Journal of Theoretical Educational Science, 15(3), 453-474, July 2022
Kuramsal Eğitimbilim Dergisi, 15(3), 453-474, Temmuz 2022
[Online]: http://dergipark.org.tr/akukeg
DOI number: http://doi.org/10.30831/akukeg.1027601

Assessment of Mathematical Reasoning Competence in Accordance with PISA 2021 Mathematics Framework

Matematiksel Muhakeme Etme Yeterliğinin PISA 2021 Matematik Çerçevesine Göre Değerlendirilmesi

Zeynep ÖZAYDIN* **Çiğdem ARSLAN****

Received: 23 November 2021 **Research Article** **Accepted:** 05 April 2022

ABSTRACT: The aim of this study is to develop a rubric to assess mathematical reasoning competence. Since the aim is to assess a competency, the frameworks of the PISA exams in the literature, which give an important place to competencies, have been examined. Due to its focus and in-depth analysis of mathematical reasoning, each of the actions expected from the mathematical reasoning process in the PISA 2021 Mathematics Framework was handled as a criterion and a rubric was created. Data were collected from 30 mathematics teachers with the data collection tool consisting of questions that were stated to be suitable for assessing mathematical reasoning skills in the resource presented as a broadened perspective for the PISA 2021 Mathematics field. The collected data were analyzed using the finalized rubric. In reliability analysis; The "percentage of agreement" among researchers was used to determine the criteria included in the rubric and to be followed with the data collection tool, "Cohen's kappa coefficient" and "Krippendorff's alpha coefficient" methods were used for the agreement between the raters after the analysis of the collected data. The findings show that the Reasoning Competence Rubric (RCR), which consists of 12 criteria, is valid and reliable.

Keywords: Mathematical reasoning, competence, mathematical reasoning rubric, mathematical literacy, PISA.

ÖZ: Bu çalışmanın amacı, matematiksel muhakeme etme yeterliğini değerlendirmeye yönelik bir değerlendirme tablosu geliştirmektir. Amaç bir yeterliğin değerlendirilmesi olduğu için yeterliklere önemli bir yer veren PISA sınavlarının literatürdeki çerçeveleri incelenmiştir. Matematiksel muhakeme etmeyi odağına alması ve derinlemesine incelemesi sebebiyle PISA 2021 Matematik Çerçevesi'nde yer alan matematiksel muhakeme etme sürecinden beklenen eylemlerin her biri kriter olarak ele alınıp değerlendirme tablosu oluşturulmuştur. PISA 2021 matematik alanı için genişletilmiş bir perspektif olarak sunulan kaynakta yer alan matematiksel muhakeme etme becerilerini ölçmeye uygun olduğu belirtilen sorulardan oluşturulan veri toplama aracı ile 30 ilköğretim matematik öğretmeninden veriler toplanmıştır. Toplanan veriler son hali verilen değerlendirme tablosu kullanılarak analiz edilmiştir. Güvenilirlik analizlerinde, değerlendirme tablosunda yer alan ve veri toplama aracı ile izlenecek olan kriterlerin tespiti için araştırmacılar arasındaki "uyum yüzdesi", toplanan verilerin analizi sonrasında puanlayıcılar arasındaki uyum için "Cohen'in kappa katsayısı" ve "Krippendorff'un alfa katsayısı" yöntemleri kullanılmıştır. Değerlendirme tablosunun geçerliğini sağlamak için uzman görüşlerine başvurulmuştur. Elde edilen bulgular 12 kriterden oluşan Muhakeme Etme Yeterliği Değerlendirme Tablosunun (MYDT) yeterli düzeyde geçerli ve güvenilir olduğunu göstermektedir.

Anahtar kelimeler: Matematiksel muhakeme, yeterlik, muhakeme etme değerlendirme tablosu, matematik okuryazarlığı, PISA.

* *Corresponding Author*: Master Student, Bursa Uludağ University, Bursa, Turkey, zeynepozaydin@uludag.edu.tr, http://orcid.org/0000-0003-1768-3963
** Assoc. Prof. Dr., Bursa Uludağ University, Bursa, Turkey, arslanc@uludag.edu.tr, http://orcid.org/0000-0001-7354-8155

Citation Information
Özaydın, Z., & Arslan, Ç. (2022). Assessment of mathematical reasoning competence in accordance with PISA 2021 mathematics framework. *Kuramsal Eğitimbilim Dergisi [Journal of Theoretical Educational Science], 15*(3), 453-474.

ISSN: 1308-1659

Reasoning is the process of supporting an idea, confirming that idea or justifying a decision (Toulmin et al., 1984). Leighton (2003) defined reasoning as the process of organizing evidence, beliefs, and thoughts about the accuracy of conclusions. It can be said that these two definitions are similar in that they see reasoning as a process. Considering the relationship of reasoning with mathematics, reasoning is the foundation of mathematics epistemologically. As science verifies through observation, mathematics relies on logic (Steen, 1999).

Bal-İncebacak and Ersoy (2016) stated that mathematics contains many skills that an individual needs to develop, and one of the most important of these skills is reasoning. There are many definitions in the literature on mathematical reasoning. Russell (1999) defined mathematical reasoning as a tool for understanding the abstract, symbolic expressions that create mathematics. Peresini and Webb (1999) stated that mathematical reasoning is a dynamic activity that includes various ways of thinking. Yackel and Hanna (2003) stated that mathematical reasoning is a common activity in which more than one mathematical skill interacts, and they present a view parallel to Peresini and Webb's (1999) definition.

Reasoning is one of the important aims of mathematics teaching (Yankelewitz, 2009). When national and international improvement movements are examined, it can be seen that mathematical reasoning plays an important role in learning mathematics (Erdem, 2015). For this reason, it was stated that understanding mathematics would be incomplete without reasoning (Ball & Bass, 2003). It is useful to examine the definitions of reasoning in the mathematics teaching programs of different countries since reasoning is important among the skills that mathematics should provide to the individual and is among the objectives of mathematics teaching.

In the secondary education program of the Ministry of National Education [MoNE] (2013), reasoning is defined as the process of acquiring new information. Again, some indicators that should be taken into account to gain reasoning skills in the secondary school mathematics course 2013 curriculum are;

- Defending the accuracy and reality of inferences
- Making logical inferences and generalizations
- Explaining and using mathematical patterns and relationships when analyzing a mathematical situation
- Making predictions about the outcome of operations and measurements using strategies such as rounding, grouping appropriate numbers, using first or last digits, or strategies they have developed.
- Making an estimation of the measurement by taking into account a certain reference point (MoNE, 2013, p. 5).

In the secondary school mathematics course 2018 curriculum, the aim of providing students with reasoning skills is "The student will be able to easily express their own thoughts and reasoning in the problem-solving process, and will be able to see the mathematical reasoning deficiencies or gaps of others (MoNE, 2018, p. 9)." is in the form. The New Jersey Mathematics Teaching Curriculum states that mathematical reasoning is the critical skill that enables a student to make use of all other math skills (New Jersey Mathematics Coalition and the New Jersey Department of Education [NJMCF], 1996, p. 44). In the Australian Mathematics Teaching Curriculum, on the other hand, reasoning is considered as a mathematical competence that needs to be developed in students and is defined as the capacity for actions such as proving, evaluating, explaining and making inferences (Australian Curriculum and Assessment Authority [ACARA], 2017).

