship is expressed by "fewer than". Several studies suggest that the word "more" is comprehended easier than the word "less" (Riley and Greeno, 1988).
- Problem 9C can be solved either on the level of (1) the number of tours or (2) the earned CZK. In both cases, additional calculations must be done. (1) – the total number of tours must be calculated from the total amount of money and the price of one tour. (2) – in calculations, the money awarded for one tour must be used.

Design of test booklets

As stated above, each problem is in four variants that differ from each other in certain parameters. Also the sequence of problems in the test booklet (the order of problems and the connection to previous problems) can affect pupils' performance (both the success rate and the selected solving procedure).

The requirement was to make the variants of the test booklets relatively equally difficult and to respect general recommendations for creation of tests, i.e. that problems at the beginning of the test should be relatively less difficult. Also the time demandingness of all test booklets should be the approximately the same. That is why the individual problems had been piloted before the main testing. In the piloting, their time demandingness, their difficulty, presence of pupils'

mistakes were studied. Based on these findings, we stated the total number of problems to be included in the test to make sure that ordinary pupils (not special needs pupils and pupils whose mother tongue was not Czech) could solve the test in 40 minutes without feeling any time pressure. The test booklets for the 9th grade, which included the problems analysed in this paper, consisted of the total of 6 problems.

When planning administration of the test booklets, we had to take into account that the variants of one problem are very similar at the first sight and always have the same result. That is why each pair of pupils sitting in one desk needed such a variant of the test booklet that looked very different from the other.

There were relatively many demands on the design of the test booklets. Based on these demands the following design of test booklets was created trying to respect all the demands (Table 3). One of them was the pupils' seating plan during testing.

Each variant of a problem was in a different test booklet. This implies that four different text booklets were created. Other four variants were created by changing the order of problems, see Tables 4 and 5.

One of the variants of each problem is the basic one. Changes in parameters usually make this basic problem cognitively more demanding. Sometimes a change of parameters makes the problem wording longer. That is why none of the test booklets could consist only of basic variants (i.e. A1, B1 etc.).

Position of the problem in the test \ number of test booklet	T1	T2	T3	T4
Problem 1	A1	A2	A3	A4
Problem 2	B4	B1	B2	B3
Problem 3	C3	C4	C1	C2
Problem 4	D2	D3	D4	D1

Table 3: Design of text booklets

With respect to the fact that the impact of a particular parameter may be affected by the position of the problem in the test booklet, for each of the four test variants a complementary variant was created. The couple of complementary variants was labelled I and II, where II means inverse order of problems. This means that the total of eight variants of test booklets was created (Table 4). Problems A and D were problems that came out as easier in piloting.

	T1_I	T1_II	T2_I	T2_II	T3_I	T3_II	T4_I	T4_II
Problem 1	A1	D2	A2	D3	A3	D4	A4	D1
Problem 2	B4	C3	B1	C4	B2	C1	B3	C2
Problem 3	C3	B4	C4	B1	C1	B2	C2	B3
Problem 4	D2	A1	D3	A2	D4	A3	D1	A4

Table 4: Variants of test booklets

Namely in test booklets for the 9th grade, the sequence of problems is given in Table 5. Bold is used for problems we focus on in this study.

	T1_I	T1_II	T2_I	T2_II	T3_I	T3_II	T4_I	T4_II
Problem 1	8B1	9D2	8B1	9D3	8B2	9D4	8B2	9D1
Problem 2	**9A1**	9B3	**9A2**	9B4	**9A3**	9B1	**9A4**	9B2
Problem 3	**9C4**	8D1	**9C1**	8D2	**9C2**	8D1	**9C3**	8D2
Problem 4	8D1	**9C4**	8D2	**9C1**	8D1	**9C2**	8D2	**9C3**
Problem 5	9B3	**9A1**	9B4	**9A2**	9B1	**9A3**	9B2	**9A4**
Problem 6	9D2	8B1	9D3	8B1	9D4	8B2	9D1	8B2

