DOI 10.1007/s10986-024-09642-2
Lithuanian Mathematical Journal

Analyzing arithmetic word problems: Blink of an eye for textbooks authors

Ieva Kilienė

Institute of Mathematics, Vilnius University, Naugarduko 24, LT-03225 Vilnius, Lithuania
(e-mail: ieva.kiliene@mif.vu.lt)

Received April 13, 2024; revised June 19, 2024

Abstract. In this paper, we study the frequency of different types of arithmetic word problems (AWP) in Lithuanian textbooks. The results show the lack of variety among types of AWP. We propose the framework for analysis of the frequency of types of AWP in a textbook and apply it to a particular set of primary school textbooks. We use a statistical method to compare the sample from the textbook rather than from the entire textbook. Also, we compare the proportions of types of AWP in Lithuanian textbooks with those in Singaporean and Spanish textbooks. The approach adopted in the paper can be used to analyze other textbooks from different countries.

MSC: 97D10, 97U20, 20F10

Keywords: quantitive reasoning, word problems, additive word problems, multiplicative word problems

1 Introduction

Primary school students' first exposures to reasoning come through word problems. In primary school, students begin to learn quantitative reasoning, which is the ability to use basic mathematics skills, such as algebra, to analyze and interpret real-world quantitative information [3].

Thompson [13] explains quantitative reasoning conception as follows:

- A prominent characteristic of quantitative reasoning is that numbers and numeric relationships are of secondary importance and do not enter into the primary analysis of a situation. What is important is relationships among quantities [13].

The quantitative reasoning is necessary to solve various arithmetic word problems (AWP) using strategies other than direct modeling or the keyword strategy. Two types of quantitative relations appear in word problems, additive and multiplicative reasoning [9]. This is a different area of mathematics for primary school students compared to what they have previously encountered [9]. However, there is a risk that word problem-solving can become merely a mechanical process. It is important to teach students various problem types to develop students' quantitative reasoning skills. Quantitative reasoning involves understanding relationships among quantities. Therefore it is important to teach various types of relationships and different types of AWP. Greater diversity in AWP of additive and multiplicative types would imply more efficient cultivation of quantitative reasoning skills. However, the frequency of AWP types across textbooks in different countries is similar

[12, 16], and some types of AWP are absent from primary school textbooks. We aim to describe a model that would enable the selection of a balanced frequency of AWP in a textbook. Also, we aim to describe a framework to study more countries and textbooks in a time-saving manner.

Ministry of Education, Science, and Sports of Lithuania states that the goal of education is continuous improvement [19]. There is still plenty of space to improve Lithuania's education system, which is the case with mathematics teaching. In the last few years, many changes have been happening regarding the mathematics curriculum: the curriculum has been updated, new textbooks are currently being written, and the necessity of avoiding previous issues has been enhanced. Lithuanian students struggle with word problems, as do many students all over the world [15]. Students solve word problems from the first grade to the graduation of school and also at final exams, but there are more questions than answers on how to teach word problems. While over the past 50 years, this issue has been widely researched in Europe, studies have explored comprehensions, solution strategies, cognitive resources, graphical representations, and the impact of teaching environment (such as textbooks, software, and teachers) on learners' word problem-solving (survey, 2020 [15]), word problems in teaching and learning are not being researched in Lithuania.

Vicente et al. [16] compare Singapore and Spain with the idea that Singapore is an example of a country with a high student performance in mathematics and Spain is an average-performing country, and some aspects of textbooks can relate to that [16]. According to Trends in Mathematics and Science Study (TIMSS) report [8], 55% 4th-grade Singaporean students achieved advanced level results, and only 4% of Spain students achieved such results; compared in Lithuania, 13% of students achieved advanced level results. Cognitive abilities at Reasoning tasks scores are also related to mathematical reasoning. Lithuanian 4th-grade students in the 2019 TIMSS had a 534 average scale score [8]. According to cognitive abilities at Reasoning tasks, Lithuania is between these two countries by results, Singapore (614 average scale score) and Spain (497 average scale score). The authors test the idea that organizational illustrations help students reach better results. This was the main difference found in the textbooks of Spain and Singapore. We are going to compare AWP with proportion of organizational illustrations in Lithuanian with proportion in Singapore (see hypothesis (3.1)).

The theoretical framework and methodology are built on Vicente et al.'s work [16] with our changes.

Similarities. We classify AWP using the same coding system: 20 additive AWP types, 14 multiplicative AWP types, and AWP with illustrations in 3 categories: figurative, informational, and organizational. We compare frequencies of types of AWP calculated for Lithuania, Singapore, and Spain.

Enhancements and additions. Our research focuses on textbooks without workbooks, whereas [16] uses data from textbooks and workbooks together. To compare countries, we used unpublished data from Vicente et al.'s research on Spain and Singapore textbooks, excluding workbooks. Our work additionally examines the types of word problems grouped by mathematical operations (addition, subtraction, multiplication, division): 7 types of AWP require addition, 13 require subtractions, 5 require multiplication, and 9 require division. We divide all textbook sets into parts by time where these operations are presented and analyze these parts separately.

We use a statistical method to compare the sample from the textbook rather than from the entire textbook. We describe a theoretical model for the proportions of AWP types compared with p_0, where p_0 is the desired proportion.

Benefits of the new approach. This method is intended to facilitate broader research by including more countries and textbooks for comparison.

There is a need for broader research involving more countries and textbooks. After examining the word problems used in their textbooks, this paper aims to reveal how different countries can compare their results with those of other countries. We do not propose an ideal textbook model; this should be done after detailed and extensive research of many textbooks from various countries. The long-term goal is to create new, higher-quality teaching materials using the existing and future research.

This study focuses on the content of Lithuanian primary school mathematics textbooks, specifically analyzing the frequency, types, and illustrations of word problems. The study aims to provide insights into the extent to which Lithuanian textbooks promote quantitative reasoning skills among primary school students.

In this study, we are going to answer the following research questions:

(Q1) What are the frequencies of AWP in Lithuanian primary school textbooks and how does this distribution compare with the frequencies of AWP types in Spain and Singapore?
(Q2) What insights can be gained by classifying AWP types according to mathematical operations (addition, subtraction, multiplication, division) or exploring their frequency, focusing on sections related to these operations within textbooks?
(Q3) Can a randomly generated sample of textbook pages be used to analyze the textbook?
(Q4) Do proportions of AWP differ in Lithuania compared to Singapore and Spain by different groups of AWP?

1.1 Theoretical framework

We define an AWP as a verbal description of problem situations in which one or more questions are raised, and the answer can be obtained by applying mathematical operations to numerical data presented in the problem statement.

AWP *types.* Arithmetic word problems in primary school textbooks can be classified into additive and multiplicative categories based on the mathematical operations involved. Both categories can be divided by subcategories and types of AWP depending on actions in word problem situations. The diversity of problem types in the textbooks is believed to enhance reasoning [14, 16, 17].

Vicente et al. [16] claim that the diversity of problem types improves mathematical reasoning. Also, we think that the diversity of problem types is related to a specific type of reasoning, quantitative reasoning. The word problem is a story about the relationship between quantities, and quantitative reasoning is the ability to analyze and interpret real-world quantitative information. Additive reasoning and multiplicative reasoning are kinds of quantitative reasoning [9], and for both these reasoning kinds, there are different categories of AWP (see Tables 1 and 2).