Mathematical Reasoning and Mathematical Literacy

The concept of mathematical literacy entered the literature with the Program for International Students Assessment [PISA] exams. PISA exams, developed in 1997 and administered for the first time in 2000, are administered by the Organization for Economic Co-operation and Development [OECD] in three-year periods to evaluate the knowledge and skills of students in the 15-year-old group. PISA exams are accepted as an important tool to evaluate the education quality of countries, and the results of PISA exams create a discussion environment where issues, such as education systems and teaching quality of countries, can be discussed (Dabic-Boricic et al., 2020). Turkey has not been able to get rid of the back ranks in the PISA exams it has participated in since 2003 (OECD, 2003, 2007, 2010, 2013). This situation has been instrumental in taking PISA results into account when deciding on education policies in Turkey, as in many participating countries (Gür et al., 2012). When considered in this context, it can be said that the PISA exams and the frameworks determined for these exams have an important place in the mathematics education literature.

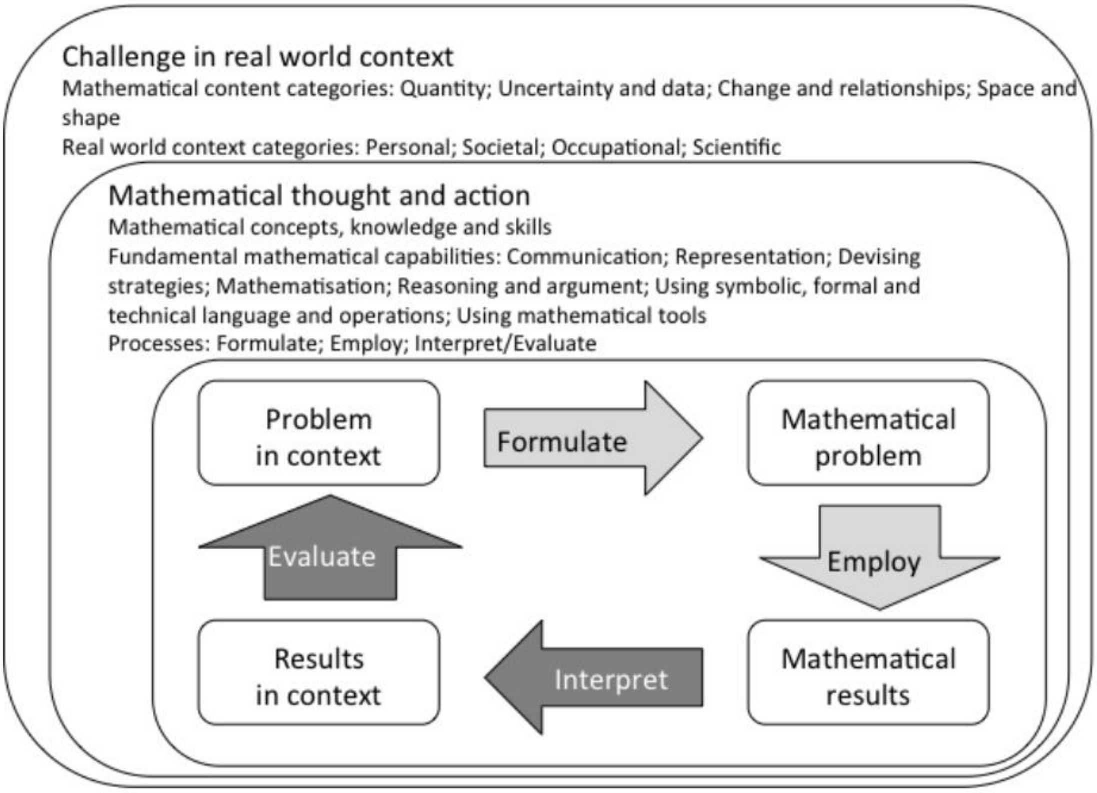
The literature on mathematical literacy is quite rich (Gatabi et al., 2012; Güzel, 2017; Matteson, 2006; Özgen, 2021; Yıldız, 2019). Many definitions of mathematical literacy have been made so far (İskenderoğlu & Baki, 2011; OECD, 2013; PISA 2021a; Steen et al., 2007). Steen et al. (2007) defined mathematical literacy as the capacity to use mathematical knowledge and understanding effectively to overcome difficulties in daily life. İskenderoğlu and Baki (2011) defined mathematical literacy as a way of using mathematics in our individual lives as well as using individual capacity and making causal defenses to understand the role of mathematics in the world. The definition made by the OECD in the PISA 2021 Mathematics Framework is as follows:

> Mathematical literacy is an individual’s capacity to reason mathematically and to formulate, employ, and interpret mathematics to solve problems in a variety of real-world contexts. It includes concepts, procedures, facts and tools to describe, explain and predict phenomena. It assists individuals to know the role that mathematics plays in the world and to make the well-founded judgments and decisions needed by constructive, engaged and reflective 21st century citizens (PISA, 2021a, p. 7).

Mathematical reasoning takes its place among the mathematical competencies both in the OECD definition of mathematical literacy (Mathematical literacy is an individual’s capacity to reason mathematically and to formulate, use and interpret mathematics to solve problems in a variety of real-world contexts) and in the mathematical literacy model specified in Figure 1.

Figure 1

PISA Mathematical Literacy Model



Note. (PISA, 2021b, p. 5)

The key to mathematical competencies is mathematical literacy (Doyle, 2007). Having mathematical competencies is closely related to mathematical literacy (Turner, 2010). Mathematical competencies are "the cognitive processes that should be activated to connect the real world in which the problem arises with mathematics and solve the problem posed" (Sáenz, 2009, p. 126). Niss (2003) defined the concept of mathematical competence, which he examined under eight different titles (thinking mathematically, posing and solving mathematical problems, modelling mathematically, reasoning mathematically, representing mathematical entities, handling mathematical symbols and formalisms, communicating in, with, and about mathematics, making use of aids and tools), as the ability to understand and use mathematics, while Niss and Højgaard (2019) defined mathematical competence as the ability to master the basic aspects and wishes of mathematics and to act effectively in this field.

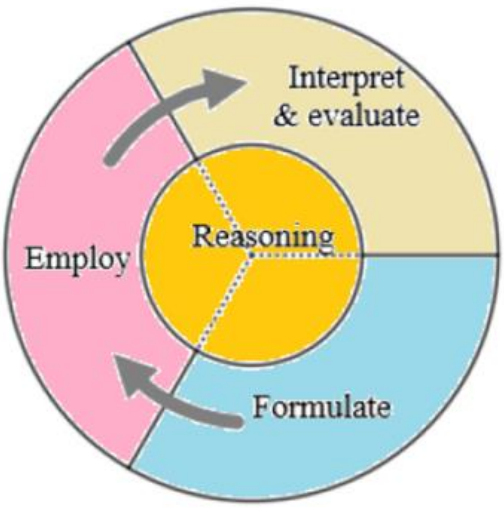
Various sources (Altun, 2020; Kilpatrick et al., 2002; Niss, 2003; OECD, 2013; PISA, 2021a) in the literature show that mathematical reasoning is among the mathematical competencies. An individual with mathematical competence can solve problems encountered in school, mathematics lessons and daily life by reasoning (Demir & Vural, 2016). Herman (2018) stated that reasoning is a very important aspect of mathematical competencies in learning mathematics. These statements reveal the importance of reasoning among the mathematical competencies.

OECD (2013) states that mathematical reasoning competence includes the processes of searching for problem elements, inferring from them, checking a rationale, searching the rationale for statements to provide solutions to problems, and logical thinking processes that link all these together. The PISA 2021 Mathematics Framework supports that reasoning is a core competence for the science of mathematics, saying that mathematics is a science about well-defined objects and concepts that can be analyzed and transformed in different ways by mathematical reasoning.