Table 5: Variants of test booklets for 9th grade

The basic scheme of the seating plan following the rule "maximum difference between test booklets of pupils sitting next to each other" is in Table 6:

T1_I	T3_II
T2_I	T4_II
T3_I	T1_II
T4_I	T2_II

Table 6: Seating plan when solving the tests

Problems 8B and 8D were used as anchor problems in IRT (Item Response Theory) analyses (Lord, 1980; Van der Linden and Hambleton, 1997). Other means of anchoring were the results from initial tests and assignment of test booklets to pupils in two subsequent rounds of testing in the GAČR project, which is shown in the following Table 7. Our problems were included in testing 1. Testing 2 consisted of different problems than testing 1.

Pupils	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Initial test from Czech	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Initial test from mathematics	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Testing 1	T1	T1	T1	T1	T2	T2	T2	T2	T3	T3	T3	T3	T4	T4	T4	T4
Testing 2	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4	T1	T2	T3	T4

Table 7: Assigning test booklets to pupils

Methods of analysis

To answer the research questions, quartets of Problem 9A and 9C were posed. The study is of a mixed methodology design, consisting of quantitative and qualitative parts.

Quantitative analysis

IRT was used both for the division of the pupils into equally able groups, each of which was solving a different variant of the word problem, and for the quantitative interpretation of data. The pupils' written solutions were analysed. The scoring rubric was as follows: 0 points (no or wrong solution), 1 point (partially correct solution), 2 points (correct problem model with a numerical mistake), 3 points (correct solution).[1]

To analyse the parameters of problems, a two-parameter logistic

---

1   The aim was to assess pupils' understanding of the problem, not their ability to carry out numerical operations and thus we considered as correct all solutions in which an appropriate problem model was created.

model was used (Lord, 1980): $P_{ij} = \dfrac{1}{1 + e^{-a_i(\theta_j - b_i)}}$, where P_{ij} is the probability that a pupil j with the ability of Θ_j will solve the problem i correctly, a_i characterises problem discrimination and b_i its difficulty. To find the latent ability of pupils, a scale was put on a z-score with Bayes estimation EAP, in an iterative way using results in the Initial Test (for which a model graded for the total result in the test was used) and Test 1 (9A) and Test 2 (9C). This allowed compensation for any inconsistencies in terms of ability grouping which may have arisen at the beginning.

The situation is visually depicted in Figure 1. Θ_j is on the *x*-axis and the value of 0 means an average result in the initial test. On the *y*-axis we can read the probability with which the pupil of a certain ability will solve the task. For example, an average pupil will solve 9A2 with the probability approximately 0.6 (Figure 1 left), 9C2 with the probability approximately 0.14 (Figure 1 right).

Qualitative analysis

Finally, a qualitative analysis of the data was carried out. The pupils' written solutions were analysed carefully for mistakes and solving strategies with the main focus on the relationships between the variants of the problem and order of information recorded in pupils' legends. In this case, we created a spreadsheet in which the order of information in the legends was recorded for each variant of the two problems. The occurrence of frequent mistakes was linked with the use of the order of information in legends.

Results

In this section, we present the results gained from our experiments. The results are divided into sections focusing on quantitative and qualitative analyses.

Quantitative analysis using IRT

Table 8 presents the relative frequencies of point distributions and the average success rate for the four variants of Problem 9A. It seems that most difficult are the variants 9A4 and 9C4, while in the variants 9A1, 9A2, 9A3, and 9C1, 9C2, 9C3 average success rates are similar. The order of numerical data played its role in case of a less familiar context, resp. formulation without "if". This corresponds to our expectations. The results confirm that Problem 9C was more difficult than 9A in three variants while in variants 9A4, 9C4 the average success is comparable and low, which was surprising. This can be explained by the fact that in both cases the pupils came across some "intricacy" that made the use of school procedures more difficult. These results are further specified by IRT evaluation and verified by the statistical methods below.