Nunes et al. [9] review three different methods for teaching quantitative reasoning, including both additive and multiplicative reasoning, using word problems. One of the teaching methods uses organizational illustrations, which we define below. We are taking into account Nunes et al.'s idea [9] that quantitative reasoning is effective when there is a variety of additive and multiplicative AWPs presented in the textbook. Additive and multiplicative AWP categories can be further classified into specific types by action. Various authors have proposed different classifications [4, 14]. We use the classification from Vicente et al. [16], which includes 20 types of additive problems (Table 1) and 14 types of multiplicative problems (Table 2). For problems involving two or more mathematical operations, we divide them into parts and assign an AWP type to each part.

Illustrations. Illustrations are part of the AWP for 1–4 grade students; some illustrations are only for beauty, some give numerical information, and some can help to solve a word problem. The Concrete-Pictorial-Abstract (CPA) approach has been used in Singapore since the 1980s [6]. The CPA approach effectively improves students' attitude and performance in mathematics [11]. There is evidence that organizational illustrations can help students solve word problems [10, 18].

The method where such illustrations are used is called the schema-based instruction, and it emphasizes both the semantic structure of the problem and its mathematical structure [7].

We also use the same classification as Vicente et al. [16] to examine the illustrations. We distinguish three types of illustrations: figurative, informational, and organizational.

- Figurative illustration in word problems is some visual object without any numeric information of word problem story (Fig. 1(a)).
- Informational illustration has some or all numeric values of the problem (Fig. 1(b)).
- An organizational illustration is a schematic illustration. It represents mathematical structure (a part or the whole) in such a way that "enables students to understand the mathematical relations between the problem sets" [16] (Fig. 1(c)).

Table 1. Additive AWP types by action (adapted from [16])

	Win/More/Whole	Lose/Less/Part
	Change	
Final	1. Goda had 5 stickers. Adas gave her 8 more stickers. How many stickers does Goda have now? **[Code: Ch+F]**	2. Goda had 13 stickers. She gave 5 stickers to Adas. How many stickers does she have left? **[Code: Ch-F]**
Change	3. Goda has 5 stickers. How many more stickers does she need to buy to have 13 stickers? **[Code: Ch+Ch]**	4. Goda had 13 stickers. She gave a few stickers to Adas. Now she has 8 stickers. How many stickers did Goda give to Adas? **[Code: Ch-Ch]**
Initial	5. Goda had some stickers. Adas gave her 5 more stickers. Now she has 13 stickers. How many stickers did Goda have initially? **[Code: Ch+Init]**	6. Goda had some stickers. She gave 5 stickers to Adas. Now she has 8 stickers. How many stickers did Goda have initially? **[Code: Ch-Init]**
	Compare	
Difference	1. Goda has 13 stickers. Adas has 5 stickers. How many more stickers does Goda have than Adas? **[Code: Com+Dif]**	2. Goda has 13 stickers. Adas has 5 stickers. How many fewer stickers does Adas have compared to Goda? **[Code: Com-Dif]**
Compared	3. Adas has 5 stickers. Goda has 8 more stickers than Adas. How many stickers does Goda have? **[Code: Com+Com]**	4. Goda has 13 stickers. Adas has 5 fewer stickers than Goda. How many stickers does Adas have? **[Code: Com-Com]**
Reference	5. Goda has 13 stickers. She has 5 more stickers than Adas. How many stickers does Adas have? **[Code: Com+Ref]**	6. Adas has 5 stickers. He has 8 fewer stickers than Goda. How many stickers does Goda have? **[Code: Com-Ref]**
	Equalize	
Difference	1. Goda has 13 stickers. Adas has 5 stickers. How many more stickers does Adas need to buy to have the same number of stickers as Goda? **[Code: Eq+Dif]**	2. Goda has 13 stickers. Adas has 5 stickers. How many stickers does Goda need to lose to have the same number of stickers as Adas? **[Code: Eq-Dif]**
Compared	3. Goda has 13 stickers. If Adas gets 5 more stickers, then he will have the same number of stickers as Goda. How many stickers does Adas have? **[Code: Eq+Com]**	4. Adas has 5 stickers. If Goda loses 8 stickers, then she will have the same number of stickers as Adas. How many stickers does Goda have? **[Code: Eq-Com]**
Reference	5. Adas has 5 stickers. If he gets 8 more stickers, then he will have the same number of stickers as Goda. How many stickers does Goda have? **[Code: Eq+Ref]**	6. Goda has 13 stickers. If she loses 5 stickers, then she will have the same number of stickers as Adas. How many stickers does Adas have? **[Code: Eq-Ref]**
	Combine	
	1. Goda has 5 red stickers and 8 blue stickers. How many stickers does she have in total? **[Code: Comb+]**	2. Goda has 13 stickers. 5 are red, and the rest are blue. How many blue stickers does Goda have? **[Code: Comb-]**

Table 2. Multiplicative AWP types by operation (adapted from [12, 16])

	Rate	
	Multiple	Simple
Product	1. If you draw 3 drawings in 6 days, the how many drawings will you draw in 4 days, drawing at the same rate? **[Code: RatePrM]**	2. You bought 3 notebooks, and each notebook costs 5 €. How much did you spend in total? **[Code: RatePrS]**
Multiplying		3. You spent 15 € on 3 notebooks. How much did one notebook cost? **[Code: RateMul-ing]**
Multiplier		4. You spent 15 € on notebooks. One notebook costs 5 €. How many notebooks did you buy? **[Code: RateMul-ier]**
	Compare	
	Times more	Times less
Difference	1. I spent 15 €, you spent 5 €. How many times more did I spend than you? **[Code: Com+Dif]**	2. I spent 15 €, you spent 5 €. How many times less did you spend than me? **[Code: Com-Dif]**
Compared	3. I spent 5 €, you spent 3 times € as much as I did. How many € did you spend? **[Code: Com+Com]**	4. I spent 15 €, and you spent 3 times less € than me. How many € did you spend? **[Code: Com-Com]**
Reference	5. I spent 15 €, I spent 3 times as much as you did. How many € did you spend? **[Code: Com+Ref]**	6. I spent 5 €. I spent 3 times less than you. How many € did you spend? **[Code: Com-Ref]**

Table 2 (*Continued from previous page*)

	Product	Measure
Cartesian product		
	1. I have 3 different pens and 5 different notebooks. How many different ways can they be combined? **[Code: CarPr+]**	2. I have 3 different pens and several different notebooks. If I can combine them in 15 different ways, then how many notebooks do I have? **[Code: CarPr-]**
Rectangular matrix		
	1. I have a big piece of paper that is 3-meters long and 5-meters wide. What is the area of the plot? **[Code: RecMat+]**	2. I have a plot of paper that is 15 m^2. The plot measures 5-meters long. How long is the plot? **[Code: RecMat-]**

(a)

(b)

(c)

Figure 1. Examples of three types of illustrations from Lithuanian textbook. (a) Figurative: a visual object without any numeric information of word problem story; (b) Informational: illustration has some or all numeric values of the problem; (c) Organizational: an illustration represents the mathematical structure (a part or the whole) of the problem.

