BABEŞ-BOLYAI UNIVERSITY

FACULTY OF PSYCHOLOGY AND EDUCATIONAL SCIENCES

„DIDACTICS. TRADITION, DEVELOPMENT, INNOVATION” DOCTORAL SCHOOL

SUMMARY OF PhD THESIS

THE CONSISTENCY EFFECT IN WORD PROBLEMS: PERSPECTIVES OF PRIMARY SCHOOL PUPILS AND TEACHERS

PhD Student: BARTALIS (MAGYARI) ÁGNES

PhD Coordinator: Prof. dr. ZSOLDOS-MARCHIS IULIANA

Cluj-Napoca

2025

Table of contents

1. Introduction

The saying “less is sometimes more” is well known. Surprisingly, this idea can also apply to the solving of word problems. Although the proverb usually refers to the beauty of simplicity, the contrast it contains can also be identified in certain mathematical word problems. How can “less” become “more” in word problems? And why is this question important? The aim of the present dissertation is to provide answers to these questions. Hopefully, by the end of the dissertation, the reader will also gain a clear understanding of the idea that has been at the center of my research over the past four years. In order to understand the relevance of this issue, it is worth briefly reviewing the characteristics of word problems and their role in school education.

The roots of school word problems can be traced back several millennia; their traces can already be found on ancient papyri (Swetz, 2009). Their importance has not diminished over time; on the contrary, word problems are often considered the benchmark of mathematical problems. There are several reasons for this. On the one hand, practicing word problems contributes to the acquisition of new concepts and to the development of various mathematical skills and abilities. On the other hand—perhaps even more importantly—word problems can also be regarded as a motivational tool, as they highlight the usefulness of mathematics by showing how it appears in everyday life and why it is worth learning mathematics. In addition, word problems offer opportunities to practice different real-life situations (Verschaffel et al., 2000).

From an educational perspective, the process of solving word problems can be considered a particularly important situation, as it provides an opportunity to demonstrate the level of understanding of acquired mathematical concepts and processes, such as mathematical modeling (Thevenot & Barrouillet, 2014). It is therefore not surprising that curricula in primary and lower secondary education worldwide include word problems or competencies related to them, that numerous intelligence tests contain such tasks, and that they also play a prominent role in international assessments. Word problems thus serve as an excellent assessment tool for mathematical modeling skills, problem-solving abilities, and, more broadly, mathematical literacy. Consequently, since the 1970s, there has been a strong research interest in word problems and their solution processes within psychology and mathematics education (Blum & Leiß, 2007; Boonen et al., 2016; Daroczy et al., 2015; Fuchs et al., 2006; Kintsch & Greeno, 1985; Lewis & Mayer, 1987; Thevenot & Barrouillet, 2014; Verschaffel et al., 2000, 2020).

However, the picture of actual practice in solving word problems is far less positive. In many respects, this area is highly problematic. Word problems are among the most difficult and complex types of tasks in elementary mathematics education (Daroczy et al., 2015; Verschaffel et al., 2020). The text of a word problem itself often belongs to the most difficult types of texts to read (Schell, 1982). Difficulties in problem solving thus arise partly from the specific characteristics of the tasks, partly from students' cognitive abilities, knowledge, and attitudes, and partly from pedagogical factors, such as teachers' competencies, instructional methods, and curricular frameworks (Cruz & Yangco, 2023).

Among the word problems commonly used in elementary education (change problems, combine problems, compare problems), compare problems can be considered the most difficult of all (Lewis & Mayer, 1987; Riley et al., 1983). Compare problems express a static, numerical relation between the values of two variables, as illustrated by the example used by Riley et al. (1983): "Anna has 6 marbles. Peter has 2 more marbles than Anna. How many marbles does Peter have?" One source of difficulty in these problems is that primary school pupils do not have sufficient experience in describing quantitative relations bidirectionally. As a result, they often struggle to understand that the same difference can be expressed using both the terms "more" and "less." Another source of difficulty is that this relational expression can be consistent or inconsistent with the mathematical operation that needs to be performed in order to solve the problem correctly. This phenomenon is referred to as the consistency effect (Hegarty et al., 1992). As an inconsistent version of the previous example, the problem would be formulated as follows: "Anna has 6 marbles. This is 2 less than the number of marbles Peter has. How many marbles does Peter have?" Thus, despite the relational term "less," which suggests subtraction, addition is required to obtain the correct solution.

If the problem solver extracts only the numbers and key elements from the problem text, an incorrect solution can easily be obtained. This commonly used problem-solving approach is referred to as the keyword strategy (Karp et al., 2019), which is closely related to the strategy described by Hegarty et al. (1995) as direct translation. It must be recognized, however, that this type of approach does not ensure correct problem solving in all cases. As can be observed in the examples presented above, while this strategy is helpful in solving the first, consistent problem, it fails in the case of the second, inconsistent problem. The schema-based approach also focuses exclusively on the quantities of the variables and the relations between them (Kintsch & Greeno, 1985). However, these problem-solving models were later subject to considerable criticism (see Blum & Leiß, 2007). Consequently, they were replaced by models in which the situation model appears as the mental representation of the events, actions, actors, and situations described in the text (Dijk & Kintsch, 1983). During problem solving, the situation model is constructed prior to the mathematical model, thereby ensuring that the representation of the real situation is incorporated into the mathematical model (Coquin-Viennot & Moreau, 2003). Although the construction of an appropriate situation model is often not simple, in practice it can greatly increase the likelihood of correct problem solving (Coquin-Viennot & Moreau, 2003; Riley & Greeno, 1988). The problem, however, is that pupils frequently do not adopt this approach and instead tend to rely on simpler strategies, such as the keyword strategy, which are easier to apply but often unreliable.

Additional difficulties are related to pupils' attitudes toward word problems and their adherence to fixed schemas. For many pupils, solving word problems often means nothing more than performing a few arithmetic operations based on the numbers appearing in the text, without any real analysis of the problem situation. As a consequence, these tasks fail to fulfill their original function, namely contributing to the development of mathematical modeling competencies (Reusser & Stebler, 1997; Verschaffel et al., 2020). In such cases, solving word problems is reduced to stripping away the context and practicing calculations, that is, problems are solved in a way that does not involve genuine understanding (Stern & Lehrndorfer, 1992). At the same time, it can also be observed that the majority of pupils "solve" a considerable proportion of unsolvable problems as well, without expressing any kind of realistic reaction, when these tasks are presented in a typical classroom context (Reusser & Stebler, 1997).

These difficulties in the domain of word problems are regularly reflected in the results of the PISA assessments in Romania, which fall well below the average performance of OECD countries (OECD, 2023). While the results between 2006 and 2012 indicated a slight improvement, the period that followed – including the results from 2022 – clearly shows a declining trend. Since the assessment contains a high proportion of word problems, the results of the mathematics test are often explained by performance on the reading literacy test, particularly among low-performing students (Ding & Homer, 2020). This suggests that students who perform poorly on reading comprehension tasks are also likely to achieve weak results in mathematics. Adequate reading comprehension can therefore be considered a fundamental prerequisite for successful performance in mathematics. However, similarly to mathematics, Romanian 15-year-old students also achieve relatively weak results in reading literacy.

Although the national curriculum takes into account national and international trends as well as the results of international assessments (Ministerul Educației Naționale, 2014), difficulties related to word problems remain a persistent problem in Romanian mathematics education (Purcar et al., 2024). While general and specific competencies related to word problems do appear in the curriculum, these tasks are fundamentally organized around problem-solving methods (such as comparison, graphical representation, inverse method, and balance method). The problem lies in the fact that textbooks adopt the same pattern. For instance, the presentation of the inverse method is almost exclusively followed by tasks that can be solved using this specific method. This instructional approach encourages the practice of algorithms rather than the development of genuine thinking and understanding.

The present dissertation therefore builds on these fundamental problems. Investigating the consistency effect in word problems is of key importance, as the difficulties primary school pupils experience in solving word problems are closely related to the issues outlined above. Accordingly, the aim of the research is, on the one hand, to explore primary school pupils' problem-solving patterns in consistent and inconsistent compare problems, and to examine how these patterns are related to performance. On the other hand, the research seeks to respond to the difficulties caused by inconsistency through the use of different instructional approaches and problem-solving methods. In addition, the dissertation aims to explore teachers' perspectives on the consistency effect. To achieve these aims, eye-tracking was used with pupils, and interviews were conducted with teachers. Since the research consists of four main parts, the research questions are presented in accordance with this structure.

Given that the present research is structured into four parts, the research questions are also presented in accordance with this structure. First, the dissertation seeks to examine how pupils' attention is distributed across text elements that are crucial for problem solving in the case of consistent and inconsistent compare word problems, and how these attentional patterns are related to performance. The second part of the research focuses on task instructions. Accordingly, the analysis centers on the question of how different instructions provided for the tasks influence the success of problem solving, as well as eye-movement patterns related to reading the problem text. Building on this line of inquiry, the third part of the research examines how different problem-solving methods affect performance and eye-movement patterns when solving inconsistent compare problems. Finally, the focus shifts to teachers. Through the analysis of interviews conducted with teachers, the study aims to reveal their views on the factors underlying pupils' difficulties and successes in solving word problems, the characteristics of their general instructional approaches to teaching word problems, and the ways in which the consistency effect influences their instructional approaches, task instructions, and judgments regarding task difficulty.

Although numerous studies have investigated the consistency effect appearing in compare word problems, the present research contributes to a more precise understanding of this phenomenon in several respects (Hegarty et al., 1992, 1995; Lewis & Mayer, 1987; Múñez et al., 2013; Pape, 2003). The existing literature has primarily focused on the effects and consequences of inconsistency (Leong & Jerred, 2001; Ng et al., 2021; van der Schoot et al., 2009a), while the exploration of possible solutions to this problem has largely remained in the background. The novelty of the present research manifests primarily in this area, as it draws teachers' attention to a problem-solving method that effectively supports solution processes in inconsistent problems, as well as in other linguistically complex word problems. At the same time, this solution-focused approach can serve as a clear example of how different theoretical frameworks can be integrated into pedagogical practice. In our research, we relied on a theoretical approach whose decisive role in the problem-solving process has been confirmed by numerous studies (Coquin-Viennot & Moreau, 2003; Stephany, 2021; Stern & Lehrndorfer, 1992), yet whose possibilities for practical application had not been evident. Through the application of this method, it is possible to create problem-solving conditions under which it becomes almost inevitable that pupils construct the appropriate situation model based on the task, that is, that they genuinely imagine the situation described in the problem. Although previous studies have attempted to reduce the consistency effect using traditional methods (De Koning et al., 2017; de Koning et al., 2022), these approaches contributed to this goal only indirectly. The significance of the present research therefore lies in the fact that the context-focused method proposed here, despite its simplicity, is able to contribute directly to the reduction of the consistency effect.