In PISA frameworks published by OECD in previous years (OECD, 2013, 2017, 2019), expected actions in the mathematical reasoning process are integrated with mathematical process skills. In the framework of PISA 2021, mathematical reasoning was handled as a fundamental aspect of mathematical literacy and focused on (Figure 2), and the expected actions in the mathematical reasoning process are listed (Figure 3) as follows:

Figure 2

Mathematical Literacy: The Relationship between Mathematical Reasoning and Mathematical Process Skills



Note. (PISA, 2021a, p. 8)

Figure 3

A Part of Expected Actions for Mathematical Reasoning

** Draw a simple conclusion
** Select an appropriate justification
** Explain why a mathematical result or conclusion does, or does not, make sense given the context of a problem
Represent a problem in a different way, including organising it according to mathematical concepts and making appropriate assumptions
Utilise definitions, rules and formal systems as well as employing algorithms and computational thinking
Explain and defend a justification for the identified or devised representation of a real-world situation
Explain or defend a justification for the processes and procedures or simulations used to determine a mathematical result or solution
Identify the limits of the model used to solve a problem
Understand definitions, rules and formal systems as well as employing algorithms and computational reasoning
…

Note. (PISA, 2021a, p. 35)

Although mathematical reasoning is a competence covered in many mathematics teaching curricula around the world, teachers struggle to understand, teach and assess mathematical reasoning (Loong et al., 2018). Loong et al. (2018) developed a rubric that deals with each of the three reasoning actions (analyzing, generalizing, justifying) at five levels (not evident, beginning, developing, consolidating, extending) assessment students' reasoning and presented it to the use of teachers. Bal-İncebacak and Ersoy (2016) used the mathematical reasoning stages of TIMSS 2003 to assess the mathematical reasoning skills of secondary school 7th grade students. Çoban and Tezci (2020), on the other hand, developed a scale consisting of multiple-choice and open-ended questions aiming to reveal the level of mathematical reasoning skills of secondary school students. On the other hand, Ersanlı et al. (2018) revealed that subject-based dimensions of mathematical reasoning (e.g., proportional reasoning, reasoning about ratio-proportionality, etc.) were addressed in more than one study, but there was a gap in research on thinking styles and perspectives in the dimension of mathematical reasoning. Teachers are the most effective people in teaching reasoning skills to students. (Altıparmak & Öziş, 2005). Despite this important role of the teacher in helping students gain reasoning competence, Herbert (2019) listed the difficulties of teachers in assessing mathematical reasoning as follows; teacher's reasoning knowledge, teacher's awareness of reasoning, students' difficulties in expressing reasoning, lack of guidance/support in curriculum documents, lack of work examples, difficulty in monitoring and reporting student progress in reasoning. When these are taken into consideration, to direct future studies and reform movements, it is considered important to assess the reasoning competencies of teachers. Therefore, the study group of this study consists of mathematics teachers. In this context, the aim of this study is to create a rubric that can be used to assess teachers' mathematical reasoning competencies, and for this, the expected actions in the reasoning process specified in the PISA 2021 Mathematics Framework were used since it is an up-to-date resource and focuses on mathematical reasoning and examines it in depth.

Method

Creation of the Rubric

A rubric is a rating system in which teachers can determine at what level a student can perform a task or demonstrate knowledge of a concept (Brualdi-Timmins, 1998). Many methods have been mentioned in the literature on the development of rubrics (Brualdi-Timmins, 1998; Goodrich-Andrade, 1997; Moskal, 2000; Moskal & Leydens, 2000; Russell & Airasian, 2001). By taking into consideration these methods, the stages of creating the rubric used in this study are discussed in detail in the following.

Determining the Criteria

Since the aim is to assess the competence of mathematical reasoning, the sources of the PISA exams in the literature, which frequently bring the concept of competence to the agenda, were scanned. Thus, due to the fact that it focuses on mathematical reasoning and examines it in detail, the expected actions in the mathematical reasoning process within the framework of PISA 2021 are primarily discussed as a list. The actions expected in the mathematical reasoning process were translated into Turkish by taking expert opinion. Expert opinions were taken in one-on-one interviews to adapt the actions to the Turkish language. The researchers consulted the opinions of two different experts about the translations they made with the expert and the points where they had dilemmas. Later, the actions were evaluated as criteria and started to be tabulated.

In PISA 2021, the first of the reasoning actions is handled in three separate actions (1a, 1b, 1c) to include students at the extreme. Since the present study was carried out on teachers, not students, unlike PISA, the third item (1c: Explain why a mathematical result or conclusion does, or does not, make sense given the context of a problem), which includes the first two, was included and the other two items (1a: Draw a simple conclusion, 1b: Select an appropriate justification) were omitted in the rubric by taking expert opinion. Only some criteria whose action was changed were combined by taking expert opinion. The combined criteria are shown in Figure 4 and included in the rubric with the codes indicated next to it.

Figure 4

Criteria Combined with Expert Opinion

Explain and defend a justification for the identified or devised representation of a real-world situation Provide a justification for the identified or devised representation of a real-world situation	Provide, explain, or defend a justification for the identified or devised representation of a real-world situation (C3).
Explain or defend a justification for the processes and procedures or simulations used to determine a mathematical result or solution Provide a justification for the processes and procedures used to determine a mathematical result or solution	Provide, explain or justify a mathematical result or solution for the processes used to determine a mathematical result or solution (C4).
Critique the limits of the model used to solve a problem Identify the limits of the model used to solve a problem	Identify or criticize the limits of the model used to solve a problem (C5).
Utilize definitions, rules and formal systems as well as employing algorithms and computational thinking Understand definitions, rules and formal systems as well as employing algorithms and computational reasoning	Utilize or understand definitions, rules, formal systems, as well as using employing algorithms, computational thinking and reasoning (C6).

In the framework of PISA 2021, it is stated that understanding numbers, operations, representations (symbols containing numbers, etc.) and how to move between representations are the basis of mathematical reasoning. It can be said that the importance of establishing the relations between the expression “moving between representations” and the contextual language and mathematical language required for representation is emphasized. Therefore, “Explain the relationships between the context-specific language of a problem and the symbolic and formal language required to represent it mathematically.” The statement “Construct or explain the relationships between the context-specific language of a problem and the symbolic and formal language required to represent it mathematically. (C9)” and included in the rubric with the code specified next to it. Other statements (C1, C2, C7, C8, C10, C11, C12) were included in the rubric without any changes. While the verbs of the criteria are transferred to the rubric, they are conjugated with the singular third person. As a result of all these steps, a rubric with 12 criteria was obtained.

Deciding on the Type of Rubric and the Levels to be Used

According to many researchers, in the scoring made by dividing a field into sub-fields, the scoring is less affected by the subjectivity of the raters and the difference between the raters decreases (Güler, 2019). Bülbül (2019) said that the categories of the criteria in the rubrics should be determined numerically and their contents should also be defined. In other words, the scoring in the rubrics should be made according to the criteria and the way these criteria are applied and found (Bülbül & Bülbül, 2021). For this reason, the scoring strategy of the rubric was made as "analytical". Analytical rubrics contain criteria that define the dimensions of a task at multiple levels (McGatha & Darcy, 2010). Each criterion in the rubric is addressed at 0-1-2 points. According to the rubric, the individual who fully demonstrates the relevant criterion gets 2 points, while the individual who does not show any gets 0 points.

Defining the Score Levels of the Criteria

While preparing the rubrics, the highest levels of the criteria are determined first. At this stage of the study, first of all, the meaning of the highest score levels (2 points) of the criteria was determined by consulting expert opinions. After the top score levels were defined, the meaning of the other score levels (0 Points and 1 Point) was shaped in line with the expert opinions.

Creating the Draft Rubric and Determining How It will be Used

At this stage, a draft version of RCR with 12 criteria and 0-1-2 score levels for each criterion was created in detail. It is not necessary to observe each of the criteria in the rubric for any question included in a data collection tool that will use RCR in its analysis. For a question in the data collection tool, it should first be determined which of the criteria in the rubric are observable. For example, if a question allows four of the criteria to be followed, the maximum score that can be obtained from that question will be eight.

