



HAL
open science

Word problems as a tool to diagnose pupils' difficulties in mathematics

Jana Slezáková, Darina Jirotková

► **To cite this version:**

Jana Slezáková, Darina Jirotková. Word problems as a tool to diagnose pupils' difficulties in mathematics. Thirteenth Congress of the European Society for Research in Mathematics Education (CERME13), Alfréd Rényi Institute of Mathematics; Eötvös Loránd University of Budapest, Jul 2023, Budapest, Hungary. hal-04422818

HAL Id: hal-04422818

<https://hal.science/hal-04422818>

Submitted on 28 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Word problems as a tool to diagnose pupils' difficulties in mathematics

Jana Slezáková¹ and Darina Jirotková¹

¹Charles University, Faculty of Education, Prague, Czech Republic;

jana.slezakova@pedf.cuni.cz

Word problems are used in school mathematics to apply mathematics to real life or develop problem-solving skills. Our paper presents word problems as a diagnostic tool to reveal pupils' difficulties in mathematics. The sample consists of 618 Grade 2 pupils who were assigned a mathematics test in which three word problems were included. Pupils' solutions were analysed in terms of the wrong results provided by them. Three categories of mistakes were distinguished (mistakes at the level of the mathematical model, mistakes due to the previous wrong results, and mistakes due to inattentiveness). The results that appear in pupils' solutions could be expected to be underlain by a pupil's rational, theoretically justifiable solving procedure and our experience as researchers and teachers. In this study, we provide possible explanations for procedures leading to these wrong results.

Keywords: Word problems, cause of mistake, pupils' difficulties.

Introduction

The study presented in the paper is part of the project *Teachers' Understanding of the Causes of School Failure and the Effectiveness of Pedagogical Interventions*, involving researchers from mathematics education, Czech language education, psychology, special education and elementary education. The project aims to identify pupils at risk of school failure and follow their teachers' work in the Czech language and mathematics lessons in Grades 2 and 3 to understand how it influences these pupils' achievements.

School failure is usually defined as a pupil's significantly below-average assessment in key (or all) subjects. In our project, we identified pupils at risk of school failure using their results in psychological tests, tests in mathematics and the mother tongue, and their teachers' appraisal. The study presented in the paper focuses on one part of the mathematics test assigned to pupils at the beginning of the project in Grade 2, consisting of three word problems (WPs). WPs were included in the test as they are an important part of the curriculum, and pupils' solutions to them can attest to their understanding of mathematics and ability to apply mathematical calculations in context. Thus, the paper aims to document how the selected WPs can be used as a diagnostic tool to diagnose pupils' difficulties in mathematics.

Theoretical Background

Research agrees that many pupils who can solve arithmetic tasks and have good reading literacy still fail to solve WPs (Daroczy et al., 2015). At least three processes are required for a successful solution of the WP: (i) linguistic understanding of the WP, (ii) carrying out the arithmetic calculation and (iii) understanding and relating the text and arithmetic calculation (Nesher & Teubal, 1975). According to Reusser (1990), a pupil uses several levels of comprehension. Text comprehension is followed by situation comprehension, which goes beyond the text. A situation model is constructed based on the pupil's real-world knowledge

and experience. Then the quantities and relationships between them are abstracted from the situation, and the mathematical model is made.

Our study deals with additive arithmetic WPs solvable by a simple arithmetic calculation (Daroczy et al., 2015). Four basic types of one-step WPs were distinguished (Riley et al., 1983) – change, compare, combine, and equalise – later subdivided into other types; Daroczy et al. (2015) mention 14 types of one-step addition and subtraction WPs. It was found that the semantic structure influences the difficulty of WPs, and different types of WPs were subject to much research. We will only mention some results relevant to our study. WPs comprising relations are significantly more difficult than those involving just quantities (Nunes et al., 2016). The difficulty depends on the position of the unknown in the text (Rosenthal & Resnick, 1974). The meta-analysis of research on WPs (Daroczy et al., 2015, Hembree, 1992; Verschaffel et al., 2020) confirmed that WPs in which events were mentioned out of chronological order (mixed order WPs) were more difficult to solve than the proper order WPs. The WPs with the starting unknown were more difficult and took longer than when the unknown was at the end of the text. Another widely researched variable is the inconsistent language in which the keywords refer to an operation opposite to the one required for the solution (e.g., Pape, 2003; Daroczy et al., 2015). WPs with inconsistent language are more difficult for pupils across age groups.