The methodological strength of the dissertation lies in the inclusion of eye-tracking technology, which serves as an innovative and reliable tool for investigating word problem solving (Strohmaier et al., 2020). Although in recent years eye-tracking has increasingly appeared alongside traditional methods (paper-and-pencil tests, think-aloud protocols, etc.) in studies examining word problem solving (Daroczy et al., 2015; De Corte et al., 1990; Dröse, 2019; Gros et al., 2020; Roth et al., 2025; Wu et al., 2021), many aspects of solving inconsistent problems have remained unclear. One of the aims of the present dissertation is to address this research gap, as we examined textual elements in the problems that had not been considered in previous eye-tracking studies, despite the fact that they may play a key role in the solution process of inconsistent problems. In the present case, such an element was the pronominal reference "this," the neglect of which can easily lead to an incorrect solution.

At the same time, the literature contains very few eye-tracking studies that have examined the consistency effect in word problems specifically among primary school pupils (van der Schoot et al., 2009). This gap justified the decision to investigate word problems designed for primary school pupils with this age group itself, as their reading patterns differ in several respects from those of older, more experienced readers (Csépe, 2006). Although, from a practical perspective, it would have been more convenient to involve adults or university students—as was the case in most previous studies—it is not certain that results obtained in this way can be generalized to children's problem-solving processes. An additional consideration was the lack of Hungarian-language studies that explicitly examine the relationship between the consistency effect and success in problem solving. For this reason, the aim of the present research was to test and refine conclusions drawn in the international literature regarding the solution of compare problems by examining primary school pupils using Hungarian-language tasks.

Finally, a review of the literature revealed that although several studies have investigated teachers' approaches to word problems (Bingolbali et al., 2011; Chapman, 2006; Pearce et al., 2013), to our knowledge there are no studies that have specifically focused on teachers' attitudes toward inconsistency. Based on this observation, in addition to exploring pupils' successes and difficulties in problem solving and instructional approaches, the present research also set out to examine how and to what extent the consistency effect appearing in word problems influences teachers' instructional approaches, task instructions, and their judgments regarding task difficulty.

In light of all these considerations, we argue that the dissertation is based on a complex and gap-filling investigation that enriches the research field concerned with word problems and the consistency effect. The following chapters of the dissertation provide evidence for this claim. After the review of the relevant literature, the four empirical studies conducted within the framework of the dissertation are presented: (1) The consistency effect in primary school pupils' word problem solving – an eye-tracking study; (2) The role of instructional guidance in pupils' performance and reading patterns in inconsistent compare word problems; (3) The effects of different problem-solving methods on primary school pupils' word problem solving in the case of inconsistent problems; (4) The consistency effect in word problems from teachers' perspectives.

2. The Consistency Effect in Primary School Pupils' Word-Problem Solving – An Eye-Tracking Study

The following chapter presents the first, pilot-type empirical investigation of the topic, in which we examined the solving process of inconsistent compare word problems among 10–11-year-old primary school pupils. Previous research has clearly indicated that inconsistent compare problems are more difficult for pupils than consistent ones (Orrantia & Múñez, 2013; Pape, 2003; Riley & Greeno, 1988). However, the majority of studies investigating the consistency effect were conducted with university students (Hegarty et al., 1992, 1995; Lewis & Mayer, 1987; Orrantia & Múñez, 2013). Eye-tracking studies involving primary school pupils have mainly focused on differences between solving various task types (A. Boonen & Jolles, 2015; De Corte et al., 1990; Riley et al., 1983), consistently showing that compare problems represent the most difficult category. Some investigations have concentrated exclusively on the problem-solving abilities of either primary school pupils (van der Schoot et al., 2009) or secondary school students when solving compare problems (A. J. H. Boonen et al., 2016; Orrantia & Múñez, 2013; Pape, 2003).

Two key factors justified the need for a deeper exploration of this research area. First, only one study has addressed, specifically, the eye-movement patterns of primary school pupils while solving compare problems (van der Schoot et al., 2009). Second, the literature has confirmed that during the mathematical modelling of word problems, the language of the task text also influences the inversion process required for solution (González-Calero et al., 2020). To date, no survey or experiment has been conducted using Hungarian-language tasks that specifically examined both the consistency effect and solution success. Therefore, we considered it important to reinvestigate the conclusions drawn in earlier international research on solving compare problems among primary school pupils, including solving processes based on Hungarian-language problem texts. This was complemented by the mapping of additional reading-based problem-solving patterns, which will be elaborated in the following sections.

2.1 Scope and Hypothesis

The scope of the research is to study how primary school pupils comprehend the text and solve compare word problems by analyzing ET data. As part of this question, this article has two main aims: (1) to compare fixation durations on different text elements in the case of IL and CL problems and (2) to compare fixation durations on different text elements by successful and unsuccessful solvers. In both cases, we had formulated a hypothesis that had not been studied before.

The first hypothesis is that the fixation duration on numbers is longer in the case of IL problems than in the case of CL problems. To solve a word problem, it is evident to use the numbers from the text. So, even if someone does not really understand the relations of variables, he/she tries to do some operation with the numbers (Hegarty et al., 1995). However, when dealing with an inconsistent task, the relational term must be inverted for the correct solution

(Lewis & Mayer, 1987), and it can be assumed that this difficulty affects the fixation duration of the numbers as well. Although, as mentioned before, unsuccessful solvers spend more time fixating on numbers (Hegarty et al., 1995), we are interested in whether this also applies to successful solvers.

Moreover, if the reversal of the operation in IL problems increases the working memory load, a second hypothesis can be formulated as the whole relational sentence is fixated for a longer time in IL problems compared to CL problems. This is the part of the task text that is the most concise and requires the greatest level of attention, while the interpretation of the relational sentence is only possible together with the preceding information.

In addition to numbers, in this sentence, there are other important data as well. A difficulty in understanding the text of the IL compare word problems is related to the presence of the pronominal reference in the relational sentence (Hegarty et al., 1995). To identify the referent of the pronoun, the solver has to search the statement sentence. Unsuccessful problem solvers could have difficulty identifying the variable name to which the pronoun refers. Thus, the third hypothesis of this study is that unsuccessful problem solvers have longer fixation duration on the pronominal reference word in the case of IL problems compared to successful problem solvers.

2.2 Methodology

2.2.1 Participants

Forty-two 10–11-year-old pupils (4th graders, 26 girls, and 16 boys) from Romania participated in the study. The participants were selected from four different schools where the language of instruction is Hungarian. Parental consent was obtained, and participation was voluntary. The study utilized convenience sampling to select the classes, taking into consideration to include both urban and rural classes. Pupils were randomly chosen from these classes. Five pupils were excluded from the analysis due to vertical drift or poor reading ability, which would result in data distortion.

2.2.2 Material

In this study, two-step compare problems were used, which were written in Hungarian language. This problem type has been used in several earlier studies (Hegarty et al., 1992, 1995; Lewis & Mayer, 1987; van der Schoot et al., 2009), therefore the tasks were built on the same schema, presented in different, pupil-familiar contexts. In the test, four problems were included, two of them CL and two of them IL problems. In the case of both CL and IL problems, in one of them, the relational term was unmarked, and in the other one, marked. The order of the items was the same for all participants. The difficulty of the problems increased gradually, while CL and IL problems were presented alternately.

2.2.3 Apparatus

To study the performance and text reading patterns of primary school pupils while solving arithmetic compare word problems, Tobii Pro Fusion hardware and Tobii Pro Lab Screen-Based Edition software were used to collect eye movement data. It samples real-time fixations at a 250 Hz sampling rate. The eye-tracker was connected to a laptop (with a 15.6-in., non-touch screen) and positioned beneath its screen. A separate room within the schools served as the location for data collection, which contributed to the comfortable and usual feeling and the ease of testing and measurement. We considered using a familiar environment to be more important than the advantages of the laboratory environment. During data collection, adequate lighting conditions and limitation of distractions were ensured.

2.2.5 Procedure

Instructions before testing consisted of first explaining to the child that four simple arithmetic word problems will appear on the screen. It was then told that position herself/ himself comfortably on the chair at the beginning because it is important to stay mostly stable during working with the tasks and not cover his/her eyes. The distance between the stimuli (the laptop screen) and the participant was about 60 cm. Before starting the study, the experimenter calibrated the eye-tracker for the participants (five calibration and four validation points were used). Participants were instructed that word problems would appear on the screen, and their task was to read the problem from the screen and write down the arithmetical calculations and the solution of each task on the answer sheet. Participants were also told that there are no time constraints and they can work at their own pace. Moving from one problem to another was managed by the experimenter. Pupils were not allowed to touch the laptop.

2.2.6 Data Analysis

For the two types of items (CL and IL), different AOIs were established. Common areas were the following: numbers, relational terms, and sentences of the problems (sentence 1 – statement sentence, sentence 2 – relational sentence, and sentence 3 – question). For the CL problems, one additional AOI was defined: the subject in the relational statement. For the IL problems, the pronominal reference word in the relational sentence was defined as additional AOI. For the established AOIs, fixation duration and number of fixations were recorded, as well as solution time for every problem. In the data analysis, descriptive statistics (mean, standard deviation) and comparisons with non-parametrical tests (Friedman test, Conover's post hoc test, Wilcoxon matched-pairs signed rank test, and Mann–Whitney's test) were performed.

2.3 Fixation Duration for Items and Performance

By analyzing the answer sheets, all solutions are considered correct where the pupil wrote the arithmetic operation correctly, even if the final result was incorrect due to a minor calculation error. The reason for this was that we did not want to assess calculation skills, but whether he/she was able to correctly interpret the task given.

The success rate varied from task to task. Nevertheless, it can be observed that on average only eight participants were unable to solve the consistent tasks correctly, while the inconsistent tasks were completed correctly in exactly half and half proportions. Task 3 proved to be the easiest, followed by task 1 (both CL type). However, the lack of a warm-up task before task 1 may have negatively affected its success. Based on the success rate of solving the tasks, tasks 2 and 4 appeared to be the most difficult, both of which were IL type.

In the following, when speaking about performance, pupils are divided into two groups based on their results: successful solvers and unsuccessful solvers. Successful solvers have solved at least three problems correctly, and unsuccessful solvers have solved 0, 1, or 2 problems correctly. Considering previous studies (Boonen et al., 2016; van der Schoot et al., 2009), grouping pupils by their arithmetical performance was necessary to identify whether there is any difference regarding their reading pattern, specifically with reference to the third hypothesis.

The performance presented above and the solution time of each problem are not clearly related. The pupils spent the least time on Item 3, which proved to be the easiest problem. Although Item 1 had a high correct solution rate, participants spent the most time on average on this problem. A Friedman test was carried out to com- pare the total solution duration for the four items. A significant difference was found between the solution duration of the items, $\chi2(3) = 15{,}800$, $p = 0.001$. Based on Conover's post hoc test, a significant difference was found between Items 1 and 3 ($p = 0.001$), Items 2 and 3 ($p = 0.001$), and Items 3 and 4 ($p = 0.010$). The solution duration for the two IL problems, Items 2 and 4, is similar. A significant negative correlation between average solution duration and performance was found: unsuccessful problem solvers spend more time solving the problems ($r = -0.439$, $p = 0.004$).

2.4 Fixation Durations for the Sentences of the Problems

To compare fixation duration on different sentences of the problems, fixation durations were weighted by the number of characters, the comparisons were made on the average fixation duration on a character from each sentence.