Table 8 and Figure 1 present the parameters of the problems according to IRT. They add more detail to the information that we can read from the average success rate. In Figure 1, graphs on the left in suggest that 9A2 has the best discrimination among 9A variants. 9A1 and 9A2 discriminations are comparable. Below-average pupils solved the variant with an unfamiliar context better, while the average and above-average pupils were not much affected by the context. 9A3 discriminates the worst pupils. Graphs on the right suggest that 9C is much more difficult than 9A – an important number of above-average pupils failed to solve its variants. 9C4 has the best discrimination. 9C2 and 9C3 discriminations are comparable.

	N	Points				Average success rate
		0	1	2	3	
Familiar context						
9A1	46	37%	9%	2%	52%	57%
9A2	46	28%	15%	9%	48%	59%
Formulated with "if-clause"						
9C1	45	42%	33%	7%	18%	33%
9C2	46	30%	43%	7%	20%	38%
Unfamiliar context						
9A3	45	29%	18%	7%	46%	57%
9A4	45	36%	20%	7%	37%	49%
Formulated without "if-clause"						
9C3	46	46%	26%	11%	17%	33%
9C4	45	51%	38%	4%	7%	22%

Table 8: Results for 9A, 9C according to success rate
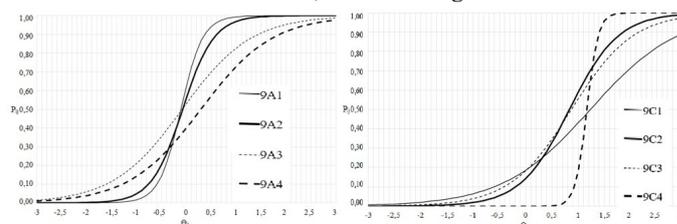
Figure 1: Two-parameter IRT model for the variants of 9A (left) and 9C (right)

Qualitative analysis

Solving strategies

The experiment was conducted with 182 pupils. Only 12 of them selected successfully as their solving strategy the use of an equation (7 used Guess – check – revise: 9A1 – 1 pupil, 9C1 – 1 pupil, 9C2 – 1 pupil, 9C3 – 3 pupils, 9C4 – 1 pupil). 6 pupils solved the problem arithmetically as a problem on division of the whole into three equal parts, which can be accounted for by their failure to understand the conditions in the text statement correctly (9A3 – 1 pupil, 9A4 – 1 pupil, 9C1 – 1 pupil, 9C2 – 2 pupils, 9C3 – 1 pupil).

Legends

In Problems 9A and 9C, the proper order of numerical data is (P1, P2, P3, W), where W describes the whole in all variants. In 9A1, 9A2 P1 refers to gymnastics, P2 basketball and P3

swimming whereas in 9A3, 9A4 P1 refers to the Liberal Club, P2 to the Conservative Club and P3 to the Green Club; in 9C P1 stands for Eva's, P2 for Karel's and P3 for David's number of tours. Table 9 and 10 present the distribution of direct and indirect legends in comparison to the proper order of numerical data in the statement. Recorded are only those cases where the pupil attempted to solve the problem and created a legend (independently of whether they solved the problem successfully). In some cases, the pupil used the whole, but they did not record it in the legend. The order of parts in legends that did not appear in the pupils' solution are not included in tables 9 and 10.

using the information from the problem statement and finally calculate the money earned by each student. The other is to work directly with earned money, then it is essential to supplement the relationship between parts on the award for one tour; and the results are in crowns. If pupils do not realize what approach they use, they may come to a wrong result.

