Education and Science tedmem

Vol 42 (2017) No 190 171-188

A New Classification Proposal for Mathematical Literacy Problems

Murat Altun [1], Işıl Bozkurt [2]

Abstract

The improvement of mathematical literacy level remains one of the fundamental problems of primary education in Turkey. In that case there is need to know students' difficulties at solving mathematical literacy problems. This study reveals the difficulties encountered by Turkish students when solving mathematical literacy problems through the factors formed based on the structural characteristics of the questions addressed. A total of 435 middle school students (8th grade) were asked mathematical literacy questions. The data obtained from student responses were subjected to factor analysis. A six-factor structure was obtained at the end of the analysis. The obtained factors were found to have adequate variance in explaining mathematical literacy. These factors were named as making algorithmic operations, mastering rich mathematical content, mathematical inference, developing mathematical proposals and interpreting a developed proposal, understanding the mathematical equivalence of real world situations, and understanding the counterpart of mathematical language in life. The students failed in the factors of mathematical inference, developing mathematical proposals and interpreting a developed proposal, and understanding the mathematical equivalence of real world situations. The results of this study can be guiding for analyzing mathematical literacy achievement, developing a mathematical literacy scale, developing mathematics curricula, and organizing its instruction.

Keywords

Mathematical literacy
Problem solving
Mathematical literacy achievement level
PISA
Factor analysis

Article Info

Received: 09.08.2016
Accepted: 03.20.2017
Online Published: 05.10.2017

DOI: 10.15390/EB.2017.6916

[1] Uludağ University, Faculty of Education, Division of Elementary Mathematics Education, Turkey, maltun@uludag.edu.tr
[2] Uludağ University, Faculty of Education, Division of Elementary Mathematics Education, Turkey, ibozkurt@uludag.edu.tr

Introduction

Eliminating the disconnection between school mathematics and everyday life remains one of the fundamental challenges of mathematics education. Many documents highlight the importance of working with actual data in primary school mathematics courses to eliminate such disconnection (OECD, 2014; National Council of Teachers of Mathematics [NCTM], 2000, 2006). This increases the importance of mathematical literacy, which refers to capacity to use mathematical knowledge acquired in school in everyday life, and makes development of mathematical literacy one of up-to-date goals. Rise in the importance of mathematical literacy has attracted attention to Programme for International Student Assessment (PISA), whose main theme is the assessment of literacy. As a result, many countries have started to take PISA as a reference when developing their own educational policies (Breakspear, 2012).

PISA is a test has been administered to students at the age of 15 once every three years since 2000 to measure the degree to which knowledge learned in primary education is reflected on everyday life and measures students' mathematical and scientific literacy as well as their language skills. The present study focuses on mathematical literacy. The aim of this study is to obtain the main components of mathematical literacy and reveal that mathematical literacy can be measured over such main components.

Mathematical Literacy

Organization for Economic Co-operation and Development (OECD, 1999, 2003, 2006, 2009a), which organizes PISA, defined *mathematical literacy* as "an individual's capacity to identify and understand the role that mathematics plays in the world, to make well-founded judgements and to use and engage with mathematics in ways that meet the needs of that individual's life as a constructive, concerned and reflective citizen. Making a partial change in this definition later on, OECD (2013, 2016) defined mathematical literacy as "an individual's capacity to formulate, employ and interpret mathematics in a variety of contexts". It also added the following explanation to this definition: "It includes reasoning mathematically and using mathematical concepts, procedures, facts and tools to describe, explain and predict phenomena. It assists individuals to recognise the role that mathematics plays in the world and to make the well-founded judgements and decisions needed by constructive, engaged and reflective citizens." (OECD, 2013, 2016). McCrone and Dossey (2007) summarize mathematical literacy as capacity to understand the role of mathematics in the daily life and use mathematics for solving the problems encountered in the daily life.

In this case, raising as equipped with mathematical literacy is a goal, and identifying mathematical literacy levels to organize works for reaching this goal is a need. Assessments for identifying mathematical literacy levels are conducted by allowing students to use their mathematical competencies through addressing them contextual, conceptual, and operational problems (Saenz, 2009). Classifications made after such assessments may offer great opportunities in the interpretation of results. The present study deals with a new classification that is different from the existing classifications and can be more useful with some features it holds.

The need focused on in this study has two sources in the related literature. The first one is about the classification (i.e. the classification of questions) that is taken as reference in determining the scope of mathematical literacy assessments and in assessing student achievement (OECD, 2016, 2013, 2010, 2009a; Saenz, 2009). The second one is about the variables collected through PISA student questionnaire and considered to have a possible partial influence on mathematical literacy (e.g. information about family and home, information about school and classroom, perception, attitude) (Eraslan, 2009; Forgasz & Leder, 2017; Özkan & Güvendir, 2014). As studies of this type are limited to the data outside the teaching content, they are likely to make only a limited contribution to explaining achievement.

The Classification of Mathematical Literacy Questions

In PISA, four different types of classification have been employed so far: subject areas, capabilities, competency clusters, and process skills. OECD (2013, pp. 33) determined four subject areas for PISA: (i) Quantity, (ii) Change and relationships, (iii) Space and shape, (iv) Uncertainty. The subject area of uncertainty was turned into uncertainty and data within the process (OECD, 2013, 2016). These titles cover all kinds of subjects contained in the primary school mathematics and do not allow ignoring any mathematical knowledge. This classification can be applied when there is a need to assess a student in any specific subject area as well. When this classification is used, the share of questions related to more than one area in assessing mathematical literacy in such specific area can be a matter of debate. A similar situation may also occur when identifying the area(s) in which a student participating in the assessment succeeds or fails.

