



(ISSN: 2602-4047)

Demir, F. (2021). The Item Pool Development In Mathematical Literacy, *International Journal of Eurasian Education and Culture*, 6(13), 1368-1427.

DOI: <http://dx.doi.org/10.35826/ijoecc.399>

Article Type (Makale Türü): Research Article

## THE ITEM POOL DEVELOPMENT IN MATHEMATICAL LITERACY

**Furkan DEMİR**

Ph.D., Dumlupınar University, Kütahya, Turkey, [furkan.demir@dpu.edu.tr](mailto:furkan.demir@dpu.edu.tr)

ORCID: 0000-0003-3740-8088

Received: 01.02.2021

Accepted: 25.05.2021

Published: 14.06.2021

### ABSTRACT

This study aimed to examine, revise, test, and evaluate the items produced in mathematical literacy in terms of language, structure, and content and create an item pool containing valid, reliable, and applicable items. The study method was based on the formative assessment of development research types. Thus, small group tests were carried out with 8th grade students, field tests with 9th and 10th graders. The small group tests consisted of 20-30 students, and 3053 and 2918 students participated in the first and second field tests. A total of 102 items were tested. Following the small group tests, student solutions were analyzed, and revision procedures were carried out accordingly. At the end of the first and second field tests, item analysis was completed through 27% upper and lower groups method. According to the first field test results, 33 items were revised, considering the item discrimination index and the accumulation rates in distracter items among the upper and lower groups. Those items were re-tested in the second field test. The structural and cognitive factors that affect the validity were included in the results related to the second field test analysis of the revised items. These factors would contribute to the revision methodology literature. At the end of the study, there were 94 items, of which the discrimination index was  $r_{jk} \geq 0,30$ . Besides, the mathematical literacy sub-dimensions measured by each of the given items were determined within this study's scope. The items, mathematical literacy sub-dimensions, and item statistics were in the appendixes. It is suggested that using the given items in mathematics education research, in-class activities, and measurement and evaluation activities will be beneficial for students to develop mathematical literacy.

**Keywords:** Item analysis, item pool, mathematical literacy, PISA, revision methodology.

---

This research was supported by the Scientific Research Projects Unit (BAP) of Dumlupınar University (Project Number: 2018-34). This study was divided into two parts and presented at these congresses in 2020 and 2021, respectively: International Congress of Eurasian Social Sciences-4, Bodrum-Muğla, Turkey and 14th National Science and Mathematics Education Congress, Burdur, Turkey.

## INTRODUCTION

The results of the Program for International Student Assessment [PISA] are considered as a national report card in many countries. In addition to the PISA results published by the Organization for Economic Cooperation and Development [OECD], each country prepares its national reports and conducts national research based on these results. The results provide countries with information about ranking and many opportunities to make criteria-based assessments. One of these criteria is mathematical literacy, which is defined as the individual's capacity to formulate, apply, and interpret mathematics in various ways in life (OECD, 2013a; OECD, 2016).

PISA categorizes mathematical literacy items according to (i) context, (ii) mathematical content knowledge, and (iii) mathematical processes for the solution. Thus, it monitors how students manage the necessary mathematical processes by using their mathematical content knowledge in solving the problem given in a context. Besides, the "Mathematical Literacy Proficiency Scale" of PISA helps countries make skill-based and competency-oriented assessments. Because the basis of this 6-level scale (OECD, 2019) consists of eight characteristic mathematical skills that are meaningful for each level. The proficiency levels describe the differences between the students' performances as follows: "After solving the questions with specific difficulty level, the basic mathematical skills required for the solution are examined." (Stacey, 2013). Then, the students' mathematical literacy proficiency levels are determined on the scale, which summarizes the mathematical literacy skills of all participant students. Thus, the participant countries have a general idea about the mathematical literacy levels of their 15-year-old contestants. All these features make the scores obtained from PISA applications more meaningful and more functional.

Many countries take PISA's mathematical literacy results as a reference to structure national curriculums (Breakspear, 2012). Also, in the last 30 years, with the researchers' suggestions (Lloyd, Herbel-Eisenmann & Star, 2011), some countries including England and the United States [USA] have started to develop context-based mathematical ideas in the curriculum (Boaler, 1993; Robinson, Robinson & Maceli, 2000). Following the reports by Tomlinson (2004) and Smith (2004), England has made mathematical literacy a key goal in its schools. In the USA standards, improving students' literacy skills in all primary branches in public schools (e.g., English, mathematical, science, social) has been the chief goal, and literacy-based teaching has been promoted (National Governors Association Center for Best Practices, 2010). According to Turkey's secondary school mathematical curriculum's first statement, invoked in students will develop mathematical literacy skills and use them effectively (Ministry of National Education [MEB], 2018). Hence, it is understood that countries take specific theoretical measures to improve students' mathematical literacy. And the practical efforts to identify and enhance individuals' mathematical literacy can be evaluated through the studies given below.

It is suggested to increase the number of literacy items in national exams (Jürges, Schneider, Senkbeil & Carstensen, 2012; Uysal & Yenilmez, 2011). Although it seems advantageous, it is directly about assessment and evaluation, and its scope should be extended to include the education and teaching because the evaluation regarding the achievement of the of primary mathematics education objectives should be acknowledged as a

complement of education (National Council of Teacher of Mathematics, 2002). Thus, mathematical literacy items are necessary for in-class activities. This study aims to produce items that will contribute to the literature. Such an initiative could reinforce the relationship between the mathematical teaching and assessment and evaluation processes. Besides, inherent to the definition of mathematical literacy can contribute to associating mathematical teaching with life.

Özgen (2013) stated that prospective mathematics teachers have difficulties in associating mathematics with life. Leikin and Levav-Waynberg (2007) also indicated that middle school teachers experience challenges in associating mathematics with life, so they prefer to give abstract examples, which might result in disconnections between the teaching material and daily life experiences (Altun & Akkaya, 2014), and lead to a deficiency in students' mathematical literacy skills. One of the efforts against the current situation is to encourage teachers and students to plan a PISA-based (contextual) learning and design teaching assessments based on PISA criteria (Kohar, Zulkardi & Darmawijoyo, 2014).

Reinke (2020) contributed to the literature regarding the use of contexts in mathematical teaching by presenting a model in which a student initially developed images by dealing with contextual problems to make sense of subsequent tasks in a standard classroom setting. Uysal and Yenilmez (2011) suggested that students should frequently experience and solve problem situations such as those found in PISA to improve mathematical literacy skills. According to Wardono and Mariani (2014), the lack of learning tools and activities to develop mathematical literacy is one reason for students' inability to improve mathematical literacy skills. Altun and Bozkurt (2017) expressed that it is necessary to discuss daily-life texts with mathematical content in learning environments, and students should be familiarized with them. He stated that this problem could be solved by involving and asking contextual items and texts with mathematical content and opening discussions in mathematical teaching. Ekawati, Susanti and Chen (2020) proposed to design learning activities that provide secondary school students with more opportunities to learn and solve mathematical literacy items. These literature suggestions highlight the need to develop valid, reliable, and practical mathematical literacy items and activities, which leads researchers to study on producing such items and activities.