To test our second hypothesis, individual sentences from consistent and inconsistent tasks were compared. The Wilcoxon matched-pairs signed rank test indicates that there is a significant difference in the average fixation duration on Sentence 2 between consistent and inconsistent problems.

2.5 Fixation Duration on the Problems' Data

In the following, the fixation durations of data are described. These are the relevant data from the problems that pupils must use to successfully solve the problem: three different numbers, relational words, reference words, and the names of the variables. The first number in the problems represents the value of the first variable, and the second number represents the difference between the first and second variables. The third number appears in the question, and this represents the quantity that must be taken from the second variable (i.e., the value of the second variable has to be multiplied by this number).

Comparing consistent and inconsistent problems, the Wilcoxon signed-rank test indicates a significantly longer fixation duration ($W = 227.000$ and $p = 0.004$) on numbers from IL problems ($M = 17783.93$) than on numbers from CL problems ($M = 12964.98$). This does not apply to consistent problems. The fixation duration on numbers takes 14.83% of the total fixation duration in the case of CL problems, and 19.41% in the case of IL problems.

After comparing fixation duration for numerical data and relational expressions, we also examined the pronominal reference ("this") in the inconsistent task. The pronominal reference from the IL problems was significantly longer fixated by unsuccessful solvers ($M = 2217.38$) than by successful solvers ($M = 1756.70$), as indicated by the Mann–Whitney test ($W = 286.000$, $p = 0.045$).

2.6 Discussion and Conclusion

Previous studies have shown that pupils are less successful in solving IL problems than CL ones (Múñez et al., 2013; Orrantia & Múñez, 2013; Pape, 2003; Riley & Greeno, 1988), which is also supported by the results of the present research. While in the case of the CL problems, the success rate was 80.95%, the IL problems were solved correctly by half of the pupils. As we mentioned before, there are a few differences with respect to successful and unsuccessful pupils.

Firstly, as Hegarty et al. (1995) pointed out, the solution time of low-ability pupils is longer than that of high- ability pupils. This is clearly true for present research as well; in addition, a negative correlation between performance and solution time can also be observed. Furthermore, the fact that low-ability pupils have more errors in the solution of IL problems than CL problems (Hegarty et al., 1995). However, it should be noted that mathematical ability levels do not necessarily coincide with problem-solving success on the test problems. Our study does not use external reverence of mathematical ability; thus, success- fulness and problem-solving ability indicators were collected on the same test items.

From previous studies (De Corte et al., 1990; Hegarty et al., 1995), it is already known that there are longer fixations on IL problems, but it was not clear that it is also true for the numbers. Our study provides converging evidence concerning the first hypothesis that fixation duration on numbers is longer in the case of IL problems than in CL ones.

The second hypothesis is connected to the first one since it refers to longer fixation duration on the relational sentence in the case of IL problems in comparison to CL ones. In this regard, we should take into account that, in order to successfully solve an IL problem, it is necessary to reverse the relational term (Lewis & Mayer, 1987) and to process the meaning of the pronominal reference word, which refers to the first number (Hegarty et al., 1992). This may explain why pupils fixate on the whole relational sentence longer than in the case of CL problems. This result indicates the evidence that translating the relational sentence generates problems among 10-year-old pupils. It is obvious that this is a key step in this problem type, because if the pupils omit this step of the solving process, they cannot solve it correctly.

The third hypothesis is related to the pronominal reference word from the IL problems. The results show that there is a significant difference in the duration of fixations on this word between the successful and unsuccessful problem solvers group. This result confirms the third hypothesis, according to which unsuccessful problem solvers have longer fixation durations on the pronominal reference word in the case of IL problems compared to successful problem solvers. It suggests that unsuccessful solvers realized the importance of this information from the text, but they did not associate it with the first variable that this pronominal reference word actually refers to. Therefore, we can also conclude that, indeed, poor text comprehension leads to the incorrect solution.

The findings of our study need to be considered in light of a number of important limitations that are discussed below. First, assessing compare word problems with ET technology among young children has more challenges than with adults. We considered taking into account the age characteristics of the children, which avoided us from using filler tasks in the measurement. Since their attention span is much shorter than that of adults, too many tasks could also have caused data distortion due to lost focus. This may explain why previous studies in this field primarily assessed college pupils or secondary school pupils. An additional observation is the fact that there was no warm-up task before the first task and the tasks were not randomly selected, so these may have influenced our results, especially regarding Item 1.

Despite these limitations, the findings of the present study show how 10–11-year-old children are likely to misinterpret inconsistent word problems compared to consistent ones and what the main differences between successful and unsuccessful solvers are. However, the result also suggests that there are substantial individual differences in children's solution process. We consider the present study to provide additional evidence that primary school pupils have difficulties solving compare word problems, and hopefully, our results contribute to further studies in this field.

Based on the results of the study, teaching methods regarding compare problems should be developed. Thus, exercises focusing on the context of the problem and graphical visualizations of the relation between variables facilitate pupils to concentrate more on other text elements besides numbers. Training pupils in comprehension skills is essential for solving compare word problems, as the result indicated: the source of the difficulty is not linked to the solution phase but to the comprehension phase that must be fostered.

3. The Role of Teaching Instruction in Pupils' Performance and Reading Patterns – In Inconsistent Compare Word Problems

In line with earlier research, the results presented in the previous chapter also confirmed that the success rate of solving inconsistent compare word problems is substantially lower than that of consistent ones (Orrantia & Múñez, 2013; Pape, 2003; Riley & Greeno, 1988). Therefore, this chapter places specific emphasis on pupils' performance and reading patterns observed while solving the more problematic, inconsistent compare problems.

The solving of word problems can be strongly supported by constructing an adequate situation model, as well as by forming a mental representation based on the magnitude relations between the variables included in the task text (Coquin-Viennot & Moreau, 2003; Orrantia & Múñez, 2013). The present study builds on the theoretical framework of these two prominent mental representation models, as—shown later in this chapter—they may play a key role in the successful solving of inconsistent compare word problems. The aim of our research is therefore to support pupils' problem-solving processes by applying instructional approaches that have a well-established theoretical background and that may contribute to solution success. One of the main objectives of this investigation is to explore how effectively the different instructional methods applied in our study operate in supporting the solving process. In addition, we also aim to gain a deeper understanding of pupils' solving processes from an internal perspective, which is why eye-tracking technology was applied.

Several earlier studies have already made use of the possibilities offered by eye-tracking devices (De Corte et al., 1990; Hegarty et al., 1992; van der Schoot et al., 2009), yet no investigation has been conducted to date that specifically assessed the effect of teaching instructions on pupils' performance in inconsistent compare word-problem solving.

3.1 Scope and research questions

The scope of the research was to find out if there are differences in solution success and reading patterns of the problems' text based on different instructions when giving the problems.
The research aims to find answers to the following research questions:

1. How does the type of instruction influence the performance?
2. How does the type of instruction influence the fixation durations on different key elements of the text of the problem?
3. How does the type of instruction influence the fixation durations on different sentences of the text of the problem?

3.2 Participants

Convenience sampling was used to obtain a sample of 56 Hungarian speaking fourth grade students (10- 11 years old), 20 girls and 36 boys from two regular schools located in Cluj-Napoca, Romania. Since the limitations of the eye-tracking studies (only one person can be recorded at a time, labor resources etc.) enabled only a convenience sampling, however, participants were selected randomly from each class. Therefore, classes participating in the investigation are from urban areas, representing the inner- city schools.

3.3 Material

The material (the test) consisted of 7 two-steps inconsistent compare word problems (tasks), containing one warm-up task and 6 test tasks. The problem sheet was written in Hungarian. The students got the problem sheet on the computer with eye-tracking technology included and solved the problems on a sheet of paper. During the development of the test tasks, we followed the schemas of earlier research (Hegarty et al., 1992, 1995; Lewis & Mayer, 1987; van der

Schoot et al., 2009a), however, the problem texts were presented in contexts familiar to pupils. An example of this problem type is: “A notebook costs 20 lei. This is five times less than the price of a book. What is the total cost of the book and the notebook together?”

After the warming-up task the problems were divided into three groups with different task instructions. In case of each instruction type two tasks were given, their order was randomly selected. In addition, after each task an empty slide was included to present the instruction and to answer some incidental questions that may arise during the test.

In the first case (task 1 and task 2) the problem text is given without any additional instruction. Thus, pupils could freely concentrate on the text elements they wanted. We will call these problems simple problems in the following.

In the second case (task 3 and task 4) students were forced to spend more time focusing on the context of the problem. For this purpose, they got the task without any number or data, the numbers were replaced by different symbols. In this case they could not actually solve the problem, however, they had time to think about it, to understand the context, to familiarize themselves with the situation and to start elaborating the solution strategy. We will call these problems as problems with symbols in the following. After this, the same task appeared with numbers instead of symbols, allowing pupils to actually solve it. These tasks will hereafter be referred to as problems with numbers after symbols. In the study, symbolic tasks and problems with numbers after symbols are analysed as separate categories, although they essentially represent two different versions of Tasks 3 and 4.

In the third case (task 5 and task 6), pupils got the problem, and they were asked not only to solve the problem, but also had to create the graphical representation of the word problem taking into account the known values of the variables and the relation between the variables. We will call these problems as problems with graphical representation.

As observable above, even if the focus was on three different instruction types, the given problems can be grouped into four types: simple problems, problems with symbols, problems with numbers after symbols, and problems with graphical representation.

3.2.3 Apparatus

The apparatus was identical to the eye-tracking equipment described in Chapter 2.

3.2.4 Procedure

The procedure was identical to the eye-tracking equipment described in Chapter 2.

3.2.5 Data analysis

For analyzing the reading pattern of pupils' while solving inconsistent compare problems, some Area of Interests (AOIs) were determined. These AOIs were associated with those key data from the problem text, the use of which were essential in order to solve the problem. These data were: the two numbers, the relational terms (more or less) and the pronominal reference word (this). In addition, different sentences of the problem were also analyzed individually: the statement sentence, the relational sentence, and the question of the problem. For the established AOIs fixation duration and number of fixations were recorded, as well as solution time for every problem. Furthermore, in order to make the numerical data easier comparable, we converted them into percentage data. Descriptive statistics (mean, standard deviation) and comparisons with repeated-measure ANOVA were performed in data analysis. In case of all the four types of problems each correctly solved task was counted (1 point for each correctly solved problem, no fractional points were assigned). The possibly obtained highest score for each problem type was 2.

3.3 Results and discussion

Below, we do not discuss the survey results in detail, as these will be presented in the following chapter in comparison with the post-test findings. In this chapter, we report only the key results and the conclusions drawn from them, focusing on how the different instructional approaches influenced students' performance in solving inconsistent compare problems, and how these instructions shaped their reading patterns, with particular attention to the key-elements and the individual sentences within the task text.

The highest score was obtained for the problems which were first given with symbols instead of numbers. This result could be explained by the fact that students were forced to read the text more carefully as they got it without numbers first. In this way they spent more time reading all the element of the text, they understood better the context of the problem, and this helped them to construct the situational model. The construction of the situational model is important for the successful solution of the word problem (Hegarty et al., 1995; Kintsch,1998).