Table 9

9A1			
Order of numerical data in the statement	**P2, P1, P3, W**		
	Position of W		Legend — Order of parts
	W as last: 30		*P2, P1, P3: 11*
	W as first: 4		P1, P2, P3: 20
			P2, P3, P1: 3
			P3, P2, P1: 0

9A2			
Order of numerical data in the statement	**W, P2, P1, P3**		
	Position of W		Legend — Order of parts
	W as last: 18		*P2, P1, P3: 11*
	W as first: 15		P1, P2, P3: 22
	Not recorded: 3		P2, P3, P1: 2
			P3, P2, P1: 1

9A3			
Order of numerical data in the statement	**P2, P1, P3, W**		
	Position of W		Legend — Order of parts
	W as last: 25		*P2, P1, P3: 29*
	W as first: 2		P1, P2, P3: 1
	Not recorded: 7		P2, P3, P1: 4
			P3, P2, P1: 0

9A4			
Order of numerical data in the statement	**W, P2, P1, P3**		
	Position of W		Legend — Order of parts
	W as last: 17		*P2, P1, P3: 26*
	W as first: 9		P1, P2, P3: 0
	Not recorded: 4		P2, P3, P1: 3
			P3, P2, P1: 1

Table 9: Distribution of direct and indirect legends for 9A

Table 9 shows that in case of proper order of numerical data in the statement most pupils use direct legend independently of familiarity of context. In case of mixed order of numerical data in the statement the situation is different for familiar and unfamiliar contexts: In case of an unfamiliar context, the proportion of pupils who changed the order of information in legend, albeit by moving W to the end or by swapping the order of recorded parts is much higher than in case of familiar context. As stated earlier in the paper, Problem 9C is more difficult than Problem 9A as it does not contain only the relationship of the whole to parts, but also the relationship between the money earned and number of tours. Thus there are two possible approaches to the solution. One is to find out how many tours were taken, find the number of tours of each of the students

Table 10

9C1			
Order of numerical data in the statement	**W, P2, P1, P3**		
	Position of W		Legend — Order of parts
	W as first: 7		*P2, P1, P3: 29*
	W as last: 21		P1, P2, P3: 3
	Not recorded: 4		P2, P3, P1: 0
Number of pupils who interchanged level incorrectly: 11			

9C2			
Order of numerical data in the statement	**P2, P1, P3, W**		
	Position of W		Legend — Order of parts
	W as first: 2		*P2, P1, P3: 28*
	W as last: 22		P1, P2, P3: 2
	Not recorded: 8		P2, P3, P1: 2
Number of pupils who interchanged level incorrectly: 21			

9C3			
Order of numerical data in the statement	**W, P2, P1, P3**		
	Position of W		Legend — Order of parts
	W as first: 10		*P2, P1, P3: 22*
	W as last: 13		P1, P2, P3: 2
	Not recorded: 5		P2, P3, P1: 4
Number of pupils who interchanged level incorrectly: 8			

9C4			
Order of numerical data in the statement	**P2, P1, P3, W**		
	Position of W		Legend — Order of parts
	W as first: 4		*P2, P1, P3: 29*
	W as last: 25		P1, P2, P3: 2
	Not recorded: 4		P2, P3, P1: 2
Number of pupils who interchanged level incorrectly: 25			

Table 10: Distribution of direct and indirect legends for 9C

Table 10 is analogical to Table 9 for Problem 9C. In addition the table shows how many pupils made mistakes when switching between the levels of the number of tours and money earned. Table 10 shows that most pupils used direct legend in variants 9C2 and 9C4; this can be explained by the order in which data are used in calculations. In the variants 9C1 and 9C3 the order of parts mostly corresponded to the problem statement but the information on the whole was often in a different place that in the statement. This could be caused by the two-level problem statement which requires from pupils to pay a lot of attention to all data from the beginning in order to understand and by their attempt to find out the total number of tours. The table also shows that in the variants with W at the end of the statement independently of the presence of "if-clause" pupils make more mistakes in switching between the number of tours and money earned. This could be explained by the fact that they can process the relationships between parts without paying attention to the money issue.

Analysis of mistakes that occurred in the solutions of problem 9A

The structure of problem 9A can be described by the equation $x + 3x + (3x + 114) = 380$. The problem can be solved in several ways, not only using an equation, but the 9th graders can be expected to be so familiar with linear equations that they will prefer this way of solving the problem.