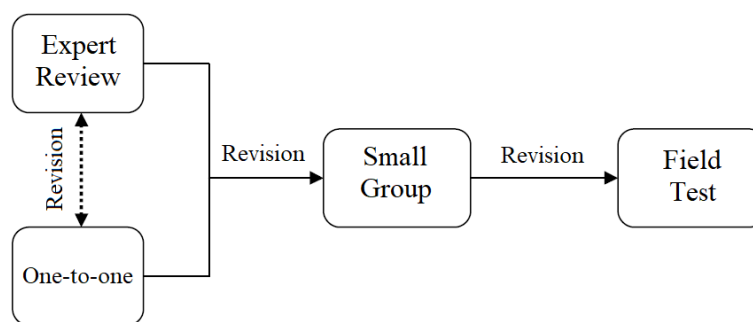
Many researchers carried out researches to develop items similar to PISA to measure students' mathematical reasoning skills (Annisah, 2011), mathematical communication skills (Mardhiyanti, Putri & Kesumawati, 2013), mathematical literacy (Oktiningrum, Zulkardi & Hartono, 2016), and problem-solving ability (Nasution, Fauzi & Syahputra, 2019). Lestariningsih, Nurhayati, Susilo, Cicinidia and Lutfianto (2020) developed mathematical literacy questions that have potential impact on students' representation skills. Besides, Silva (2011) produced items in the uncertainty category and categorized them according to the mathematical literacy competence sets. Kamaliyah, Zulkardi and Darmawijoyo (2013) created the sixth level of PISA-like mathematical literacy problems. In the last decade, although there have been many international studies listed above, Turkey scored low in math literacy in PISA (Acar, 2012; Altun & Akkaya, 2014; Aydin, Sarier & Uysal, 2014; OECD, 2013b). It might stem from

the lack of items and activities to support mathematical literacy (Wardono & Mariani, 2014), and there are still a few national studies to produce mathematical literacy items and activities.

As mentioned in the research results and suggestions above, an item pool, including contextual questions that can be used for in-class activities designed to contribute to mathematics education, is an essential need today. Thus, the current study aims to produce valid, reliable, and applicable mathematical literacy questions. With this purpose, the mathematical literacy questions obtained from Demir and Altun (2018) studies and Demir (2019; 2020) with prospective teachers were evaluated, modified, applied, and analyzed in terms of language, structure, and content.

## METHOD

The study aimed to analyze, revise, test, and evaluate the item pool in mathematical literacy in terms of language, structure, and content and obtain valid, reliable, and functional items. The study applied the formative assessment design, including one-to-one student interviews, expert reviews, small groups, and field tests (Tessmer, 1993). The steps of formative assessment are shown in Figure 1.



**Figure 1.** The Steps of Formative Assessment (Tessmer, 1993)

Tessmer's (1993) method has been used in various item development researches (e.g., Oktiningrum et al., 2016; Nasution et al., 2019). However, "one-to-one" student interviews are generally applicable for studies in which the number of items is limited to a few questions. Since the study aimed to test 102 items with students and the "one-to-one" step that involves interviewing several students for each question was dismissed, which can be a limitation of this study. However, a measure was taken to reduce the limitation's impact, which is explained 4th stage of the method below. The method stages are described in detail below.

- 1) Mathematical literacy items were compiled from the literature.
- 2) Expert reviews were performed in terms of language, structure, and content of the compiled items, and revision procedures were carried out based on the experts' comments and suggestions.
- 3) Open-ended item pool and answer key were created.
- 4) Open-ended questions were tested in small groups. The reason for using open-ended questions was to studying with small groups rather than "one-to-one" interviews proposed by Tessmer (1993). It was assumed

that students' answers to the open-ended questions could provide some data, which is thought to reduce the study's limitation.

5) Based on the data from small groups, item options were written, and the optimal time for the answer was determined approximately. The expert evaluations of the items with multiple-choice were then completed in terms of language, structure, and content. The comments and suggestions of the experts were received, and necessary revisions were made.

6) The item difficulty indices were determined according to the data obtained from small groups. The mathematical literacy sub-dimensions (context, mathematical content, mathematical process) of questions were determined. Hence, each item was labeled with its sub-dimension and estimated item difficulty index.

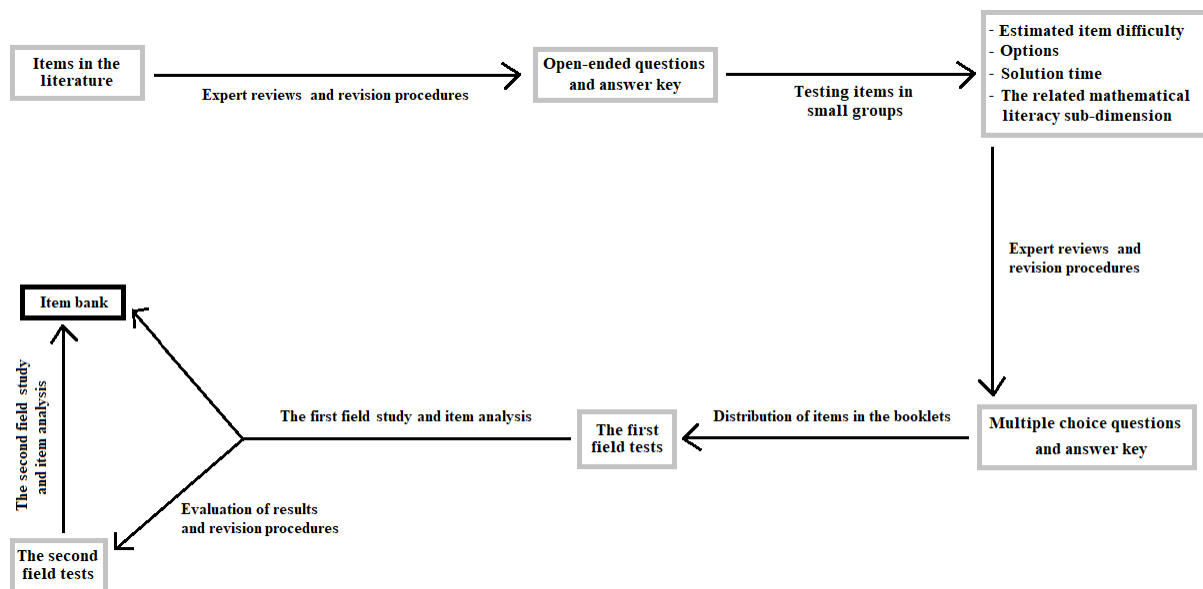
7) Draft multiple-choice item pool and answer keys were created.