The second highest score was obtained for the problems where students were asked to make a graphical representation of the data from the problem. Processing the numbers involves a mental representation of numerical magnitude (Orrantia & Múñez, 2013). In case of word problems not only the numerical magnitude of each number is important, but also the relation between the variables. While working on the graphical representation students read the text of the problem more carefully and they concentrated on the magnitude-based representation of the variables' values and the relation between these variables. The results show that creating the magnitude- based model helps them in selecting correctly the operation needed for the solution.

Solution time was the highest in case of the problem with graphical representation, which was expected as during this time students not only solved the problem but also created a magnitude-based model. The second highest solution time was obtained for problems where symbols are given instead of numbers. This type of problem was unusual for students thus they needed more time to read the problem and think about the answer. However, it is surprising that students also need a significantly longer time than in case of simple problems to solve the same problem as read before given with numbers instead of symbols now. The hypothesis would be that they already know the context of the problem, they already constructed the situational model, they just need to process the numbers for solving the problem.

As regards percentage of times fixating on key elements of the text, different key elements were fixated in higher percentage in case of different type of problems: number 1 in the case of problems with numbers after symbols, number 2 in the case of problems with symbols, the relational word in the case of problems with graphical representation, and the pronominal reference word in the case of simple problems. It is not surprising that the percentage of time fixating on number 1 is high in case of problems with numbers after symbols, as students already read the text of the problem before and when got the problem with numbers, fixated more on them. The number 2 (actually, the symbol which replaced number 2) was fixated the highest percentage of time in case of problems with symbols. Number 2 is part of the relational sentence (sentence 2), which was also fixated the highest percentage of time in the case of this type of problem. Sentence 2 also contains the relational word and the pronominal reference word. As the requirement was to identify the operation needed for the solution, it is not surprising that in the case of problems with symbols students spent the highest percentage of time on sentence 2 in general, and, in particular, on number 2. The second highest percentage of time fixating on number 2 is in the case of problems with numbers after symbols, which can be explained, as in case of number 1, with the fact that students already knew the problem, they just needed the numerical data for the solution. The relational word is fixated the highest percentage of time in case of problems with graphical representation. In this type of problem students needed to represent in a drawing the relations between variables, which made them concentrate more on the relational word. The second highest percentage of time was in case of problems with symbols, where students needed to determine the operation required for the solution thus, they concentrated on the relational word.

In conclusion we can say that the instruction type influenced both performance and reading pattern. When students are forced to spend more time reading the entire text (problems given first with symbols), the solution success rate is higher. Graphical representation of the problem's data also contributes to a better performance. There are also differences in percentages of reading times spent on fixation different key elements of the text or different sentences of the problem. The key element fixated for a higher percentage of time depends on the type of instruction.

The results of this paper have an important message to the primary school teachers: when solving mathematical word problems more approaches should be used for a better understanding of the problem's context and of a more accurate modelling of the problem's data.

4. Investigating the Effect of Different Solving Methods in the Case of Inconsistent Compare Problems

Our research presented in the previous chapter clearly demonstrated that instructional cues alone play a decisive role in students' problem-solving performance and in their reading patterns. For this reason, we considered it important to further develop these instructions. In this chapter, we introduce an instructional method specifically designed to respond to the challenges that 10–11-year-old students face when solving inconsistent compare problems.

While previous research has proposed methods facilitating the resolution of compare tasks, our decision to develop a new method is grounded in three key considerations. Firstly, despite existing methodologies, recent research indicates that these particular tasks remain among the most challenging for students (Bartalis et al., 2023; Múñez et al., 2013; Orrantia & Múñez, 2013; Riley & Greeno, 1988). Secondly, our approach is motivated by the absence of a method explicitly designed for solving inconsistent tasks, recognizing the crucial step of reversing the relational term as essential for accurate problem resolution. In this context, a nuanced understanding of the task's context emerges as particularly vital for successful interpretation and problem- solving. Thirdly, in our previous study, we concluded that the instructions associated with the inconsistent, compare problems can also influence the success of the solution. Therefore, it seemed worthwhile to further develop and devise a specific method (Bartalis & Zsoldos-Marchiș, 2023).

4.1 Research goals

The purpose of this study is to assess the effectiveness of two different teaching methods that facilitate the process of solving two-step, inconsistent compare tasks among elementary school students. The first of the two methods is the above-mentioned bar charts model, while the second method was developed by us, based on the previously presented situational model. The essence of the method is that in the first round, children receive the text of the task in such a way that numerical data is replaced with different symbols. In this stage they are prompted to focus on the context, that is the situation presented in the text, and when they understand it, only then are they given the task with numerical data. In order to exclude all other factors and only compare the problem-solving methods themselves, we also included the control group in the intervention. The three groups participating in the research (two treatment and one control group) received the same set of tasks during the intervention, but in each of the groups these tasks were practiced using different methods. In the control group the children were allowed to solve the tasks at their own discretion (there was no direct method applied), in the first treatment group the problem-solving method was based on magnitude-based mental representation (use of bar diagram), and finally the second treatment group used the method based on the episodic situation model (use of context-focused tasks). The main question of our research therefore is whether practicing with different problem-solving methods changes the students' performance and focus, and if so, in which direction does the performance change, and how is the children's attention divided between the context of the text and the key elements.

4.2 Methodology

4.2.1 Participants

A total of n=50 fourth grade students (10-11 years old) participated in the study, 31 of whom were boys and 19 were girls. They all study in Hungarian-speaking regular education classrooms in Cluj-Napoca, Romania. Three classes were conveniently selected. From each class randomly selected students formed a group. To each group the treatment type was randomly assigned. There were the two treatment groups and the control group. 17 students were included in each treatment group, while there were 16 students in the control group.

4.2.2 Materials and procedure

The test assessing problem-solving ability of inconsistent compare word problems was conducted in the spring of 2023 for all three groups. All subsequent procedures were identical to the research protocol described in the previous chapter. The measurement tool used in the pre-test, along with the associated data-collection procedure, instructions, and the AOIs applied in the data-analysis phase, have already been presented earlier. The post-test differed from this only in the wording of the tasks themselves.

4.2.3 The intervention

After the pre-test, a six-session intervention was conducted in all three groups, during which all students solved the same set of tasks (consisting of 30 word problems). This included tasks of the same type that were used during the initial assessment. Since the aim of the research was to compare the effectiveness of the different teaching methods, we considered it important that practice should not only take place in the treatment groups (with two different methods), but students in the control group also encounter these tasks the same way.

In the case of the latter, the students did not receive any direct instructions, they solved the problems at their own discretion, spontaneously or with the help of previously learned procedures. In comparison, the students in the first treatment group solved all 30 tasks using graphical representation. In the second treatment group, students received all their tasks at first in a way that the text contained symbols instead of numbers (thereby prompting them to focus on the context of the task, to think carefully about "what the story of the task is"), and only afterwards were they given the problems with numbers to solve them.

4.3 Pupils' performance

In accordance with our research question, the presentation of results is divided into two parts. In the first part, the performance of the students is presented, while in the second part, the time spent on the text of the task, more precisely, the proportion of time the students focused on the key elements and the context within the text of the task. Accordingly, the performance was analyzed using the points obtained for correctly/incorrectly solving the tasks, while attention focus was examined by comparing the fixation times spent on the AOIs. Descriptive statistics, t-test and repeated-measures ANOVA were used in the data analysis.

For simplicity, the names of the task types in the results section will be abbreviated as follows: P_simple (simple tasks) P_graph (tasks with graphical representation) and P_sym (tasks with symbols). The names of the groups will be marked as follows: Exp_graph group, Exp_sym group and Control group.

When scoring students' performance, a task was considered correctly solved if the appropriate operation was chosen regardless the correctness of the final numerical result. This is because we were interested in whether they had been thinking in accordance with the task, and not just about executing the operation itself. Out of a total of 7 tasks, the first, warm-up task was not scored, apart from that, each correctly solved task was worth 1 point. By categorizing the tasks according to the instructions, a total of 2 points on two tasks could be obtained within each category. Students were able to score a maximum of 6 points in total.

In the first part of this section comparisons between performances on pre-test and post-test are presented for each group. The Shapiro-Wilk Test indicated that the normality assumption was violated in case of each group and each type of problem, thus, the Wilcoxon matched-pairs signed rank test was used to compare pre-test and post-test results.

The Exp_graph_group showed a significant performance increase on simple compare problems in the post-test compared to the pre-test. In the Exp_graph group, a significant improvement was observed across all three task types. In the control group, the intervention did not result in any meaningful change in performance.

The performance on different types of problems in the case of the three groups were compared with ANOVA for the pre-test respectively post-test. The homogeneity assumption is violated on the post-test in case of P_simple ($p = .006$) and P_sym ($p = .004$) tasks, in these two cases the Brown-Forsythe homogeneity correction was applied. There is a significant difference between the results of the three groups in case of the P_sym tasks: the exp_sym group obtained significantly higher scores than the control group ($t = 2.554$, $p = .037$).

4.4 Percentage of the reading time fixated on key elements (PRTKE)

Our research question also touched on how different problem-solving methods affect students' attention focus on key elements and context. To examine this, we used the fixation duration as a basis. The fixation duration of the key elements was determined by merging the AOIs marked for the key information: fixation duration on number 1, number 2, relational term ("less" or "more") and pronominal reference word ("this"). The sum of these forms the key elements variable.

Furthermore, in order to determine the proportion in which students focused on these key elements while solving the problems, we compared these values with the total reading time.

The total reading time was obtained by marking one AOI for each sentence of the text and summing these values for each task. During data analysis, the total reading time variable is considered as 100% and the percentage of the reading time fixated on key elements is calculated. The higher the percentage value of the variable fixation duration on key elements, the lower the percentage value of the context, and vice versa. We used these percentage values because they illustrate the students' attention focus better than the fixation duration values would do alone.

Pre-test and post-test values for PRTKE are compared in the case of each group for each type of problem. The Shapiro- Wilk Test indicated that the normality assumption was met, thus paired t-test was used. In the case of each group the PRTKE has significantly decreased for P_simple and P_graph problems. The differences in PRTKE for different problems are statistically significant ($p <.001$). The Holm post-hoc test shows that the PRTKE for P_sym tasks is significantly higher than the PRTKE for P_simple ($p < .001$) and P_graph ($p < .001$) tasks.

4.5 Discussion

The aim of this study was to present a problem-solving method that can effectively help 10-11-year-old students solve inconsistent, compare word problems. The essence of this method, rooted in the episodic situation model, lies in its aim to guide students in observing the context of the text before extracting key elements and engaging in any arithmetic operations. This approach deliberately discourages students from solely focusing on isolated key elements, instead prompting them to direct their attention to the broader context of the text initially. To assess the effectiveness of this developed method, we conducted a comparative analysis of the results with another well-regarded method, the bar diagram. This comparison serves as a valuable means to evaluate the efficacy of the episodic situation model-based approach in contrast to the established effectiveness of the bar diagram method.