2 Materials and methods

This study follows Vicente et al. [16] comparison between two countries' textbooks. We did the same research for Lithuania and compared it to the existing results. Additionally, regarding the author's ideas, we described recommendations for the frequency of types of arithmetical word problems in textbooks in Lithuania. We made a wider analysis of the frequency between groups of types by mathematical operations and in different textbook segments.

2.1 Materials

We analyze the Lithuanian mathematics textbook set "TAIP" for primary grades 1–4 (2017–2020 years of edition). It consists of 12 books, and each of the four grades has 3 textbooks. We do not include workbooks in our analysis. It is worth noting that workbooks are not uniformly employed in every school in Lithuania; also, there are many different workbooks used, and some institutions utilize online or other activities as alternatives. Consequently, our research prioritized a comprehensive understanding of textbook properties and aimed to provide valuable recommendations for textbook authors, given that other educational materials in Lithuania predominantly rely on textbooks.

2.2 Method

2.2.1 Analysis of activities: AWPs vs. other mathematical activities (OMA)

We employed a method used by Vicente et al. [16]. We coded all activities in the textbook. In this study, an *activity* is defined as a task or set of related tasks that constitute a separate instructional activity on a textbook's page, as indicated by the heading, number, or instruction on top of the activity or by any other layout aspect [16]. An activity refers to a task or explanation that describes how to solve a problem and obtain the answer. It is associated with students' mathematical knowledge and calculations. All activities were divided into two categories: (i) arithmetical word problem (AWP) activities and (ii) other mathematical activities (OMA).

An AWP activity may consist of several AWPs coded in more detail.

2.2.2 Categories of arithmetic word problems

AWP describes a real situation that can be modeled using one of the mathematical operations (addition, subtraction, multiplication, division). We classified AWPs by mathematical operations into additive AWPs (which can be modeled by addition or subtraction) and multiplicative AWPs (which can be modeled by multiplication or division).

1. *Additive AWPs*. These problems are solved using addition or subtraction. There are four subcategories of situations: change, compare, combine, and equalize. We derived 20 AWP types from these subcategories depending on the unknown set and relation (addition, subtraction) as demonstrated in Table 1.
2. *Multiplicative AWPs*. These problems are solved using multiplication or division. There are four subcategories: rate, Cartesian product, compare, and rectangular matrix. We derived 14 types from these subcategories depending on the unknown set and relation (multiplication, division) as demonstrated in Table 2.

2.2.3 Categories of analysis: Illustrations

Three new categories, including illustrations, were created for AWPs.

1. *Figurative*. Illustrations of part of the situation or the entire situation without numerical data.
2. *Informational*. Illustrations of part of the situation or the entire situation, including part or all of the numerical data of the problem.
3. *Organizational*. Schematic illustrations that depict part or the entire situation and help to understand the relations between problem sets.

2.2.4 Assumptions for statistical method

Let $\mathbf{S}$ be a set of all Lithuanian mathematics textbooks for 1–4 grades. Let $A(s)$ be the number of all activities in the textbook s. The variable s can represent either a single textbook or a set of textbooks, depending on the hypothesis we want to test.

If we want to analyze specific textbook sections, then we can partition the textbook into parts $s_1, s_2, \dots, s_n$ and examine each part individually.

Consider the population

$$y_1(s),\ y_2(s),\ \dots,\ y_{A(s)}(s),$$

where i corresponds to the ith activity in the textbook s. Each $y_i \in B = \{0, 1\}$, and B is the set of activity types: 0 – other mathematical activity, 1 – AWP activity.

Then $\mathrm{AWP} = \{y_i(s) = 1\}$ is a set of all AWP activities, and one or more AWPs belong to one AWP activity.

Let $N(s)$ be the number of all AWPs in the textbook s.

Consider the population

$$x_1(s),\, x_2(s),\, \ldots,\, x_{N(s)}(s),$$

where i corresponds to the ith AWP in the textbook s. There each $x_i \in D = \{1, 2, \ldots, 34\}$, and D is the set of possible problem types (20 additive types and 14 multiplicative types as described in Tables 1 and 2).

We consider several characteristics of these populations:

$p_{\mathrm{AWP}}(s)$, proportion of AWP activities, and $\mathrm{AWP} = \{y_i(s)\colon\, y_i(s) = 1\}$;
$p_{\mathrm{OMA}}(s)$, proportion of OMA, and $\mathrm{OMA} = \{y_i(s)\colon\, y_i(s) = 0\}$;
$p_{\mathrm{AWP}(a)}(s)$, proportion of additive AWPs and $\mathrm{AWP}(a) = \{x_i(s)\colon\, x_i(s) \leqslant 20\}$;
$p_{\mathrm{AWP}(m)}(s)$, proportion of multiplicative AWPs, and $\mathrm{AWP}(m) = \{x_i(s)\colon\, x_i(s) > 20\}$;
$p_{\mathrm{AWP}(i)}(s)$, proportion of AWPs that include illustrations; $p_{\mathrm{AWP}(io)}(s)$ corresponds to organizational illustrations. $\mathrm{AWP}(io)$ is the subset of x_i that includes organizational illustrations.

We group types of AWPs and consider the following characteristics:

$p_{\mathrm{AWP}(+)}(s)$, proportion of AWPs that requires addition, and $\mathrm{AWP}(+) = \{x_i(s)\colon\, x_i(s) \leqslant 7\}$;
$p_{\mathrm{AWP}(-)}(s)$, proportion of AWPs that requires subtraction, and $\mathrm{AWP}(-) = \{x_i(s)\colon\, 7 < x_i(s) \leqslant 20\}$;
$p_{\mathrm{AWP}(*)}(s)$, proportion of AWPs that requires multiplication, and $\mathrm{AWP}(*) = \{x_i(s)\colon\, 20 < x_i(s) \leqslant 25\}$;
$p_{\mathrm{AWP}(:)}(s)$, proportion of AWPs that requires division, and $\mathrm{AWP}(*) = \{x_i(s)\colon\, 25 < x_i(s)\}$.

Mathematically, as an example,

$$p_{\mathrm{AWP}(a)}(s) = \frac{1}{N(s)} \sum_{i=1}^{N(s)} \mathbf{1}_{\mathrm{AWP(a)}}(x_i)(s).$$

We need to estimate these proportions. For this, we have several possibilities. We can fix a textbook and involve a Census, a sample survey that attempts to include the entire population in the sample. Then, page by page, we categorize all the activities in the textbook. It is very time consuming. Another way is to use random sampling from the population. To this aim, from the book s we randomly select $M(s)$ pages and obtain s samples $\hat{y}_1(s), \ldots, \hat{y}_{a(s)}(s)$ and $\hat{x}_1(s), \ldots, \hat{x}_{n(s)}(s)$, where $a(s)$ is the number of activities in selected pages, $n(s)$ is the number of AWPs in selected pages, and $\hat{y}_i(s) \in B$, $\hat{x}_i(s) \in A$.