8) The items were distributed in the booklets, considering the estimated item difficulty index and related mathematical literacy dimensions.

9) The first field study was carried out, and item analysis and revisions were made accordingly.

10) The other field study was carried out, and the item analysis was done based on the data.

The first seven stages above continued as the new items were added from the literature. In other words, the new questions underwent the first seven processes. Then, those questions were saved until the necessary consents were received for field studies. Ethical rules were followed at every stage of the method.



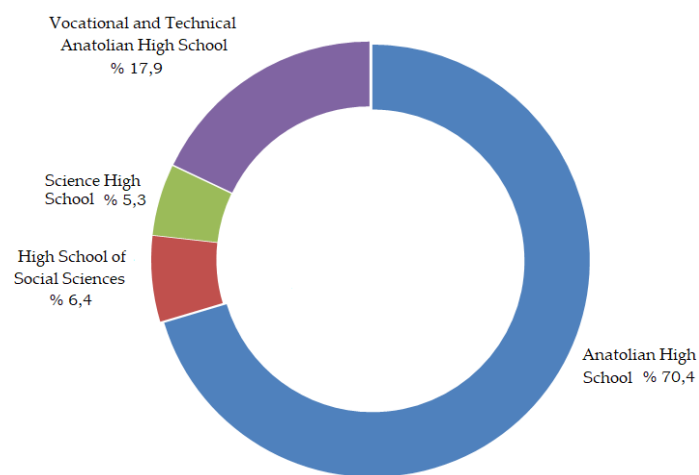
**Figure 2.** Model of the Processes Followed in the Method

Figure 2 shows the processes followed in the study. Similar models were also used for the item pool/question bank development studies in the literature (e.g., Wongpanya, Pativisan & Ngamman, 2018; Yuentrakulchai, Kamtet & Dechsri, 2018).

### Study Sample

There were two groups in this research. The first group was at the fourth stage of the study and involved two 8th grade classes in which the number of students ranged between 20-30 on application days. Classes were selected from a public school in a district in the Mediterranean region. The population of the district is about 160 000. The socioeconomic status of students was medium. The reason for choosing these two classes can be explained as follows: The most crucial aim of the 4th stage was to test the items (whether they were overlapped with what the student understood from the question). Thus, the questions were asked open-ended, and students were encouraged to find solutions. The validity and reliability of the data on the item clarity depend on the students' focus on the solution, and their teachers play a crucial role in encouraging them to concentrate on the solution. The most significant factor in choosing the students of these two classes for the solutions with small groups was that teachers agreed to support the research in this regard.

The second group was in the 9th and 10th stages of the method and included students who participated in both field studies. The second group consisted of 9th and 10th grade students at 12 high schools in a city center in the Aegean region. It consists of 110 classes, and the students in those classes participated in both field tests. 3053 and 2918 students participated in the field tests, respectively. The difference arose from the absent students on the application day. 55% of the participants were 9th graders, and 45% were 10th graders. The reason for preferring only 9th and 10th class levels was that 96.5% of Turkey sample in PISA-2018 consisted of 9th and 10th grade students (MEB, 2019).



**Figure 3.** Distribution of Students by School Types in Field Tests

8 Anatolian, 2 Vocational Technical Anatolian, 1 Science, and 1 Social Sciences High Schools participated in the study. The percentage distribution of the students by their school types in field tests is shown in Figure 3. 81.4% of Turkey sample in PISA-2018 involved the students from 4 school types (MEB, 2019).

#### **Data Collection Tools, Data Collection and Analysis**

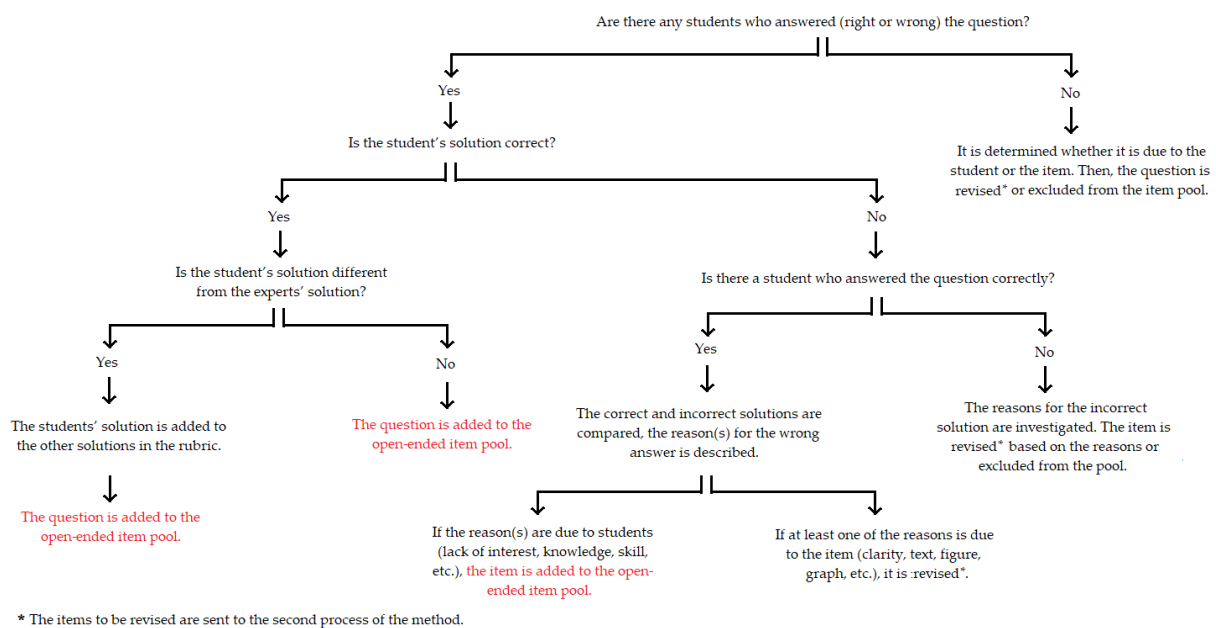
The research method was explained in 10 processes. The first process was compiling mathematical literacy items in the literature (Demir & Altun, 2018; Demir 2019; Demir, 2020). A total of 102 open-ended questions were compiled at the end of the process.

The second stage reflected the expert assessments of the collected items regarding language, structure, and content. The assessments were made by two Turkish education experts, a mathematical teacher (teacher of the 8th graders in small groups), a mathematics education researcher, and an assessment and evaluation specialist. Turkish education experts were two research assistants at the doctoral dissertation stage. The mathematical teacher had eight years of experience. The assessment and evaluation specialist also had a degree in mathematics education and has published academic studies on PISA and mathematical literacy. Turkish education experts evaluated the questions in terms of language. Other experts were asked to answer the questions, evaluate the items' suitability for 15-year-old students, and express their comments and suggestions regarding text, figure, and clarity. Revisions were made according to the feedback from the experts.