The results of the performance comparison show that in the group where problem solving was facilitated by graphical representation, only in case of the simple problems did they perform significantly better in the post-test compared to the pre-test. In addition, in the same group, unexpectedly, the performance did not improve in the case of graphical representation tasks. In the control group, where during the intervention no direct methods were used, and the students performed the tasks at their own discretion, the results of the pre-test and post-test show no significant difference. In terms of performance, the results of the symbolic group show the greatest improvement, as they significantly outperformed their pre-test results across all task types in the post-test.

Although several studies have proven the effectiveness of graphical representation and its role in a successful resolution of inconsistent tasks (de Koning et al., 2022; Kozhevnikov et al., 2005), our investigation nevertheless allows us to conclude that the method based on the episodic situation model proves to be more effective in terms of performance. Thus, these results confirm the significance of creating the episodic situational model. Even though the creation of the correct situational model is acknowledged as the challenging step in problem-solving (Riley & Greeno, 1988), it remains a crucial factor that may hold the key to achieving the correct solution (Coquin-Viennot & Moreau, 2003). Recognizing its significance, our study aligns with the insights from previous research, underscoring the importance of this step in the overall effectiveness of problem-solving methodologies.

However, this kind of difference does not show itself in the attention focus of the students. When we examined the percentages, we found that the students' focus of attention was similar in all three groups. As a result of the intervention, the students' attention shifted from the key elements to the context of the task, since the fixation time of the key elements is significantly shorter based on the results of the post-test compared to the pre-test. Certainly, the tasks with symbols deviate from this pattern, as in this case, the students were already familiar with the context of the task, requiring only the key elements to solve the assigned tasks. Thus, it can be concluded that the method of solution is irrelevant in terms of attentional focus, as students in each group recognized the importance of concentrating on the context. In addition, this result also leads back to the importance of creating an episodic situational model, which assumes attention to the context and, as we have seen, contributes to the increase in performance.

The conclusions drawn from our study must be viewed in the context of several significant limitations outlined below. Initially, evaluating compare word problems using Eye-Tracking technology among young children presents greater challenges compared to adults. We took into consideration the age-specific traits of the children, leading us to refrain from incorporating filler tasks in the measurement process. Another limitation is the sampling, as we could only include students in the research who were allowed by the teacher and whose parents also provided consent. This fact further restricted our possibilities.

Despite these limitations, the findings of the present study underscore the significance of emphasizing the task context in inconsistent compare word problems. Comparing the effectiveness of different problem-solving methods, our study provides converging evidence that the pupils' performance is higher in case when they focus on the context, than use a graphical representation. It is noteworthy, however, that these methods prove inconsequential concerning attentional focus, as students across all groups acknowledged the crucial importance of concentrating on the context.

5. Consistency Effect in Word Problems – From a Teacher's Perspective

The previous chapters focused on students and their characteristics. In the following, we will approach the solution of inconsistent word problems from a teacher's perspective. The question arose as to how teachers teaching these students relate to the phenomenon of the consistency effect in word problems.

Several studies have examined how teachers relate to problems related to word problems (Bingolbali et al., 2011; Chapman, 2006; Pearce et al., 2013), but we have no prior knowledge of any study that has examined their relationship to inconsistency in word problems. Therefore, the aim of this research is to explore what Hungarian teachers in Romania believe to be the main reasons behind students' problem-solving problems and successes, and what characterizes teachers' teaching approaches to word problems in general. Our most important question, however, is whether the consistency effect influences all of these. Accordingly, in our research, we undertook to examine how and to what extent the consistency effect in word problems influences teachers' teaching approaches, instructions and their judgments regarding the difficulty level of the problem.

5.1 Research Objectives

The research aims to explore teachers' views and teaching approaches to word problems. In addition, we aim to examine to what extent and in what way these views and approaches are influenced by the consistency effect in word problems. Based on the research objectives, we formulated the following questions:

- What is behind students' problem-solving problems and successes according to teachers?
- What generally characterizes teachers' teaching approaches to word problems?
- How and to what extent does the consistency effect in word problems influence teachers' teaching approaches, instructions and judgments regarding the difficulty level of the problem?

5.2 Methodology

5.2.1 Participants

Our interview research involved 33 primary school teachers, teaching in Hungarian-medium elementary classes in state schools in Transylvania. The participants came from a total of 13 schools, 11 settlements and 4 counties (Harghita, Cluj, Mureș and Timiș). Convenience sampling was used; therefore, it cannot be said that our sample reflects the entire population of Hungarian-medium elementary class teachers in Transylvania. However, we tried to reach teachers from Hungarian and mixed-medium schools in different small (12 teachers), medium-sized (7 teachers) and large (14 teachers) settlements, as far as possible.

5.2.2 The interview

The data collection was carried out using a structured interview, which provided an opportunity to explore teachers' views and teaching approaches to solving word problems in more depth, while at the same time ensuring that the data obtained were comparable. The nature of the research questions therefore justified the use of a qualitative approach.

The set of questions consists of three units in terms of content: demographic questions, questions exploring general views and teaching approaches regarding word problems, and questions related to the consistency effect. To assess attitudes towards the consistency effect, we also included two simple comparative tasks in the interview. Both tasks are two-step tasks based on a similar mathematical structure, one with consistent and the other with inconsistent wording. We formulated the same questions for these two tasks, which relate to the difficulty of the task, the steps of solving the task, and the resolution of difficulties arising during the task solution.

The interview protocol contained 18 open-ended questions. The data collection was carried out by the author, personally, through individual questioning. These mostly took place in schools, in a closed, undisturbed room. The audio recordings were recorded by telephone, and their duration was between 13 and 20 minutes. The audio recordings were transcribed on the Maxqda online interface, followed by checking the transcripts and correcting any errors or distortions, and finally the content analysis, which was also done using the Maxqda software.

5.3 Results and conclusions

The main aim of our research was to explore how and to what extent inconsistency in word problems influences teachers' teaching approaches. To do this, we first mapped their general approach to word problems (use of strategies, supporting comprehension) and the reasons behind the teachers' perceived student problems and successes. We then examined how the consistency effect influences teachers' perceptions of the difficulty level of the problem, the steps of problem solving, and the difficulties encountered during the problem solving process.

Our results clearly indicated that teachers primarily attribute students' difficulties in word problems to psychological factors – in particular, poor comprehension skills. This is consistent with the literature, which states that text comprehension is a fundamental condition for solving text tasks, especially when the task is linguistically complex or semantically complex, in this case inconsistent (van der Schoot et al., 2009; Boonen et al., 2016). Furthermore, the fact that teachers attribute most of the problems to students is also consistent with the conclusions of Bingolbali et al. (2011), according to which the psychological characteristics of students are mostly behind the difficulties in solving text tasks, based on the views of the teachers she examined. However, it contradicts the results of Pearce et al. (2013), who, mapping the views of teachers, found that the difficulties and their causes in this area can be more related to the curriculum, the examination system, textbooks, teacher training and the previous teacher. According to the results of our research, these latter factors (which we classified as pedagogical factors) hardly arose among the teachers, just as the difficulties arising from the task do not pose a problem for them (epistemological factors), but very few mentioned other socio-economic reasons.

According to their confession, the teachers' general teaching approaches are dominated by the step model, which fits well with Pólya's (2004) classical problem-solving model. However, several applied strategies appear in the comprehension stage. According to their explanations, most of them use the filtering of key elements when solving text tasks, which, if applied alone, can be identified with the keyword strategy presented by Karp et al. (2019). In many cases, however, this step was combined with other strategies, such as reading the text of the task several times, raising awareness of the task question, and creating external representations (pictures, diagrams) also proved to be popular, the effectiveness of which has been reported in several studies (de Koning et al., 2022; Ott, 2020; Purcar et al., 2024). These strategies help to develop adequate mental representations, which is especially important in the case of inconsistent tasks, where linguistic references and mathematical operations are not consistent (Hegarty et al., 1995). The step following the comprehension stage, creating a plan, proved to be less popular, as it was highlighted by less than half of the teachers. Finally, checking or reflecting after the solution was also mentioned by only one third of them during the task solution. Overall, however, teachers clearly find the step-by-step approach useful, for which we have seen several different models above (Richardson and Morgan, 2002; Roe et al., 2001; Ryder and Graves, 2008). At the same time, they expressed their conviction that the routine and algorithm they have developed contributes to correct task solving.

In the case of inconsistent tasks, understanding the context of the task is particularly important (van der Schoot et al., 2009), so when we asked teachers about the task solving process, we considered it important to highlight this phase and the possibilities for facilitating it. Nearly a third of teachers recommended rereading or reading the text of the task several times, and some – Ott (2020) and Purcar et al. (2024) research – they suggested the creation of external representations, as well as leading questions, the division of the task into parts, but also the reformulation, as Haghverdi (2012) and Vicente et al. (2008) also pointed out their importance.

Teachers' reactions to the difficulty level of consistent and inconsistent tasks showed clear differences. The majority of teachers considered the consistent task easy, and they mainly saw the possible student difficulties in comprehension or operational deficiencies. In contrast, three-quarters of teachers considered the inconsistent task more difficult, and several of them emphasized the importance of comprehension. In the case of the inconsistent task, however, several of them emphasized linguistic factors as a possible source of error. Although they did not explicitly mention the consistency effect, two-thirds of the teachers clearly pointed out to the phenomenon.

It should also be emphasized that although half of the teachers would not modify the applied task-solving steps in the case of the inconsistent task (compared to the consistent task), the other half would apply a more targeted, more detailed approach. In this case, several of them consider it necessary to support the task-solving with more explanations, rephrases, identifying more and less quantities, and highlighting the reference word "this". This suggests that half of the teachers surveyed perceive the importance of the consistency effect and try to adapt their methods accordingly, especially in the field of interpretation and illustration. This is in line with the didactic consideration that when solving inconsistent tasks, it is worth emphasizing situational model building instead of direct translation (Staub and Reusser, 1995). This was also confirmed by the fact that teachers only responded to problems arising during the task solution with explanations and illustrations in the case of consistent tasks, while in the case of inconsistent tasks, in addition to illustrations, rewording, practice, leading questions, highlighting of reference words and situational games appeared. Overall, we find that while three-quarters of teachers consider inconsistent tasks more difficult than consistent ones, only half would take targeted steps during the task solution.

However, we can only reach all these conclusions by taking into account certain limitations that affect the validity and generalizability of the results. Although we tried to expand the sampling, due to the available sampling, the participants do not fully represent the entire population of Hungarian teachers in Romania. On the other hand, data collection took place in the form of structured interviews, which - although it allowed for comparison - could have limited the spontaneous expression of opinions by the participants. However, the interview does not always accurately reflect the classroom practice of teachers, since – in contrast to direct observation – it is dominated by the subjective judgment of the teachers. Finally, the study of the consistency effect was based on only two comparative task types, which narrows the range of situations that can be examined and does not cover the full spectrum of text tasks.