Obtaining Expert Opinion on the Rubric

An expert opinion form was used to consult the opinions of the experts regarding the RCR draft form. For an expert opinion, three academicians (one professor, two associate professors) who are experts in the field of mathematics education were asked for their opinions. Attention has been paid to the fact that the experts are experienced in the fields of mathematical competences and reasoning, have studied in related fields or taught courses in the relevant fields. Two of the experts conducted a graduate-level mathematical reasoning course, and all of the experts conducted a thesis on mathematical reasoning and mathematical literacy. They were asked to express their opinions and make suggestions in terms of the suitability of being able to be used without being tied to any mathematics subject, and the suitability of the way the criteria were handled at 0-1-2 score levels (appropriate, not suitable/suggestion). The expert opinion form was also open to expert suggestions, apart from these mentioned issues.

Performing the Validity and Reliability Analyzes of the Rubric

After expert opinions, RCR (Appendix 2) was revised, and its validity was ensured and made ready for reliability analysis. Two researchers used the rubric developed independently of each other in the analysis of the data obtained from the data collection tool. The findings of the analyzes are given in the findings section.

Data Collection Tool

As a data collection tool, a Reasoning Competence Test (RCT) consisting of two open-ended questions, in which the RCR will be used in the analysis, was used. In the resource presented as a broadened perspective for the PISA 2021 Mathematics field, there are sample questions that are stated to be suitable for measuring reasoning skills (PISA, 2021b). Among these sample questions, two questions in Appendix 1 were chosen because they have different content areas and different real-life contexts. The content categories and real-life contexts of the questions are given in Table 1.

Table 1

Content Categories and Real-Life Contexts of the Questions in RCT

Questions	Mathematical Content Categories	Real-Life Context Categories
Multiplication	Quantity	Personal
Tree Leaves	Space and Shape	Scientific

The naming of the questions belongs to the researchers. Sub-questions items were added to the questions by taking the opinion of a mathematical education expert so that the questions were aimed to gain depth, and two open-ended questions became seven questions together with their sub-items.

Data Collection

At all stages of the research, data were collected online due to the COVID-19 Pandemic period. The study group of the research consists of 30 mathematics teachers (for 11 to 14 years old pupils) working in schools under the Ministry of National Education in different cities. It was stated to the teachers that participation in the research was not compulsory, it was on a voluntary basis, and their names would be kept confidential. Demographic characteristics of the teachers participating in the study, such as gender, professional experience and education level, are given in Table 2 below.

Table 2

Demographic Characteristics of the Study Group

	Professional Experience			Education Level		
	1-5 year	6-10 year	11 year and more	Bachelor's	Master's	Doctorate
Female	20	1	2	19	4	-
Male	2	5	-	6	-	1

RCT was represented to the participating teachers online, and their feedback was received online. For this reason, uncontrollable variables such as the inability to control the environment in which the questions were solved may have had an impact on the findings, since the researcher could not actually be with them while the teachers were solving the RCR.

Analysis of Data

In the first stage of data analysis, it is necessary to decide which of the two questions in RCT allows monitoring which of the criteria in RCR. For this, the percentage of agreement between the choices made by the researchers on the basis of the criteria was examined. The percentage of agreement was calculated using the reliability formula proposed by Miles and Huberman (1994) in order to ensure the reliability of the research. This formula is as follows:

[Reliability = Agreement / (Agreement + Disagreement) x 100]

Secondly, after the criteria had been determined, the collected data were analyzed independently by the researchers using RCR. Sample analyzes are included in Appendix 3. In the literature, it is seen that many techniques such as inter-rater correlation coefficient, Cohen's kappa coefficient, Krippendorff's alpha coefficient, ANOVA based on the difference between the scorers are used to determine the reliability of agreement between raters. Cohen's kappa coefficient was developed to determine the degree of agreement between two raters, and Krippendorff's alpha coefficient was developed to determine the degree of agreement between two or more raters (Cohen, 1960; Krippendorff, 1995).

In this study, "Cohen's kappa coefficient" and "Krippendorff's alpha coefficient" were calculated to determine the reliability of agreement between raters since there were two raters. Since both analysis methods were suitable for this study, both were applied to increase reliability. SPSS for Windows 22.0 program was used to calculate Cohen's kappa coefficient. In order to calculate Krippendorff's alpha coefficient, the data obtained were uploaded to the website prepared by Freelon (2010) to calculate this coefficient and the results were tabulated. In this study, one of the methods used in the examination of the reliability of the Reasoning Competence Rubric (RCR) is Krippendorff's alpha coefficient. Having a sample size of at least 30 is a sufficient size for the Kripendorff alpha fit coefficient to accurately estimate the parameter (Kanik et al., 2010).

Ethical Approval

This study ethics committee's approval was received with the session date and number 27.07.2021/2020-05 from Bursa Uludağ University Research and Publication Ethics Committees, Social and Human Sciences Research and Publication Ethics Committee. Scientific, ethical and citation rules were followed during the writing process of this study. No falsification was made on the collected data.

Findings

The findings related to the RCR, which was developed to assess mathematical reasoning, are presented under two headings: findings related to validity and findings related to reliability.

Findings Regarding the Validity of RCR

Expert opinions were sought in order to evaluate whether the rubric developed to assess the mathematical reasoning proficiency is a valid assessment tool. In this direction, the opinions of the experts on the use of the criteria in the rubric without depending on sampling and any mathematics subject, and the appropriateness of the handling of the scoring levels in the rubric were taken. All of the experts stated that the criteria in the rubric could be used regardless of sampling and any mathematics subject. Two of the experts suggested that the handling of the scoring levels in the evaluation table was appropriate, and the third expert suggested verbal changes in some of the level 1 expression. Changes were made in one-on-one interviews with the expert. Thus, it was observed that the opinions of the experts as appropriate/not suitable were mostly similar. As the criteria were taken from the reasoning section of the PISA 2021 Mathematics Framework, it was not questioned whether they served the scope of mathematical reasoning.

Findings Regarding the Reliability of RCR

In order for RCR can be used in the data analysis of research, it is first necessary to determine which of the criteria in the RCR allow the monitoring of the questions in the data collection tool. For this reason, the reliability analyzes of the RCR, which was developed to assess the mathematical reasoning competence, started from here. The researchers independently decided which of the criteria in RCT allowed the questions in RCR to be followed. As a result, the percentage of agreement (compliance) between the choices made by the researchers on the basis of criteria was calculated as %87.5.

The criteria agreed between the researchers and allow the questions in RCT to be watched in RCR are given in Table 3.