Another classification intended to reveal the scope of mathematical literacy practices is the classification based on mathematical capabilities activated during problem solving. Capabilities are cognitive processes that should be activated when a situation that is presented in context is to be combined with mathematics to solve a problem (Saenz, 2009). These capabilities include mathematical thinking and reasoning; mathematical argumentation; modelling; using symbolic, formal and technical language and operations; and devising strategies for solving problems (Dossey, McCrone, Turner, & Lindquist, 2008; Saenz, 2009). When each of these capabilities is included in the practice, a difficulty may be encountered in determining the volume of any capability in the practice. Also, as capabilities may be activated at different levels in different problems, a confusion may occur in identifying the capabilities in which students have difficulty while solving the problems. For instance, capability in modelling is needed in solving a problem that requires expressing a situation with variables, whereas it might not be needed in a problem that requires inferences derived from a formula or in interpreting a formula (Dossey et al., 2008). Thus, at times it can be very difficult to evaluate capabilities. This being the case, it might be better to define mathematical literacy skill sets by combining questions with similar degrees of difficulty and then evaluating different competency clusters (OECD, 2013). These competency clusters are known as *reproduction, connections*, and *reflection* competencies. *Reproduction* competencies involve understanding mathematical processes and types of problems as well as performing routine operations. Routine problems that require direct use of an algorithm or a formula fall under this category. *Connections* competencies are required for problems that are non-routine and that ask students to interpret various situations and connect them with one another. Generally, problems requiring these competencies are of medium difficulty. *Reflection* competencies require identifying and using one’s knowledge and competencies for problems in which the mathematical knowledge and competencies needed to solve the problem are not directly and explicitly identified in the given situation. These ways of use require reorganization of the thought. Problems requiring reflective competencies are more difficult than other types of problems. (OECD, 2013, pp. 28). Moreover, according to Saenz (2009), that a problem requires reproduction, connections, and reflection competencies indicates the problem’s high degree of complexity, and the degree of complexity explains more than 50% of the variance in providing correct answers to questions. The classification of problems in this way, as indicated by Blum (personal communication, December 27, 2013), may be useful in creating an item pool, but this is a theoretical classification that does not explain why and where failure occurs.

Since PISA offers no classification system that meets all the relevant requirements, the OECD has continued its search. In PISA 2012, competency clusters lost their importance. Instead, the questions were classified by considering the points outstanding in mathematical processes (OECD, 2013). This classification is described as follows in OECD sources: (i) formulating situations mathematically, (ii) employing mathematical concepts, facts, procedures and reasoning, and (iii) interpreting, applying and evaluating the mathematical outcomes (OECD, 2013, pp. 28). The solution of any problem involves all these competencies, but the problems are classified based on the competency that is most prominent in their solution. However, in this classification also, hesitations come out in regard to the categories under which questions are to be put. Despite all the classifications summarized above, difficulty in thoroughly explaining mathematical literacy keeps the seek for classification on the agenda. The present study aims to fulfill this need by offering a classification that can be more useful.

The Reflections of Mathematical Literacy Assessments on Mathematics Curricula

Since 2000 when PISA tests on mathematical literacy were started, many studies have been conducted to compare different countries within the scope of PISA. Some have compared countries according to several variables and identified those variables that influence mathematical literacy. Most of these studies compare one country with other countries whose mathematical literacy is higher. For instance, Wood (2007) compared the USA and Finland in terms of self-regulation skills (belief, motivation, and learning strategies); to represent two different cultures**,** Ross (2008) compared the USA, Britain, and Canada with Japan, Korea, and Hong-Kong, respectively, in terms of motivation and academic achievement; Satıcı (2008) compared Turkey and Hong-Kong in terms of familial characteristics and learning environments; Lydia and Wilson (2009) compared the USA and Hong Kong in terms of mathematics achievement; Akyüz and Pala (2010) compared Turkey, Greece, and Finland in terms of learning environments; and Liang (2010) compared the USA, Canada, and Finland in terms of learning environments.

Some studies in the literature are about mathematical literacy level. Saenz (2009) classified mathematical literacy questions into contextual, procedural, and conceptual and investigated in which category Spanish pre-service teachers had difficulty. In that study, contextual questions were found more difficult than other questions, and it was concluded that PISA classification of reproduction, connections, and reflection is an indicator of the complexity of questions. İlbağı (2012) compared different types of schools and regional achievement levels by using open-access PISA questions and reported that increased social and economic welfare level makes a positive contribution to achievement. Altun and Akkaya (2014) tried to estimate the questions most difficult for Turkish students by directing questions to primary mathematics teachers and explored teachers' recommendations for how to eliminate these difficulties. The teachers mostly stated that students are unfamiliar with mathematical literacy questions, and these kinds of questions must be included in textbooks. Arıkan (2014) explained the factors that influence mathematics achievement through regression analysis by using the data of Turkish students who participated in the 2012 PISA. Beliefs and motivation levels regarding mathematics were found to have a stronger influence on achievement than other variables.

As understood from the summary above, previous studies have mostly been carried out based on the data obtained from OECD processes. They have been limited to the family and study environments of participating students and measurements via psychological tests. Studies confronting students with questions directly are not sufficient. This study differs from other studies in that mathematical literacy questions were addressed to students, and the answers they gave were subjected to factor analysis, thereby re-forming a set out of them. Factor analysis is an analysis that forms a set of the interrelated variables in a single dataset based on the correlations between them and presents them as factors. In this way, it becomes possible to interpret the distribution over fewer variables (as many as the number of factors) compared to the number of questions (Tabachnick & Fidell, 2013, pp. 612). Subjecting the questions to a new classification based on this analysis may offer many new opportunities. For example, this study is process-oriented. A process-oriented classification may be more useful than the existing classifications for planning mathematical literacy education and evaluating its results.

Purpose of the Study

The purpose of this study is to analyze the answers given by students to mathematical literacy questions and by factoring the questions with the motions from these answers, thereby determining the main components, which are fewer compared to the number of the questions, that will be used for explaining mathematical literacy. The object of curiosity here is whether or not the questions will have a new structure/classification depending on the answers the students will give to them. If they do, the components of this structure can be used for explaining mathematical literacy levels.

Method

A survey method was employed in this study. Survey models are approaches to research that describe a situation in the past or present as it was or is. The characteristics under analysis cannot be changed by any means (Creswell, 2014; Fraenkel, Wallen, & Hyun, 2011). Mathematical literacy problems were presented to eighth grade students without any instruction. Mathematical literacy problems were presented to eighth grade students without any instruction. We determined the students' percentages of correctly responding to the questions.