However, since many of our conclusions resonate with the results of previous research, it emerged that the majority of teachers perceive the difficulties arising from tasks with inconsistent wording, and about half of them adjust their teaching instructions accordingly. The consistency effect is therefore a relevant factor in terms of classroom practice, which can encourage teachers to use strategies more consciously, emphasize reading comprehension support, and provide differentiated instructions. In light of this, it can be recommended that during the training and further education of teachers, greater emphasis be placed on the semantic analysis of text tasks, awareness of linguistic-mathematical connections, and the practice of applying situation models.

6. General conclusions and recommendations

Our research focused on the consistency effect in word problems. We examined the topic from both a student and a teacher perspective. Therefore, on the one hand, we undertook to explore the relationship between the problem-solving patterns of elementary school students and their performance, in the case of consistent and inconsistent problems, and to respond to problems arising from inconsistency using different instructions and problem-solving methods. On the other hand, we examined the attitude of the teachers teaching them to the consistency effect, answering the question of what, based on their insights, lies behind the problem-solving problems and successes of elementary school students; what generally characterizes their teaching approaches to word problems; and how and to what extent the consistency effect in word problems influences teachers' teaching approaches, instructions, and judgments regarding the difficulty level of the problem. The results were presented in the previous chapters, below we would like to draw attention to only the most important conclusions and their connections.

Our first pilot study contributed to the foundation of further studies in several aspects. First, it confirmed results related to the consistency effect that we did not know were also valid among Hungarian 10-11 year old primary school children. For example, inconsistent tasks are more difficult to solve than consistent ones; unsuccessful task solvers spend more time solving the task than successful ones, and unsuccessful ones make more mistakes in solving inconsistent than consistent tasks. Second, it pointed out details related to the consistency effect that have not been taken into account so far: the fixation time for numbers is longer in the case of inconsistent than in the case of consistent tasks; the relational sentences of inconsistent tasks are characterized by a longer fixation time than those of consistent tasks; and finally, unsuccessful problem solvers fixed on the demonstrative pronoun “this” for a longer period of time than successful ones. All these results suggest that this age group has difficulty coping with the consistency effect in word problems, the main reason for which is the problem of translating the relational expression appearing in the relational sentence. We were convinced that due to the resulting cognitive load, students are unable to mentally represent the problem situation appropriately for the task, and thus fail to solve the task. At the same time, it is likely that unsuccessful problem solvers choose the direct translation strategy or the keyword strategy during problem solving, instead of creating a situation model appropriate for the task.

Based on this, we set ourselves the goal of focusing specifically on the difficulties associated with inconsistent tasks. To do this, we first designed different instructions for solving inconsistent tasks, and then, seeing their effectiveness, we further improved them. Thus, we examined methods that, in our opinion, can effectively support the solution of inconsistent tasks in early school age. The results of the intervention confirmed the part of our assumptions that related to the performance of the students. Although the problem solving supported by the representation also partially contributed to the correct solution of the tasks, the method we developed significantly supported the performance in all task types. However, this difference was not typical in the reading patterns of the students, since by the end of the intervention, the recognition of the importance of the context was typical of all student groups. Overall, our context-focused method, which directly aims at the creation of the situation model in the problem-solving process, effectively helps the solution of inconsistent tasks at this age. At the same time, it excludes the possibility that the students use a keyword strategy during the solution. Finally, we could not ignore the teachers' point of view in this area. The structured interview conducted by them revealed that teachers attribute the difficulties encountered in solving tasks primarily to the students, especially their poor reading comprehension skills. Another clear conclusion we had about them was that their teaching approaches are dominated by the step model, their classroom practice is determined by the keyword strategy, which is often accompanied by multiple readings of the task text, awareness of the question, graphic representation, and less often by planning and reflection. Regarding the consistency effect, it was outlined that the majority of teachers perceive the difficulties associated with inconsistent tasks, yet half of them would not change their teaching instructions and methods compared to solving tasks with consistent language.

All in all, we can say that inconsistent tasks are not routine classroom tasks that can be solved without thinking using the method just learned, or that fit into the tasks that can be solved with the keyword strategy. The semantic structure of inconsistent tasks requires a task-solving approach that assumes a deeper understanding on the part of the problem solver. Our context-focused method offers a solution to this, which can effectively facilitate the mental representation of inconsistent tasks. At the same time, there is a connection between the difficulties of students and the teaching approach of teachers. If teachers also often choose the keyword strategy, then it is easily integrated into the students' task-solving strategies, and therefore they will also apply it. On the other hand, if teachers do not react in any way to solving linguistically more complex text tasks, and textbooks do not contain such tasks, it is not surprising that students do not perform well in solving them either.

In light of these, we have formulated several suggestions, which are primarily addressed to teachers, but also affect university lecturers, textbook writers, and curriculum writers. First, we would like to draw attention to the conscious use of strategies and the conscious selection and variation of text tasks. To this end, we would consider it important, on the one hand, that teachers not only practice a single solution algorithm, but also present as many diverse task-solving strategies as possible and encourage students to apply them. They should also choose approaches that support the creation of a situation model during problem solving, so that students – regardless of the text task in question – can imagine the problem situation and select their solution strategy accordingly. On the other hand, it is worth carefully considering what text tasks we give children and how and along what logic they are connected to each other. If we want to teach the reverse path method, we must of course practice it with a few tasks, but in the next step it should no longer be obvious that the next task "must" be solved with the same method. On the other hand, we find it worthwhile to integrate linguistically more complex tasks into the education, since realistic task solving is not based on traditional text tasks either. Inconsistent tasks could also be integrated in this way, through which students are not forced to simply solve routine tasks, but rather acquire strategies that are based on interpreting the text, developing a situation model, and critical thinking. However, all these suggestions do not only concern teachers, but also their preparation in university education, as well as textbook writers.

Our results also raised several new questions in the area of the consistency effect, which may point to further research directions. On the one hand, eye tracking could be applied more broadly in the study of the consistency effect: by including variables other than fixation and task-solving time, since, as we have already mentioned, they can offer many other perspectives and provide new evidence on the factors behind successful or unsuccessful task solving. Furthermore, we consider it worthwhile to examine teachers' teaching methods using the observation method, in a classroom environment, where we would probably get a more realistic and nuanced picture of their practices when teaching text tasks. Overall, one of the most important conclusions of the research is that the solution of text tasks with inconsistent language can be effectively supported among 10–11-year-old students with targeted, context-focused methods. In addition to the theoretical framework, the dissertation also offers relevant results from the point of view of practical applicability. In this way, it can serve as a guide for both researchers and elementary school teachers in the future, for whom our results can provide tangible assistance in teaching word problems and in identifying and solving related difficulties. The consistent presence of the consistency effect, the methodological novelties of eye tracking, and the examination of teachers' attitudes and strategies all contributed to obtaining a more complex picture of elementary school students' word problem solving in the case of consistent and inconsistent tasks.

References

Alamian, V., & Khani Baseri, L. (2020). The Effect of Manipulation on Reducing the Errors of Fourth-Grade Elementary Students while Solving Fraction Word Problems. *Archives of Pharmacy Practice*, *11*(S1), 80–86.

Andersson, U. (2007). The contribution of working memory to children's mathematical word problem solving. *Applied Cognitive Psychology*, *21*(9), 1201–1216. https://doi.org/10.1002/acp.1317

Bartalis, Á., Péntek, I., & Zsoldos-Marchiș, I. (2023). A Pilot Study on Investigating Primary School Students' Eye Movements While Solving Compare Word Problems. *Open Education Studies*, *5*(1), 20220207. https://doi.org/doi:10.1515/edu-2022-0207

Bartalis, Á., & Zsoldos-Marchiș, I. (2023). How task instruction influences the performance and the text reading pattern in case of inconsistent compare problems in primary school. *Acta Didactica Napocensia*, *16*(1), 252–266. https://doi.org/10.24193/and.16.1.17

Bartalis, Á., Zsoldos-Marchiş, I., & Gálfi, B. (2024). Investigating the Effect of Different Solving Methods in the Case of Inconsistent Compare Problem. *International Journal for Technology in Mathematics Education*, *31*(2), 45–52. https://doi.org/10.1564/tme_v31.2.02

Barton, M., Heidema, C., & Jordan, D. (2002). Teaching Reading in Mathematics and Science. *Educational Leadership*, *60*.

Bingolbali, E., Akkoç, H., Ozmantar, M. F., & Demir, S. (2011). Pre-Service and In-Service Teachers' Views of the Sources of Students' Mathematical Difficulties. *International Electronic Journal of Mathematics Education*, *6*(1), 40–59. https://doi.org/10.29333/iejme/260

Björn, P. M., Aunola, K., & Nurmi, J.-E. (2016). Primary school text comprehension predicts mathematical word problem-solving skills in secondary school. *Educational Psychology*, *36*(2), 362–377. https://doi.org/10.1080/01443410.2014.992392

Blum, W., & Borromeo Ferri, R. (2009). Mathematical Modelling: Can It Be Taught And Learnt? *Journal of Mathematical Modelling and Application*, *1*.

Blum, W., & Leiß, D. (2007). How do Students and Teachers Deal with Modelling Problems? In *Mathematical Modelling* (pp. 222–231). Elsevier. https://doi.org/10.1533/9780857099419.5.221

Boonen, A. J. H., de Koning, B. B., Jolles, J., & van der Schoot, M. (2016). Word Problem Solving in Contemporary Math Education: A Plea for Reading Comprehension Skills Training. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00191

Boonen, A., & Jolles, J. (2015). Second Grade Elementary School Students? Differing Performance on Combine, Change and Compare Word Problems. *International Journal of School and Cognitive Psychology*, *2*(2). https://doi.org/10.4172/2469-9837.1000122

Brysbaert, M., & Vitu, F. (1998). Word Skipping. In *Eye Guidance in Reading and Scene Perception* (pp. 125–147). Elsevier. https://doi.org/10.1016/B978-008043361-5/50007-9

Carpenter, T. P., & Moser, J. M. (1984). The Acquisition of Addition and Subtraction Concepts in Grades One through Three. *Journal for Research in Mathematics Education*, *15*(3), 179. https://doi.org/10.2307/748348

Chapman, O. (2006). Classroom Practices for Context of Mathematics Word Problems. *Educational Studies in Mathematics*, *62*(2), 211–230. https://doi.org/10.1007/s10649-006-7834-1

Clark, H. H. (1969). Linguistic processes in deductive reasoning. *Psychological Review*, *76*(4), 387–404. https://doi.org/10.1037/h0027578

Coquin-Viennot, D., & Moreau, S. (2003). Highlighting the role of the episodic situation model in the solving of arithmetical problems. *European Journal of Psychology of Education*, *18*(3), Article 3. https://doi.org/10.1007/BF03173248

Cramer, K., Bezuk, N., Behr, M., Laing, R. A., & Channell, D. E. (1989). Activities: Proportional Relationships and Unit Rates. *The Mathematics Teacher*, *82*(7), 537–540. https://doi.org/10.5951/MT.82.7.0537

Cruz, A. B. D., & Yangco, R. T. (2023). Epistemological, Psychological, and Pedagogical Sources of Difficulties of Senior High School Students on Problem-Solving Tasks in General Mathematics: Basis for Teaching Guide Recommendations. *Psychology and Education: A Multidisciplinary Journal (Psych Educ)*, *12*, 530–539. https://doi.org/10.5281/zenodo.8260635

Csépe, V. (2006). *Az olvasó agy*. Akad. K.