The keyword strategy (or a direct translation strategy) is usually mentioned as an important cause of pupils' failure in the WPs with the mixed order of data and inconsistent language (e.g., Nesher & Teubal, 1975). Young pupils especially tend to solve WPs by direct modelling (Nunes et al., 2016), thus avoiding the creation of the situation model.

One-step WPs are usually easier to solve than multiple-step ones in which a chain of two or more calculations is present; the result of the first calculation is needed to carry out the second, etc. (Verschaffel et al., 2020). Mental processing of these problems places high demands on working memory as the pupil has to retain information and relationships (Cowan, 2010), and there is a danger of cognitive overload (Sweller, 2010). Much research showed that WP solving is associated with general cognitive resources such as working memory capacity and inhibitory skills (Verschaffel et al., 2020).

WPs are used in school mathematics worldwide. They are seen, among others, as a way to apply mathematics to real life and to develop problem-solving skills. In our paper, we want to show WPs as a diagnostic tool by addressing the following questions: What mistakes do pupils make when solving selected WPs? What are their possible causes?

Methodology

The study, which is the focus of this paper, was conducted in Grade 2 in the school year 2020/2021. The schools were selected from several regions of the Czech Republic. All of them were ordinary primary schools. Teachers selected for the study were teachers with a good reputation with at least three years of teaching experience at an elementary school level. A total of 669 pupils from 29 classes from 22 schools were involved in the research.

A team of mathematics educators designed a didactic test from mathematics, consisting of six problems that covered the knowledge and skills expected in mathematics at the beginning of

the 2nd grade (word problem solving, numerical skills, patterns in number sequences, shapes on the grid). The test was planned for the autumn of 2020, but due to the covid lockdown, it was not administered until March 2021 (N = 618).

Research Tool

Part of the mathematics test consisted of three arithmetic WPs, graded in difficulty. We will present them with their *a priori* analysis.

- a) Ada is 7 years old. In how many years will she be 11?
- b) Daniel is 1 year older than Ada. In how many years will Daniel be 11?
- c) Ada is 2 years younger than Kuba. In how many years will Kuba be 11?

One-step WP a) belongs among the change problems. The unknown is the change (CH: “In how many years?”), and the start and result states ($S_S = 7$ years, $S_R = 11$ years) are given. The semantic structure is $S_S + \underline{CH} = S_R$. The unknown change is underlined. The order of information complies with the given structure.

WPs b) and c) are two-step WPs. The first parts of b) and c) are compare problems: $S_A + C = \underline{S_D}$; ($\underline{S_K}$). The initial state (Ada’s age $S_A = 7$) is given, but it must be taken from WP a). The operators (“1 year older”, “2 years younger”, $C = 1; 2$) are given. The result state is given ($\underline{S_D}$; $\underline{S_K}$) – Dan’s and Kuba’s ages. The second parts of the two WPs are change problems: $S_D + \underline{CH} = S_{RD}$, $S_K + \underline{CH} = S_{RK}$. While in b), the unknown is the compared quality in c), the referent quality is unknown. That is why c) is more difficult; it includes lexically inconsistent language. The word “younger” mostly refers to subtraction, while addition is needed for the correct solution.

Next, we will consider three components of difficulty: (i) the linguistic complexity of the WP text, (ii) the numeric complexity of the arithmetic calculation and (iii) the relationship between (i) and (ii) (Daroczy et al., 2015).

(i) The text is short and does not include factors complicating understanding the sentences (Daroczy et al., 2015). The question is formulated in the same way in all the WPs. A semantic factor makes the WPs more difficult: the problems belong among the dynamic WPs (Nesher, 1982) with two time points (‘now’, ‘then’). ‘Now’ means for Ada 7 years, Dan 8 years and Kuba 9 years. The time point ‘then’ (11 years) is included in the question. The pupil’s direct experience does not support the considerations about what will happen after some time; abstract thinking is required. The presence of ‘passing the time’, the process, and the fact that all the numbers include units of time make the visualisation more difficult. It might adversely affect the creation of the situation model, which is not supported by a static picture. The availability of models and/or reference objects and pictures makes WP solving easier, especially for younger pupils (Riley et al., 1983). The WPs in our study thus require that the pupil connects the process and concept (Riley et al., 1983; Gray & Tall, 1994).

The dynamic WPs on age are rare in Czech mathematics textbooks. We selected such WPs for the test so pupils cannot solve them by keyword or other superfluous strategies and must create a situation model to find a solution.

(ii) The numeric aspect of the WPs is not a complicating factor. Whole numbers up to 11 are required, which is appropriate for pupils at the beginning of Grade 2.