Table 11 presents the mistakes pupils made when solving the problem. It does not contain erroneous solutions where the used solving procedure could not be identified (3 cases in 9A2, 1 case in 9A3 and 9A4). Apart from *Division of the whole into three parts* (procedure I), the given procedures in the variants were used only by one or two pupils. That is why we do not present the numbers in brackets in case of solving procedures II to XII.

I	Division of the whole into three equal parts	9A1 (3), 9A2 (4), 9A3 (3), 9A4 (3)
II	$x + 3x - 114 = 380$	9A1
III	$x + 3x + 3(x - 114) = 380$	9A1
IV	$x + x + 3x + (3x + 114) = 380$	9A2
V	$3x + x = 114 \times 3$	9A4
VI	$x + 3x + 114 = 380$	9A1, 9A4
VII	$3x + x + (x + 114) = 380$	9A1, 9A3, 9A4
VIII	$(380 - 114): 2$	9A4
IX	$3x + 114 = 380$	9A3, 9A4
X	$3x - 114 = 380$	9A3, 9A4
XI	$x + 3x + (3x + 114) = x$	9A2
XII	$3x - 114 + x + y = 380$	9A4

Table 11: Types of mistakes in variants of problem 9A

Table 11 shows that there were 12 different erroneous procedures in pupils' solutions. Only procedure I can be come across in all variants. Procedure VII was used in three variants, 3 procedures (VI, IX, X) were used in two variants.

The very list of erroneous procedures and variants in which they occurred shows clearly that pupils found variant 9A4 where 8 different wrong procedures were used the most difficult. In variants 9A3, 9A4, i.e. in variants with an unfamiliar context, the same erroneous procedure was used in two cases (IX, X). Procedures VI, VII can be come across once in variants with familiar context and once in a variant with unfamiliar context. Figure 2 shows an example of solution using procedure I (division of the whole into three equal parts).
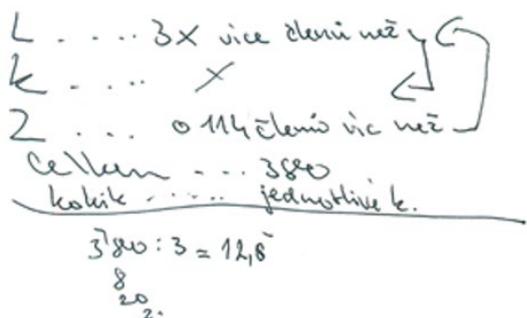
Figure 2: Variant 9A3, procedure I

Division of the whole into three equal parts is a mistake that is often to be come across in this type of problems. (Novotná, 2010) The strategy is used most often by those solvers who have problems with comprehending the text. Since they cannot grasp the relations between unequal parts, they look for salvation in division into equal parts.

The use of other erroneous procedures implies that the pupils did not give up on dividing the whole into unequal parts but failed to mathematize the conditions from the problem wording correctly. The number of solving procedures used signals that mathematization of conditions from the problem wording