Csíkos, C. (2003). Matematikai szöveges feladatok megértésének problémái 10-11 éves tanulók körében. *Magyar Pedagógia*, *103*(1), 35–55.

Csíkos, C., Szitányi, J., & Kelemen, R. (2012). The effects of using drawings in developing young children's mathematical word problem solving: A design experiment with third-grade Hungarian students. *Educational Studies in Mathematics*, *81*(1), 47–65. https://doi.org/10.1007/s10649-011-9360-z

Daroczy, G., Artemenko, C., Wolska, M., Meurers, D., & Nuerk, H.-C. (2025). Are text comprehension and calculation processes in word problem solving sequential or interactive? An eye-tracking study in children. *Canadian Journal of Experimental Psychology / Revue Canadienne de Psychologie Expérimentale*, *79*(2), 206–211. https://doi.org/10.1037/cep0000366

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*, *06*. https://doi.org/10.3389/fpsyg.2015.00348

Davis-Dorsey, J., Ross, S. M., & Morrison, G. R. (1991). The role of rewording and context personalization in the solving of mathematical word problems. *Journal of Educational Psychology*, *83*(1), 61–68. https://doi.org/10.1037/0022-0663.83.1.61

de Corte, E., & Verschaffel, L. (1985). Beginning first graders' initial representation of arithmetic word problems. *The Journal of Mathematical Behavior*, *4*(1), 3–21.

De Corte, E., Verschaffel, L., & Pauwels, A. (1990). Influence of the semantic structure of word problems on second graders' eye movements. *Journal of Educational Psychology*, *82*(2), Article 2. https://doi.org/10.1037/0022-0663.82.2.359

de Koning, B. B., Boonen, A. J. H., Jongerling, J., van Wesel, F., & van der Schoot, M. (2022). Model method drawing acts as a double-edged sword for solving inconsistent word problems. *Educational Studies in Mathematics*, *111*(1), 29–45. https://doi.org/10.1007/s10649-022-10150-8

De Koning, B. B., Boonen, A. J. H., & Van Der Schoot, M. (2017). The consistency effect in word problem solving is effectively reduced through verbal instruction. *Contemporary Educational Psychology*, *49*, 121–129. https://doi.org/10.1016/j.cedpsych.2017.01.006

Dewolf, T., Van Dooren, W., Hermens, F., & Verschaffel, L. (2015). Do students attend to representational illustrations of non-standard mathematical word problems, and, if so, how helpful are they? *Instructional Science*, *43*(1), 147–171. https://doi.org/10.1007/s11251-014-9332-7

Dijk, T. A. van, & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic press.

Ding, H., & Homer, M. (2020). Interpreting mathematics performance in PISA: Taking account of reading performance. International Journal of Educational Research, 102, 101566. https://doi.org/10.1016/j.ijer.2020.101566

Doherty-Sneddon, G., Bruce, V., Bonner, L., Longbotham, S., & Doyle, C. (2002). Development of gaze aversion as disengagement from visual information. *Developmental Psychology*, *38*(3), 438–445. https://doi.org/10.1037/0012-1649.38.3.438

Dröse, J. (2019). Comprehending mathematical problem texts – Fostering subject-specific reading strategies for creating mental text representations. In U. T. Jankvist, M. van den Heuvel-Panhuizen, & M. Veldhuis (Eds.), *Eleventh Congress of the European Society for Research in Mathematics Education* (Vol. TWG06, Issue 10). Freudenthal Group. https://hal.science/hal-02408765

Dröse, J., Prediger, S., Neugebauer, P., Delucchi Danhier, R., & Mertins, B. (2021). Investigating Students' Processes of Noticing and Interpreting Syntactic Language Features in Word Problem Solving through Eye-Tracking. *International Electronic Journal of Mathematics Education*, *16*(1), em0625. https://doi.org/10.29333/iejme/9674

Duchowski, A. (2007). *Eye Tracking Methodology*. Springer London. https://doi.org/10.1007/978-1-84628-609-4

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., Schatschneider, C., & Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology*, *98*(1), 29–43. https://doi.org/10.1037/0022-0663.98.1.29

Fuchs, L. S., Gilbert, J. K., Fuchs, D., Seethaler, P. M., & N. Martin, B. (2018). Text Comprehension and Oral Language as Predictors of Word-Problem Solving: Insights into Word-Problem Solving as a Form of Text Comprehension. *Scientific Studies of Reading*, *22*(2), 152–166. https://doi.org/10.1080/10888438.2017.1398259

García, A. I., Jiménez, J. E., & Hess, S. (2006). Solving Arithmetic Word Problems: An Analysis of Classification as a Function of Difficulty in Children With and Without Arithmetic LD. *Journal of Learning Disabilities*, *39*(3), 270–281. https://doi.org/10.1177/00222194060390030601

Glenberg, A. M., & Langston, W. E. (1992). Comprehension of illustrated text: Pictures help to build mental models. *Journal of Memory and Language*, *31*(2), 129–151. https://doi.org/10.1016/0749-596X(92)90008-L

González-Calero, J. A., Berciano, A., & Arnau, D. (2020). The role of language on the reversal error. A study with bilingual Basque-Spanish students. *Mathematical Thinking and Learning*, *22*(3), 214–232. https://doi.org/10.1080/10986065.2020.1681100

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). DISCOURSE COMPREHENSION. *Annual Review of Psychology*, *48*(1), 163–189. https://doi.org/10.1146/annurev.psych.48.1.163

Greer, B. (1997). Modelling reality in mathematics classrooms: The case of word problems. *Learning and Instruction*, *7*(4), 293–307. https://doi.org/10.1016/S0959-4752(97)00006-6

Gros, H., Sander, E., & Thibaut, J.-P. (2020). *The cognitive cost of semantic recoding: An eye-tracking study on arithmetic word problem solving*. PsyArXiv. https://doi.org/10.31234/osf.io/gbk6n

Haghverdi, M. (2012). *The Influence of Familiarization and Conceptual Rewording in Facilitating the Students' Performance in Solving Types of Mathematics Word Problems*. *2*, 3711–3718.

Hegarty, M., Mayer, R. E., & Green, C. E. (1992). Comprehension of arithmetic word problems: Evidence from students' eye fixations. *Journal of Educational Psychology*, *84*(1), Article 1. https://doi.org/10.1037/0022-0663.84.1.76

Hegarty, M., Mayer, R. E., & Monk, C. A. (1995). Comprehension of arithmetic word problems: A comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology*, *87*(1), Article 1. https://doi.org/10.1037/0022-0663.87.1.18

Hudson, T. (1983). Correspondences and Numerical Differences between Disjoint Sets. *Child Development*, *54*(1), 84. https://doi.org/10.2307/1129864

Ibarra, C. G., & Lindvall, C. M. (1982). Factors Associated with the Ability of Kindergarten Children to Solve Simple Arithmetic Story Problems. *The Journal of Educational Research*, *75*(3), 149–156. https://doi.org/10.1080/00220671.1982.10885372

Jitendra, A. K., Griffin, C. C., Deatline-Buchman, A., & Sczesniak, E. (2007). Mathematical Word Problem Solving in Third-Grade Classrooms. *The Journal of Educational Research*, *100*(5), 283–302. https://doi.org/10.3200/JOER.100.5.283-302

Jõgi, A., & Kikas, E. (2016). Calculation and word problem-solving skills in primary grades – Impact of cognitive abilities and longitudinal interrelations with task-persistent behaviour. *British Journal of Educational Psychology*, *86*(2), 165–181. https://doi.org/10.1111/bjep.12096

Kail, R., & Hall, L. K. (1999). Sources of developmental change in children's word-problem performance. *Journal of Educational Psychology*, *91*(4), 660–668. https://doi.org/10.1037/0022-0663.91.4.660

Kaiser, G. (2017). *The teaching and learning of mathematical modeling* (J. Cai, Ed.; Haplo Repository; pp. 267–291). The National Council of Teachers of Mathematics, Inc. https://acuresearchbank.acu.edu.au/item/8658w/the-teaching-and-learning-of-mathematical-modeling

Karp, K. S., Bush, S. B., & Dougherty, B. J. (2019). Avoiding the Ineffective Keyword Strategy. *Teaching Children Mathematics*, *25*(7), 428–435. https://doi.org/10.5951/teacchilmath.25.7.0428

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* (pp. xvi, 461). Cambridge University Press.

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review*, *92*(1), 109–129. https://doi.org/10.1037/0033-295X.92.1.109

Knoblich, G., Ohlsson, S., & Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory & Cognition*, *29*(7), 1000–1009. https://doi.org/10.3758/BF03195762

Kozhevnikov, M., Kosslyn, S., & Shephard, J. (2005). Spatial versus object visualizers: A new characterization of visual cognitive style. Memory & Cognition, 33(4), 710–726. https://doi.org/10.3758/BF03195337

Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S. W.-Y., Lee, M.-H., Chiou, G.-L., Liang, J.-C., & Tsai, C.-C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, *10*, 90–115. https://doi.org/10.1016/j.edurev.2013.10.001

Leiss, D., Plath, J., & Schwippert, K. (2019). Language and Mathematics—Key Factors influencing the Comprehension Process in reality-based Tasks. *Mathematical Thinking and Learning*, *21*(2), 131–153. https://doi.org/10.1080/10986065.2019.1570835

Leong, C. K., & Jerred, W. D. (2001). Effects of consistency and adequacy of language information on understanding elementary mathematics word problems. *Annals of Dyslexia*, *51*(1), 275–298. https://doi.org/10.1007/s11881-001-0014-1

Lewis, A. B., & Mayer, R. E. (1987). Students' miscomprehension of relational statements in arithmetic word problems. *Journal of Educational Psychology*, *79*(4), 363–371. https://doi.org/10.1037/0022-0663.79.4.363

Lilienthal, A. J., & Schindler, M. (2019). *Current Trends in the Use of Eye Tracking in Mathematics Education Research: A PME Survey*. https://doi.org/10.48550/ARXIV.1904.12581

Liversedge, S. P., Gilchrist, I. D., & Everling, S. (2011). *The Oxford handbook of eye movements*. Oxford university press.

McConkie, G. W., & Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society*, *8*(5), 365–368. https://doi.org/10.3758/BF03335168

Miller, I. (1986). Techniques for Teaching Word Problems. *Lifelong Learning*, *9*(4), 29–30.

Ministerul Educației Naționale. (2013). Programa școlară pentru disciplina Matematică. Clasele a III-a – a IV-a. București: Ministerul Educației Naționale.