The analyses were made over the scores obtained by the students in each question. Thus, descriptive research model was used. Descriptive research aims to define, interpret, classify, and explain a situation examined in depth or identify the relationships between specific situations (Cohen, Manion, & Morrison, 2011, pp. 256; Glass & Hopkins, 1984).

Study Group

The study group includes 435 students of 10 teachers who attended a seminar (30 hours) within the context of a project[3] on selecting and writing questions to test mathematical literacy. Since the teachers were accepted for the seminar not based on any specific criteria but rather upon their application, there was no biasness in student selection. Based on the average mathematics scores obtained from the Transition from Primary to Secondary Education Exam (TEOG) conducted in Turkey for placing students in high schools, the ranking distribution of the schools where the teachers worked was as follows within the ranking across Bursa province: three schools ranked in the top 100 (i.e. 29, 31, 54); four schools ranked between 100[th] and 200[th] (i.e. 101, 146, 153, 200); and three schools ranked between 300[th] and 400[th] (i.e. 336, 361, 362). Given the fact that 476 schools participated in the above-mentioned exam from Bursa, it can be said that the study involved a sample that was relatively above the Bursa population.

Research Process

The questions (to be introduced below in the section on data collection tools) were given to the teachers to be administered to the students for one hour under their control at the beginning of the seminar, after which the responses were collected. The students received no instruction, and no intervention was made before the test. The data obtained through the administration of the mathematical literacy questions to 435 middle school eighth grade students were subjected to factor analysis, thereby making an attempt to identify the variables (main component) explaining achievement. Factor analysis is based on maximizing the variance of measurement projections (components) in a certain axis or certain axes through rotation of the axis system in the analytic system which explains the data (Cohen et al., 2011, pp. 678). In this case, total variance remains the same, and the projections in other axes shrink. The system disintegrates due to expanding variances, and it becomes easier to analyze it. Thus, the variables which are associated with each other form a set and turn into a single variable (factor). Hence, data group can be expressed through fewer variables (factors) compared to the number of variables used for measurement (Cohen et al., 2011).

Data Collection Tool

The data collection tool used in the study is the Mathematical Literacy Test (MLT). The MLT consists of 10 questions, 2 (Deputy, Paint) of which were prepared by the researchers and the remainder taken from PISA materials. Some of these questions include two or three sub-questions depending on the item roots. When each item is considered a separate question, the MLT includes 17 questions in total. Each question used in the study was named, and the questions are included in the text under these names. Five of these 17 questions are multiple-choice while the rest are open-ended. The questions written by the researchers (Appendix 1) were chosen from among many questions included in a project being carried out based on their quality of attracting attention during teacher training. One of the

[3] The project, "The Influence of PISA Mathematical Literacy Education Given to Mathematics Teachers on Student Achievement," was conducted through the participation of the Bursa Directorate of National Education and Uludag University, Project No: KUAP(E)-2015/26.

questions (Deputy) was about social life, and the other (Paint) was about personal life (Appendix 1). Other questions can be found in sources about PISA (Ministry of National Education of Turkey [MEB], 2012; OECD, 2013). Before the study, the questions were administered to 232 eighth grade students not included in the study in order to test clarity and appropriateness for the students' grade level. It was determined that the item test correlations of all questions were over 0.20, which is accepted as the critical value for significance (Tan, 1998). Moreover, most of the questions had been used in the PISA before, which indicates a high level of internal validity.

When selecting questions, the existing classifications about mathematical literacy in PISA were taken into consideration. In this regard, attention was paid to apply the weights close to those used in PISA practices whose main theme is to measure literacy level in terms of competency clusters, subject areas, mathematical processes, contexts, and item types (OECD, 2010, 2013, 2014, 2016). The distribution of the questions in terms of competency cluster, contents, processes, and contexts is given in Table 1.

Table 1. The Categories in Which the Questions are Classified

	Competency Clusters			**Mathematical Content**				**Mathematical Processes**			**Situations / Contexts**			
Questions	Reproduction	Connections	Reflection	Space and Shape	Change and Relationships	Uncertainty and Data	Quantity	Formulating	Employing	Interpreting	Scientific	Societal	Occupational	Personal
The Best Car 1	x						x		x			x		
The Best Car 2			x		x			x				x		
Deputy 1	x						x		x			x		
Deputy 2			x				x			x		x		
Heartbeat 1		x			x				x		x			
Heartbeat 2		x			x			x			x			
Test Scores		x				x				x			x	
Earthquake			x			x				x	x			
Paint		x					x			x				x
Skateboard1	x						x			x				x
Skateboard2	x						x		x					x
Skateboard3		x					x			x				x
Carpenter		x		x					x					x
Height 1	x					x			x				x	
Height 2			x			x				x			x	
Height 3			x			x			x				x	
Oil Spill			x	x					x		x			

Data Analysis

The answers given by 435 participating students were read within the framework of the rubrics published by PISA. Based on the levels specified in such rubrics, one of the following points was assigned to each answer: 0 (null or wrong), 1 (partly true), and 2 (true). As an example, the rubric of the Test Scores question (OECD, 2009b; MEB, 2011) is annexed (Annex 2). The answers were evaluated by the researchers separately. The answers graded differently were re-examined, and an agreement was reached on them. Then the student scores were subjected to factor analysis to determine the variables explaining achievement. Before conducting factor analysis, whether or not distributions were appropriate for parametric statistical methods was investigated. Therefore, mean, median, mode, kurtosis, and skewness values were calculated, and whether or not each group of data had normal distribution was explored. Moreover, a sphericity test was conducted for all data (Can, 2012; Pallant, 2001). Analysis was made as the data structure was found to be fit for factor analysis ($X^2_{(136)}$= 808.679; $p<0.05$) (Cohen et al., 2011). To test the stability of the factors, the sub-groups chosen from the data group were subjected to factor analysis again. After the factors were decided, the researchers went on with naming them. The question texts and the factor loadings in Table 4 were given to five academics from the field. They were asked to offer their recommendations about the factor names in written. The final versions of the names were determined in the joint session held later on. The academics taking part in the joint session had research in the field of mathematics education, had received the Advanced Quantitative Statistical Analyses course, and were knowledgeable of mathematical literacy.