was far from easy. Figures 3 and 4 show examples of pupils' solutions of variant 9A4 and illustrate how the pupil was trying to handle all relevant conditions unsuccessfully. Figure 5 shows a solution where the pupil managed to find the correct solution after several unsuccessful attempts.
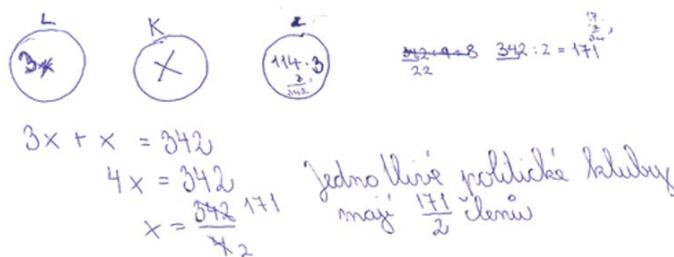
Figure 3: Variant 9A4, procedure V
(in Czech 114. 3 stands for 114 x 3)
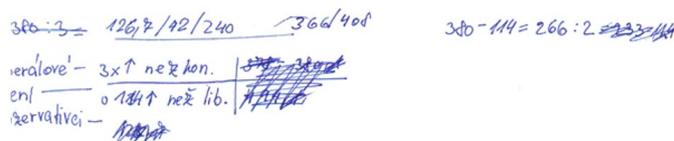
Figure 4: Variant 9A4, several unsuccessful attempts without reaching the correct procedure
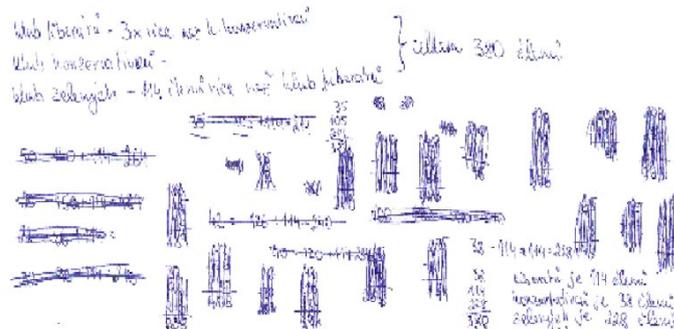
Figure 5: Variant 9A3, the correct procedure reached after several unsuccessful attempts

The last row in Table 11 shows the use of two unknowns in the legend and solution (see Figure 6). This is probably the result of the pupil's failure to mathematize all the relations between the sought values, in consequence of which he tried to map the situation using two different letters for two different unknowns. This is not a mistake. The problem is the pupil did not manage to express the relation between the two sought values correctly.
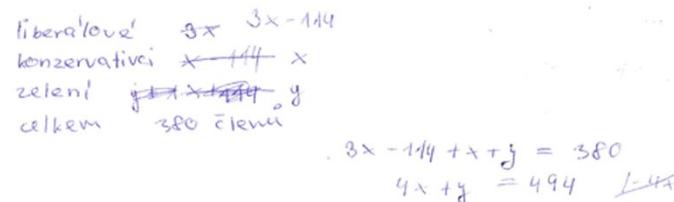
Figure 6: Variant 9A4, procedure XII

Table 12 shows the number of pupils who did not even start solving problem 9A.

Variant 9A1	Variant 9A2	Variant 9A3	Variant 9A4
10	4	10	7

Table 12: Numbers of pupils who did not solve problem 9A

Table 12 shows that the order of information had major impact on this indicator. As we mentioned already in the description of the foursome of variants of problem 9A, the order of information in the problem wording was the same in the couples of variants 9A1, 9A3 and 9A2, 9A4. At the same time the number of pupils

who did not start solving the problem in variants 9A1, 9A3 is significantly higher than in variants 9A2, 9A4. It seems the consequence of different order of information presented in the wording of problems 9A1, 9A3 was that these problems were much more difficult for pupils than variants 9A2, 9A4. A higher number of pupils who did not start solving in variant 9A4 than 9A2 implies that the unfamiliar context could have been relevant. However, in the research sample the same does not hold for variants 9A1, 9A3, where the number of pupils who did not start solving is balanced.

Discussion

This study of the impact of the order of presenting data in the wording of a mathematics problem brings new knowledge into the area of word problem solving. It comes out of an analysis of pupils' written solutions exclusively. Thus, the causes of the mistakes remain unclear in many cases, as written solutions do not provide enough information. If we want to get information on the impact of the order of data that is more precise, we should supplement the analysis of pupils' written solutions with interviews with the pupils. So far, several interviews with pupils in which the tested problems were discussed have been conducted with pupils not involved in the testing. These interviews reinforce the results gained from the testing. The gained results comply with the results in (Vicente, Orratia, Verschaffel, 2008), where a significant impact of the order of information in wording of additive problems with two transformations was detected.

Downton and Sullivan (2017: 303) indicate that 'posing of appropriately complex tasks may actually prompt the use of more sophisticated strategies'. This was not the case of Problems 9A and 9C. The success rate in Problem 9C was smaller than in Problem 9A (see Table 8). This can be explained by a higher level of complexity of Problem C. The used strategies were comparable.