The third stage involved the development of the answer key (rubric) of each item. At the end of the process, the researcher obtained an open-ended item pool and answer keys (rubrics) consisting of questions and answers given by experts.

The fourth stage involved testing the items in small groups of 20-30 participants. The teacher encouraged students to focus on the solution to ensure validity and reliability (e.g., by using encouraging expressions, using in-class performance grades, arousing interest for the items). It is thought that students' focus on the solution increases the probability of valid and reliable results about the items. At least 7-8 students solved each question. At the end of the process, students' solutions were clustered.

In the fifth stage, student solutions were analyzed. The algorithm in the analysis is presented in Figure 4. As a result of the analysis, the options based on student solutions in small-group applications were written for the open-ended items.



**Figure 4.** The Algorithm in Analysis of Student Solutions of Open Ended Questions

At the end of the fifth stage, the experts assessed the multiple-choice items in terms of language, structure, and content, and gave feedback. The revision was made based on the feedback. Besides, it was determined to give students 45 minutes to solve 10-11 problems.

The sixth stage included identifying the mathematical literacy sub-dimensions (context, mathematical content, mathematical process) considering the matched items and expert assessments and measuring the item difficulty index based on the data obtained in small groups. The estimated item difficulty was determined as an integer in the range 0-10 for each item. The values closer to 10 represent the difficulty of the item. The explanations about the sub-dimensions in PISA reports and the item classifications of the published PISA questions were accepted as criteria in matching the item with the mathematical literacy sub-dimensions. In this sense, the researcher carried out the matching procedures, and the measurement and evaluation specialist separately. The items were labeled according to the discussion results based on the PISA arguments on the matching procedures. As a requirement of the seventh stage, 102 items completing the six stages above were added to the draft multiple-choice question pool and answer keys.

The eighth stage involved distributing the items into the booklets, considering the item difficulty index and the mathematical literacy dimensions they matched. It was aimed to ensure that the difficulty levels of the items in the booklets were similar and that each booklet included approximately the same questions in every mathematical literacy sub-dimension. The goal was to create similar booklets, thus ensuring the data's validity and reliability from field studies. Thus, 98 questions were shared in 10 different booklets, each consisting of 10 questions for the first application. Since four items did not complete the above processes yet, they were transferred to the second field study. The number of mathematical literacy questions that students could solve in 45 minutes was considered in determining the number of items (10) per booklet.



The ninth stage included the first field study, analysis of the data, and revisions. In this process, interviews were done with the relevant institutions, and the necessary official permissions were obtained for both applications. Then, the ninth and tenth-grade students at all schools covered by the permit were determined. The first application was made on Wednesday, 6 March 2019.

#### ***The Validity and Reliability Studies in the First Field Study***

1. Booklets of each class were placed in different envelopes in which each of the ten booklet types was included. For example, for a class of 32 students, at least 3 of each booklet type were placed in the envelope, which contributed to the study's validity and reliability in two ways. First, since the booklets' items were different, the students were prevented from communicating regarding the answers during the exam. Second, the booklets were distributed homogeneously throughout the sample, and the maximum number of students solved each problem.
2. The instructions and issues to be considered by students and teachers were written on the exam documents in detail (exam envelopes, booklets, and testing forms).
3. During the exam documents distribution, the responsible principal was informed that the proctor teachers should read the exam rules on the envelopes.
4. The exams were simultaneously held in the third-class hour by the proctor teacher in all schools and classrooms on Wednesday, 6 March 2019.
5. 50 minutes were given for the exam. Students were given 45 minutes to solve ten mathematical literacy questions in the booklet. The remaining 5 minutes were separated for the students' order in the classroom to distribute and collect booklets and testing forms.
6. After the exam, the exam papers were submitted to the school principals by the proctor teachers and then were received by the researcher.

#### ***Data Analysis of the First Field Study***

The testing forms of 49 students were not included in the analysis as they could not sign the form correctly or did not sign the wrong booklet type. Also, the item "H10" (H: Booklet type, 10: Item sequence) with the "72B" ID number was canceled and not included in the analysis because the options were misspelled. Thus, after the elimination of the missing data, the data of 3004 students and 97 mathematical literacy items were analyzed using the "upper and lower groups method (27%)" (Baykul, 2000) and "Test Analysis Program" (TAP). The analysis results brought item statistics related to item difficulty index ( $p_i$ ), discrimination index ( $r_{jx}$ ), and distractors in the options. Item difficulty index is the ratio of the correct answers to all respondents (Atılğan, Kan & Doğan, 2017). Therefore, it takes a value between 0 and 1 ( $0 \leq p_i \leq 1$ ). The value approaches 0 as the question becomes difficult and 1 as the question becomes easy. The item discrimination index, also called the item validity coefficient (Atılğan et al., 2017), is calculated by subtracting students who answer incorrectly in the lower group from the

ratio of students who answer correctly in the upper group (Başol, 2019). Therefore, it takes a value between -1 and 1 ( $-1 \leq r_{jk} \leq 1$ ).

Determination of Item: Item difficulty index ( $p_j$  value) may vary according to the test's difficulty. According to Crocker and Algina (1986), if the item discrimination index ( $r_{jk}$  value) is

- higher than 0.3 or equal to 0.3 ( $r_{jk} \geq 0.3$ ), items (questions) function well and can be kept in the test without making any changes.
- Between  $0.20 \leq r_{jk} \leq 0.29$ , the item can be revised and used accordingly.
- $r_{jk} < 0.20$ , the item should not be used or should be modified.

The analysis was not limited to the item discrimination index ( $r_{jk}$ ). Also, the data on the options of each item were examined separately. In item analyzes on upper and lower groups, distractors generally attract students in the lower group rather than in the upper group (Tekin, 2019). In this sense, the items that met one of the two conditions listed below were revised.

- i.  $r_{jk} < 0.3$  (29 items).
- ii. Despite  $r_{jk} \geq 0.3$ , the students' ratio from the upper group was higher than the ones in the lower group (4 items).

Table 1 shows the locations, ID numbers, and reasons for revisions of the 33 items.

