Error analysis of logarithm problems

KLÁRA KELECSÉNYI, ÉVA OSZTÉNYINÉ KRAUCZI

GAMF Faculty of Engineering
and Computer Science, Kecskemét
John von Neumann University
Hungary
kelecsenyi.klara@nje.hu
osztenyine.eva@nje.hu

Abstract

Error analysis can be a tool to predict future errors made by students, to explain how and why students make them, and to help us to resolve misconceptions in our teaching). In the paper we categorize errors committed by first year undergraduate students in logarithm problems in a written class test. We have categorized the errors according to two dimensions, namely the type and the source of errors. By using this two- dimensional grouping, we were able to classify the majority of typical logarithm errors listed in the literature. We were able to clearly categorize the errors made during the post-test carried out later into classes based on the originally given categorization. The distribution of post-test errors was not significantly different from the original distribution.

Keywords

Error analysis, logarithm, undergraduate education

Résumé

L'analyse des erreurs peut servir à anticiper les erreurs futures commises par les étudiants, à expliquer comment et pourquoi ils les commettent, et à nous aider à corriger les idées fausses dans notre enseignement. Dans cet article, nous classons les erreurs commises par des étudiants de première année de licence dans des exercices sur les logarithmes lors d'un contrôle écrit en classe. Nous avons classé les erreurs selon deux critères, à savoir leur type et leur origine. Grâce à ce regroupement en deux dimensions, nous avons pu classer la plupart des erreurs typiques liées aux logarithmes répertoriés dans la littérature. Nous avons ainsi pu classer clairement les erreurs commises lors du test de vérification réalisé ultérieurement en différentes catégories, sur la base de la classification

initialement établie. La répartition des erreurs observées après le test ne différait pas de manière significative de la répartition initiale.

Mots-clés

Analyse des erreurs, logarithme, enseignement de premier cycle

Cite this article

Kelecsényi, K., & OszténYiné Krauczi, É. (2026). Error analysis of logarithm problems. *Review of Science, Mathematics and ICT Education*, *20*(1), 59-80. https://doi.org/10.26220/rev.5517

Introduction

Students arrive with various math skills at university. Remedial mathematics courses aim to bridge the gap between (often superficial) high school knowledge and prerequisites of university mathematics courses. These courses are offered in various formats, such as providing a study matrial with the prerequisites to help the transition between secondary and tertiary education (Clark & Lovric, 2008) or using game based learning (Kelecsényi et al., 2023), as educational games were proven effective to provide motivation and build a positive attitude towards the subject (Giannakoulas & Xinogalos, 2024). From a teaching point of view, it is good to get an idea of the mathematical skills with which first-year undergraduate students arrive at university. What are the areas of these skills that are not "strong" enough to build on the university mathematics material. The ability to use logarithms is a typical example. The use of logarithm is inevitable for students studying science or economics for calculating pH-values, producing and interpreting log-scaled graphs, determining interest rates, etc. According to our practice understanding logarithm causes several difficulties as it was noted by various researchers as well (Chua & Wood, 2005; Kelecsényi et al., 2025; Kenney & Kastberg, 2013; Mulqueeny, 2012; Weber, 2016). We think that examining the errors in students' solutions can be a useful method for understanding the source of these difficulties.

Therefore, error analysis can be a tool to predict future errors made by students, to explain how and why students make them, and to help us to resolve misconceptions in our teaching (Olivier, 1989). We hope that error analysis can help us to identify the difficulties students and instructors face during remedial university courses known to be ineffective in many cases (Clark & Lovric, 2008), while we are aware that university success depends on various factors (Asli & Zsoldos-Marchis, 2023).

We often see common patterns in the mistakes students make when solving logarithm problems. However, in the literature we could not find appropriate sources to categorize and explain the possible sources of these common errors, as the answers were either too general or very specific to the actual problems presented. Thus, in the paper we suggest a systematic method for categorizing the errors committed by students in logarithm problems.

Literature review

One approach for the classification of errors came from the direction of mathematics education, namely the error analysis of mathematical problems, identifying the type of errors based on the paper of Movshovitz-Hadar et al. (1987). They identified general types of errors applicable in different areas of secondary mathematics, such as geometry, algebra, proofs, etc. They categorized the errors as the following: (i) misused data, (ii) misinterpreted language, (iii) logically invalid inference, (iv) distorted theorem or definition, (v) unverified solution and, lastly (vi) technical error. Ganesan and Dindyal identified (i) misused data, (iv) distorted theorem or definition, (v) unverified solution and (vi) technical error as common errors in students' work related to logarithm problems (Ganesan & Dindyal, 2014). Later Rafi and Retnawati (2018) confirmed their results.

The other approach for classifying the errors students made came from mathematics itself, thus categorizing errors based on their source. Olivier presented some misconceptions during problem solving and their sources, i.e. general *causes* (Olivier, 1989). Among the works on error analysis Chua and Wood dealt specifically with logarithms (Chua & Wood, 2005). They constructed a theoretical framework, categorized common *causes* of errors, and showed examples for errors, and error types associated with the *causes* identified by the authors. They identified levels of understanding and categorized the tasks accordingly, namely: Knowledge or Computation, Understanding, and Application. They identified three main sources of errors, such as those coming from (i) over-generalization of concepts and rules, (ii) the deficient mastery of concepts, rules and prerequisite skills and a third (iii) miscellaneous category. "Errors due to deficient mastery of concepts, rules and prerequisite skills; due to over-generalization; and, miscellaneous (indecipherable errors, slips due to incorrect arithmetic computations, incorrect algebraic manipulations, errors due to guessing, and bad handwriting)" (Chua & Wood, 2005, p. 61).

Other articles identified students' levels of understanding of logarithms and listed errors. Hoon et al. gave a detailed list of 15 errors to each category; they listed errors regarding misconceptions such as the concept of antilogarithm and logarithm rules and errors regarding operations such as incorrect changes of base or cancelling out log notations (Hoon et al., 2010). De Gracia listed 15 alternative conceptions among high aptitude- mathematically challenged high school students causing systematic errors in the solution of exponential and logarithm problems, nine of these refer to logarithm (de Gracia, 2016). These include wrong concept on antilogarithm; incorrect change of logarithm to index form and vice versa; linearized rules on logarithm; misconceptions about the logarithm of a product, quotient and power; misuse of inverse and one-to-one property of logarithm; and simply cancelling out logarithmic notation or treating logarithm as a variable.