Then we set

$$\hat{p}_{\mathrm{AWP}(a)}(s) = \frac{1}{n(s)} \sum_{i=1}^{n(s)} \mathbf{1}_{\mathrm{AWP}(a)} \sqrt[n]{\big(\hat{x}_i(s)\big)}$$

as an estimator of $p_{\mathrm{AWP}(a)}(s)$ and $p_{\mathrm{OMA}}(s)$. Then the 95% confidence interval for population parameter $p_{\mathrm{AWP}(a)}(s)$ is

$$\left(\hat{p}_{\mathrm{AWP}(a)} - 2\sqrt{\frac{\hat{p}_{\mathrm{AWP}(a)}(1 - \hat{p}_{\mathrm{AWP}(a)})}{n}},\, \hat{p}_{\mathrm{AWP}(a)} + 2\sqrt{\frac{\hat{p}_{\mathrm{AWP}(a)}(1 - \hat{p}_{\mathrm{AWP}(a)})}{n}}\right),$$

where $\hat{p} = \hat{p}_{\mathrm{AWP}(a)}(s)$ as an example, and $n = n(s)$. The same holds for all other proportions.

If we believe that, in a sense, effective proportion is, say, p_0, then we can test the hypothesis

$$\text{(H0)}\ \ p(s) = p_0 \quad \text{versus} \quad \text{(H1)}\ \ p \neq p_0\ \ (p < p_0 \text{ or } p > p_0).$$

We selected the Lithuanian textbook set s. It consists of 12 books. There are a total of 902 pages in the entire textbook set (excluding the introductory and end pages, which do not contain mathematical problems). A total of 50 pages were selected for a sample. If the number of pages in a book is m, then the number of pages chosen as a random sample is x, so that $m/902 = x/50$. For example, the first part of the textbook has activities on pages 8–83, that is 76 pages with activities; we need to select $50 \cdot 76/902 \approx 4$ pages from this textbook. We get randomly selected pages from this textbook. After these steps, all 50 randomly selected pages were obtained, and their activities were analyzed.

To test hypotheses of proportions equality, we first wanted to check if the sample matches the entire s. We test the hypothesis that the sample distribution of s matches the empirical distribution of total s of AWP types. To test this hypothesis, we use the chi-square test.

Our data satisfy the following assumptions:

- Data are obtained randomly.
- Data are quantitative.
- The groups of AWP types we compare are incompatible.
- The data values are independent.

We consider the following hypotheses:

(H0) The sample distribution of s matches the empirical distribution of the total s of AWP types.
(H1) The sample distribution of s does not match the empirical distribution of the total s of AWP types.

We select a group of AWP types and test these hypotheses against their distribution using the chi-square test.

In the textbook topics of addition, subtraction, multiplication, and division go one after the other, and therefore it is useful to look at the set of textbooks not only as a single object but also to examine its parts separately. This was not done by Vicente et al. [16], and therefore in this work, we use a different representation: we observe the textbook as a vector consisting of AWPs and provide the type of the AWPs.

Considering the curriculum in Lithuania, we observe that in the textbooks of the first grade, there should be only additive types of word problems without multiplicative ones. From the second to the fourth grade, the problems should be both additive and multiplicative types. We divided primary grades textbooks into five separate parts:

- Before introducing addition.
- After introducing addition and before introducing subtraction.
- After introducing subtraction and before introducing multiplication.
- After introducing multiplication and before introducing division.
- After introducing division.

We analyze these parts. This helps to understand whether the whole textbook is not equally distributed by AWP types or only some parts of it. It helps to understand the textbook authors' position. Are they choosing some AWP types by chance, or are their decisions related to the mathematical operations students learned?

We can identify specific time moments significant for the frequency of problem types:

- z_1, the time moment when the addition topic is introduced;
- z_2, the time moment when the subtraction topic is introduced;
- z_3, the time moment when the multiplication topic is introduced;
- z_4, the time moment when the division topic is introduced.

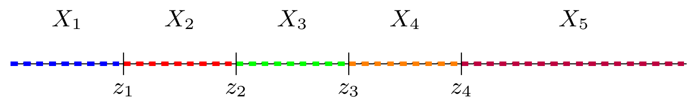

Figure 2. Vector X with time moments z_1, z_2, z_3, z_4 and vectors X_1, X_2, X_3, X_4, X_5.

Then we split the vector $X = (x_1, x_2, x_3, \dots, x_n)$ into five vectors:

- X_1, where we do not expect any specific frequency. Also, this vector should not be long;
- X_2 with the expectation having seven different types of additive problems (where only addition is used);
- X_3, in which we expect a balanced frequency of all 20 additive types (7 addition and 13 subtraction);
- X_4, in which we expect a balanced frequency of all additive types (20 of them) and multiplicative types (5 of them) where multiplication is used;
- X_5, in which we expect a frequency of all 34 AWP types.

The results describe the hypothetical balanced frequency of the vector X. Also, the hypothetical AWP model has a balanced frequency of each X_1, X_2, X_3, X_4, and X_5 separately.

2.2.5 Comparative analysis of the textbooks

The unpublished data of Vicente et al.'s research [16] shared by the authors and the data from our research allow us to compare the textbooks of the three countries in the proportions of certain types of activities in the textbooks. We use a two-sample Z test of proportions to compare Lithuania with Singapore and Lithuania with Spain by proportions.

There are two assumptions for a correct application of the test, independent observations and sufficient sample sizes. According to Agresti and Franklin [1], for both samples, we must have $p{\cdot}n > 10$ and $(1-p)\cdot n > 10$, where n is the size of the sample, and p is the proportion of categorical data in the sample.

We test the hypothesis that the proportions of certain types of activities in the textbook s and the proportion of Singapore coincide:

(H0) $p(s) = p_0$.
(H1) proportions differ.

Here p_0 is an exemplary proportion we would like Lithuania to achieve, and we choose the Singapore proportion as p_0.

3 Results

3.1 Frequency of AWP

The Lithuanian textbook comprised a total of 2626 mathematical activities, of which 879 activities (33.47%) were classified as AWP activities. One AWP activity can consist of several AWPs.

A total of 1966 AWPs were analyzed in the Lithuanian textbook; 53% of them were additive AWPs.

3.1.1 Frequency of additive AWPs

It is interesting to look deeper into the frequency of additive AWP types. We can see from the bar chart (Fig. 3) that in the Lithuania textbook, as in the other two countries, there is no balanced frequency of additive AWP types. The horizontal lines show the average numbers of AWP types of the countries.

In Lithuania, three types of additive AWPs, Com+, Com-, and Ch-F, collectively constitute 68.40% of all additive AWPs (Fig. 3). According to [9, 16], these types can be classified as either low-difficulty (Com+, Ch-F) or medium-difficulty (Com-) problems.

When examining the types of problems that are rarely found or completely absent in the textbook, we can see that Ch+Init and Change-Init are quite rare problems, and there are no Com+Init, Com-Init, and Equalize problems, or there is only one problem of that type in all textbooks of four grades (Fig. 3).

3.1.2 Frequency of multiplicative AWPs

Multiplicative AWPs are another category we can look at in textbooks. There are 14 types by the action of multiplicative AWPs. The same as in the additive types, we can see that there is no balanced frequency (Fig. 4) in all three countries. The Lithuanian textbook also did not provide students with a balanced frequency of multiplicative AWP types.

The same as additive AWPs, multiplicative AWPs also have the most common types of AWPs: there were 81.69% simple rate problems: RatePrM, RatePrS, and RateMul-Ing (Fig. 4). Some types are not used in the material of all four grades.