**Table 1.** Items and Revision Reasons

Location	Item ID	Revision Reason	Location	Item ID	Revision Reason
A1	1	i	F2	60	i
D4	6	ii	D8	61	i
D2	7	ii	F7	65	i
I1	9	ii	D3	79	i
A2	10	i	F10	83	i
E1	11	i	J5	84	i
I2	13	i	B1	19A	i
C9	22	i	B2	19B	i
B6	27	i	C2	23A	i
J3	29	i	H3	49B	i
C4	31	i	H5	66A	i
I6	34	i	H6	66B	i
B7	39	i	G10	70B	i
B8	40	i	I4	76B	i
H1	51	ii	I7	82A	i
H7	53	i	I8	82B	i
A10	58	i			

Every item in Table 1 was reviewed and revised in terms of form, text, content, and options by a researcher and a measurement and evaluation specialist who had a degree in mathematics education considering the first application data. Some items were revised more than once, and a total of 44 revised items were obtained. Thus, those revisions increased the number of functional items. Each of the revised versions was labeled "R1", "R2," and the ID number. For example, three different revised versions of the item "60" were labeled as "60R1", "60R2," and "60R3" for the second field study.

The 33 items in Table 1 were not included in the booklets of the second field study. Forty-four items (revised versions) derived from the 33 items were placed in the second field study's booklets. Besides, four new items that were not tested in the first application were added to the booklets and draft multiple-choice question pool following the method's previous stages. The item distribution to the booklets was completed according to the criteria given in the 8th stage. However, different revised versions of the same item were distributed to different booklet types in the second field study. Thus, booklets for the second field study were completed.

The tenth stage of the method involved the second field study and analysis of the data. Similar validity and reliability measures in the first application were also taken in the second application. The second field study was held on Friday, 3 May 2019, with the same classes and schools of the first application.

### ***Data Analysis of the Second Field Study***

The testing forms of 72 students were not included in the analysis because they made an incorrect sign or did not write the booklet type in the testing form. Thus, after the elimination of the missing data, the data obtained from 2846 students were analyzed using the "27% upper and lower groups method" (Baykul, 2000) and "Test Analysis Program" (TAP). The analysis results gave item statistics related to item difficulty index ( $p_j$ ), discrimination index ( $r_{jx}$ ), and distractors in the options were obtained.

Depending on the  $p_j$  and  $r_{jx}$  indices from the first and second applications (Başol, 2019), standard deviations ( $s_j = \sqrt{p_j \cdot (1 - p_j)}$ ) and reliability coefficients of all items ( $r_j = r_{jx} \cdot s_j$ ) were calculated. The values are presented in the findings of the study (Appendix-1).

Similar to the assessment procedures of the first field study, the items were divided into three categories according to the criteria of Crocker and Algina (1986): "reliable," "reliable with revisions," and "unreliable or must be completely revised." At the end of the data analysis, the item discrimination indices in the range of  $0.29 < r_{jx} < 0.295$  were rounded up to  $r_{jx} = 0,29$ ; and the index of  $0.295 \leq r_{jx} < 0.3$  were rounded up to  $r_{jx} = 0.3$ , which led three items to change categories (item "27" with  $r_{jx}$  value of 0.2296; item "66B" with  $r_{jx}$  value of 0.2298, and item "82A" with  $r_{jx}$  value 0.296). The three items were included in the "reliable" category.

### **FINDINGS**

At the end of the first, second, and third stages of the method, 102 open-ended items were obtained. One of them (Cargo Company) is presented in Figure 5, with the findings obtained in the fourth and fifth stages of the method. Figure 5 shows the open-ended version of the item with a small group, an example of a student solution, and the multiple-choice version obtained due to the revisions.

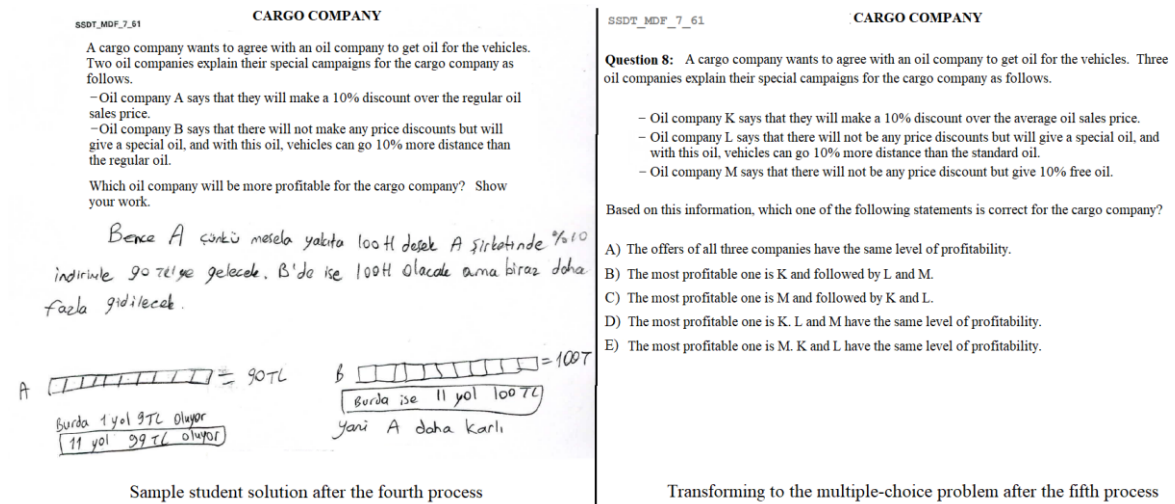


Figure 5. The Sample Item at the End of the Third, Fourth and Fifth Stages

Following the analysis of the student solution above (left), it is seen that what the student understood overlapped with the item content. Hence, the item completed the 4th stage successfully. Only two companies were presented that could be preferred in the open-ended version of the item, making it challenging to write five options for the multiple-choice version. Therefore, in transforming the problem into multiple-choice (in Stage 5), the experts suggested adding a new company to the item. This item was chosen as an example as it underwent a more comprehensive revision than other items. At the end of the fifth stage, the problem's final version is demonstrated in Figure 5 (right).

The sixth stage of the method brought findings related to the item difficulty index and the mathematical literacy sub-dimensions. The number of items at each difficulty level is presented in Table 2. Values range between 0-10. The item is represented with higher index value as it gets challenging.

Table 2. Item Distribution by Difficulty Levels

Difficulty Level	1	2	3	4	5	6	7	8	9	10	Total
Item Number	0	0	0	17	19	26	20	12	2	6	102

In Table 2, the item difficulty levels varied between 4-10. The items of medium difficulty (4, 5, 6) were in the majority. Thus, the findings suggest that the average difficulty levels of the multiple-choice question pool items were 6.21.

The number of items matching the mathematical literacy sub-dimensions is shown in Table 3. It is seen that the number of items to measure quantitative content knowledge in a personal context was higher than in others. The table in which the items were precisely matched with the sub-dimensions is given in Appendix-1.

**Table 3.** The Number of Items per Sub-Dimension

Dimension	Context*				Math. Content**				Math. Process ***		
	SCI	PER	OCP	SOC	UD	CR	QNT	SS	FRM	EMP	INT
Sub-dimension											
Item Number	11	48	28	15	10	27	49	16	41	36	25
Total				102				102			102

\*Contexts: SCI: scientific, PER: personal, OCP: occupational, SOC: societal.