Other researchers have attempted to identify problem-solving strategies and have also identified major categories of errors. Aziz et al. (2017) identified the common mistakes and strategies used by college students by solving logarithmic problems. They identified different strategies used by students while solving logarithm problems, such as (j) processing base, (ii) holding the rule, (iii) separating, (iv) jumping, and (v) conditioning. The main source of errors was caused by students' inflexibility of the conception of logarithm, mistakes of arithmetical calculation, and misuse of algebra concept. They identified 4 types of errors such as Definition of Logarithm, Symbolic Manipulation, Logarithm Equations, Graph of Logarithm and Logarithmic Function.

Theoretical framework

Olivier (1989) distinguished between slips, errors, and misconceptions. He identified slips as a. result of carelessness due to processing. They are not systematic, can easily be detected and corrected. In contrast, errors are due to planning and are systematic, i.e. appearing regularly in the same situations. Errors are signs of misconceptions, as they are rooted in systematic conceptual errors in the underlying cognitive structure.

First, we have sorted out the mistakes due to carelessness, which we call slips according to the terminology of Olivier (1989). We did not look for the reasons for these, since they cannot be given on the basis of existing information. Such errors are due to carelessness, e.g., writing 1/2 instead of the multiplication factor of 1/3 specified in the task.

In the following two subsections, we present the dimensions of our two-dimensional classification: the types of errors and the sources of errors by filling in the cells of Table 1.

Table 1

Two-dimensional classification

	Possible sources of error			
Type of error	*Distortion of the concept*	*Problems with accommodation*	*Previous erroneous schemes*	*Unidentified causes*
Distorted definition or theorem				
Misused data				
Misinterpreted language				
Technical errors				

The types of errors

The first dimension of our error-analysis is based on the categorization of Movshovitz-Hadar et al. (1987). Due to the nature of our task, only 4 categories appeared out of the 6 classes the authors distinguished. These are errors due to *distorted theorem or definition, misused data, misinterpreted language, and technical errors*.

- *Distorted theorem or definition* denotes errors due to distortion of a specific principle or definition, such as applying a theorem outside its conditions. Similarly applying the distributive property in a case of a non-distributive function or operation. Citing a theorem or a definition incorrectly also belongs to this category.
- *Misused data* refers to situations when there is a discrepancy between the data given and data treated by students. Students often use data and properties of related concepts different from those given in the problem. Such errors are using data that do not follow from the information given; utilizing inconsistent information; adding irrelevant data; neglecting, or misusing information given in the problem, thus endowing a thing with a quality that is not its own. The authors included errors caused by incorrectly copying information, however, we coded these as slips.
- Errors caused by *misinterpreted language* are originated from the incorrect translation between languages, one of those are usually symbolic, such as errors regarding expressions translated incorrectly between a natural and mathematical language or using incorrect mathematical formalism.
- *Technical errors* refer to computational errors regarding concepts of elementary mathematics as well as the manipulations with elementary algebraic symbols, such as leaving parentheses out.

The possible sources of errors

In addition to the types of errors, we used a second dimension of analysis, namely we have also tried to reveal the possible causes of errors. Olivier states that errors often happen, because students try to apply their prior knowledge in a new situation (Olivier, 1989). Although the application of prior schemas (the structure of interrelated concepts) can support the acquisition of new knowledge, there are cases when these previous schemas interfere with the new idea. According to constructivist learning theories, learning involves two types of interrelated interaction between the learners' prior schemas and the new ideas, such as assimilation and accommodation. Olivier argues that misconceptions often happen because of overgeneralizing properties of previously known concepts. Instead of the necessary accommodation of prior schemas, the new concept is simply assimilated into previously existing schemas. According to this idea, the causes for errors by learning a new concept can happen because of the following reasons: (1) difficulties with the new concept itself, (2) difficulties with incorporating the new concept into the learner's prior schemas, (3) previously incorrect schemas or (4) some unknown reasons. Thus, we identified four categories as the source of errors. These are distortions of the concept and its properties, distortions of the inclusion of the concept in previous schemas, previous erroneous schemas, and unidentified causes.

- The first category of possible causes is the *distortion of the concept* and its properties. This type of error comes from the concept of logarithm itself. In these cases, the error was caused because of the interpretation and use of the new concept, furthermore, we could not find any connection between the error and previous schemas. Thus, we assumed here that the error originated from the incorrect interpretation of the concept and properties of the logarithm itself.
- The second category refers to *problems with accommodation* of the new concept into previous schemas. We have categorized errors into this group, when an over-generalization of previously correct schemas could be noticed, as mentioned by Chua and Wood (2005). This over-generalization refers to both the concept and the properties of logarithm. Olivier (1989, p. 5) explained this over-generalization as the existing schemas are not accommodated according to the new concept and properties of logarithm, instead students simply assimilate the new information without reorganizing the system of their previous schemas. Consequently, the application of previous schemas regarding algebraic properties leads to incorrect results in the case of logarithm.
- The third category contains misconceptions originated from *previous erroneous schemas*. In our study, these faulty schemas refer to basic operations, linguistic formalism, and the incorrect use of the exclusion principle. Obviously, this does not mean that students do not hold other faulty schemas related to the concept, only that the erroneous solutions of the given task revealed only these categories.
- The fourth category included *unidentified causes* where the source of error was not recognizable for us, was unambiguous or could have been caused by various reasons.

According to the above categories we have formalized the following research question: *How can we categorize the errors in logarithm problems, such that not only the types but also the sources of errors are identified?*

As a result of our investigation, we try to categorize the errors made by first-year undergraduate students for logarithmic assignments based on a two-dimensional categorization by classifying the type of error and the possible sources of these errors.

Method

We had students test their knowledge of the concept of a logarithm and its identities. This students are first-year undergraduate students majoring engineering and computer science at John von Neumann University, Kecskemét, Hungary. The class consists of 173 students in Autumn semester 2019/20 and 156 students in Autumn semester 2022/23. Students were enrolled in a remedial mathematics class accompanying the compulsory Calculus I module. The remedial mathematics class was aimed to review high school mathematics necessary to accomplish the first-year Mathematics and Physics requirements, thus the topics address mainly Algebra and Trigonometry. As an assessment, students wrote individually bi-weekly 10-minute-quizzes on the actual material covered. Anecdotal evidence shows that one of the most problematic topics for students is logarithm.