An important issue is the dependence of the order of information in the wording of a problem on the solvers' age. In the project GAČR 16-06134S *Context problems as a key to the application and understanding of mathematical concepts*, the impact of the order of information is analysed with pupils from several age categories. So far, unpublished analyses show that the impact of the order of information decreases with growing age of the solvers. For example, analyses of solutions of 4th graders show that the differences when changing the order of data are statistically significant. This does not hold for 9th graders, which can be accounted for both by pupils' greater experience with problem solving, by their higher proficiency in reading but also by their use of more advanced mathematical solving strategies, namely equations. Setting up an equation or a set of equations describing the structure of the problem wording is characterized by parallel grasping of information from the wording (Novotná, 2010) when the solver first registers all relevant information and then creates the appropriate mathematical model. Then the order of information becomes less significant for successful solution of the problem.

As shown in the analysis of legends created by the tested pupils, not all pupils make a legend while solving. In fact, creation of a legend is coding of data from the problem wording in a new form that is more comprehensible or clearer to the pupil. Gelfman and Kholodnaya (1999) consider ways of coding data to be a subjective means that help us project the world around us into our own experience. This also holds for processing of data from the word problem wording. If the pupil is convinced that they grasp the structure of the problem, they do not need to visualize it and do not make a legend. However, if the structure of the problem wording is harder to grasp for any reason, creation of a legend may help the pupil get insight into the structure of the problem. A well-created legend can make solving of a problem easier for the solver, an inconveniently selected legend can make the solution much harder (Novotná, 2010).

Conclusions

The variables analysed here for 14-16-year old pupils are studied for pupils 8 to 16 in the GAČR project. The research in the GAČR project has important characteristics. The variations of the problems are carefully formulated so that only one variable was changed, the other being controlled (as requested by Daroczy et al. (2015)). Great care is paid to dividing pupils into equally able groups using IRT. The quantitative analysis is accompanied by a qualitative analysis to better understand the influence of investigated variables.

Analogical analyses of pupils' solutions in which various mathematical, psychological and language variables are changed allow us to get deeper insight into pupils' difficulties when solving word problems. The presented study has its limitations: a relatively small number of participants, one type of word problems, analyses based on written solutions only. However, it brought very interesting results important for mathematics education as well as test creation. The study brings more light to the role of word problems in mathematics education. As Mason (2018: 334) states: [Pupils should be taught to] 'recognise and express relationships in the situation using whatever support devices and modes of presentation are recommended for this purpose'.

Understanding the impact of the order of data on the difficulty of problems is important not only for teachers but also for authors of tests, curricular documents and textbooks. If problems with *mixed* order of data are used in teaching, due attention must be paid to efficient strategies allowing pupils to understand relations among them. Authors of tests as well as others should be aware of the fact that a change in the order of data in the problem wording might be seemingly small but might affect the difficulty of the problem and its discriminative property with an impact on the test reliability significantly.

Acknowledgement

References

Daroczy, G., Wolska, M., Meurers, W. D. and Nuerk, H. C. (2015) 'Word problems: a review of linguistic and numerical factors contributing to their difficulty', *Frontiers in Psychology*, vol. 6, no. 348, pp. 22–34. http://dx.doi.org/10.3389/fpsyg.2015.00348

Downton. A. and Sullivan, P. (2017) 'Posing complex problems requiring multiplicative thinking prompts students to use sophisticated strategies and build mathematical connections', *Educational Studies in Mathematics*, vol. 95, no. 3, pp. 303-328. http://doi.org/10.1007/s10649-017-9751-x

Gelfman, E. and Kholodnaya, M. (1999) 'The role of ways of information coding in student's intellectual development', *Proceedings of the First Conference on the European Society for Research in Mathematics Education*, vol. 2, pp. 38-48, [Online], Available: http://www.mathematik.uni-dortmund.de/~erme/doc/cerme1/cerme1_proceedings_part2.pdf. [15 Nov 2018].