3.1.3 Frequency of appearance of AWP types in the textbook

If we suggest a balanced frequency of AWP types throughout the textbook, the frequency of all types should be $N(s)/34$, where $N(s)$ is the number of all AWPs in the textbook. On the other hand, the four main topics associated with AWP types are addition, subtraction, multiplication, and division. The Lithuanian textbook vector of AWP types, denoted as X, was divided into five different vectors X_1, X_2, X_3, X_4, and X_5, corresponding to the sections of the textbook before these topics are introduced and the sections where these topics are presented.

The obtained results are as follows (refer to Figs. 3 and 4):

- The vector X_1 contains no AWP.
- In the vector X_2, only 26 AWPs were found, all of the “Comb+” type, and no other types of problems were present.
- For the vector X_3, 20 different additive types are expected. Notably, the most common types observed in the textbook set are also prevalent in this vector. Combining the most common types, “Comb+” and “Ch-F,” accounts for 69.49% of all AWPs in this vector. Additionally, one multiplicative AWP was found in this vector.
- For the vector X_4, 20 additive types and 5 multiplicative types are expected (in the latter, only multiplication, not division, is used). However, this part of the Lithuanian textbook is quite short, containing only 34 AWPs: 2 additive and 34 multiplicative (30 “RatePrS” and 4 “Com+Com”).
- The vector X_5 is the largest part of the textbook, where all topics have been introduced. This vector is expected to have 20 additive types and 14 multiplicative types. Similar templates to those observed throughout the entire textbook are visible in this vector.

Figure 4 shows the hypothetical frequency of AWP types for the Lithuanian textbook we examined. Each column represents the number of tasks corresponding to one type of AWPs in the textbook. The first 7 columns provide information about AWPs for addition, the next 13 columns give information on AWPs for subtraction, the next 5 columns give information about AWPs for multiplication, and the next 9 columns give information about AWPs for division. The colors in the chart show the level of mathematical knowledge of a student used to solve AWPs. For example, the first column displays all tasks of the first AWP type. Orange indicates the number of AWPs where a student knows only addition, gray where a student knows addition and subtraction, yellow where a student knows multiplication, and blue indicates the number of tasks where the student can already perform all four mathematical operations. As long as the student knows only addition (orange color), we expect him to solve all 7 types of addition problems, so the number of problems marked in orange is divided into 7 equal parts. When a student already knows both addition and subtraction, we expect him to solve 20 different types of AWPs (addition and subtraction), so we divide all problems marked in gray into 20 equal parts. We do the same for multiplication and division AWP types. In Lithuania, students are taught division immediately after multiplication, so only a few tasks are marked in yellow in the diagram. The horizontal line represents the average of all 34 types of AWPs.

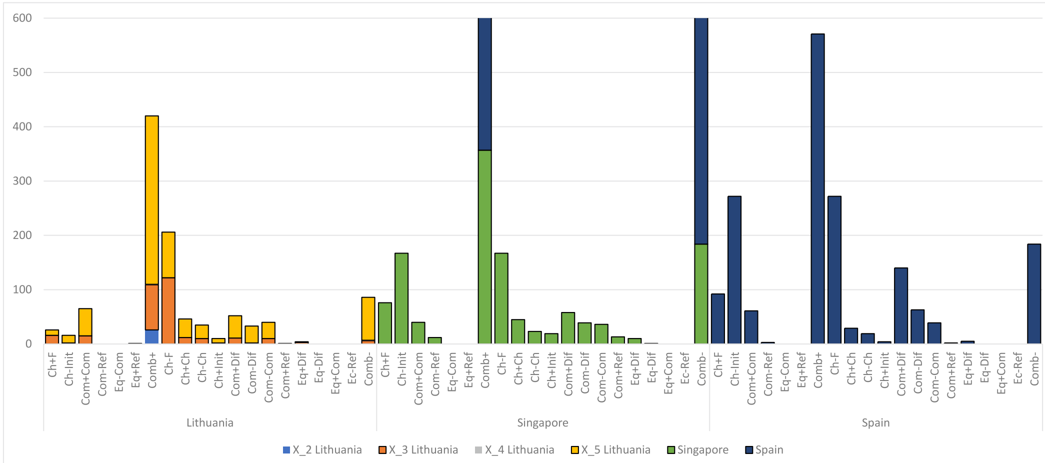

Figure 3. Distribution of additive types in Lithuanian, Singapore, and Spain textbooks.

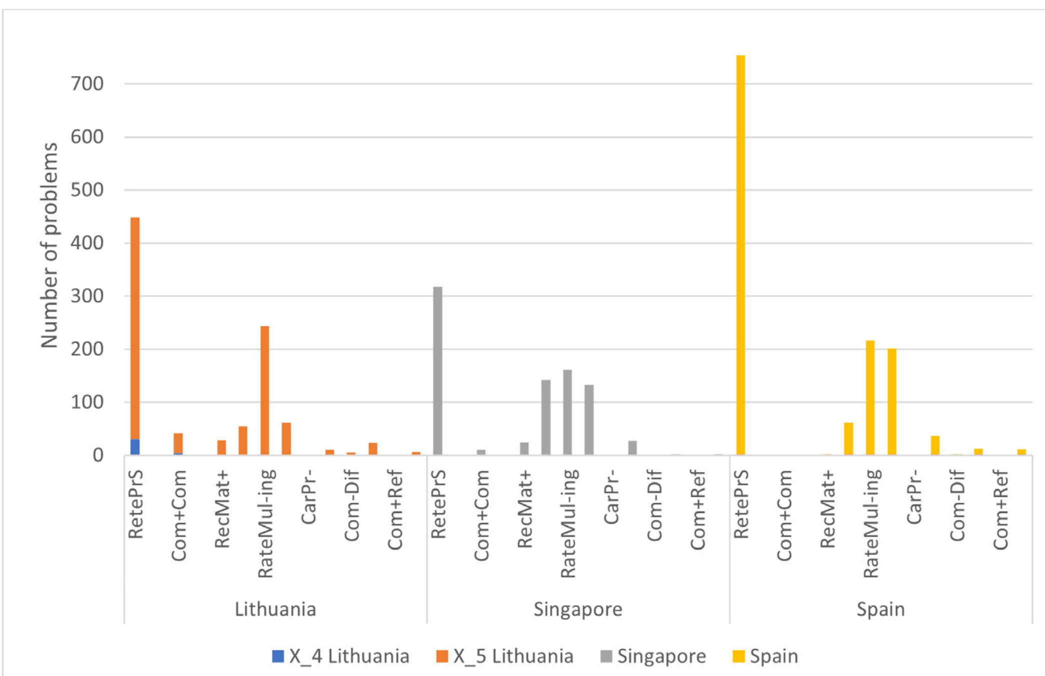

Figure 4. Distribution of multiplicative types in Lithuanian, Singapore, and Spain textbooks.