In Autumn semester 2019/20 the task was designed according to the following principles:

- The task could be solved by ordering the arguments.
- To order the logarithmic expressions in the task, the use of logarithm identities couldn't be avoided.
- The exact value of expressions cannot necessarily be calculated without a calculator.
- The task should be short.

Due to technical reasons two different problems were given to students enrolled in different groups. The students wrote only one of the tasks. The two 10-minutes tasks were the following (The corresponding task is shown in Figure 1 and Figure 2):

Task 1

Without using a calculator, evaluate the following expressions and join them with the corresponding value on the number line.

FIGURE 1

$$\log_5 30 - \log_5 6; \quad \frac{1}{3}\log_5 \frac{27}{125} - \log_5 \frac{3}{5}; \quad 2\log_5 \frac{1}{5}; \quad \log_5 5 + \frac{1}{2}\log_5 25$$

The number line problem given to students

Task 2

Arrange the following expressions in increasing order without using a calculator.

FIGURE 2

$$\log_3 7+\log_3 2; \quad \log_3 17-\log_3 2; \quad 2\log_3 4; \quad \frac{1}{2}\log_3 25-\frac{1}{3}\log_3 8$$

The ordering problem given to students

Out of the class of 173 students, 156 students solved the task independently (four classes were given the first, and two the second tasks). We used all solutions for the categorization. Categorization was conducted independently by the authors.

During the analysis, we first identified and removed errors labelled as slips, as they did not influence our categorization. The remaining errors were coded systematically according to the framework suggested by Movshovitz-Hadar et al. (1987). In our investigation four of the Movshovitz-Hadar et al. (1987) categories could be identified. Four each of these categories we aimed to identify the underlying sources of errors. To support this process we also examined students' responses to other task if it was necessary to better understand the underlying reasons of the mistakes. The coding of the types and sources of errors were done independently by two researchers. In case of disagreement, the codes were constantly refined. In the few cases when redefining the categories did not result in consensus among the researchers, the source was coded as unidentifiable. The categorization was verified on a similar set of problems to check. The results section gives examples of how the categorization was done.

To check our two-dimensional classification, we carried out a second survey. In Autumn semester 2022/23 the task was designed according to the following principles:

- We kept the ones that were interesting from the previous one.
- None of the sub-tasks of the problems to be calculated should be noteworthy. But work with values close to notable values.
- Let there be less than 5 and greater than 5 arguments between problems.
- In the number equation, only the values 0 and 1 should be ticked. Checking the values triggers the use of the estimation method.
- Do not ask for an estimate directly.
- Estimation should be done at the end of the task so that you have experience using logarithm by then.

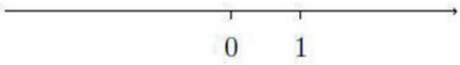
The tasks were the following (Figure 3).

Task 3

Without using a calculator, calculate and then mark the positions of the following terms in the equation. Justify your solution.

FIGURE 3

1. $\log_5 30 - \log_5 6$
2. $\frac{1}{3}\log_5 \frac{27}{125} - \log_5 \frac{3}{5}$
3. $2\log_5 \frac{1}{5}$
4. $\log_5 3 + \frac{1}{2}\log_5 49$



The modified problem given to students

Out of the class of 156 students, 139 students solved the task independently.

RESULTS (STUDY 2019/20) AND DISCUSSION

In this chapter, we present the categorization of errors found in the 156 solutions received in Autumn semester from 156 students. A total of 101 students completed Task 1. Overall, 61% of these students gave the correct answer. In Task 1, part 1 was solved correctly by 76 students (75%), part 2 by 39 students (39%), part 3 by 54 students (53%), and part 4 by 77 students (77%). A total of 55 students solved Task 2. Of these students, 55% gave the correct answer. part 1 of Task 1 was solved correctly by 31 students (56%), art 2 by 31 students (56%), part 3 by 35 students (64%), and part 4 by 25 students (45%). The distribution of errors is also presented in Table 2.

TABLE 2

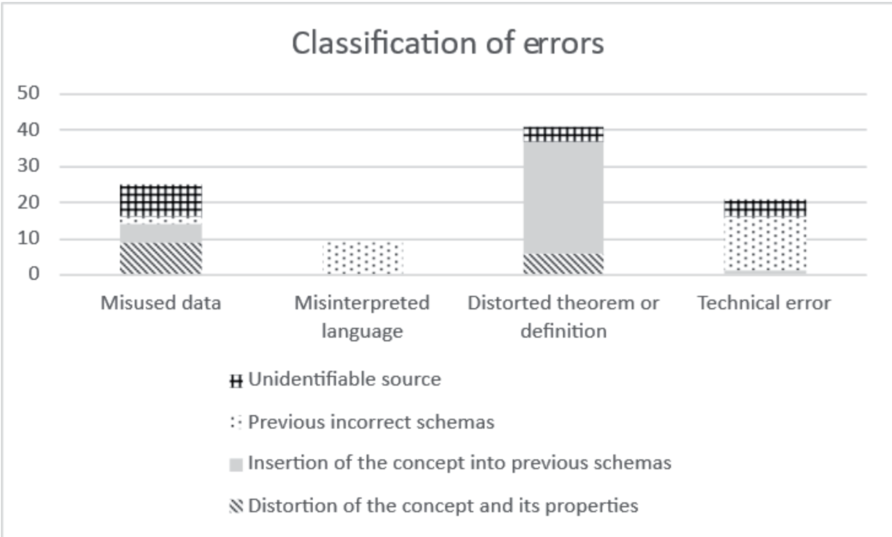
Pearson's Correlations between error types in 2019/20

Variable		DD/DT	MD	ML	TE
1. DD/DT	Pearson's r	-			
	p-value	-			
2. MD	Pearson's r	0.189	-		
	p-value	0.018	-		
3. ML	Pearson's r	0.219	0.050	-	
	p-value	0.006	0.534	-	
4. TE	Pearson's r	0.163	0.440	0.098	-
	p-value	0.042	< .001	0.224	-

Summary of errors

The 156 students made a total of 96 mistakes. We also found 9 slips, which were removed from the error analysis. The four types of errors we used from the six Movshovitz-Hadar et al. (1987) categories are the following: incorrect data use, errors due to interpretation, errors due to incorrect definition or item use and technical errors. The error rates (proportions) for these four error types are (in order) the following 25(26.04%), 9(9.37%), 41(42.71%) and 21(21.88%). The distribution of causes of errors according to error types is shown in Figure 4.