Table 3

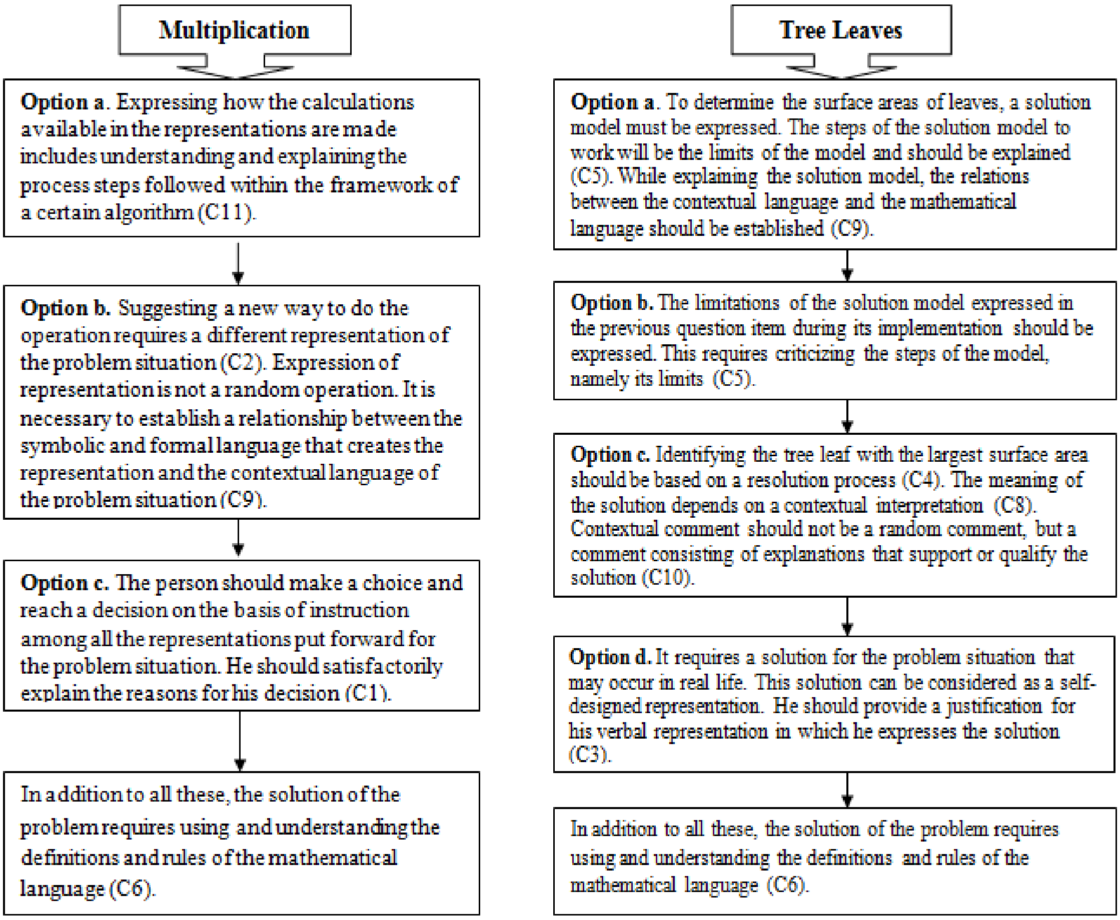
Criteria by which the Questions Allowed to be Tracked

Question	Option a	Option b	Option c	Option d	Throughout the question
Multiplication	C11	C2-C9	C1	-	C6
Tree Leaves	C5-C9	C5	C4-C8-C10	C3	C6

The rationale for the selection of criteria is shown in Figure 5.

Figure 5

The Rationale for Selection of Criteria



The researchers scored the data collected using RCT independently of each other according to the criteria they agreed on in Table 3. “Cohen’s kappa coefficient” and “Krippendorff’s alpha coefficient” were calculated to determine the reliability of agreement between raters. The values obtained as a result of the analysis are shown in Table 4 on the basis of criteria.

Table 4

Cohen’s Kappa Coefficient and Krippendorff’s Alpha Coefficient for the Reliability of Agreement between Raters of the RCR

Criteria	Cohen's Kappa Coefficient	Krippendorff’s Alpha Coefficient
C1	.886	.888
C2	.634	.705
C3	.774	.778
C4	.774	.778
C5	.750	.758
C6	.688	.686
C7	-	-

C8	.774	.774
C9	.750	.753
C10	.750	.750
C11	.837	.839
C12	-	-

When Table 4 is examined, Cohen's kappa coefficient values calculated to determine the reliability of agreement between raters are between .634 and .886. While the lowest agreement was obtained for the criterion coded C2, the highest agreement was obtained for the criterion coded C1. When Table 4 is examined, Krippendorff's alpha coefficient values calculated to determine the reliability of agreement between raters are between .686 and .888. The lowest agreement was obtained for the criterion coded C6, while the highest agreement was obtained for the criterion coded C1.

Discussion and Suggestion

In this study, it is aimed to develop a valid and reliable rubric to assess mathematical reasoning competence. Expert opinions were taken to ensure the validity of the developed RCR. The fact that the opinions expressed by the experts were mostly the same showed that the RCR was structurally valid. The fact that the criteria in RCR were taken from the sections of PISA resources related to mathematical reasoning without any changes is seen as sufficient evidence for the validity of RCR in scope.

Considering the findings regarding its reliability, the percentage of agreement between the choices made by the researchers on the basis of criteria regarding which of the criteria in the RCR allows monitoring of the questions in RCT is 87.5%. According to Miles and Huberman (1994), the percentage of agreement must be higher than 70% in order for the inter-rater evaluation results to be considered reliable. Considering the findings of the analysis, it can be said that the selections of the researchers on the basis of the criteria are consistent.

Cohen's kappa coefficient values, which were calculated to determine the reliability of agreement between raters in the data analyzed by the researchers, ranged between .634 and .886. Kappa values between .61 and .80 were considered significant, and values between .81 and 1 were accepted as an indicator of a very high level of agreement (Landis & Koch, 1977). Calculated Krippendorff's alpha coefficient values ranged between .686 and .888. According to Krippendorff (1995), an alpha coefficient greater than .60 indicates that the assessment tool is reliable. Bıkmaz-Bilgen and Doğan (2017) revealed in the reliability analysis performed by calculating Cohen's kappa coefficient and Krippendorff's alpha coefficient that the highest reliability values were obtained when there were two raters, and the reliability gradually decreased as the number of raters increased. In this study, as a result of the analyzes carried out on two raters and the findings obtained, it was decided that the RCR is a reliable tool.

As a result, a rubric was developed for the assessment of mathematical reasoning competence in this study. When the studies on the assessment of mathematical reasoning are examined; it is thought that the rubric developed by Loong et al. (2018) is suitable for use in the data analysis part of a research, and the scale developed by Çoban and Tezci (2020) is suitable for use in the data collection part of a research. Considering that concepts such as rubric and scale can be used with the same meanings (Goodrich-Andrade, 1997), there is no harm in comparing related studies. So, considering the use of RCR in research, which was created in line with the purpose of the current study, while it is similar to the rubric developed by Loong et al. (2018), it differs from the scale developed by Çoban and Tezci (2020). Bal-İncebacak and Ersoy (2016) used the reasoning stages introduced by TIMSS 2003 while evaluating 7th grade students' reasoning skills. The criteria in the RCR created as a result of the current study, on the other hand, include the mathematical reasoning actions set forth by the PISA 2021 framework. Therefore, it can be said that since PISA and TIMSS are international student assessment exams, the two studies show similarities in terms of analysis. The fact that RCR is a tool to assess the mathematical reasoning competence of the individual suggests that it will be useful in eliminating the difficulties in assessing the reasoning put forward by Herbert (2019).

In order for the use of RCR to become more functional in a study, it is recommended that researchers who will use RCR first determine the criteria in the RCR that they want to follow and then choose a question appropriate to those criteria. Since the questions in the RCT used in this study do not contain the criteria C7 and C12, the aforementioned criteria could not be analyzed in this research. Studies on the relevant criteria are also continuing.

Statement of Responsibility

The authors contributed equally to the related research. Therefore, each author is equally responsible.

Conflicts of Interest

The authors declare that they have no conflict of interest.

Author Bios

Zeynep Özaydın, graduated from Bursa Uludağ University, Department of Elementary Mathematics Teaching in 2019. She started his master's education at Bursa Uludağ University in 2019 and is currently in the thesis period. In April 2020, she started to work as a scholar in the TUBITAK project "Increasing the Level of Mathematics Literacy with the Dual-Focused Teaching Model" and continues to work in this position. She has studies on mathematical reasoning and mathematical literacy.