Let us describe the amount of problems for any textbook set. We denote the numbers of vector components as follows: n_{X_1} for the vector X_1 should be 0 (it should contain no AWPs), n_{X_2} for the vector X_2, n_{X_3} for the vector X_3, n_{X_4} for the vector X_4, and n_{X_5} for the vector X_5. In this context, if we expect the same frequency for each AWP type in vectors, then we can count the number of each type of AWP. For AWPs that can be solved with addition (indicated from 1 to 7), the number of problems for each type should be

$$\frac{n_{X_2}}{7} + \frac{n_{X_3}}{20} + \frac{n_{X_4}}{25} + \frac{n_{X_5}}{34}.$$

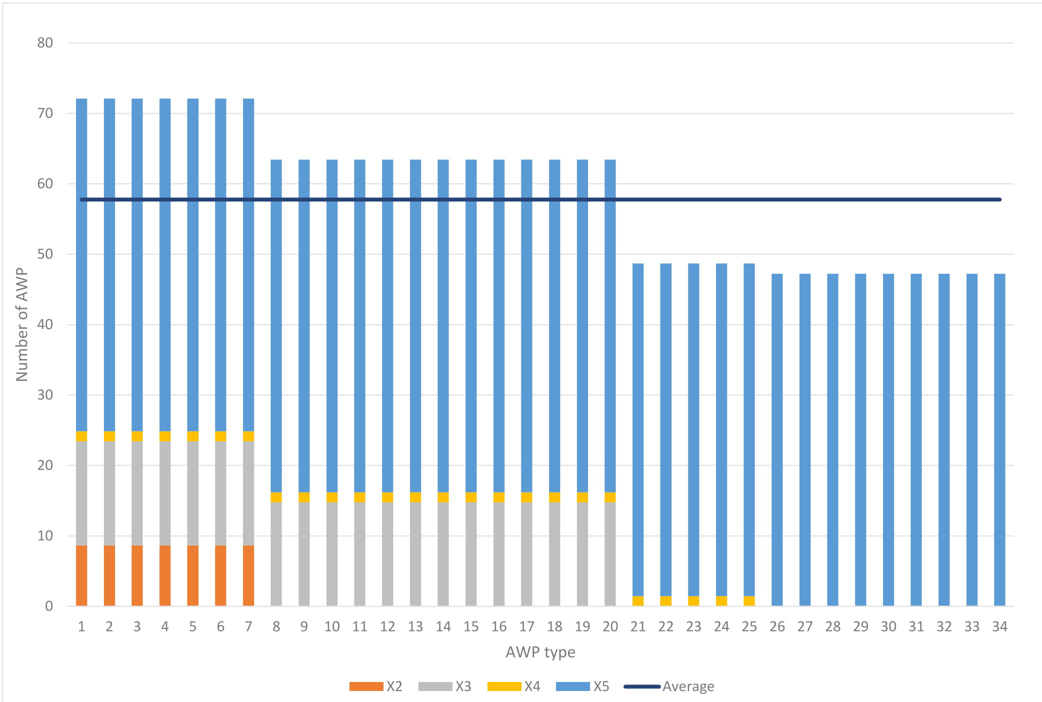

Figure 5. Hypothetical distribution of types in additive AWPs.

For AWP types that can be solved with subtraction (indicated from 8 to 20), the number of problems for each type should be

$$\frac{n_{X_3}}{20} + \frac{n_{X_4}}{25} + \frac{n_{X_5}}{34}.$$

Similarly, for AWP types that require multiplication(indicated from 21 to 25), the number of problems for each type should be

$$\frac{n_{X_4}}{25} + \frac{n_{X_5}}{34}.$$

Finally, for AWP types that require multiplication (indicated from 26 to 34), the number of problems for each type should be

$$\frac{n_{X_5}}{34}.$$

Figure 5 represents the suggested distribution of all types. In Figs. 3 and 4, we can see the real distribution in the Lithuanian textbook.

3.1.4 AWP distribution by the mathematical operation

We divide all AWP types into four groups based on mathematical operations of addition, subtraction, multiplication, and division. A notable frequency pattern emerges in all three countries, as depicted in Fig. 6; the sizes of the groups are quite similar.

These findings suggest that the textbook authors emphasize operations like addition, subtraction, multiplication, and division when presenting AWP activities. However, there seems to be a lack of attention given

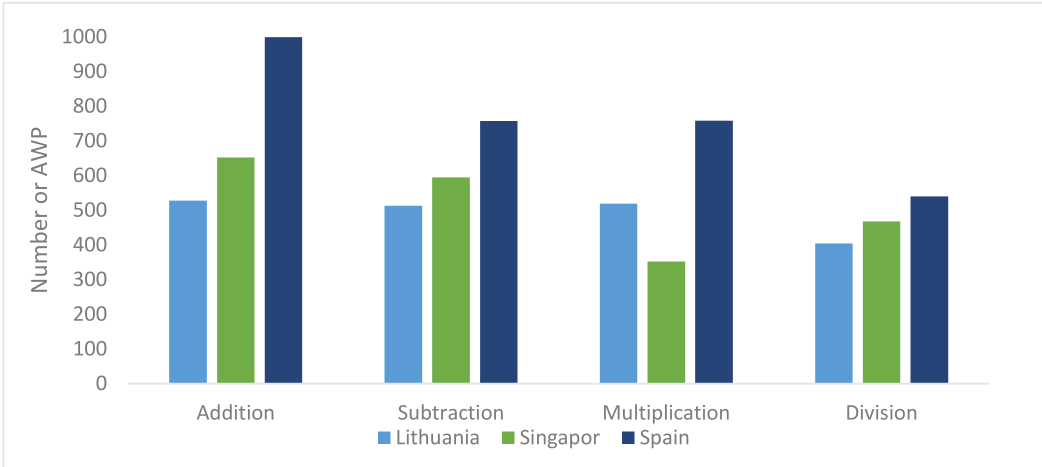

Figure 6. Numbers of AWPs by mathematical operation and introduced topics in Lithuania textbooks.

to the diverse AWP types that can be addressed using these operations. It is important to note that several different types of AWP problems correspond to each mathematical operation, and all of them should be used in textbooks.

3.2 Matching the random sample to the entire textbook

A random sample of 50 pages from the textbook was generated. 110 mathematical activities were found, of which 32 (29.09%) were AWP activities. 66 AWPs were analyzed, of which 39 (59.09%) were additive.

We compare the sample distribution of AWP types and the empirical distribution of s of AWP types.

(H0) The sample distribution of s matches the empirical distribution of the total s of AWP types.
(H1) The sample distribution of s does not match the empirical distribution of the total s of AWP types.

First, we categorize the activities into OMA and AWP and test the hypothesis that the sample distribution and the s distribution match. After applying the chi-square test, we get the P value $p = 0.633559$, $p > 0.05$; therefore, based on these data, we have no reason to reject the hypothesis H_0, which states that the distributions match.

Repeating the same procedure, we test the sample corresponding to s by the distribution of additive and multiplicative AWPs and the distribution of various mathematical operations AWPs. In both cases, we get that the distribution of samples according to AWP types corresponds to the distribution of s.

We can see that the random sample corresponds to the textbook in various aspects: it has the same ratio of AWPs and OMAs, the same ratio of additive and multiplicative AWPs, and the same ratio of AWP groups according to mathematical operations.

Therefore, when conducting textbook research, creating a random textbook sample saves time and helps to make wider research from more countries and textbooks within those countries, ultimately helping to identify high-quality benchmarks in textbooks worldwide.