Hembree, R. (1992) 'Experiments and relational studies in

problem solving: A meta-analysis', *Journal for Research in Mathematics Education*, vol. 23, no. 3, pp. 242-273. http://dx.doi.org/10.2307/749120

Lord, F. M. (1980) *Applications of item response theory to practical testing problems*, Hillsdale, N.J.: Lawrence Erlbaum Associates.

Mason, J. (2018) 'Structuring structural awareness: A commentary', *Building the foundation: Whole Numbers in the Primary Grades,* Cham: Springer International Publishing, pp. 323-340.

Nesher, P., Hershkovitz, S. and Novotná, J. (2003) 'Situation model, text base and what else? Factors affecting problem solving', *Educational Studies in Mathematics*, vol. 52, no. 2, pp. 151-176. https://doi.org/10.1023/A:1024028430965

Novotná, J. (2010) *Study of solving word problems in teaching of mathematics. From atomic analysis to the analysis of situations*, Saarbrücken: LAP LAMBERT Academic Publishing.

Novotná, J. (2018) 'Impact of order of data in word problems on division of a whole into unequal parts', *ERIE 2018*, Prague: CULS, pp. 257-264.

Novotná, J., Eisenmann, P., Přibyl, J., Ondrušová, J. and Břehovský, J. (2013) 'Heuristic strategies in problem solving in school mathematics', *ERIE 2013*, Prague: CULS, pp. 461-468.

Novotná, J. and Vondrová, N. (2017) 'Pupils' strategies for missing value proportional problems', *Proceedings of ERIE 2017*, Prague: CULS, pp. 279-286.

Přibyl, J. and Eisenmann, P. (2014) 'Properties of problem solving strategies', *Proceedings of ERIE 2014*, Prague: CULS, pp. 623-630.

Riley, M. S. and J. G. Greeno (1988) 'Developmental Analysis of Understanding Language about Quantities and of Solving Problems', *Cognition and Instruction*, vol. 5, no. 1, pp. 49-101. http://dx.doi.org/10.1207/s1532690xci0501_2

Searle, B.W., Lorton, P. Jr. and Suppes, P. (1974) 'Structural variables affecting CAI performance on arithmetic word problems of disadvantaged and deaf students', *Educational Studies in Mathematics*, vol. 5, pp. 371-384, http://dx.doi.org/10.1007/BF01424555

Soneira, C., González-Calero, J.A. and Arnau D. (2018) 'An assessment of the sources of the reversal error through classic and new variables', *Educational Studies in Mathematics*, vol. 99, no. 1, pp. 43-56. https://doi.org/10.1007/s10649-018-9828-1

Van der Linden, W.J. and Hambleton, R.K. (1997) *Handbook of Modern Item Response Theory,* New York: Springer, http://dx.doi.org/10.1007/978-1-4757-2691-6

Vicente, S., Orrantia, J. and Verschaffel, L. (2008) 'Influence of situational and mathematical information on situationally difficult word problems', *Studia Psychologica*, vol. 50, no. 4, pp. 337-356.

Vondrová, N., Novotná, J. and Havlíčková, R. (2019) 'The influence of solving and situational information on pupils' achievement in additive word problems with several states and transformations', *ZDM Mathematics Education,* pp. 1-15, https://doi.org/10.1007/s11858-018-0991-8

Wijaya, A., van den Heuvel-Panhuizen, M. And Doorman, M. (2015) 'Opportunity-to-learn context-based tasks provided by mathematics textbooks', *Educational Studies in Mathematics*, vol. 89, no. 1, pp. 41-65. https://doi.org/10.1007/s10649-015-9595-1

Zohar, A. and Gershikov, A. (2008) 'Gender and Performance in Mathematical Tasks: Does the Context Make a Difference?', *International Journal of Science and Mathematics Education*, vol. 6, no. 4, pp. 677-693, http://dx.doi.org/10.1007/s10763-007-9086-7