After graduating from Akdeniz University, Faculty of Science, Department of Mathematics in 1996, Assoc. Prof. Çiğdem Arslan worked as a teacher at various levels in the Ministry of National Education until 2001. She completed his master's and doctorate education in Uludağ University Faculty of Education in 2007, which she started as a research assistant in 2001. She worked at Istanbul University Hasan Ali Yücel Faculty of Education between 2008-2020. She has been working as an associate professor at Bursa Uludağ University since 2020. Her main areas of study include mathematical reasoning, problem solving, teaching geometry, mathematics textbooks, teacher education and mathematical literacy.

References

Altıparmak, K., & Öziş, T. (2005). An investigation upon mathematical proof and development of mathematical reasoning. *Ege Journal of Education, 6*(1), 25-37.

Altun, M. (2020). *Matematik okuryazarlığı el kitabı.* Alfa Aktüel Akademi Yayıncılık.

Australian Curriculum & Assessment Authority [ACARA]. (2017). Australian Curriculum: Mathematics. Retrieved from https://www.australiancurriculum.edu.au/f-10-curriculum/mathematics/rationale/

Bal-İncebacak, B., & Ersoy, E. (2016). Analysis of mathematical reasoning ability of the grade 7 students according to TIMSS. *Journal of International Social Research*, *9*(46), 474-481.

Ball, D. L., & Bass, H. (2003). Making mathematics reasonable in school. J. Kilpatrick, W. G. Martin., & D. Schifter. (Eds.), *A research companion to principles and standards for school mathematics* (pp 227-236). National Council of Teachers of Mathematics.

Bıkmaz-Bilgen, Ö., & Doğan, N. (2017). The comparison of interrater reliability estimating techniques. *Journal of Measurement and Evaluation in Education and Psychology, 8*(1), 63-78. https://doi.org/10.21031/epod.294847

Brualdi-Timmins, A. C. (1998). Implementing performance assessment in the classroom. *Practical Assessment, Research, and Evaluation, 6*(2). https://doi.org/10.7275/kgwx-6q70

Bülbül, B. Ö. (2019). How to create and use rubrics? In S. Çepni & Y. Kara (Eds.), *Current studies in educational measurement evaluation* (pp. 257-268). Paradigma Akademi.

Bülbül, S., & Bülbül, B. Ö. (2021). Assessment of secondary school textbooks: A rubric development study. *YYU Journal of Education Faculty*, *18*(1), 882-911. https://doi.org/10.33711/yyuefd.938727

Çoban, H., & Tezci, E. (2020). Developing the mathematical reasoning skills assessment scale. *OPUS International Journal of Society Researches*, *15*(24), 2805-2837. https://doi.org/10.26466/opus.610197

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37-46. https://doi.org/10.1177/001316446002000104

Dabic-Boricic, M., Vulic, I., & Videnovic, M. (2020). Mathematical literacy and assessment: Differences between the PISA study paradigm and mathematics teachers' conceptions. *Malta Review of Educational Research*, *14*(S), 101-121.

Demir, G., & Vural, R. A. (2016). The examination of teacher opinions on mathematical competence and skills of secondary mathematics curriculum. *Adnan Menderes University, Journal of Institute of Social Sciences, 4*(1), 118-139. https://doi.org/10.30803/adusobed.309074

Doyle, K. (2007). The teacher, the tasks: Their role in students' mathematical literacy. In J. Watson & K. Beswick (Eds.), *Mathematics: Essential research, essential practice. Vol. 1-2. Proceedings of the 30th annual conference of the Mathematics Education Research Group of Australasia* (pp. 246-254). MERGA.

Erdem, E. (2015). *The effect of enriched learning environment on mathematical reasoning and attitude* [Unpublished doctoral dissertation]. Atatürk University.

Ersanlı, E., Gencoglu, C., & Duran, V. (2018). A meta-synthesis study on investigations of dissertations on education regarding reasoning skills in turkey. *International Online Journal of Educational Sciences, 10*(5), 222-243. https://doi.org/10.15345/iojes.2018.05.015

Freelon, D. G. (2010). ReCal: Intercoder reliability calculation as a web service. *International Journal of Internet Science*, *5*(1), 20-33.

Gatabi, A. R., Stacey, K., & Gooya, Z. (2012). Investigating grade nine textbook problems for characteristics related to mathematical literacy. *Mathematics Education Research Journal*, *24*(4), 403-421. https://doi.org/10.1007/s13394-012-0052-5

Goodrich-Andrade, H. (1997). Understanding rubrics. *Educational Leadership, 54*(4), 14-17.

Güler, N. (2019). *Eğitimde ölçme ve değerlendirme* (14. baskı). Pegem Akademi. https://doi.org/10.14527/9786053641247

Gür, B. S., Çelik, Z., & Özoğlu, M. (2012). Policy options for turkey: A critique of the interpretation and utilization of PISA results in Turkey. *Journal of Education Policy, 27*(1), 1-21. https://doi.org/10.1080/02680939.2011.595509

Güzel, S. (2017). *The evaluation and development of maths lesson teaching program at 6th grades from the proficiencies of mathematical literacy* [Unpublished master's thesis]. Uludağ University.

Herbert, S. (2019). Challenges in assessing mathematical reasoning. *Mathematics Education Research Group of Australasia*, 348-355. https://files.eric.ed.gov/fulltext/ED604186.pdf

Herman, T. (2018). Analysis of students' mathematical reasoning. *In Journal of Physics: Conference Series* (Vol. 948, No. 1). IOP Publishing. https://doi.org/10.1088/1742-6596/948/1/012036

İskenderoğlu, T., & Baki, A. (2011). Classification of the questions in an 8th grade mathematics textbook with respect to the competency levels of PISA. *Education and Science*, *36*(161), 287-301.

Kanik, E. A., Temel, G. O., & Kaya, I. E. (2010). Effect of sample size, the number of raters and the category levels of diagnostic test on Krippendorff alpha and the fleiss kappa statistics for calculating interrater agreement: A simulation study. *Türkiye Klinikleri Biyoistatistik Journal of Biostatistics*, *2*(2), 74-81.

Kilpatrick, J., & Swafford, J., & Findell, B. (Eds.). (2002). *Adding it up: Helping children learn mathematics.* National Academy Press.

Krippendorff, K. (1995). On the reliability of unitizing continuous data. *Sociological Methodology*, *25,* 47-76. https://doi.org/10.2307/271061

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159-174. https://doi.org/10.2307/2529310

Leighton, J. P. (2003). Defining and describing reasoning. In J. P. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning* (pp. 3-11). Cambridge.

Loong, E., Vale, C., Widjaja, W., Herbert, S., Bragg, L. A., & Davidson, A. (2018). Developing a rubric for assessing mathematical reasoning: A design-based research study in primary classrooms. Hunter, J., Perger, P., & Darragh, L. (Eds.), *Making waves, opening spaces* (pp. 503-510). MERGA.

Matteson, S. M. (2006). Mathematical literacy and standardized mathematical assessments. *Reading psychology*, *27*(2-3), 205-233. https://doi.org/10.1080/02702710600642491

McGatha, M. B., & Darcy, P. (2010). Rubrics at play. *Mathematics Teaching in the Middle School*, *15*(6), 328-336.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. SAGE Publications.

Ministry of National Education [MoNE]. (2013). *Secondary school mathematics curriculum (5th, 6th, 7th and 8th grades)*. Ministry of National Education.

Ministry of National Education [MoNE]. (2018). *School mathematics curriculum (Primary and Secondary School Grades 1, 2, 3, 4, 5, 6, 7 and 8)*. Ministry of National Education.

Moskal, B. M. (2000). Scoring rubrics: What, when and how? *Practical Assessment Research and Evaluation,7*(3). https://doi.org/10.7275/a5vq-7q66

Moskal, B. M., & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment, Research & Evaluation, 7*(10), 1-6. https://doi.org/10.7275/q7rm-gg74

New Jersey Mathematics Coalition and the New Jersey Department of Education [NJMCF]. (1996). *New Jersey mathematics curriculum framework: The first four standards, standard 4- reasoning, K-12 overview.* State of New Jersey Department of Education.