3.3 Comparison of Lithuanian and Singaporean and Lithuanian and Spanish textbooks

From our study we have the numbers of different types of AWPs, and Vicente et al. [16] provided the numbers of different types of AWPs from Singapore and Spain textbooks without workbooks (see Tables 3–5) of their research.

We do not have a definitive textbook benchmark, and no studies unequivocally recommend specific proportions of AWP types. However, we can identify characteristics that could lead to establishing such a benchmark when comparing textbooks from different countries. Assuming that the Singaporean textbook is an exemplary model, we will compare the proportions of AWP types in Lithuanian and Singaporean textbooks. Similarly, we will perform the same comparison if the aspiration is to achieve the proportions found in Spanish textbooks.

We test the hypothesis that the proportions of AWP types of two countries

1. Lithuania and Singapore and
2. Lithuania and Spain

are equal:

$$(\text{H0})\ \ p_1 = p_2 \quad \text{and} \quad (\text{H1})\ \ p_1 \neq p_2.$$

3.3.1 AWP proportion

We test the hypothesis that the proportions of AWP in the textbooks of both countries are equal.

We performed a Z-test comparing the proportion of AWP problems in the Lithuanian and Singaporean textbooks. We obtained the Z value $Z = 6.9$. If the Z-test assumptions are met, then Z approximately follows a standard normal distribution. From this we get $P(2-\text{tailed}) = 0.01$, and comparing Lithuanian and Spanish textbooks, $Z = 10.59$, and $P(2-\text{tailed})$ is close to 0.

In both cases, we accept hypothesis (H1) that the averages differ. To determine the reasons for the differences, it is necessary to examine in detail the mathematics curricula of the countries, as well as the culturally established attention paid to AWPs. More detailed studies with a larger sample of textbooks and countries are needed to assess whether a higher proportion of AWPs in the textbook benefits the student.

3.3.2 Additive AWP proportion

When comparing the countries according to the proportion of additive problems in textbooks, we see that we have no reason to reject the hypothesis that the parts of additive problems in Lithuanian and Spanish textbooks are the same: $Z = 0.014$, $P(2-\text{tailed}) = 0.9886$, so we can say that Lithuanian and Spanish textbooks are similar in this aspect, but the Lithuanian textbook differs from the Singaporean textbook ($Z = -2.63$, $P(2-\text{tailed}) = 0.009$), which contains a larger part of additive problems. In this aspect, Lithuanian textbooks are more similar to Spanish than to Singapore.

3.3.3 Proportions by mathematical operation

We also compared Lithuania with Singapore and Spain according to the proportion of AWPs that use a certain mathematical operation (addition, subtraction, multiplication, division). We have no reason to reject the

Table 3. Numbers of AWPs and OMAs in Lithuania, Singapore and Spain textbooks

	Lithuania	Singapore	Spain
OMAs	1747	3167	3681
AWP activities	879	1097	1047
Total	2626	4264	4728

Table 4. Numbers of additive and multiplicative AWPs in Lithuania, Singapore, and Spain textbooks

	Lithuania	Singapore	Spain
Additive	1042	1096	1520
Multiplicative	924	820	1349
Total	1966	1916	2869

Table 5. Numbers of AWPs of different mathematical operations in Lithuania, Singapore, and Spain textbooks

	Lithuania	Singapore	Spain
Addition	528	1057	1046
Subtraction	513	1211	1045
Multiplication	519	736	1184
Division	404	828	648
Total	1964	3832	3923

Table 6. Total number of organizational, informational, and figurative illustrations in Lithuania, Singapore, and Spain

	Lithuania	Singapore	Spain
Organizational	108 (10.76%)	190 (36.4%)	21 (1.8%)
Informational	392 (39.04%)	244 (46.7%)	348 (52.1%)
Figurative	504 (50.20%)	88 (16.9%)	308 (46.1%)

hypothesis that the parts of the AWP assigned to the addition operation are the same, neither in the case of Lithuania and Singapore ($Z = -0.57$, $P(2 - \text{tailed}) = 0.57$) nor in the case of Lithuania and Spain ($Z = 0.18$, $P(2 - \text{tailed}) = 0.86$). In the case of Lithuania and Singapore, the parts of the subtraction AWPs differ ($Z = -4.32$). There is no reason to reject the hypothesis that Lithuania and Spain parts of the subtraction AWP are the same ($Z = -0, 424478445$, $P(2 - \text{tailed}) = 0.67$).

In the case of multiplication, we reject the hypotheses with both countries: for Lithuania and Singapore, $Z = 6.32$, and for Lithuania and Spain, $Z = -3$.

In the case of division, we do not reject the hypothesis that Lithuania and Singapore parts of the division AWP differ ($Z = 0.91$, $P(2 - \text{tailed}) = 0.36$), but reject this hypothesis for Lithuania and Spain ($Z = 3.82$).

3.3.4 Illustrations

We analyzed the illustrated AWPs in a Lithuanian textbook. We identified the quantities of Figurative, Informational, and Organizational illustrations within Lithuanian textbooks and compared them with the results of Singapore and Spain textbooks [16].

The main factor discussed by the authors is the number of Organizational illustrations, and we can see that Lithuania is between Singapore and Spain due to the percentage of Organizational illustrations. Lithuania is also between Singapore and Spain by the results of 4th-grade students achieving advanced level and cognitive abilities at Reasoning tasks according to Trends in Mathematics and Science Study (TIMSS) report [8] as mentioned above. The obtained result does not contradict Vicente et al.'s idea that a larger percentage of organizational illustrations in a textbook positively affects the performance of reasoning tasks [16].

We compared AWPs with organizational illustrations proportion in Lithuanian with proportion in Singapore. We test the hypothesis

$$\text{(H0)}\ \ p_{\mathrm{AWP}(i)}(s) = p_0 \quad \text{versus} \quad \text{(H1)}\ \ p_{\mathrm{AWP}(i)}(s) \neq p_0\ (p < p_0 \text{ or } p > p_0), \tag{3.1}$$

where p_0 is the Singapore proportion of AWPs with organizational illustrations, $p_{\mathrm{AWP}(i)}(s)$ is the proportion of AWPs with organizational illustrations in Lithuanian textbook s. This means that we accept that the Singapore textbook is the desired level for organizational illustrations and aim for the Lithuanian textbook to have the same proportion of organizational illustrations. To test the hypothesis, we conducted a Z-test. The z value is -5.1755. We reject hypotheses about proportion equality.

4 Discussion

The Lithuanian textbook exhibited a lower total count of mathematical activities in comparison to the Singaporean and Spanish textbooks. 2 626 mathematical activities were found, and 879 (33.47%) of them were AWP-solving activities; the Singaporean textbook included 4 264 mathematical activities, with 1 097 activities (25.73%) designated as AWP-solving activities, and the Spanish textbook contained 4 729 activities, among which 1 047 (22.14%) were AWP-solving activities [16]).