(iii) The WPs do not include semantic cues (Hembree, 1992; Nesher & Teubal, 1975), which would highlight semantic relationships and directly lead to the required operation. If the pupils do not make a situation model, they might focus on the numbers and relational terms to choose an operation. Thus, they might fail.

Data Analysis

Trained helpers assessed the pupils' solutions. They entered all the wrong results into an Excel matrix. Absolute frequencies of such results were found. The authors then analysed pupils' written solutions in which wrong results appeared. The analysis aimed to find possible causes which led to the wrong results. Only some pupils wrote down partial results in WP b) and WP c), while most pupils only recorded the final result (the calculations were simple). Thus, the authors used their *a priori* analysis of the WPs and their experience with elementary pupils' solutions to such WPs to infer possible causes. They worked independently and then met and negotiated any discrepancies concerning pupils' possible solving process. In repeated discussions, three categories of mistakes originated.

Results

The maximum number of points for the three WPs was 14. The average number of points was 11.4; thus, the success rate was 81.5 %. It is rather high. The reason might be that the test was assigned to pupils towards the end of Grade 2 rather than at its beginning (due to covid). In total, 388 pupils solved all the WPs correctly, and 230 made mistakes (for example, 45 pupils did not solve any WP, and 35 pupils committed mistakes in b) and c)).

This paper focuses on the mistakes and their possible causes. Tables 1, 2 and 3 present the wrong results for WP a) to c), ordered according to their frequencies. In WP a), 61 pupils made 12 different types of mistakes. In WP b), 117 pupils committed 16 different mistakes. In WP c), 208 pupils committed 19 types of mistakes. From the statistical point of view, the frequencies of the mistakes are rather small, but they are important for diagnostic purposes. Only the results in grey cells will be focused on below.

Table 1: Wrong results for WP a) and their frequencies

Result	18	5	3	9	11	6	8	15	7	10	16	28
Frequency	29	7	6	4	4	3	3	1	1	1	1	1

Table 2: Wrong results for WP b) and their frequencies

Result	10	5	12	19	4	2	1	8	11	6	9	29	20	17	30	16
Frequency	38	17	12	9	8	7	7	4	3	2	2	2	2	2	1	1

Table 3: Wrong results for WP c) and their frequencies*

Result	6	9	1	5	3	7	13	20	4	31	15	8	17	10
Frequency	50	35	19	16	14	12	9	8	5	3	3	3	2	2

*Results 41, 27, 19, 16 and 14 were only reached by a single pupil each.

We distinguished three categories of mistakes. Each will be described separately.

Mistakes at the level of the mathematical model

In this category, we included the solutions in which pupils created a different mathematical model than the correct one or a partial one. It happened especially in WPs b) and c) requiring chains of operations. The cause of mistakes is skipping the stage of the situation model and making a mathematical model based on the surface structure of the WP or misunderstanding the relationships between numbers. This group consists of results 18 (WP a), 4, 5, 10, 12, 19 (WP b), and 6, 9, 1, 5, 13, 20 (WP c). The wrong results will be in bold.

WP a): Result **18** can be reached by adding 7 and 11, using a keyword strategy: “There are two numbers, there is a keyword for addition (‘in how many years will she be’) and thus, I will add $7 + 11$.” The question was transformed into “How old will Ada be in 11 years?”. If the pupil was aware of this, then they solved an easier WP with a different structure: $S + S$ (combine). Such WPs are far more frequent in Czech mathematics textbooks.

WP b): Result **19** has the same cause as above. The pupils correctly found that Dan was 8 years old and answered the question “How old will Dan be in 11 years?” ($11 + 8$).

Result **10** was probably reached by the calculation $11 - 1$. The pupils transformed the problem into an easier one similar to WP a): “Dan is 1 year old. In how many years will he be 11?” The same reason might have led to result **12**, but this time the pupils added $11 + 1$. A two-step problem was simplified to a one-step problem.

Result **5** probably means that the pupils correctly used 4 (the number of years until Ada is 11) and added 1 year (rather than subtracted). The word “older” denotes addition rather than subtraction, so this might be the reason for this mistake.

WP c): Result **6** was probably reached by adding $4 + 2$. Number 4 was taken from WP a): “Ada will be 11 in 4 years.” Ada is two years younger than Kuba; that is, Kuba is two years older than Ada, which points to the addition of 2.

Some pupils correctly added 2 to Ada’s age from WP a) (7) to get Kuba’s age (**9**), but they did not proceed to find out in how many years he will be 11.