FIGURE 4



Distribution of error types and distribution of sources of errors within each error type in 2019/20

The highest proportion of errors was due to the use of a distorted theorem or definition. We have found 41 such errors. In our analysis, errors related to the concept of logarithm were well separated from errors due to the use of theorems related to logarithm, with 9 and 32 errors, respectively. In the case of errors due to distortion of the definition, three were due to distortion of the concept and its properties, three to distortion of the inclusion of the concept in previous schemas and three to unidentified causes. For errors resulting from the use of a distorted item, the use of identity for sum or difference was incorrect in 26 cases. In three cases the use of the item itself was distorted and in 23 cases the insertion of the item into the previous schemas was distorted. Incorrect use of the identity of the exponent occurred in six cases. Of these, five errors were due to distortions in the insertion of the concept into previous schemas and 1 was due to an unidentified cause.

The second most common type of error is the category of incorrectly used data with 25 errors. Within this category, a new group of errors was found to be the group of errors due to estimation. We found four such errors due to unidentified causes. Therefore, further research is needed on errors due to estimation in or- der to identify their source. Nine errors were coded in the category omitting the logarithm. Of these, four errors belong to the category of distortion of concept and its properties, four errors to the category of distortion of the inclusion of the concept in previous schemes, and one to the category of unidentified cause. The category "Other errors" include 12 items. All four reasons (Distortion of the concept, Accommodation problems, Previous erroneous schemas, Unidentified causes) occurred as sources of error, in 5, 1, 2 and 4 cases respectively. We found interesting, unique errors here, which we plan to describe in detail in a future publication.

The third group, technical errors, contains 21 items. Out of these we have found 13 computational errors, of which 9 were due to previous faulty schemas and 4 were unidentifiable. An additional 6 cases were due to missing brackets. The source of the remaining two errors was a distortion of the inclusion of the concept in previous schemas and an unidentified cause.

We identified nine errors resulting from erroneous interpretation. The distribution of these was as follows: five formal errors in the use of theorems and four incorrect formalisms of sub-calculations. All nine were due to the use of previously incorrect schemes.

Table 2 shows the correlations between the different types of errors. Based on these correlations, there is no connection between the occurrence of different types of errors.

The distribution of the errors by the four causes (distortion of the concept and its properties, distortion of the incorporation of the concept into previous schemas, previous faulty schemas and unidentifiable causes) are the following: 15(15.63%) errors belong to the first, 37(38.54%) to the second, 26(27.08%) to the third and 18(18.75%) errors to the fourth category, respectively. The absence of an interview can explain the high proportion of unidentifiable causes. The distribution of errors by type and causes is shown in Table 3.

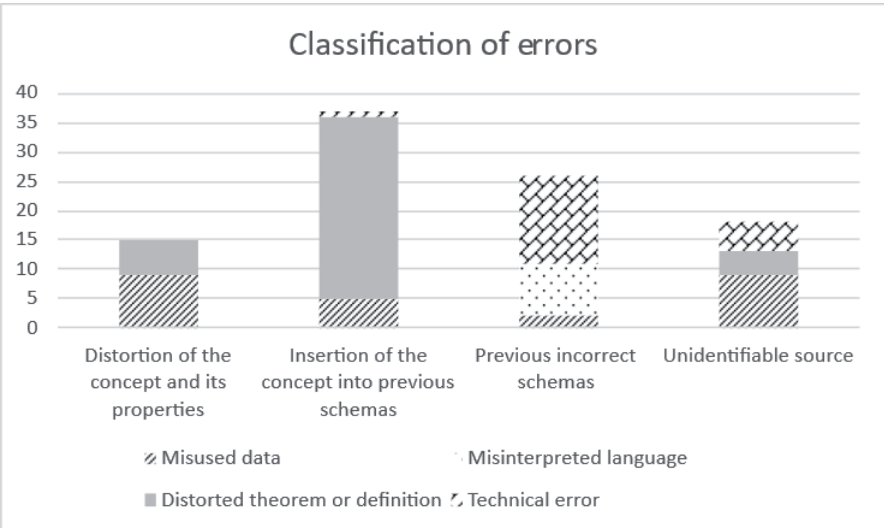
TABLE 3

Distribution of students' errors in 2019/20

Type of error	Possible sources of error				
	Distortion of the concept	*Problems with accommodation*	*Previous erroneous schemes*	*Unidentified causes*	*Total*
Distorted definition or theorem	6	31		4	41
Misused data	9	5	2	9	25
Misinterpreted language			9		9
Technical errors		1	15	5	21
Total	15	37	26	18	96

The distribution of error types within possible sources is shown in Figure 5.

FIGURE 5



Distribution of error types and distribution of sources of errors within each error type in 2019/20

Detailed analysis

In the following we give a detailed description the errors belonging to the categories presented. We describe the errors by type. Admittedly, the sources of these errors often cannot be identified exactly.

Distorted definition or theorem

First, let's investigate errors due to *distorted definition or theorem*. Not surprisingly, this category contains the most errors. Even though the concept of logarithm was known for the students previously, many of them had to relearn the concept again.

Errors caused by the *distortion of the concept* are due to the erroneous scheme of logarithm or the properties of logarithm. Many students tried to solve the problem by using the definition and find the corresponding exponential form, but in various cases the change between the logarithm and the exponential form was incorrect, e.g. for evaluating $\log_5(5^{-2})\log_5(5^{-2})$ a student used $5^{-2} = x^5 5^{-2} = x^5$. Distortion of the concept regarding theorems caused one of the most common types of errors when logarithm rules were used incorrectly. In these cases, the sum or difference of the logarithmic expressions has been rewritten as the product or quotient of the expressions. . This error appeared mostly in the case of sums and differences such as $\log_5 30 - \log_5 6 = \frac{\log_5 30}{\log_5 6}\log_5 30 - \log_5 6 = \frac{\log_5 30}{\log_5 6}$.