Results

This section first presents descriptive statistics concerning the MLT administered to the students participating in the study. Second, findings concerning the factor analysis applied to the results obtained from the MOT are provided.

Table 2. Descriptive Statistics about MLT Questions

Questions	Mean	Standard Deviation
The Best Car 1	1,71	,709
Height 1	1,32	,949
Skateboard1	1,23	,898
Skateboard2	,92	,998
Earthquake	,92	,998
Skateboard3	,72	,962
Carpenter	,70	,836
Height 3	,61	,921
The Best Car 2	,49	,857
Test Scores	,38	,783
Deputy 1	,34	,608
Oil Spill	,28	,696
Paint	,27	,677
Heartbeat 1	,20	,598
Height 2	,17	,454
Deputy 2	,06	,319
Heartbeat 2	,06	,328

Table 2 shows the mean scores obtained by the students per question as well as relevant standard deviations. The mean scores vary between 0.06 and 1.71 out of 2. As is seen in Table 2, a level of achievement (correct answer) above 50% was obtained in three questions. The level of achievement was 50% to 25% in six questions and less than 25% in the remaining eight questions. It is noteworthy that level of achievement was too low in the Deputy2 and Heartbeat2 questions.

In factor analysis, it does not matter whether or not every variable in the dataset has a distribution that conforms to the normal (Tabachnick & Fidell, 2013, pp. 618). Because of the questions with a high level of difficulty, it may not be possible for the answers given to all questions to display a normal distribution. On the other hand, the solution is stronger when the covariables have a normal distribution (Tabachnick & Fidell, 2013, pp. 618). The closeness of the mode (0.625), median (0.625), and arithmetic mean (0.605) values calculated for the mean scores to each other and the fact that skewness (0.409) and kurtosis (0.130) coefficients were in the range of 0-1 (Morgan, Leech, Gloeckner, & Barret, 2004) showed the appropriateness of the data for normal distribution. The X^2 value calculated for the sphericity of the data ($X^2_{(136)}$=808.679; p<0.05) also indicated that the data were fit for factor analysis. While the item-test correlations calculated on the basis of the questions varied between 0.20 and 0.47, the Cronbach's alpha values again calculated on the basis of the questions ranged from 0.70 to 0.73. These statistical values indicate that each of the questions and the entire test are valid and reliable (Çokluk, Şekercioğlu, & Büyüköztürk, 2014).

At the end of the analysis, the Carpenter question, which was one of the 17 questions in the MLT, was accepted to be overlapping as it had a load value higher than the accepted level (0.32) in more than one factor and the difference(s) between such load values was(were) less than 0.10 (Table 4) (Tabachnick & Fidell, 2013), and so it was ignored when determining the factors (Çokluk et al., 2014). The results of the analysis performed with other questions yielded a six-factor structure as the calculated eigenvalues were higher than 1. It was determined that six factors explained 53.59% of the total variance. The factors and the amount of variance explained by them are given in Table 3.

Table 3. Total Variance Explained (Variance Values of the Factors)

Component	**Rotation Sums of Squared Loadings**		
	Total	**% of Variance**	**Cumulative %**
1	2,018	11,873	11,873
2	1,625	9,561	21,434
3	1,410	8,292	29,726
4	1,390	8,177	37,903
5	1,384	8,139	46,042
6	1,283	7,548	53,590

Factor load value refers to correlation value between the variable and the factor (Can, 2012). In order to consider a variable as the component of the relevant factor, its contribution to the factor is expected to be not less than 10%. Such a contribution can be obtained through a factor load of 0.33 or higher (0.33^2=0.10) (Tabachnick & Fidell, 2013; Tatlıdil, 2002). When the values not more than 0.30, which are considered to indicate weak factor loading, are deleted to make factor components more explicit, it is seen that mathematical literacy is explained by a total of six factors, and each factor has minimum two components (Table 4).

Table 4. Factors and Their Components

Questions	Factors*					
	1	2	3	4	5	6
Height 1	**,714**					
Skateboard1	**,624**					
Height 3	**,592**					
Skateboard3	**,491**	,312				
Test Scores		**,696**				
Deputy 1		**,601**				
Skateboard2		**,523**				
Carpenter		,425	,417			,417
Height 2			**,696**			
Oil Spill			**,680**			
Deputy 2				**,729**		
Heartbeat 1				**,680**	,312	
The Best Car 2	,303			**,549**		,311
Paint					**,721**	
Heartbeat 2					**,648**	
Earthquake					,302	**,761**
The Best Car 1	,342					**,525**

* The factor loads of the questions taken as factor component are bold.

The six factors obtained at the end of the analysis explain approximately 54% of the total achievement. This value conforms to the expectation that the variance should be between 40% and 60% in multi-factor designs (Tavşancıl, 2014). The analysis results (Table 4) revealed a six-factor structure for mathematical literacy. Except for one factor, all the factors have a value not below 0.50, which indicates strong components.

As mentioned under the title of "Data Analysis", the factors were named through an academic group work, during which the contents and the factor loads of the questions were taken into consideration. The content of the items with a bigger factor load and the difference between the item and the other components of the factor were influential on naming. The following recommendations were put forward as the factor names, respectively: "basic operation skill, operational competencies, making algorithmic operations" for the first factor; "making sense out of the text, reading comprehension, mastering mathematical content, mastering rich mathematical content" for the second factor; "mathematical inference, interpreting the mathematical knowledge in the text, expressing the interpretation mathematically" for the third factor; "mathematical modelling, developing a mathematical model, interpreting the mathematical model, developing mathematical proposals and interpreting a developed proposal" for the fourth factor; "making a change in the mathematical model, interpreting the mathematical result for life, understanding the mathematical equivalence of real world situations" for the fifth factor; and "understanding the counterpart of mathematical language in life" for the sixth factor. Upon the discussions on the names recommended, a consensus was reached on the following names: (i) Making algorithmic operations, (ii) mastering rich mathematical content, (iii) mathematical inference, (iv) developing mathematical proposals and/or interpreting a developed proposal, (v) understanding the mathematical equivalence of real world situations, (vi) understanding the counterpart of mathematical language in life.