Niss, M. (2003). Mathematical competencies and the learning of mathematics: The Danish KOM project. In *3rd Mediterranean conference on mathematical education* (pp. 115-124).

Niss, M., & Højgaard, T. (2019). Mathematical competencies revisited. *Educational Studies in Mathematics, 102*(1), 9-28. https://doi.org/10.1007/s10649-019-09903-9

Organization for Economic Co-operation and Development [OECD]. (2003). *The PISA 2003 assessment framework – mathematics, reading, science and problem solving knowledge and skills.* OECD Publishing.

Organization for Economic Co-operation and Development [OECD]. (2007). *PISA 2006: Science competencies for tomorrow's world.* OECD Publishing.

Organization for Economic Co-operation and Development [OECD]. (2010). *PISA 2009 Results: What students know and can do: Student performance in reading, mathematics and science.* OECD Publishing.

Organization for Economic Co-operation and Development [OECD]. (2013). *PISA 2012 assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy.* OECD Publishing.

Organization for Economic Co-operation and Development [OECD]. (2017). *PISA 2015 assessment and analytical framework: Science, reading, mathematic, financial literacy and collaborative problem solving.* OECD Publishing.

Organization for Economic Co-operation and Development [OECD]. (2019). *PISA 2018 assessment and analytical framework.* OECD Publishing.

Özgen, K. (2021). Checklist on question design for mathematical literacy. *Pegem Journal of Education and Instruction, 11*(1), 259-298. https://doi.org/10.14527/pegegog.2021.008

Peresini, D., & Webb, N. (1999). Analyzing mathematical reasoning in students' responses across multiple performance assessment tasks. In L. V. Stiff and F. R. Curcio (Ed.), *Developing mathematical reasoning in grades K-12* (pp. 156-175). National Council of Teachers of Mathematics.

Program for International Students Assessment [PISA]. (2021a). PISA 2021 mathematics framework draft. https://pisa2022-maths.oecd.org/files/PISA%202022%20Mathematics%20Framework%20Draft.pdf

Program for International Students Assessment [PISA]. (2021b). PISA 2021 mathematics: A broadened perspective. https://one.oecd.org/document/EDU/PISA/GB(2017)17/en/pdf

Russell, M. K., & Airasian, P. W. (2001). *Classroom assessment: Concepts and applications.* McGraw-Hill.

Russell, S. J. (1999). Mathematical reasoning in the middle grades. In L. V. Stiff & F. R. Curcio (Ed.), *Developing mathematical reasoning in grades K-12* (pp. 1-12). National Council of Teachers of Mathematics.

Sáenz, C. (2009). The role of contextual, conceptual and procedural knowledge in activating mathematical competencies (PISA). *Educational Studies in Mathematics, 71*(2), 123-143.

Steen, L. A. (1999). Twenty questions about mathematical reasoning. In Stiff L. (Ed.), *Developing mathematical reasoning in grades K-12* (pp. 270-285). National Council of Teachers of Mathematics.

Steen, L. A., Turner, R., & Burkhardt, H. (2007). Developing mathematical literacy. Blum, W., Galbraith, P. L., Henn, H. W., & Niss, M. (Ed.), *Modelling and applications in mathematics education* (pp. 285-294). Springer.

Toulmin, S., Rieke, R., & Janik, A. (Ed.). (1984). *An introduction to reasoning: Second edition.* Collier Macmillan Publishers.

Turner, R. (2010). Exploring mathematical competencies. *Research Developments, 24*(24), 5.

Yackel, E., & Hanna, G. (2003). Reasoning and proof. In J. Kilpatrick, G. Martin, & D. Schifter (Ed.), *A research companion to principles and standards for school mathematics* (pp. 227-236). National Council of Teachers of Mathematics.

Yankelewitz, D. (2009). *The development of mathematical reasoning in elementary school students' exploration of fraction ideas*. Rutgers The State University of New Jersey-New Brunswick.

Yıldız, H. (2019). *Investigation of the challenges of the seventh grade students encountered the solution of mathematical literacy questions* [Unpublished master's thesis]. Bursa Uludağ University.

Appendices

Appendix 1

1. MULTIPLICATION

A group of children are learning how to multiply 2-digit numbers in different ways. In the examples below they are calculating 47 x 36.

The ways in which students represent this process in different ways are shown below.

Representation 1

X	40	7	
6	240	42	282
30	1200	210	1410
			1692

Representation 2

Th	H	T	O
		4	7
	X	3	6
		4	2
	2	4	
	2	1	
1	2		
1	6	9	2

a) Explain how the calculation is done in Representation 1 and Representation 2.
b) Suggest a different way for this group of students to do the operation. Justify your suggestion. (If you have more than one suggestion, you can specify).
c) Which computational way do you think is more instructive, including the way you propose? Base your decision on a logical reason.

2. TREE LEAVES

OAK HORNBEAM MAPLE

a) The drawings of the leaves of the three trees are shown above. How do you measure the surface area of leaves? Your measurement should be based on a mathematical reasoning.
b) Does this measurement method create any problems or limitations? Explain your reasoning causally.
c) As the surface area of tree leaves increases, the amount of photosynthesis increases. Which of the leaves illustrated above can be said to produce more oxygen, ignoring other biological features? Your answer should be based on a mathematical solution process.
d) These are just particular leaves from the three tree kinds of trees. The actual leaves on a tree vary somewhat in size from one to the next. Suppose you had a sample of, says, 100 leaves from each tree. How would you estimate the average surface area of leaves of each type? Your solution process should be based on a mathematical reasoning.

Appendix 2

MATHEMATICAL REASONING CRITERIA'S		0	1	2
C1	Explains why a mathematical result or conclusion does, or does not, make sense given the context of a problem.	The result/conclusion is not available or is available but the reasons are not explained.	The reasons for the result/ conclusion are not adequately explained/incomplete.	The reasons for the result/ conclusion are adequately explained.
C2	Represents a problem in a different way, including organising it according to mathematical concepts and making appropriate assumptions.	Representation is not available.	Representation is available but not organized according to mathematical concepts/appropriate assumptions are not made.	There is an accurate representation that is organized according to mathematical concepts/made appropriate assumptions.
C3	Provide, explain, or defend a justification for the identified or devised representation of a real-world situation.	No justification is provided for the devised representation or there is no devised representation.	A representation is devised but not justified.	Justification provided/explained or defended for defined or devised representation.
C4	Provide, explain or justify a mathematical result or solution for the processes used to determine a mathematical result or solution.	The mathematical solution process is not determined.	The mathematical solution process is determined, but the rationale for the process is not provided.	The mathematical solution process is determined, the rationale is provided/explained or defended for the process.
C5	Identify or criticize the limits of the model used to solve a problem.	The limits of the model used to solve a problem have not been identified/criticised.	The limits of the model used to solve the problem have been identified, but the limits are inaccurate/missing or the criticisms are not sufficient for the limits.	The limits of the model used to solve the problem have been correctly identified/criticised adequately.
C6	Utilize or understand definitions, rules, formal systems, as well as using employing algorithms, computational thinking and reasoning.	Definitions, rules, and formal systems have been ignored or misused.	Definitions, rules and formal systems have been taken into account, but incorrectly/incompletely put to work.	Definitions, rules and formal systems have been taken into account and correctly/completely put to work.
C7	Reflects on mathematical arguments, explaining and justifying the mathematical result.	The mathematical argument is not reflected.	The mathematical argument has been reflected, but the mathematical result is not explained/justified in line with the mathematical argument.	The mathematical argument has been reflected and the mathematical result is explained/justified in line with the mathematical argument.
C8	Interprets a mathematical result back into the real-world context in order to explain the meaning of the results.	No interpretation is available in a real-world context.	An interpretation is available in a real-world context but the interpretation does not depend on the mathematical result or is not correct.	An interpretation in a real-world context depends on the mathematical result and is correct.
C9	Explains the relationships between the context-specific language of a problem and the symbolic and formal language needed to represent it mathematically.	The relationship between the context-specific language and the symbolic and formal language that constitutes the representation has not been established or explained.	The relationships between the context-specific language and the symbolic and formal language that constitutes the representation are incorrectly/incompletely established or explained incorrectly/incompletely.	The relations between the context-specific language and the symbolic and formal language that constitutes the representation have been correctly established or explained correctly.
C10	Reflects on mathematical solutions and creates explanations and arguments that support, refute or qualify a mathematical solution to a contextualised problem.	A mathematical solution is not reflected for contextualised problem.	A mathematical solution has been reflected, but no explanation/argument has been created that supports, refutes or qualify a solution to a contextualised problem.	A mathematical solution is reflected, and an explanation/argument is created that supports, refutes, or qualify a solution to a contextualised problem.
C11	Explains how a simple algorithm works and to detects and corrects errors in algorithms and programs.	How the algorithm works is not explained or errors in the algorithm/program are not detected.	How the algorithm works is explained but explanations are missing, or errors in the algorithm/program have been detected but corrections is not explained.	How the algorithm works is explained correctly/fully, or errors in the algorithm/program are detected and corrections are explained.
C12	Analyses similarities and differences between a computational model and the mathematical problem that it is modelling.	The similarities/differences between a computational model and the mathematical problem that it is modelling have not been analyzed.	The similarities/differences between a computational model and the mathematical problem that it is modelling have not been adequately analyzed.	The similarities/differences between a computational model and the mathematical problem that it is modelling have been adequately analyzed.