\*\*Mathematical Contents: UD: uncertainty and data, CR: change and relationships, QNT: quantity, SS: space and shape

\*\*\*Mathematical Processes: FRM: formulating situations mathematically, EMP: employing mathematical concepts, facts, procedures, and reasoning, INT: Interpreting, applying and evaluating mathematical outcomes

For the example of the "Cargo Company" question given in Figure 5, the estimated item difficulty was "7" out of 10; the context was "occupational"; the mathematical content was "change and relationships"; the mathematical process was "formulating situations mathematically." The item was labeled with these features and added to the multiple-choice question pool with the correct answer "D" (The most profitable one is K. L and M have the same profitability level.) It has been. In the seventh stage of the method, a draft multiple-choice question pool was created by labeling all items with the features given in this sample. In the eighth stage, item distribution to the booklets was made according to the findings in Table 2 and Table 3 and the matching "item- mathematical literacy sub-dimension" in Appendix-1.

The first and second field studies' findings were obtained from the ninth and tenth stages, respectively. The findings were evaluated for the items to be used or revised. The item discrimination indices regarding the field studies were considered in decision making. Ninety-eight mathematical literacy items were tested in the first field study. Thirty-three of those items were revised. Forty-four items were obtained from the 33 revised items (more than one revised version was derived from certain items). Findings related to the revised items were examined in two categories of reason. The *first category* contained 29 items, and the *second category* included four items. Explanations about the categories are provided below.

The *first category* included the revised items ( $r_{jx} < 0.3$ ). Table 4 shows the  $r_{jx}$  values of the first application's 29 items and the 40 revised items of the second application.

**Table 4.** The Comparison of Item Discrimination Indices in the First Category

Item ID	$r_{jx}$ (1 <sup>st</sup> Application)	Item ID (revised version)	$r_{jx}$ (2 <sup>nd</sup> Application)
1	0.127	1R1	0.255
10	0.087	10R1	0.107
10	0.087	10R2	0.118
11	0.244	11R1	0.442
13	0.293	13R1	0.458
22	0.234	22R1	0.327
27	0.296	27R1	0.313
29	0.158	29R1	0.245
31	0.154	31R1	0.244
34	0.124	34R1	0.23
34	0.124	34R2	0.584
34	0.124	34R3	0.225
34	0.124	34R4	0.317
39	0.124	39R1	0.216

40	0.176	40R1	0.45
53	0.29	53R1	0.381
58	0.277	58R1	0.269
58	0.277	58R2	0.118
60	0.257	60R2	0.255
60	0.257	60R1	0.262
60	0.257	60R3	0.277
61	0.228	61R1	0.301
65	0.237	65R1	0.312
79	0.079	79R1	0.328
83	0.202	83R1	0.145
84	0.195	84R2	0.163
84	0.195	84R3	0.13
84	0.195	84R1	0.092
19A	0.247	19AR1	0.373
19A	0.247	19AR2	0.593
19B	0.19	19BR1	0.345
19B	0.19	19BR2	0.588
23A	0.281	23AR1	0.315
49B	0.225	49BR1	0.15
66A	0.152	66AR1	0.28
66B	0.298	66BR1	0.225
70B	0.238	70BR1	0.26
76B	0.196	76BR1	0.142
82A	0.296	82AR1	0.597
82B	0.258	82BR1	0.293

As seen in Table 4, the 14 revised versions out of 29 items in the first category (11, 13, 22, 27, 34, 40, 53, 61, 65, 79, 19A, 19B, 23A, 82A) met the criteria of  $r_{jk} \geq 0.3$ . Two revised versions of 3 items among the 14 items (34, 19A, 19B) met the requirement of  $r_{jk} \geq 0.3$ . Thus, there were 17 items met the  $r_{jk} \geq 0.3$  conditions at the end of the revision.

The 19A and 19B items in the first category are given below to exemplify the revisions. Those items were chosen as examples since the revision gave the most significant statistical results for those items. The changes were scanned on the item. Also, difficulty and discrimination indices and reliability coefficients of the items before and after revision are presented in Figure 6.

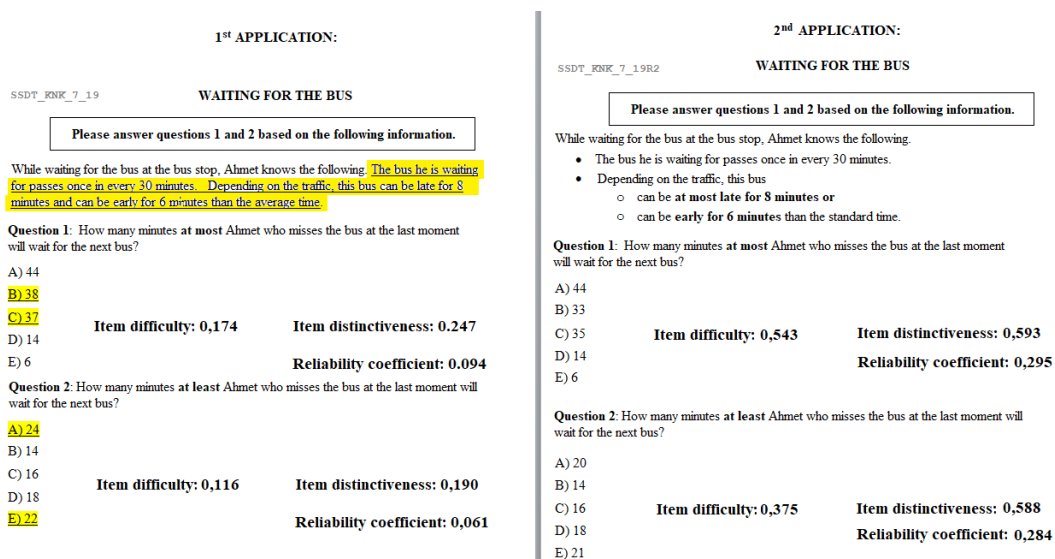


Figure 6. Revision Procedures and Item Statistics for Items 19A and 19B

When the statistics of 19A and 19B items are examined in Figure 6, it can be seen that the revisions significantly increased the item discrimination index and met the  $r_{jx} \geq 0.3$  conditions. According to the item difficulty analysis, both items became easy and approached 0.5 level. Both changes in the  $r_{jx}$  and  $p_j$  values positively affected the reliability coefficient. Those findings statistically proved that the revisions successfully ensured the validity and reliability of the item.

Although the *second category* was  $r_{jx} \geq 0.3$  based on the first application data, more students in the upper group accumulated in the wrong option/s than the lower group. There were four items in this context. The  $r_{jx}$  values of the items in this category are given in Table 5.