Obviously, certain problem types are favored over others. Simpler problem types tend to receive more emphasis compared to difficult ones. All three countries have similar patterns when examining the types of problems that are the most found in the textbooks: in Lithuania, three types of additive problems, Comb+,

Comb-, and Ch-Ch, collectively constitute 68.40% of all additive Arithmetic Word Problems (Fig. 3). Similarly, in Singaporean and Spanish textbooks, these types account for 64.6% and 69.3% of additive AWPs. Also, in all three countries, some types of AWPs are rarely found or completely absent in textbooks, such as Ch+Init and Ch-Init, Com+Ref and Com-Ref, and all types of Equalize problems. It remains unclear whether these problems are challenging, leading to their exclusion from textbooks, or if they prevent students from encountering them, making them appear difficult. It is important to develop textbooks that offer students a well-balanced frequency of AWPs to help them meet various word problems and mathematical operations in various situations.

Looking at multiplicative word problems, we can find the same similarities: in the Lithuania textbook, there were 81.69% simple rate problems: RatepRm, RatePrS, RateMul-Ing (Fig. 4). This number amounted to 75.8% and 86.8% of the multiplicative AWPs in Singaporean and Spanish textbooks, respectively [16]. As in the additive AWPs, there is the same template: easier AWPs are more common in textbooks than more difficult.

We noticed that Cartesian product problems are absent in Lithuanian textbooks because probability theory is not part of the elementary school curriculum. However, the question of whether these types of problems should be introduced in the early grades arises. It is important to mention that when evaluating the effectiveness of textbooks of each country, it is also necessary to consider their curricula.

It is important to acknowledge that both Lithuania and Spain are European countries, potentially sharing cultural similarities, which might influence the results. Future studies could explore comparisons with countries from different regions, such as the Middle East, Africa, or America. We also suggest future studies with samples from the textbooks due to research coverage of more countries and textbooks.

We compare the proportions with the desired p_0. There are multiple ways to determine the desired p_0. One approach is to compare the proportions of textbooks in countries with superior mathematics teaching results to those in other countries. Another method involves interviewing expert teachers to ascertain the desired p_0. Further research and work are necessary to identify the desired properties of textbooks.

We recommend a variety of types of both additive and multiplicative word problems in the textbook. Considering that initially students only know addition and can only solve problems that require the mathematical operation of addition and later they learn subtraction only after multiplication and division, we presented a hypothetical frequency of problem types in the textbook.

Each type of AWPs has different components. Jaffe et al. [5] name nine linguistic factors that affect AWP solving, but we do not discuss them in this paper. Also, numerical factors [2, 15] are important for AWP difficulty, and it is important to look deeper into AWPs and their integration in textbooks. This is one of the steps to have better textbooks. We suggest a hypothetical model of how AWP types can be distributed in a textbook. This model can be improved and improved.

The findings presented in this study support the hypothesis by Vicente et al. [16] regarding the benefits of incorporating organizational illustrations in arithmetic word problems. These results underscore the potential applicability of this hypothesis when designing new educational materials, such as textbooks. Educators may facilitate more effective approaches to solving word problems by integrating organizational illustrations into instructional materials, thereby enhancing student understanding.

This study contributes to the ongoing discourse on effective mathematics education and curriculum development, specifically focusing on the role of word problems and the search for textbook benchmarks.

Acknowledgment. I want to express my deepest gratitude to Prof. Rimas Norvaiša for his invaluable patience and feedback. I am also thankful to Prof. Alfredas Račkauskas, who generously provided his knowledge and expertise.

I thank Dr. Santiago Vicente Martín (Universidad de Salamanca) for providing unpublished data from his research on Spanish and Singaporean textbooks, excluding workbooks. We also appreciate his valuable suggestions and comments, which greatly improved our research.

Additionally, I am grateful for the generous assistance from Prof. Marijus Radavičius and his suggestions for future research directions.

References

1. A. Agresti, C. Franklin, and B. Klingenberg, *Statistics: The Art and Science of Learning from Data,* Books à la Carte ed., Addison-Wesley, Reading, MA, 2008.

2. G. Daroczy, M. Wolska, W.D. Meurers, and H.-C. Nuerk, Word problems: A review of linguistic and numerical factors contributing to their difficulty, *Front. Psychol.*, **6**:348, 2015.

3. S. Elrod, Quantitative reasoning: The next "Across the Curriculum" movement, *Peer Review*, **16**(3):4–8, 2014.

4. J.I. Heller and J.G. Greeno, Semantic processing in arithmetic word problem solving, in *Annual Meeting of the Midwestern Psychological Association, Chicago*, 1978.

5. J.B. Jaffe and D.J. Bolger, Cognitive processes, linguistic factors, and arithmetic word problem success: A review of behavioral studies, *Educ. Psychol. Rev.*, **35**(4):105, 2023.

6. Y.H. Leong, W.K. Ho, and L.P. Cheng, Concrete-pictorial-abstract: Surveying its origins and charting its future, 2015.

7. S.P. Marshall, Schema-based instruction, in N.M. Seel (Ed.), *Encyclopedia of the Sciences of Learning*, Springer, Boston, MA, 2012, pp. 2945–2946.

8. I.V.S. Mullis, M.O. Martin, P. Foy, D.L. Kelly, and B. Fishbein, TIMSS 2019 international results in mathematics and science, 2020, `https://timssandpirls.bc.edu/timss2019/international-results`.

9. T. Nunes, B. Vargas Dorneles, P.-J. Lin, and E. Rathgeb-Schnierer, *Teaching and Learning about Whole Numbers in Primary School*, Springer, Cham, 2016.

10. I. Polo-Blanco, M.J. González López, A. Bruno, and J. González-Sánchez, Teaching students with mild intellectual disability to solve word problems using schema-based instruction, *Learn. Disability Q.*, **47**(1):3–15, 2024.

11. N. Salingay and D. Tan, Concrete-pictorial-abstract approach on students' attitude and performance in mathematics, *Int. J. Sci. Technol. Res.*, **7**(5):90–111, 2018.

12. J.W. Stigler, K.C. Fuson, M. Ham, and M.S. Kim, An analysis of addition and subtraction word problems in American and Soviet elementary mathematics textbooks, *Cognition Instruct.*, **3**(3):153–171, 1986.

13. P.W. Thompson, Quantitative reasoning, complexity, and additive structures, *Educ. Stud. Math.*, **25**(3):165–208, 1993.

14. L. Verschaffel, E. De Corte, and S. Lasure, Realistic considerations in mathematical modeling of school arithmetic word problems, *Learn. Instr.*, **4**(4):273–294, 1994.

15. L. Verschaffel, S. Schukajlow, J. Star, and W. Van Dooren, Word problems in mathematics education: A survey, *ZDM Math. Educ.*, **52**:1–16, 2020.

16. S. Vicente, L. Verschaffel, R. Sánchez, and D. Múñez, Arithmetic word problem solving: Analysis of Singaporean and Spanish textbooks, *Educ. Stud. Math.*, **111**(3):375–397, 2022.

17. Y.P. Xin, Word problem solving tasks in textbooks and their relation to student performance, *J. Educ. Res.*, **100**(6): 347–360, 2007.

18. Y.P. Xin, The effect of a conceptual model-based approach on 'additive' word problem solving of elementary students struggling in mathematics, *ZDM Math. Educ.*, **51**:139–150, 2019.

19. State Education Strategy for the Years 2013–2022, Republic of Lithuania, 2003, `https://www.nsa.smm.lt/wp-content/uploads/2018/04/Valstybine-svietimo-strategija-2013-2020svietstrat.pdf`.