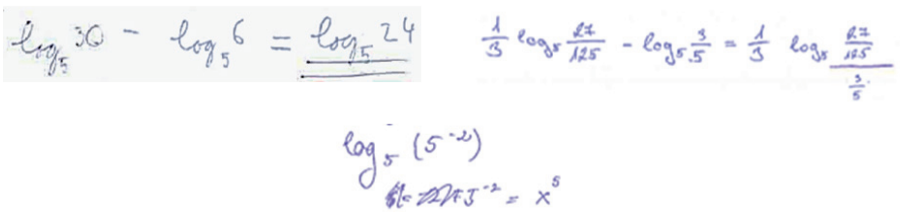
Apart from the use of formerly correct algebraic schemas in the situation of logarithm, thus the source of errors were *problems with accommodation,* students tried to interpret logarithm similarly to previously known algebraic concepts, such as square root (e.g. the interpretation of logarithm as $\log_3 4 = \sqrt{2}. \log_3 9 = \sqrt{3},\ \log_3 25 = \sqrt{5}$ $\log_3 4 = \sqrt{2}, \log_3 9 = \sqrt{3},\ \log_3 25 = \sqrt{5}$), or interpreting $\log_3$.$\log_3$. as an operation for the division by 3. In other cases, the application of logarithm properties was accommodated incorrectly, such as $\frac{1}{2}log_3 25 = \sqrt{log_3 25}\frac{1}{2}log_3 25 = \sqrt{log_3 25}$. The remaining four cases affected exponents, e.g. $\frac{1}{2}log_3 25 = log_3\left(\frac{1}{2} * 25\right)\frac{1}{2}log_3 25 = log_3\left(\frac{1}{2} * 25\right)$. Two distinct errors were included in this category. The first was when the identity was used for rewriting the problem from two logarithmic terms to one term. Then the operation to be performed between the arguments remained the same as the one between the terms.

The log was factored out such as in $\log_5 30 - \log_5 6 = \log_5 24$ $\log_5 30 - \log_5 6 = \log_5 24$ (13 cases). This problematic accommodation appeared also in the case of the order of operation, when a multiplication of on log arithmic term was applied to all terms of the sum, e.g. $\frac{1}{3}\log_5 \frac{27}{125} - \log_5 \frac{3}{5} = \frac{1}{3}\log_5 \frac{\frac{27}{125}}{\frac{3}{5}}$ $\frac{1}{3}\log_5 \frac{27}{125} - \log_5 \frac{3}{5} = \frac{1}{3}\log_5 \frac{\frac{27}{125}}{\frac{3}{5}}$ (This error was quite typical and appeared 10 times, thus we refer to this later on in the paper as “1/3 multiplier phenomenon”). In these errors, we see a *problem with accommodation* of the logarithm into the sequence of operations scheme. Performing logarithm identity takes precedence over the multiplying factor of 1/3. In the cases regarding other operations such as roots or exponents we did not experience similar error patterns.

We found 4 errors where we could not identify the cause. The reason for the incorrect use of definition or theorems were *unidentifiable* to us (e.g. $\frac{1}{2}\log_5 25 = \frac{1}{2}5$ $\frac{1}{2}\log_5 25 = \frac{1}{2}5$).

Figure 6 shows some solutions containing such errors.

FIGURE 6



Student errors related to the concept of logarithm

Misused data

The errors corresponding to misused data refers to discrepancies between the properties of data given and the properties of data treated by students. These errors are quite common in the case of logarithm due to the restrictions on the domain of the logarithm function. However, our tasks did not address this issue.

The errors corresponding to *distortion of the concept and its properties* are mainly those when logarithm disappeared at some point during the solution without any valid reasons. This error type was referred to cancelling logarithm notation and had been identified by other researchers as well (de Gracia, 2016; Hoon et al., 2010). One possible explanation can be that the monotonicity property of the logarithm function is distorted. By solving logarithmic equations students learn that, because of the monotonicity, in the case of logarithm with the same base it is possible to compare the arguments only. However, they used this idea for evaluation as well, e.g. $\log_5 30 - \log_5 6 = \frac{30}{6} = 5$ $\log_5 30 - \log_5 6 = \frac{30}{6} = 5$ or $2\log_5 \frac{1}{5} = \frac{1^2}{5} = \frac{1}{25}$ $2\log_5 \frac{1}{5} = \frac{1^2}{5} = \frac{1}{25}$. Apart from cancelling the logarithm notation, we also found other errors in this category, such as incorrect algebraic manipulations in order to be able to use logarithmic identities. We did see this error while equating $30 = 6^5$ $30 = 6^5$ makes possible to use the logarithm rule for exponents and the student finally wrote $\log_5 30 = \log_5 6^5 = 5\log_5 6$ $\log_5 30 = \log_5 6^5 = 5\log_5 6$. Students tried to apply the "step" they expected to use by solving a logarithm problem, e.g. write 30 in a form of exponent and then apply the appropriate logarithm rule similarly to the situations described in Hardy (2009).