**Table 5.** Comparison of Item Discrimination Indices in the Second Category

Item ID	1 <sup>st</sup> Application $r_{jx}$	Item ID (revised version)	2 <sup>nd</sup> Application $r_{jx}$
6	0.325	6R1	0.315
7	0.372	7R1	0.409
9	0.326	9R1	0.458
51	0.411	51R1	0.57

As given in Table 5, the discrimination indices of items 7, 9, and 51 ID in the second category increased due to the revisions. The statistics of item 9 ID ("Dam") and its revised version (9R1) are presented in Table 6. Besides, the revision processes are shown on the item in Figure 7.

**Table 6.** Item Statistics about the First and Second Application of the "Dam" Problem

Item ID	Group	Options									
		*N <sub>A</sub>	A**	N <sub>B</sub>	B	N <sub>C</sub>	C***	N <sub>D</sub>	D	N <sub>E</sub>	E
9	Total	42	0.151	76	0.273	73	0.263	27	0.097	57	0.205
9	Upper	17	0.128	27	0.203	57	0.429	8	0.06	21	0.158
9	Lower	12	0.154	34	0.436	8	0.103	8	0.103	16	0.205
9	Difference	5	-0.026	-7	-0.233	49	0.326	0	-0.042	5	-0.047
9R1	Total	37	0.144	76	0.296	66	0.257	37	0.144	38	0.148
9R1	Upper	10	0.145	9	0.13	35	0.507	5	0.072	9	0.13
9R1	Lower	15	0.185	32	0.395	4	0.049	16	0.198	13	0.16
9R1	Difference	-5	-0.04	-23	-0.265	31	0.458	-11	-0.125	-4	-0.03

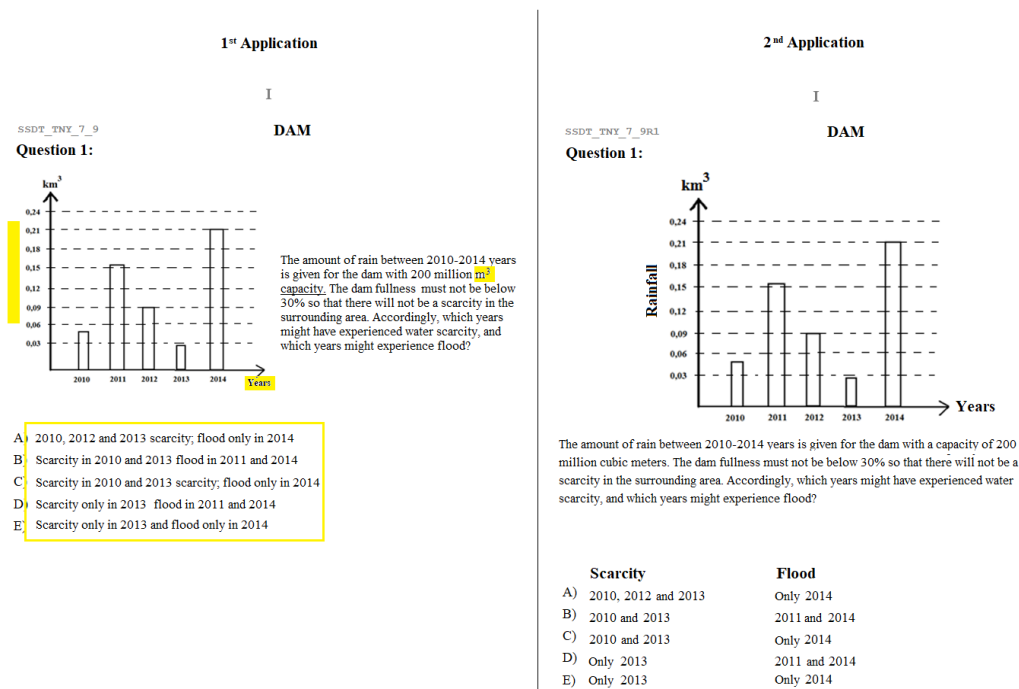
\*The number of students that mark the option

\*\*The ratio of students that mark the option to the total student number

\*\*\*Correct option

According to the Table 6, it is seen that the ratio of students who marked A, D and E options from the upper and lower groups is close to each other according to the first application data of the question. These data suggest that these 3 options may be problematic, however, the number of options that can be derived from the content of the question is limited and only similar options can be derived. Accordingly, it was thought that the problem might not arise from the options themselves, but from the fact that the options were written close each other. Therefore, instead of changing the options, it was decided to write the options of the question more discretely (like a table). The item before and after the revision is presented in Figure 7. Those items were chosen as

examples since they were visually revised among the items in the category. The changes are scanned or framed on the item.



**Figure 7.** Revisions for Item 9 ID

As seen in Figure 7, besides the options, the graphic and text in the item were written and placed obviously on the page. Following the revisions, according to the data from the second application (see 9R1 data in table 6), the item difficulty index remained almost the same (0.263 and 0.257). The discrimination index increased from 0.326 to 0.458, and the reliability coefficient increased from 0.143 to 0.2.

The items that were used to show the revision process was chosen considering the item diversity. In this regard, the findings related to the most comprehensive revision process (Figure 5), one of the revisions reviling the most statistically significant difference (Figure 6), and a revision of visual elements (Figure 7) were explained above. The similar revision procedures described above were performed for the 33 items. Other revision procedures, in which the items' validity were proved statistically and exemplified in seven items, are included in Appendix-3.

**CONCLUSION, DISCUSSION and RECOMMENDATIONS**

As a result of the study, 94 items were evaluated in the category of "Reliable" based on the item discrimination index. Ten of them were the revised versions of the same item. In other words, 10 of the 94 items were two different versions (6-6R1, 7-7R1, 9-9R1, 19AR1-19AR2, 19BR1-19BR2, 27-27R1, 34R1-34RR2, which were considered "reliable." 51-51R1, 70A-70AR1, 82A-82AR1) version. Hence, eighty-four different items were obtained. Item difficulty indices varied between 0.158 and 0.845, and reliability coefficients ranged between 0.113 and 0.364. Those items were added to the multiple-choice item pool (item bank) of the study. The statistics related to all items are presented in Appendix-1, and 7 items are exemplified in Appendix-2.



The items with discrimination index between 0.3 and 0.2 ( $0,20 \leq r_{jx} < 0,30$ ) were evaluated in the category "must be revised". The items with a discrimination index less than 0.2 ( $r_{jx} < 0,2$ ) were evaluated in the category of "Unreliable or should be completely revised." These items can be revised in the light of statistical data and tested in future applications.

The revision procedures based on the data of the first field tests can be grouped under four headings: the revisions in the items' (i) texts, (ii) images, (iii) options, and (iv) data. The successful revision procedures for any item included only one or two of these headings. No revision procedure involved three or four headings. Although there were items that could be recovered by changing the text, image, or options, no items could only be recovered by changing the data. The technically detailed revision procedures revealed the following results that increased or decreased the validity of multiple-choice questions.