According to the findings of the present study, the students' rates of achievement in these factors are as follows: 54% in the factor of making algorithmic operations, 29% in the factor of mastering rich mathematical content, 14% in the factor of mathematical inference, 13% in the factor of developing mathematical proposals and interpreting a developed proposal, 8% in the factor of understanding the mathematical equivalence of real world situations, and 66% in the factor of understanding the counterpart of mathematical language in life.

Discussion and Conclusion

This section discusses the study in three aspects: (1) characteristics of the statistical method and measurement tool used; (2) naming the factors; (3) relationships between the factors (main components) and competency clusters.

Characteristics of the Statistical Method and Measurement Tool Used

The literature contains studies about mathematical literacy that show similarity to the present study in terms of the method employed. For instance, Akyüz and Pala (2010) revealed the powers of variables that influence mathematical literacy in Turkey, Greece, and Finland (family and classroom environment) through factor analysis, whereas Arıkan (2014) presented the factors influencing mathematical literacy through regression analysis. Saenz (2009) reported that there is a need to conduct more detailed statistical analyses regarding problem solving processes in order to help students transform mathematical information into something more functional. In this sense, we can say that statistical analysis is consistent with the literature.

While identifying the scope of the measurement tool (MLT), the subject areas specified in the PISA (Space and Shape, Change and Relationships, Uncertainty and Data, Quantity) were considered individually. Moreover, special attention was paid to the numbers of questions in each area in order to thoroughly represent each subject (Table 1). In addition, special attention was paid to ensure the representation of each competency cluster (i.e. reproduction, connections, and reflection (OECD, 2010) in questions. Five of the questions in the test were about reproduction capabilities; nine were about connections capabilities; and three were about reflection capabilities. Each of the process skills (formulating, employing, and interpreting) was included in the MLT. Thus, it is possible to say that the MLT meets the criterion of content validity.

Normality and sphericity test results indicated the appropriateness of factor analysis for this study and thus the reliability of its results. Furthermore, the number of data is important for factor analysis. Studies with small samples are advised to include at least 10 times as many variables (Çokluk et al., 2014; Pallant, 2001). The number of variables in this study is 17 and the number of data is 435, which indicates the satisfaction of both criteria. Some studies (Tatlıdil, 2002) indicate that reliable results can be obtained if the number of data, regardless of the number of variables, is at least 300. The parameters of this study meet this criterion.

When the number of variables is than 40 and the sample size is large (>200), the factors composed of components whose levels of variance are considered fair, good, very good, or excellent are reliable (Tabachnick & Fidell, 2013, pp. 649). This is consistent with the parameters of this research. In addition, the same procedure was applied to the sub-groups consisting of 300-350 data that were randomly chosen from among 435 data (Tabachnick & Fidell, 2013). No significant change was observed in the factors obtained. The rate of total variance explained in this process ranged from 0.49 to 0.54. Some of the samples displayed five factors instead of the six factors obtained from this analysis. When the five-factor structure was compared to the six-factor structure, we observed that the first and the second factors combined as a single factor while other factors remained the same in all trials. When the contents of the first and the second factors are compared, the questions in these factors show similarity in that they can be solved via algorithmic operations. The questions of the second factor are based on longer texts compared to the questions of the first factor. Reading and understanding them before the solution can create difficulty for some students. Taking this similarity into account, the combination and separation of the first and the second factors can be explained. Despite this partial change, preservation of the other factors is seen as an indicator of non-fragility of the factorial structure (Tabachnick & Fidell, 2013).

When evaluating the factor structures, the components with factor loads over 0.32 are accepted interpretable (Tabachnick & Fidell, 2013). According to what (Tabachnick & Fidell, 2013, pp. 654) cite from Comrey and Lee (1992), factor loads exceeding 0.71 are considered excellent; factor loads that are not less than 0.63 are considered very good; factor loads not less than 0.55 are considered good; and factor loads not less than 0.45 are considered fair. Considering the factor loads in the present study, it is evident that the levels of variance explained by them can be considered excellent, very good, good, or fair (Çokluk et al., 2014). This is attributed to the power of the questions and indicates that the content of each question has to be taken into account while naming the factors (Cohen et al., 2011). In light of these evaluations, factor naming was conducted as explained below:

Naming the Factors

Factor 1: The four questions forming this factor were about how an arithmetic mean is calculated, the influence of data on the mean, identification of the most expensive and the cheapest options from a given price list, and the best purchase option from a given price list with a certain amount of money (limited budget). The types of operations needed to solve these questions were easily understandable. These questions required sequential operations. The achievement levels of the students answering these questions (which constitute the components of this factor) were high. This result is in line with the result offered by Saenz (2009), asserting that pre-service teachers are more successful in operational questions than in contextual and conceptual questions. The components of the factor are seen to have excellent, very good, good, and fair (Table 4) connections with the factor (Tabachnick & Fidell, 2013, pp. 654). This factor's contribution to the total variance, which was nearly 54%, was 12%. It was called *making algorithmic operations*.

Factor 2: This factor covered Test Scores, Deputy 1, and Skateboard 2 questions. The question Test Scores involved comparing graphics and revealing their differences. Deputy 1 involved reading and understanding an election principle and finding a result by using operations. Skateboard 2 involved ordering and listing. This factor explained nearly 10% of the variance in achievement. Though the questions of this factor were not very difficult, either the question texts or the answers were long. Moreover, question texts covered notations that were irrelevant to the solution of the questions. Briefly, they had rich mathematical content. The connections of the questions with the factor are at fair, good, and very good levels (Tabachnick & Fidell, 2013, pp. 654). This factor was called *mastering rich mathematical content*.