Múñez, D., Orrantia, J., & Rosales, J. (2013). The Effect of External Representations on Compare Word Problems: Supporting Mental Model Construction. *The Journal of Experimental Education*, *81*(3), 337–355. https://doi.org/10.1080/00220973.2012.715095

Muth, K. D. (1984). Solving arithmetic word problems: Role of reading and computational skills. *Journal of Educational Psychology*, *76*(2), 205–210. https://doi.org/10.1037/0022-0663.76.2.205

Ng, C.-T., Lung, T.-C., & Chang, T.-T. (2021). Operation-Specific Lexical Consistency Effect in Fronto-Insular-Parietal Network During Word Problem Solving. *Frontiers in Human Neuroscience*, *15*. https://www.frontiersin.org/articles/10.3389/fnhum.2021.631438

Niss, M. A. (2003). Mathematical competencies and the learning of mathematics: The Danish KOM project. In A. Gagatsis & S. Papastavridis (Eds.), *3rd Mediterranean Conference on Mathematical Education—Athens, Hellas 3-4-5 January 2003* (pp. 116–124). Hellenic Mathematical Society.

Oakhill, J., Cain, K., & Elbro, C. (2014). *Understanding and Teaching Reading Comprehension* (0 ed.). Routledge. https://doi.org/10.4324/9781315756042

OECD (2023). PISA 2022 Results (Volume I): The State of Learning and Equity in Education. PISA, OECD Publishing, Paris. https://doi.org/10.1787/53f23881-en.

Orrantia, J., & Múñez, D. (2013). Arithmetic word problem solving: Evidence for a magnitude-based mental representation. *Memory & Cognition*, *41*(1), 98–108. https://doi.org/10.3758/s13421-012-0241-1

Ott, B. (2020). Learner-generated graphic representations for word problems: An intervention and evaluation study in grade 3. *Educational Studies in Mathematics*, *105*(1), 91–113. https://doi.org/10.1007/s10649-020-09978-9

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

Passolunghi, M. C., De Blas, G. D., Carretti, B., Gomez-Veiga, I., Doz, E., & Garcia-Madruga, J. A. (2022). The role of working memory updating, inhibition, fluid intelligence, and reading comprehension in explaining differences between consistent and inconsistent arithmetic word-problem-solving performance. *Journal of Experimental Child Psychology*, *224*, 105512. https://doi.org/10.1016/j.jecp.2022.105512

Passolunghi, M. C., & Siegel, L. S. (2001). Short-Term Memory, Working Memory, and Inhibitory Control in Children with Difficulties in Arithmetic Problem Solving. *Journal of Experimental Child Psychology*, *80*(1), 44–57. https://doi.org/10.1006/jecp.2000.2626

Pearce, D. L., Bruun, F., Skinner, K., & Lopez-Mohler, C. (2013). What Teachers Say About Student Difficulties Solving Mathematical Word Problems in Grades 2-5. *International Electronic Journal of Mathematics Education*, *8*(1), 3–19. https://doi.org/10.29333/iejme/271

Pólya, G. (2004). *How to solve it: A new aspect of mathematical method* (Expanded Princeton Science Library edition). Princeton University Press.

Pongsakdi, N., Kajamies, A., Veermans, K., Lertola, K., Vauras, M., & Lehtinen, E. (2020). What makes mathematical word problem solving challenging? Exploring the roles of word problem characteristics, text comprehension, and arithmetic skills. *ZDM*, *52*(1), 33–44. https://doi.org/10.1007/s11858-019-01118-9

Purcar, A.-M., Bocoș, M., Pop, A.-L., Roman, A., Rad, D., Mara, D., Crișan, C., Răduț-Taciu, R., Mara, E.-L., Todor, I., Muntean-Trif, L., Neacșu, M.-G., Costache Colareza, C., Maier, M., Tăușan-Crișan, L., Triff, Z., Baciu, C., Marin, D.-C., & Triff, D.-G. (2024). The Effect of Visual Reasoning on Arithmetic Word Problem Solving. *Education Sciences*, *14*(3), 278. https://doi.org/10.3390/educsci14030278

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K. (1999). What have we learning about eye movements during reading? In *Converging methods for understanding reading and dyslexia*. The MIT Press.

Rayner, K., & Morris, R. K. (1992). Eye movement control in reading: Evidence against semantic preprocessing. *Journal of Experimental Psychology. Human Perception and Performance*, *18*(1), 163–172.

Reusser, K. (1990). *From text to situation to equation: Cognitive simulation of understanding and solving mathematical word problems*. https://doi.org/10.5167/UZH-261821

Reusser, K., & Stebler, R. (1997). Every word problem has a solution—The social rationality of mathematical modeling in schools. *Learning and Instruction*, *7*(4), 309–327. https://doi.org/10.1016/S0959-4752(97)00014-5

Richardson, J. S., & Morgan, R. F. (2002). *Reading to Learn in the Content Areas*. Wadsworth Publishing Company.

Riley, M. S., & Greeno, J. G. (1988). Developmental Analysis of Understanding Language About Quantities and of Solving Problems. *Cognition and Instruction*, *5*(1), 49–101. https://doi.org/10.1207/s1532690xci0501_2

Riley, M. S., Greeno, J. G., & Heller, J. I. (1983). Development of Children's Problem-Solving Ability in Arithmetic. In *The Development of Mathematical Thinking* (H. Ginsburg, pp. 153–196). Academic Press.

Roe, B. D., Stoodt, B. D., & Burns, P. C. (2001). *Secondary school literacy instruction: The content areas* (7th ed). Houghton Mifflin Co.

Roth, L., Nuerk, H.-C., Cramer, F., & Daroczy, G. (2025). *Can't help processing numbers with text: Eye-tracking evidence for simultaneous instead of sequential processing of text and numbers in arithmetic word problems*. PsyArXiv. https://doi.org/10.31234/osf.io/3uvt8_v2

Ryder, R. J., & Graves, M. F. (2008). *Reading and learning in content areas*. Recording for the Blind & Dyslexic.

Schell, V. J. (1982). Learning Partners: Reading and Mathematics. *The Reading Teacher*, *35*(5), 544–548. JSTOR.

Schnotz, W. (2005). An Integrated Model of Text and Picture Comprehension. In R. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (1st ed., pp. 49–70). Cambridge University Press. https://doi.org/10.1017/CBO9780511816819.005

Staub, F. C., & Reusser, K. (1995). The role of presentational structures in understanding and solving mathematical word problems. In *Discourse comprehension: Essays in honor of Walter Kintsch.* (pp. 285–305). Lawrence Erlbaum Associates, Inc.

Steklács, J. (2019). A szemkamerás vizsgálati módszer lehetőségei a pedagógiai szempontú kutatásokban. In *Szemkamerás vizsgálatok a pedagógiai kutatásokban* (pp. 5–25). Kaposvári Egyetem Pedagógiai Kar.

Stephany, S. (2021). The influence of reading comprehension on solving mathematical word problems: A situation model approach. In A. Fritz, E. Gürsoy, & M. Herzog (Eds.), *Diversity Dimensions in Mathematics and Language Learning* (pp. 370–395). De Gruyter. https://doi.org/10.1515/9783110661941-019

Stern, E., & Lehrndorfer, A. (1992). The role of situational context in solving word problems. *Cognitive Development*, *7*(2), 259–268. https://doi.org/10.1016/0885-2014(92)90014-I

Strohmaier, A. R., MacKay, K. J., Obersteiner, A., & Reiss, K. M. (2020). Eye-tracking methodology in mathematics education research: A systematic literature review. *Educational Studies in Mathematics*, *104*(2), 147–200. https://doi.org/10.1007/s10649-020-09948-1

Stuart, C., & Thurlow, D. (2000). Making It Their Own: Preservice Teachers' Experiences, Beliefs, and Classroom Practices. *Journal of Teacher Education*, *51*(2), 113–121. https://doi.org/10.1177/002248710005100205

Swanson, H. L., Cooney, J. B., & Brock, S. (1993). The Influence of Working Memory and Classification Ability on Children′s Word Problem Solution. *Journal of Experimental Child Psychology*, *55*(3), 374–395. https://doi.org/10.1006/jecp.1993.1021

Tapiero, I. (2007). *Situation Models and Levels of Coherence* (0 ed.). Routledge. https://doi.org/10.4324/9780203810040

Terry, P. W. (1921). The Reading Problem in Arithmetic. *Journal of Educational Psychology*, *12*(7), 365–377. https://doi.org/10.1037/h0076019

Thevenot, C., & Barrouillet, P. (2014). Arithmetic Word Problem Solving and Mental Representations. In R. Cohen Kadosh & A. Dowker (Eds.), *The Oxford Handbook of Numerical Cognition* (1st ed., pp. 158–179). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199642342.013.043

van der Schoot, M., Bakker Arkema, A. H., Horsley, T. M., & van Lieshout, E. C. D. M. (2009). The consistency effect depends on markedness in less successful but not successful problem solvers: An eye movement study in primary school children. *Contemporary Educational Psychology*, *34*(1), 58–66. https://doi.org/10.1016/j.cedpsych.2008.07.002

Van Dooren, W., & Inglis, M. (2015). Inhibitory control in mathematical thinking, learning and problem solving: A survey. *ZDM*, *47*(5), 713–721. https://doi.org/10.1007/s11858-015-0715-2

Vanderlinde, L. F. (1964). Does the Study of Quantitative Vocabulary Improve Problem-Solving? *The Elementary School Journal*, *65*(3), 143–152. JSTOR.

Verschaffel, L., & De Corte, E. (1993). A decade of research on word problem solving in Leuven: Theoretical, methodological, and practical outcomes. *Educational Psychology Review*, *5*(3), 239–256. https://doi.org/10.1007/BF01323046

Verschaffel, L., Greer, B., & De Corte, E. (2000). Making Sense of Word Problems. *Educational Studies in Mathematics*, *42*(2), 211–213. https://doi.org/10.1023/A:1004190927303

Verschaffel, L., Schukajlow, S., Star, J., & Van Dooren, W. (2020). Word problems in mathematics education: A survey. *ZDM*, *52*(1), 1–16. https://doi.org/10.1007/s11858-020-01130-4

Vessonen, T., Dahlberg, M., Hellstrand, H., Widlund, A., Söderberg, P., Korhonen, J., Aunio, P., & Laine, A. (2025). Individual characteristics associated with elementary school children's mathematical word problem-solving skills: A systematic review and meta-analysis. *Review of Education*, *13*(1), e70045. https://doi.org/10.1002/rev3.70045

Vicente, S., Orrantia, J., & Verschaffel, L. (2008). Influence of mathematical and situational knowledge on arithmetic word problem solving: Textual and graphical aids. *Infancia Y Aprendizaje*, *31*(4), 463–483. https://doi.org/10.1174/021037008786140959

Vilenius-Tuohimaa, P. M., Aunola, K., & Nurmi, J. (2008). The association between mathematical word problems and reading comprehension. *Educational Psychology*, *28*(4), 409–426. https://doi.org/10.1080/01443410701708228

Wu, C.-J., Liu, C.-Y., Yang, C.-H., & Jian, Y.-C. (2021). Eye-movements reveal children's deliberative thinking and predict performance on arithmetic word problems. *European Journal of Psychology of Education*, *36*(1), 91–108. https://doi.org/10.1007/s10212-020-00461-w

Yeatts, K. L. (1991). *Manipulatives: Motivating Mathematics*. Dade Public Education Fund. https://eric.ed.gov/?id=ED355097