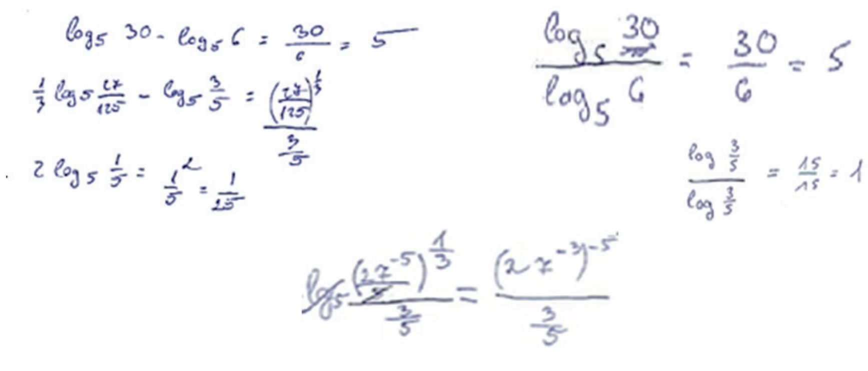
The other possible cause for this type of error was the *distortion of the accommodation of a concept in previous schemes, thus problems with accommodation*. In the case of these errors, an overgeneralization of previous schemes is found. A detailed description of such errors was given by Chua and Wood (2005). Students "simplify with logarithm" in some form as if it were an algebraic factor in a multiplication, e.g. $\frac{\log_5 30}{\log_5 6} = \frac{30}{6} = 5$ $\frac{\log_5 30}{\log_5 6} = \frac{30}{6} = 5$. In other cases, simplification is done in a completely different way. If the argument contains the base of the logarithm, then students simplified with the loga rithm using the $\log_a a = 1\log_a a = 1$ identity, such as in $\log_5 \frac{\left(\frac{27^{-5}}{5}\right)^{\frac{1}{3}}}{\frac{3}{5}} = \frac{\left(27^{-3}\right)^{-5}}{\frac{3}{5}}$ $\log_5 \frac{\left(\frac{27^{-5}}{5}\right)^{\frac{1}{3}}}{\frac{3}{5}} = \frac{\left(27^{-3}\right)^{-5}}{\frac{3}{5}}$ by crossing the $\log_5\log_5$ and the 5 in the numerator's denominator or $\log_5(5^{-1})^{\frac{1}{3}}$ $=\log_5(5^{-1})^{\frac{1}{3}} = (1^{-1})^{\frac{1}{3}}(1^{-1})^{\frac{1}{3}}$. This suggests that the concept of logarithm is not well accommodated into the sequence of operations schema. In another case, students overgeneralize the properties of equations and simplify by a common factor of arguments. In this situation the scheme for solving equations is applied for the logarithm as if the new concept made it possible to apply such algebraic manipulations, e.g. $\log_5 30 - \log_5 6 = \log_5 5 - \log_5 1\log_5 30 - \log_5 6 = \log_5 5 - \log_5 1$.

Previous erroneous schemes could be the cause of a surprising solution when a student applies linear interpolation for estimating the value of an exponential expression, as used a previously correct schema in a new situation, when it was not applicable.

Misused data coming from *unidentified causes* were mainly errors of estimations or other errors. Errors of estimation refer to errors students committed while approximating the value of logarithmic expressions. For example, a student used estimations such as $\log_3 3 = \frac{1}{5}\log_3 3 = \frac{1}{5}$ $\log_3 8 = 2\log_3 8 = 2$. Surprisingly, quite a few students used estimation, a problem-solving strategy they hardly did see in the case of high school logarithm problems. Most probably the type of task triggered this approach. The issue of estimating the value of logarithmic expressions is an interesting question in itself and could be of interest for further investigations.

In Figure 7, typical errors are shown.

FIGURE 7

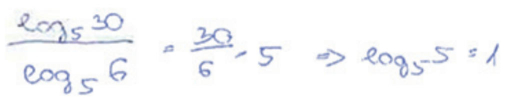


Students' incorrect solutions in which they omit the logarithm

Misinterpreted language

Errors due to misinterpreted language in our case were caused by *previously incorrect schemes* regarding algebraic manipulations. One of such errors were the formally incorrect use of logarithm theorems, such as the identity e.g. $log_5 30 - log_5 6 = \frac{log_5 30}{log_5 6} = log_5 5 log_5 30 - log_5 6 = \frac{log_5 30}{log_5 6} = log_5 5$ or other incorrect formalism, where the equal sign simply denotes that "there is a calculation in progress", e.g. $log_5 30 - log_5 6 = \frac{30}{6} = 5 = log_5 5 . log_5 30 - log_5 6 = \frac{30}{6} = 5 = log_5 5$. Figure 8 shows the co-occurrence of both errors.

FIGURE 8



Error due to previously incorrect schemes

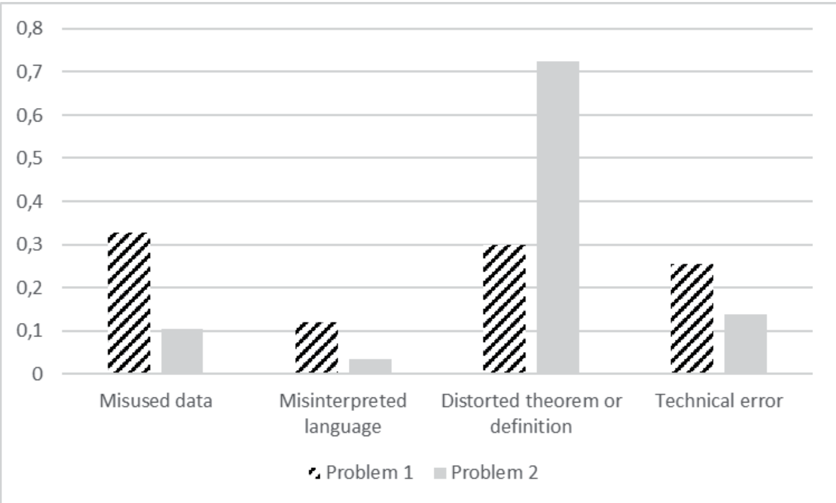
Technical errors

Technical errors include calculation errors due to various reasons. Most of these happen because of formerly incorrect schemes regarding basic algebra (for example $2 * (-1) = 12 * (-1) = 1$ or $3 * \frac{1}{3} = \frac{3}{9} 3 * \frac{1}{3} = \frac{3}{9}$) or incorrect formalism (e.g. incorrect use of parentheses, for example $\log_5 \frac{27}{125}^{\frac{1}{3}} \log_5 \frac{27}{125}^{\frac{1}{3}}$). In some cases, we could not identify the source of the errors, for example: $\frac{3}{6}\log_5 75 - \frac{2}{6}\log_5 16 = \frac{1}{6}\log_5 \frac{75}{16}$ $\frac{3}{6}\log_5 75 - \frac{2}{6}\log_5 16 = \frac{1}{6}\log_5 \frac{75}{16}$.

The distribution of errors

Task design influences the thinking processes (Stein et al., 1996), and thus the errors committed as well. Not surprisingly, the distribution of errors in tasks also depends on the nature of the task. In our study, we examined the solution of two different tasks, one in which the values had to be placed in numerical order and the other in which they had to be sorted, see Figures1 and 2. The differences in the distribution of each type of error are summarized in Figure 9. For the sake of comparability, the relative frequencies of errors are presented on the graph, as 101 students solved Task 1 and 55 students solved Task 2. In total, 67 errors occurred in the first case and 29 in the second.

FIGURE 9



Distributions of error types for the two problems

Reasons of differences in the distribution

We did see that the distribution of errors was different in the case of Task 1 and 2. The reasons for this can be varied, however, we identified three sources of the differences corresponding to the design of the tasks although we did not expect such differences beforehand. Two of these were corresponding to the values used in the problems, and the third one was due to the explicit use of the number line.