Factor 3: This factor covered Height 2 and Oil Spill questions. Its share of total variance was 8%. The first component (question) involved the context of arithmetic mean. Therefore, it required identification of whether the inferences were correct or incorrect. The second question involved prediction of a non-geometric shape's area by using a scale on a map. Geometric shapes could vary during the prediction process and there were many ways of finding the solution. These two questions tested the ability to make inferences about an objective (based on one's current mathematics background) and to evaluate the inferences. Therefore, this factor was called *mathematical inference*. The connections of the questions with the factor are at very good level (Tabachnick & Fidell, 2013, pp. 654).

Factor 4: This factor consists of three questions. Among these three questions, the Deputy2 question was about making an amendment in the D'Hont electoral system rules, which are taken as basis in determining the number of deputies elected from an electoral area, to serve the desired purpose. The Best Car 2 question was about making an amendment in the scoring formula weighted based on cars' features to enable a selected brand to come first. The third question, Heartbeat 1, involved explaining a decision after comparing two different equalities involving variables. Taking these characteristics into account, this factor was called *developing mathematical proposals and interpreting a developed proposal*. The connections of the questions with the factor are at good, very good, and excellent levels (Tabachnick & Fidell, 2013, pp. 654).

Factor 5: There were two questions in this factor. The question of Paint required selecting the appropriate option for the parameters in the question root out of many easy-to-find solutions. The students were expected to make an evaluation of the solution for everyday life by considering the parameters and leaving traditional school mathematics behind. Heartbeat 2 involved reflecting a change regarding a specified objective on a given mathematical formula. In terms of factor loads, the connections of the questions with the factor were found to be at excellent and very good levels

(Tabachnick & Fidell, 2013, pp. 654). This factor was called *understanding the mathematical equivalence of real world situations*.

Factor 6: The factor covered two questions. In the Earthquake question, a mathematical rate was given about the probability of the occurrence of an earthquake, and it was requested to find the correct one among the various meanings that could be inferred from it. The interesting part of the question was that it required the skill to transform mathematical language into vital language. The question The Best Car 1 also supported this factor but with a lower factor load. This question required making weighted score calculations. What distinguishes these questions from regular algorithmic questions is that they contain codes that may be unfamiliar to the students, but these codes do not have any influence on the content of calculation. The first question had an excellent factor load whereas the latter had a fair factor load (Tabachnick & Fidell, 2013, pp. 654). Taking the kurtosis of the first question into account, this factor was called *understanding the counterpart of mathematical language in life*.

When factors are considered together, it is clear that the factors given in Table 1 show a different classification from the classifications existing in the mathematical literacy literature (competency clusters, subject areas, mathematical processes, contexts). This shows that the questions were considered from a different perspective, and a structure associated with the mental actions required by their solutions was introduced in the present study. Considering the statistical analysis underlying factor analysis (i.e. main component analysis), this structure can be named as "*main components of mathematical literacy*". These components include (i) making algorithmic operations, (ii) mastering rich mathematical content, (iii) mathematical inference, (iv) developing mathematical proposals and interpreting a developed proposal, (v) understanding the mathematical equivalence of real world situations, and (vi) understanding the counterpart of mathematical language in life. When the main components (factors) obtained here and the classifications about mathematical literacy in the literature are considered together, it seems that this structure consisting of main components has *discrete borders that can be determined more easily* compared to the classifications taking competency clusters (reproduction, connections, reflection) and mathematical processes (formulating, reasoning, interpreting, and evaluating) as basis. In this regard, this classification can be more useful than other classifications in planning mathematical literacy practices, writing mathematical literacy questions, and assessing the success level of students (in which components they are successful and in which components they are unsuccessful). The teaching content can be supported in terms of mathematical literacy by considering the components involving failure. This kind of a contribution can be made to the literature.

Relationships between the Factors (Main Components) and Competency Clusters

One of the classifications about mathematical literacy questions is the classification made based on the competencies required by the solution. They are, as indicated by OECD (2009a, 2010), "using symbolic, formal and technical language and operations, argumentation, mathematical problems and solving (devising strategies for solving problems), reasoning and argument, modelling, presentation, communication, use of aids and tools". Evaluating the factors based on the capabilities defined by OECD (2009a, 2010) may yield certain obvious results. For example, over 50% of the students were successful in the questions of the first (making algorithmic operations) and the sixth (understanding the counterpart of mathematical language in life) factors. These questions conform to the expressions "being able to making sense of the expressions which involve symbols and formulas", "being able to use variables", and "being able to solve equations and make calculations" (OECD, 2009a, 2010) which fall under the capability of *using symbolic, formal and technical language and operations* regarding mathematical literacy. It is obvious that the students had acquired, at a reasonable level, the knowledge and skills necessary for the questions in these factors. The questions falling under the second factor (mastering rich mathematical content), in which the achievement levels of students can be considered low, conform to the expression "following and evaluating the chain in various mathematical hypotheses" (OECD, 2009a, 2010), which falls under the capability of *argumentation-enquiry*. This result indicates that including rich mathematical narratives in instruction and questions and making students familiar with such texts is a need. This need can be fulfilled by adding contextual questions and texts having mathematical content into mathematics textbooks and asking questions and starting discussions about them.

In relation to the third, fourth and fifth factors, which yielded low achievement levels on average, we offer the interpretations below: The Oil Spill question, under the third factor, conforms to the expression "identifying various types of *mathematical problems and solving* different types of questions via different methods" (OECD, 2009a, 2010), within the *devising strategies for solving problems* capability. The Height 2 question conforms to the expressions "distinguishing different definitions from each other" and "understanding the limits and aspects of current mathematical concepts" (OECD, 2010), which fall under the capability of *reasoning and argument*. Consequently, it is possible to say that the questions under this factor measure the capabilities of *reasoning and argument* and of *devising strategies for solving problems*. These questions and the related capability expressions reveal that students have skill deficits in predicting which mathematical knowledge and skills are used in certain cases, using them, and understanding which usages are correct. This difficulty can be overcome through practices highlighting the skill of "using knowledge to reproduce a piece of information" (Altun & Yılmaz, 2008), which is the second type of consolidation expressed by "+C" in the RBC+C model of formation, abstraction, and consolidation of knowledge.