Appendix 3

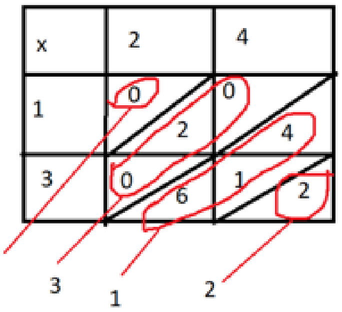
Example answers given to item b of the "multiplication" question

The answer given to item b of the "Multiplication" question must be an answer that meets the criteria coded K2 and K9.

**Example of an answer 1: [6 x (40 + 7)] + [30 x (40 + 7)] Multiplication without making a table can be done by using the distributive property over the addition operation.*

The teacher's answer of "[6 x (40 + 7)] + [30 x (40 + 7)]" can be considered as a representation, but this representation is an incomplete representation in terms of difference when compared to the existing representations in the question considering the operations it contains. The teacher gets 1 point from the K2 coded criterion. While obtaining the representation, the relations between the contextual language and the symbolic language were established correctly and explained as "…it can be done by using the distributive property of multiplication over addition". The teacher gets 2 points from the K9 coded criterion.

**Example of an other answer:*



In this method, after substituting the numbers in the table, we divide the empty squares between them as in the figure and multiply the numbers. If the resulting number is a single digit, we write 0 in the upper section. If it has two digits, we write the ones digit at the bottom with the tens digit at the top. We collect the results as I marked in red. If the addition operation is handed, we add this to the next sum. In this way, our process is finished.

The path suggested by the teacher-representation- is organized according to mathematical concepts and is different from the paths in the text of the question. The teacher gets 2 points from the K2 coded criterion. The verbal explanations he makes after the representation are the explanations in which the relations between mathematical language and contextual language are established correctly. The teacher gets 2 points from the K9 coded criterion.

Example answers given to item c of the "multiplication" question

The answer given to item c of the "Multiplication" question must be an answer that meets the criterion coded K1.

**Example of an answer 1: Representation - 1, Representation - 2, and when I think about the way I propose, I find Representation - 2 more useful and more instructive. In Representation – 2, numbers are placed in the digit table according to the number of digits. In the digit table, operations are performed according to the place value of each number. With this method, students can see where the phrase "shift one step to the left", which has been said in schools for years, comes from while the subject of multiplication is being taught. In this way, instead of going to memorization by taking what is given by the way of presentation, as in teaching, and conditioning himself by saying "I must not forget to move one step"; He finds and sees where this rule comes from. Since it learns based on logical reasons, the knowledge becomes more permanent, the connection with new knowledge can be established more easily in the future, and the learning situation can be taken to a higher level.*

The teacher made a decision by choosing representation 2 in the question item and explained the reasons for making this decision by saying, "The student finds and sees where this rule comes from instead of conditioning himself so that he should not forget to move one step. For logical reasons… it can take it to a higher level." explained in terms. The reasons explained by the teacher for his decision are sufficient. The teacher gets 2 points from the K1 coded criterion.

**Example of an other answer 2: I think Representation 2 is more instructive. We can multiply any two-digit number.*

The teacher made a decision by choosing representation 2. However, he did not adequately explain the reasons for his decision. We can also multiply any two-digit number by the representation 1. The reasons for his decision are not enough. The teacher gets 1 point from the K1 coded criterion.

Sample answers to item c of the question "tree leaves"

The answer given to item c of the "Tree Leaves" question must be an answer that meets the criteria coded K4, K8 and K10.

**Example of an answer: In advanced classes, integral curves can be calculated with parabolic curves.*

The teacher talked about the mathematical solution process without reaching a mathematical conclusion, but did not justify the process. It gets 1 point from the K4 coded criterion. Since there is no mathematical solution in the answer, there are no explanations supporting the solution (K10) and no contextual interpretation (K8) explaining the meaning of the result because there is no result. The teacher could not get points from the K8 and K10 coded criteria. If the teacher had made qualitative or supportive explanations for the solution he thought (K10) and then made a contextual comment as "… the leaf produces more oxygen" after the solution he put forward (K8), he could get full points from the relevant criteria.

**Example of an other answer: I took the printout of the models as they were in the research pdf, drew squares of the same size in each leaf model and combined the leaf parts outside the square and placed them in another model that included all 3 leaves as a control model. When I place the square models that I placed in the models and the remaining parts into the control model and investigate, I think that the hornbeam leaf has more area and produces more oxygen.*

The teacher talked about the mathematical solution process in detail. The fact that the teacher will reach the amount of oxygen production through field knowledge can be accepted as a justification for the solution process. The teacher gets 2 points from the K4 coded criterion. The teacher's comment as "Horbee produces more oxygen" is a real-world interpretation based on the mathematical result. The teacher gets 2 points from the K8 coded criterion. In addition, the teacher's statement "because the hornbeam leaf has more area ..." is an explanation for the contextual problem that characterizes its mathematical solution. The teacher gets 2 points from the K10 coded criterion.

Sample answers to item d of the question "tree leaves"

The answer given to item d of the "Tree Leaves" question must be an answer that meets the K3 coded criterion.

**Example of an answer: I would choose the largest and smallest leaves and average them.*

The teacher designed a verbal representation for the solution. However, he did not justify the representation he designed. The teacher gets 1 point from the K3 coded criterion. If the teacher had written an answer such as "I would choose the largest and smallest leaves to represent different leaf samples and take the average", he would have justified the solution process and received full marks from the question.

**Example Answer 2: I would find the square of the leaves on the assumption that many events in nature conform to the normal distribution. Because in a normal distribution mode=median=mean, I would calculate the area of the median leaf.*

The teacher designed a verbal representation for the solution of this problem that may exist in the real world. He provided a justification for the representation he designed as "many events in nature are normally distributed... since mode=median=mean in a normal distribution...". The teacher gets 2 points from the K3 coded criterion.