In the first case the values in the problem worked as distractors. Namely, the arguments of the logarithm suggested the applicability of incorrect identity: in our case, in the second set of problems the first expression was $\log_3 7 + \log_3 2 = \log_3 9$ $\log_3 7 + \log_3 2 = \log_3 9$. Here the sum of the arguments was a power of the base of the logarithm suggesting that the arguments should be added to obtain the correct result. In total, we found 9 such errors out of 21 errors in the distorted theorem in Task 2.

The second difference, the "one-third multiplier" phenomenon was encountered in the setting of Task 1, with the expression $\frac{1}{3}\log_5\frac{27}{125} - \log_5\frac{3}{5}\frac{1}{3}\log_5\frac{27}{125} - \log_5\frac{3}{5}$ in Task 1, compared to the similar $\frac{1}{2}\log_3 25 - \frac{1}{3}\log_3 8\frac{1}{2}\log_3 25 - \frac{1}{3}\log_3 8$ in Task 2. The unla beled singular coefficient in the 2nd member of Task 1 resulted that some students placed the operation of logarithm before multiplication. In Task 1, we identified 10 of these types of errors out of 20 distorted theorem errors.

The third difference is due to the explicit use of the number line in the problem. The problem posing itself resulted the appearance of estimates, since students wanted to place the value to each expression on the number line. If they could not determine the exact, an approximation was attempted. In Task 1, 10 people solved the problem using estimation, 3 of these solutions were incorrect, while in Task 2, two people estimated the value of the terms, one of them incorrectly. These errors modified the proportions of the Misused data errors.

Results (study 2022/23) and discussion

This chapter presents the results of the second survey, which took place in the autumn semester of 2022/23. 139 students completed Task 3. Overall, 23% of these students gave the correct answer. In Task 1, part 1 was solved correctly by 51 students (37%), part 2 by 16 students (12%), part 3 by 43 students (31%), and part 4 by 18 students (13%). The 139 students made a total of 172 errors. The distribution of errors is shown in Table 4.

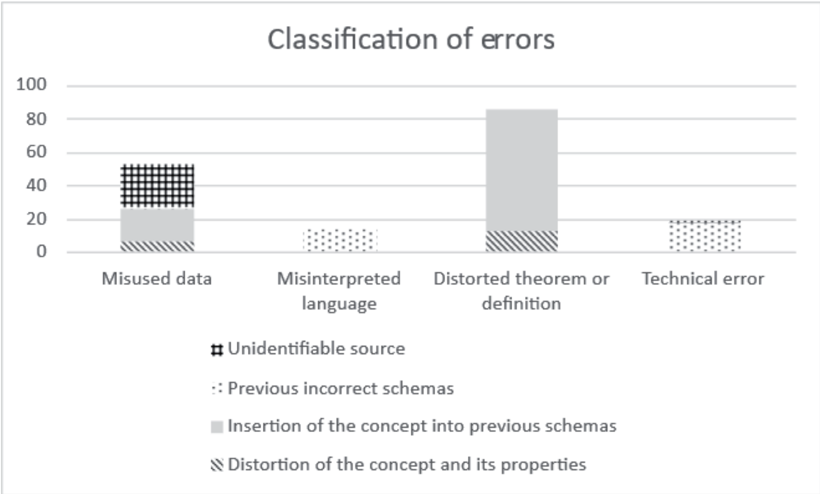
Table 4

Distribution of students' errors in 2022/23

	Possible sources of error				
Type of error	*Distortion of the concept*	*Problems with accommodation*	*Previous erroneous schemes*	*Unidentified causes*	*Total*
Distorted definition or theorem	13	73			86
Misused data	7	19	1	26	53
Misinterpreted language			14		14
Technical errors			18	1	19
Total	20	92	33	27	172

The number of errors (percentages) by type of error: 53 errors (30.81%) were due to misused data, errors due to misinterpreted language 14 (8.14%), errors due to distorted definition or theorem 86 (50%) and technical errors 19 (11.05%). The distribution of causes of errors according to error types is shown in Figure 10.

FIGURE 10

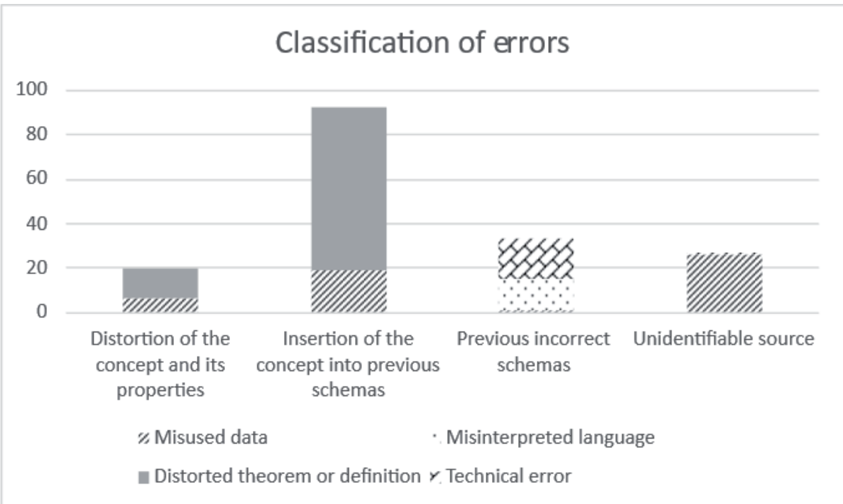


Distribution of error types and distribution of sources of errors within each error type in 2022/23

Distribution of errors by sources (percentages): Distortion of concept and its properties 20 (11.63%), problems with accommodation 92 (53.49%), the previous erroneous schemes 33 (19.19%) and unidentifiable sources 27 (15.70%).

The distribution of error types within possible sources is shown in Figure 11.

FIGURE 11



Distribution of error types and distribution of sources of errors within each error type in 2022/23

Table 5 shows the correlations between the different types of errors. It suggests that there is no connection between the occurrences of different types of errors. This reinforces the results of our earlier investigation.