The fourth factor covered the Deputy 2, The Best Car 2, and Heartbeat 1 questions, under the capability of *modelling* in relation to mathematical literacy. They conform to the expressions "constructing situations by modelling them" and "working with mathematical models, exploring and reflecting the model's accuracy, analyzing the model, and criticizing the results (with limitations)" (OECD, 2009a, 2010). We can assert that these questions measure the capability of modelling. It is obvious that the students were not able to make mathematical proposals to solve a vital problem or to express their proposals in mathematical language. This issue, proposing a mathematical solution to a problem, is unfamiliar to Turkish students (İskenderoğlu & Baki, 2011). Therefore, we can say that mathematics textbooks need contextual questions that are appropriate for developing mathematical proposals.

The questions covered by the fifth factor are Paint and Heartbeat 2. They conform to the expression "transforming the reality into mathematical structures" (OECD, 2009a, 2010), under the capability of *modelling*. In other words, the students were incompetent in finding a mathematical counterpart to the situations expressed in natural (vital) language or were incompetent in understanding the accuracy of a proposed mathematical counterpart. All in all, this result also leads us to a deficiency in modelling competency, as was the case in the fourth factor. This result supports Saenz's (2009) finding that expressing mathematical discussions in natural language is harder than expressing with numbers and formulas. This deficiency can partially be eliminated by enriching textbooks with modelling questions. Moreover, Polya (1957) mentions "looking back" as a part of the problem solving process. Activities making use of the item "express another similar model of the problem and solve it," as stated by Polya (1957), may be useful in eliminating the aforementioned deficiency.

The students' rates of achievement in each factor indicate that they had difficulty in the main components of understanding the mathematical equivalence of real world situations, developing mathematical proposals and interpreting a developed proposal, and mathematical inference, respectively. They were followed by mastering rich mathematical knowledge. The main components in which they had achievement, on the other hand, were understanding the counterpart of mathematical language in life and making algorithmic operations. These evaluations show that there is no need to be contended with the problems whose mathematical expressions are given during instruction in these kinds of groups. Students should be confront with contexts in which problems emerge or are likely to emerge, and in instructional activities, a coverage should be given to the identification and expression of problems through working on texts with mathematical content. Moreover, discussions should be allowed with regards to in which conditions the solutions obtained are meaningful; with which mathematical notations the result can be expressed (modelling) if any pattern or connection has been detected; what other kinds of problems the problem worked on leads to; and how such other problems can be solved.

The main components of mathematical literacy showed in this study offer a new classification with a perspective different from the current classifications used in mathematical literacy practices/assessments and in the identification of students' mathematical literacy. This classification made based on factors can offer high quality opportunities for both usage areas. The present study is limited to 17 items from four subject areas. With similar research to be conducted with a larger number and variety of questions, more stable structures can be introduced for the main components of mathematical literacy questions. These stable structures may contribute to developing a mathematical literacy scale as well.

References

Akyüz, G., & Pala, N. M. (2010). PISA 2003 sonuçlarına göre öğrenci ve sınıf özelliklerinin matematik okuryazarlığına ve problem çözmeye etkisi. *İlköğretim Online, 9*(2), 668-678.

Altun, M., & Akkaya, R. (2014). Matematik öğretmen adaylarının PISA matematik okuryazarlık beceri düzeylerinin incelenmesi. *Hacettepe Üniversitesi Dergisi, 29*(1), 19-34.

Altun, M., & Yılmaz, A. (2008). Lise öğrencilerinin tam değer fonksiyonu bilgisini oluşturma süreci. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi, 41*(2), 237-271.

Arıkan, S. (2014). A regresion model with a new tool: IBD analyzer for identifying factors predicting mathematics performance using PISA 2012 indices. *US-China Education Review A, 4*(10), 716-727.

Breakspear, S. (2012). The policy ımpact of PISA: An exploration of the normative effects of international benchmarking in school system performance. *OECD Education Working Papers,* No. 71, OECD Publishing.

Can, A. (2012). *SPSS ile bilimsel araştırma sürecinde nicel veri analizi*. Ankara: Pegem Yayıncılık.

Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education* (7th ed.). New York: Routledge.

Creswell, J. W. (2014). *Araştırma deseni* (2nd ed.) (Demir, S. B., Trans. Ed.). Ankara: Eğiten Kitap.

Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2014). *Sosyal bilimler için ok değişkenli istatistik SPSS ve LISREL uygulamaları* (3rd ed.). Ankara: PegemA Yayıncılık.

Dossey, J., McCrone, S., Turner, R., & Lindquist, M. (2008). PISA 2003 mathematical literacy and learning in the Americas. *Canadian Journal of Science, Mathematics and Technology Education, 8*(2). 140-152.

Eraslan, A. (2009). Finlandiya'nın PISA'daki başarısının nedenleri: Türkiye için alınacak dersler. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi 3*(2), 238-248.

Fraenkel, R. R., Wallen, N. E., & Hyun, H. H. (2011*). How design and evaluate research in education.* New York: Mc Graw Hill.

Forgasz, H. J., & Leder, G. C. (2017). Persistent gender inequities in mathematics achievement and expectations in Australia, Canada and the UK. *Mathematics Education Research Journal*, 1-22.

Glass, G. V., & Hopkins, K. D. (1984). *Statistical Methods in Education and Psychology* (2nd ed.). Boston: Allyn and Bacon.

İlbağı, E. A. (2012). *PISA 2003 matematik okuryazarlığı soruları bağlamında 15 yaş grubu öğrencilerinin matematik okuryazarlığı ve tutumlarının incelenmesi* (Unpublished doctoral dissertation). Atatürk University, Erzurum.

İskenderoğlu, T., & Baki, A. (2011). İlköğretim 8.sınıf matematik ders kitabındaki soruların PISA matematik yeterlilik düzeylerine göre sınıflandırılması, *Eğitim ve Bilim*, *36*(161), 287-301.

Liang, X. (2010). Assessment use, self-efficacy and mathematics achievement: Comparative analysis of PISA 2003 data of Finland, Canada and the USA. *Evaluation & Research in Education, 23*(3), 213-229.