Result **1** was reached by subtracting 2 from the correct result of WP b) (3). The subtraction was motivated by the keywords “2 years younger”. Result **5** might have the same origin as result 1, but the pupils realised that “younger” did not mean subtraction.

Result **13** could be reached by adding 2 and 11, and result **20** could be reached by adding 7 as Ada’s age from WP a). The reason for using them in this calculation may be the same as for result 18 in WP a).

Mistakes due to the previous wrong results

The three WPs we used are connected. The result of a) must be used to find the correct answer of b) and c). However, using the relationships between the ages of the children or the times which pass until they are 11 is evidence of pupils’ understanding of the text of the WPs and the arithmetic model. These pupils might have failed to solve WP a) or WP b) for various reasons

(see the mistakes in the first and third categories) and did not skip the stage of the situation model in at least one WP in which they were successful.

WP b): Result 4 could originate by using the wrong result (5) of WP a) in a correct consideration for WP b): as Dan is one year older, the time till he is 11 is one year less. The same applies to the origin of result 2 related to the wrong result of 3 of WP a).

WP c): We can see two possible reasons for result 3. Some pupils found the result 4 in a), 3 in b) and 3 in c). That is, only c) was solved incorrectly. Their cognitive capacity might have been exhausted from solving a) and b). Other pupils incorrectly solved a) (5), but their results for b) and c) would have been otherwise correct.

Mistakes due to inattentiveness (numerical mistakes)

It is impossible to say which mistakes might be due to pupils' inattentiveness. For example, result 5 in WP a) could be reached by a numerical error in the correct mathematical model $11 - 7 (= 5)$. Similarly, for result 3: $11 - 7 (= 3)$.

Discussion

We have identified three categories of mistakes. The mistakes in the first category could originate already on the level of creation of the situation model (Reusser, 1990) as the pupils could not create a suitable representation for "passing time". They could also originate due to the lexical inconsistency of the text in the next stage of solution in which the quantities and relationships are abstracted from the situation model; that is, the numbers and relations are transformed into an operation (Neef et al., 2003). Numerical mistakes can be due to inattentiveness or lack of cognitive resources allocated to the solution of the WPs – the pupil concentrates on creating the situation and mathematical model and does not pay enough attention to the calculation. Some mistakes (for example, result 1 in WP b) or 3, 4, 7 in WP c) could be attributed to the lack of working memory (Cowan, 2010).

Some pupils used the keyword strategy, looking for quantities and relational words in the WP. It is not surprising as this is the strategy identified in many international studies (Nesher & Teubal, 1975; Nunes et al., 2016). For these pupils, there is a problem on a metacognitive level as they approach the solution of the problem by skipping the stage of a situation model. There may be different reasons. Both textbooks and teachers can support this strategy (Depaepe et al., 2010). Another reason might be that modelling is cognitively demanding; thus, avoiding the situation model can prevent cognitive overload (Sweller, 2010).

Some of the mistakes might be attributed to the structure substitution of the WP (Polotskaia et al., 2015) in which pupils misinterpret the WP and implicitly change its structure into another one, usually easier (for example, the WP $S_S + \underline{CH} = S_R$ was changed into $S + S$). The reason might be inattentive reading or the keyword strategy.

Some pupils only solved one part of the WP, for example, finding Kuba's age in WP c) but not answering the question, which required the second step. This mistake often appears when solving multiple-step problems; pupils terminate their solution after the first step. The reason might again be a lack of cognitive resources (cognitive overload) or an insufficient working

memory. Another cause might be a lack of experience with such problems in Czech Grades 1 and 2 in which one-step problems prevail.

A relatively frequent result of 18 in WP a) can indicate that a pupil does not use their real-life knowledge as they do not find it strange that this number is far bigger than the required 11 years of age. The final stage of WP solving (semantic verification) is missing. However, this might not be the case. Pupils might see that the number is unrealistic but may not be able to find a different strategy or play what is so aptly called “a word problem game” (e.g., Greer et al., 2002). Based on their school experience, pupils might be convinced that it is possible to ignore inconsistencies with their experience or intuition in mathematics.

Conclusion

Our study's sample is sufficient to claim that some wrong results provided by the pupils are more frequent, and thus, it is necessary to deal with them. Our main assumption was that such results could be expected to be underlain by a pupil's rational, theoretically justifiable solving procedure. We explained such possible procedures based on the pupils' written solutions and our experience as researchers and teachers and related them to research literature. An obvious limitation of our study is that we could not verify our considerations with the pupils due to organisational reasons.