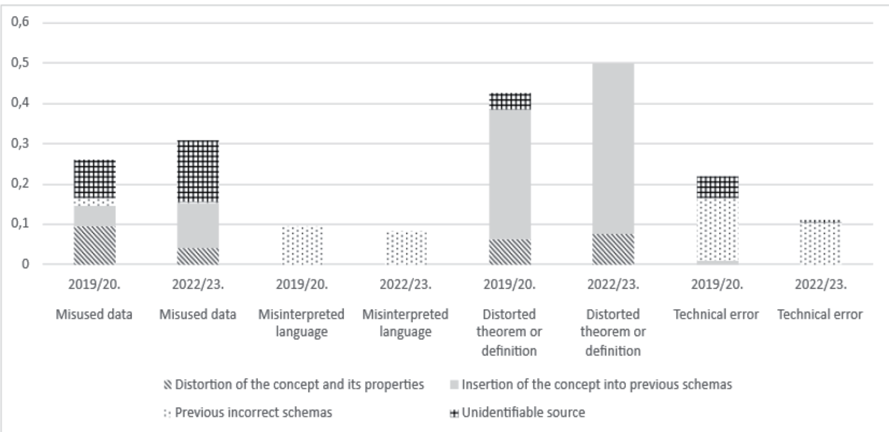
TABLE 5

Pearson's Correlations between error types in 2022/23

Variable		DD/DT	MD	ML	TE
1. DD/DT	Pearson's r	-			
	p-value	-			
2. MD	Pearson's r	0.148	-		
	p-value	0.099	-		
3. ML	Pearson's r	0.153	0.041	-	
	p-value	0.087	0.648	-	
4. TE	Pearson's r	0.274	0.179	0.217	-
	p-value	0.002	0.045	0.014	-

Finally, we compared the error distributions obtained in the two surveys. The corresponding $\chi^2\chi^2$-test showed no significant difference between the error distributions of the two surveys ($\chi_3^2 = 6.12; p = 0{,}1058\chi_3^2 = 6.12; p = 0{,}1058$). A graphical comparison of the distributions is shown in Figure 12.

FIGURE 12



Comparison of the error distributions of the two tests

Based on the above, we were able to carry out the categorization addressed in our original research question. Namely, we could clearly and concisely categorize student errors in logarithm problems according to a two-dimensional framework based on types and sources of the errors. A follow-up study confirmed that the categorization could reliably applied in a second similar error analysis.

Conclusion

Error analysis offers a tool to understand learning through misconceptions. In the paper we tried to characterize the misconceptions regarding logarithm and categorize the corresponding errors using a two-dimensional categorization. Admittedly, the obtained results need refinement, thus we plan to conduct interviews and expand the types of problems, e.g. problems involving logarithm functions, change of base formula.

The errors in the problems we studied did not cover all the previously identified logarithm-related error types. In terms of algebraic interpretation, this was mainly due to the fact that the task did not ask for a rewriting of the base of the logarithm unlike previous studies (Hoon et al., 2010). Apart from this type of error we did not find other categories not addressed in this publication, and it is likely that these errors could also be categorized using our method. However, in the case of errors related to the function concept of logarithm, the current categorization may need to be extended. Nevertheless, during this exploratory study we were able to reveal some common types of errors not investigated previously, especially in the misused data category. Apart from the errors analyzed we did find a few interesting strategies students used, such as estimation and interpolation.

References

Asli, A., & Zsoldos-Marchis, I. (2023). The effect of an intervention with teaching applications of mathematics on students' attitudes and achievement. *Review of Science, Mathematics and ICT Education*, *17*(2), 27-45. https://doi.org/10.26220/rev.4480

Aziz, T. A., Pramudiani, P., & Purnomo, Y. W. (2017). How δo College students solve logarithm questions? *International Journal on Emerging Mathematics Education*, *1*(1), 25. https://doi.org/10.12928/ijeme.v1i1.5736

Chua, B. L., & Wood, E. (2005). Working with logarithms: Students' misconceptions and errors. *The Mathematics Educator*, *8*(2), 53-70.

Clark, M., & Lovric, M. (2008). Suggestion for a theoretical model for secondary-tertiary transition in Mathematics. *Mathematics Education Research Journal*, *20*(2), 25-37.

de Gracia, R. S. (2016). *High school seniors with alternative conceptions on exponents and logarithms*. 13th National Convention on Statistics (NCS) EDSA Shangri-La Hotel, Mandaluyong City.

Ganesan, R., & Dindyal, J. (2014). An investigation of students' errors in logarithms. *Mathematics Education Research Group of Australasia.*

Giannakoulas, A., & Xinogalos, S. (2024). Studying the effects of educational games on cultivating computational thinking skills to primary school students: A systematic literature review. *Journal of Computers in Education*, *11*(4), 1283-1325.

Hardy, N. (2009). Students' perceptions of institutional practices: the case of limits of functions in college level Calculus courses. *Educational Studies in Mathematics*, 72(3), 341-358. https://doi.org/10.1007/s10649-009-9199-8

Hoon, T. S., Singh, P., & Ayop, S. K. (2010). Working with logarithms. *Malaysian Education Dean's Council Journal*, *6*(6), 121-129.

Kelecsényi, K., Osztényiné Krauczi, É., & Végh, A. (2025). Understanding logarithm in a mathematical card game environment. *International Journal of Mathematical Education in Science and Technology*, *56*(4), 646-666.

Kenney, R., & Kastberg, S. (2013). Links in learning logarithms. *Australian Senior Mathematics Journal*, *27*(1), 12-20.

Movshovitz-Hadar, N., Zaslavsky, O., & Inbar, S. (1987). An empirical classification model for errors in High School Mathematics. *Journal for Research in Mathematics Education*, *18*(1), 3. https://doi.org/10.2307/749532

Mulqueeny, E. (2012). *How do students acquire an understanding of logarithmic concepts?* Kent State University.

Olivier, A. (1989). Handling pupils' misconceptions. *Pythagoras*, *21*, 10-19.

Rafi, I., & Retnawati, H. (2018). What are the common errors made by students in solving logarithm problems? *Journal of Physics: Conference Series*, *1097*(1), 012157. https://doi.org/10.1088/1742-6596/1097/1/012157

Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal*, *33*(2), 455-488.

Weber, C. (2016). Making logarithms accessible–operational and structural basic models for logarithms. *Journal Für Mathematik-Didaktik*, *37*(1), 69-98.