Lydia, L. O., & Wilson, M. (2009). Gender differences and similarities in PISA 2003 mathematics: A comparison between the United States and Hong Kong. *International Journal of Testing*, *9*(1), 20-40.

Ministry of National Education of Turkey. (2011). *Pisa Türkiye.* Retrived from http://pisa.meb.gov.tr/wp-content/uploads/2013/07/PISA-kitab%C4%B1.pdf.

Ministry of National Education of Turkey. (2012). *PISA Türkiye***.** Ankara: Yenilik ve Eğitim Teknolojileri Genel Müdürlüğü.

McCrone, S. S., & Dossey, J. A. (2007). Mathematical literacy - it's become fundamental. *Principal Leadership, 7*(5), 32-37.

Morgan, G. A., Leech, N. L., Gloeckner, G. W., & Barret, K. C. (2004). *SPSS for introductory statistics: Use and interpretation* (2nd Ed.). London: Lawrence Erlbaum Associates.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, Va: NationalCouncil of Mathematics Teachers.

National Council of Teachers of Mathematics. (2006). *Curriculum focal points for preindergarten through grade 8 mathematics.* Reston, Va: NationalCouncil of Mathematics Teachers.

OECD, (1999). *Measuring student knowledge and skills. A new framework for assesment.* Paris: OECD Publishing.

OECD, (2003). *The PISA 2003 assesment framework – mathematics, reading, science and problem solving knowledge and skills.* Paris: OECD Publishing.

OECD, (2006). *Assessing scientific, reading and mathematical literacy. A framework for PISA 2006.* Paris: OECD Publishing.

OECD, (2009a). *PISA 2009 Assesment Framework. Key competencies in reading, mathematics and science.* Paris: OECD Publishing.

OECD, (2009b). *Take the test-sample questions from OECD's PISA assessments.* Paris: OECD Publishing.

OECD, (2010). *PISA 2012 mathematics framework to OECD, 30.* Paris: OECD Publications.

OECD, (2013). *PISA 2012 assessment and analytical framework. Mathematics, reading, science, problem solving and financial literacy.* Paris: OECD Publishing.

OECD, (2014). *PISA 2012 result: Creative problem solving. Students' skills in tackling real-life problem (Vol. V).* OECD Publishing.

OECD, (2016). PISA 2015 assessment and analytical framework. Science, reading, matematic and financial literacy. Paris: OECD Publishing.

Özkan, Y. Ö., & Güvendir, M. A. (2014). Socioeconomic factors of students' relation to mathematic achievement: Comparison of PISA and ÖBSS. *International Journal of Educational Sciences, 6*(3), 776-789.

Pallant, J. (2001). *SPSS survival manual.* Philadelphia: Open University Press.

Polya, G. (1957). *How to solve it?* Gaeden City, New York*: Doubleday Company.*

Ross, S. (2008). *Motivation correlates of academic achievement: Exploring how motivation influences academic achievement in the PISA 2003 dataset* (Unpublished doctoral dissertation). University of Victoria, Canada.

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

Satıcı, K. (2008). *PISA 2003 sonuçlarına göre matematik okuryazarlığını belirleyen faktörler: Türkiye ve Hong Kong-Çin* (Unpublished master's thesis). Balıkesir Üniversitesi, Balıkesir.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistic* (6th ed.). USA:Pearson.

Tan, Ş. (1998). *Unweighted scoring versus weighted scoring methods for likert type data* (Unpublished doctoral dissertation). Loyola University Chicago, Chicago.

Tatlıdil, H. (2002). *Uygulamalı çok değişkenli istatistiksel analiz.* Ankara: Akademi Matbaası.

Tavşancıl, E. (2014). *Tutumların ölçülmesi ve SPSS ile veri analizi* (5th ed.). Ankara: Nobel Yayıncılık.

Wood, B. (2007). *the impact of students' skills in self-regulated learning on mathematics literacy and problem solving scales as measured by PISA 2003: A comparison of the United States and Finland* (Unpublished doctoral dissertation). California State University, Long Beach, the USA.

Appendix 1. The Questions Written by the Researcher

QUESTION 2: DEPUTY

The D'Hont system, which is used in elections both in our country and in many others to determine the distribution of deputies among parties, is as follows: The total number of votes a party gets in an election region is divided by 1, 2, 3, 4, and 5, respectively, and written one under another. The values in the obtained table of numbers are listed in a descending order. The deputies are distributed to the parties one by one starting from the highest value.

In an election region which elects ***five*** deputies, four parties participating in the elections got the votes given below:

Party A	Party B	Party C	Party D
300	660	120	420

Question 2.1: DEPUTY

Specify how many deputies each party gets.

Question 2.2: DEPUTY

What kind of a change do you offer in order to make more parties represented in the assembly? Explain.

QUESTION 6: PAINT

***Question 6.1**: PAINT*

A type of paint is released to the market in 2-liter and 5-liter cans. A 2-liter can costs 8 Turkish liras while a 5-liter can costs 15 Turkish liras.

What is the minimum amount of money a person who is in need of 16 litters of paint has to spend to meet his need?

Appendix 2. Scoring Rubric for the Test Scores Question

TEST SCORES - RUBRIC (OECD, 2009b; MEB, 2011)

Full (2) credit:

One valid argument is given. Valid arguments could relate to the number of students passing, the disproportionate influence of the outlier, or the number of students with scores in the highest level.

- More students in Group A than in Group B passed the test.
- If you ignore the weakest Group A student, the students in Group A do better than those in Group B.
- More Group A students than Group B students scored 80 or over.

No (0) credit:

Other responses, including responses with no mathematical reasons, or wrong mathematical reasons, or responses that simply describe differences but are not valid arguments that Group B may not have done better.

- Group A students are normally better than Group B students in science. This test result is just a coincidence.
- Because the difference between the highest and lowest scores is smaller for Group B than for Group A.
- Group A has better score results in the 80-89 range and the 50-59 range.
- Group A has a larger inter-quartile range than Group B.
- There is not answer.