For the teacher to be able to use a pupil's mistake as a learning opportunity, they must understand its origin. We believe that the WPs used as our research tool have a high diagnostic potential. They enable the teacher to diagnose pupils' possible problems based on the numerical results they reach. Whether their causes correspond to our explanations could be verified by an interview with the pupil.

Acknowledgement

The paper was financially supported by the project *Teacher's understanding of the causes of school failure and the effectiveness of pedagogical interventions*, CZ.02.3.68/0.0/0.0/19_076/0016390.

References

- Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, 19(1), 51–57. <https://doi.org/10.1177/0963721409359277>
- Daroczy, G., Wolska, M., Meurers, W.D., & Nuerk, H.C. (2015). Word problems: a review of linguistic and numerical factors contributing to their difficulty. *Frontiers in Psychology*, 6(348), 23–34. <https://doi.org/10.3389/fpsyg.2015.00348>
- Depaepe, F., De Corte, E., & Verschaffel, L. (2010). Teachers' approaches towards word problem solving: Elaborating or restricting the problem context. *Teaching and Teacher Education*, 26(2), 152–160. <https://doi.org/10.1016/j.tate.2009.03.016>
- Greer, B., Verschaffel, L., & De Corte, E. (2002). “The answer is really 4.5”: Beliefs about word problems. In Leder, G.C., Pehkonen & E., Törner, G. (Eds.), *Beliefs: A hidden variable in mathematics education?* (pp. 271–292).

- Gray, E., & Tall, D. (1994). Duality, ambiguity and flexibility: A proceptual view of simple arithmetic. *Journal for Research in Mathematics Education*, 25(2), 116–141. <https://doi.org/10.5951/jresmetheduc.25.2.0116>
- Hembree, R. (1992). Experiments and relational studies in problem solving: A meta-analysis. *Journal for Research in Mathematics Education*, 23(3), 242–273. <https://doi.org/10.5951/jresmetheduc.23.3.0242>
- Neef, N. A., Nelles, D. E., Iwata, B. A., & Page, T. J. (2003). Analysis of precurent skills in solving mathematics story problems. *Journal of Applied Behavior Analysis*, 36(1), 21–33. <https://doi.org/10.1901/jaba.2003.36-21>
- Nesher, P. (1982). Levels of description in the analysis of addition and subtraction word problems. In T. P. Carpenter et al. (Eds.), *Addition and subtraction: A Cognitive perspective*, (pp. 25–38). Routledge. <https://doi.org/10.4324/9781003046585>
- Nesher, P., & Teubal, E. (1975). Verbal cues as an interfering factor in verbal problem solving. *Educational Studies in Mathematics*, 6(1), 41–51.
- Nunes, T., Dorneles, B. V., Lin, P. J., & Rathgeb-Schnierer, E. (2016). *Teaching and learning about whole numbers in primary school. ICME-13 Topical Surveys*. Springer.
- Pape, S. J. (2003). Compare word problems: Consistency hypothesis revisited. *Contemporary Educational Psychology*, 28(3), 396–421. [https://doi.org/10.1016/S0361-476X\(02\)00046-2](https://doi.org/10.1016/S0361-476X(02)00046-2)
- Polotskaia, E., Savard, A., & Freiman, V. (2015). Investigating a case of hidden misinterpretations of an additive word problem: Structural substitution. *European Journal of Psychology of Education*, 31(2), 135–153. <https://doi.org/10.1007/s10212-015-0257-6>
- Reusser, K. (1990). Understanding word arithmetic problem. Linguistic and situational factors. *Annual Meeting of the American Educational Research Association. Boston, MA (April 16–20, 1990)*. https://files.eric.ed.gov/full_text/ED326391.pdf
- Riley, M. S., Greeno, J. G., & Heller, J. I., (1983). Development of children’s problem-solving ability in arithmetic. In H. Ginsburg (Ed.), *The development of mathematical thinking* (pp. 152–196). Academic Press.
- Rosenthal, D. J. A., & Resnick, L. B. (1974). Children’s solution processes in arithmetic word problems. *Journal of Educational Psychology*, 66(6), 817–825.
- Sweller, J. (2010). Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational Psychology Review*, 22(2), 123–138. <https://doi.org/10.1007/s10648-010-9128-5>
- Verschaffel, L., Schukajlow, S., Star, J., & Van Dooren, W. (2020). Word problems in mathematics education: A survey. *ZDM – Mathematics Education*, 52, 1–16. <https://doi.org/10.1007/s11858-020-01130-4>