1. If the text states that different situations occur in different conditions, noting every condition and outcome (see for comparison Appendix-3, item ID: 34 → 34R1) increases the validity.
2. If the item includes explanations both above and below the figure/table, the emphasized statement should be presented after the figure/table (see for comparison, Appendix-3, item ID: 22 → 22R1). In other words, the presentation of the emphasized content at the end (or towards the end) of the item text increases the validity.
3. There can be more than one visual (e.g., object, photograph, figure, picture) in an item. In such items (see Appendix-3, item ID: 65 → 65R1), the visuals should be ordered to reflect the relationship between them (i.e., process, flow) and should include expressions related to this order in the text (e.g., by a particular criterion or from right to left or top to bottom), which also increases validity. Besides, if there are texts or numbers on the image, the item text should involve what they represent, which indicates the validity.
4. Other factors that increased validity in this study included using bold letters to emphasize particular words (only one or two words) and lining up the options instead of side by side (see for comparison, appendix-3, item ID: 53 → 53R1).
5. The removal of expressions like "What is given in the item is not enough for a result" from the options increases validity (see appendix-3, item ID: 11 → 11R1 and 34R1 → 34R4).
6. The exclusion of the most uncomplicated option increases the validity (see for comparison appendix-3, item ID: 40 → 40R1).
7. Revisions to reduce the number of procedures in the items that require only separate operations increase the validity (see for comparison, appendix-3, item ID: 79 → 79R1).

The first four of the results above overlap with the literature findings on writing multiple-choice test items (e.g., clarity and consistency between text and images) (Turgut & Baykul, 2010; Atılgan et al., 2017; Tekin, 2019). However, current results involve more specific and technical details and expressions that can guide the item authors.

The fifth, sixth and seventh results above are related to the items' cognitive characteristics. The fifth finding results from a random number of students from the upper group (enough to reduce the  $r_{jx}$  value below 0.30)

choose the option when the given items are insufficient for the solution. If that option is not the correct answer, it significantly and negatively affects the item validity. If the option is longer than others (Turgut & Baykul, 2010), it may indicate a non-cognitive and structural factor that partially directs students to choose that option. However, although the "A" option was longer than the other options in both 34R1 and 34R4, no student from the upper group selected the option "A" in the 34R4 ID number, which shows the unavailability of such a structural factor here.

The sixth result stemmed from a random number of students from the upper group (enough to reduce the rjx value below 0.30) choose the expected option. Accordingly, if the option is not the correct answer, it significantly and negatively affects the item's validity. The fifth and sixth findings are peculiar to the current study and are not found in the literature. These findings would contribute methodologically to the literature on preparing and revising multiple-choice mathematics items. If there are no other negative factors identified for the question, in both of these two cases, it can be ensured that the problem reaches a usable level of validity by removing only this type of option and replacing it with an ordinary option.

The seventh finding refers to the repetition of many operations in solving the item. This situation increased the item's difficulty and significantly reduced its validity (enough to reduce the rjx value below 0.30). This finding overlaps with Başol's (2019) results indicating that students should avoid too tricky items, and the items that require only separate operations for a solution have low validity.

In the literature on developing mathematical literacy items, the number of studies on the term "PISA-like questions" and especially the research on the PISA items is high (Kamaliyah et al., 2013; Oktiningrum et al., 2016; Murdiyani, 2018; Nasution et al., 2019). However, it may reduce the number of producing original mathematical literacy items and the possibility of providing a wide variety of contexts for the produced items. On the contrary, addressing the description, aim, and scope of mathematical literacy increases the originality, diversity, and validity of the items as the tests, including items written in various contexts, can measure students' mathematical literacy. The study findings suggest that the criterion is not the similarity between PISA items and other but appropriateness to the definition, goal, and scope of mathematical literacy. This distinction is an essential feature that distinguishes the items in the current study from the literature items. The fact that many items were tested increases the variety and validity. Thus, the items that can measure students' performance in each sub-dimension of mathematics literacy (OECD, 2013c) have been presented.

Appendix-1 includes the item statistics and the mathematical literacy sub-dimensions. There are at least nine items for measuring each sub-dimension in the multiple-choice item pool. By combining the items obtained from the current study with different study results, mathematical literacy tests can be produced in different sub-dimensions (context, mathematical content, mathematical process) and level (easy, medium, difficult). In this sense, test development studies can be performed.

Among the 102 items, the number of the ones to measure mathematical content knowledge in personal context and quantity was higher than others (see table 3). According to Demir and Altun (2018), "association" is considered as an opportunity by item writers (i.e., writing an item based on a daily life event/situation). It can be said that item writers encounter problems in daily life that require quantitative mathematical content knowledge for a solution, which should be limited to the authors of the studies in which the items are compiled. However, it is possible to generalize with research seeking answers to the question: How are the situations encountered in daily life and need mathematical knowledge in solution to be distributed to these sub-dimensions? Such research can provide data on distributing items into sub-dimensions for validity in a mathematical literacy test.

Although multiple-choice questions are more suitable for measuring large samples, they offer fewer opportunities to reflect students' mathematical literacy skills (e.g., reasoning, communication, representation) and to provide researchers with fewer opportunities in this regard than open-ended questions. The items in the open-ended item pool created in the third stage of the study method can be evaluated. The competencies required by each problem can be determined. Thus, items can be used as a tool in studies that involve measuring students' mathematical literacy proficiency.

The items obtained in this study are of significant use for in-class activities and textbooks. Textbook exercises often do not provide significant opportunities for reasoning and communicating skills related to mathematical literacy, because most textbook exercises generally include the practice and implementation of routine methods (Fan & Zhu, 2000; Stacey & Vincent, 2009; Thompson, Senk & Johnson, 2012). The studies in the national literature also found the inadequate use of mathematical literacy items in lessons and textbooks (Altun & Akkaya, 2014; Çakır, 2009; Dede & Yaman, 2005; İskenderoğlu & Baki, 2011), and studies are indicating that mathematics education at school is disconnected from the real world. According to Begel (1973; cited by Wong and Low, 2020), the textbook's content is a variable that "can be both manipulated and can affect student learning." Thus, the items obtained from this research can be used as a tool in item and activity development studies in mathematical literacy. Teachers can be suggested to use these items in mathematics teaching and the assessment and evaluation activities. The suggestions of Wong and Low (2020) regarding "teacher's role" and "nature of the class" can be taken into consideration in order to develop mathematical reasoning and communication skills, which are among the essential mathematical literacy skills.

According to the objective of middle school mathematical curriculum enforced in 2018, "students can improve and perform mathematical literacy skills effectively" (MEB, 2018). It is suggested to increase literacy items in national exams (Jürges et al., 2012; Uysal & Yenilmez, 2011). It can be seen that these studies complement each other regarding national assessment and evaluation activities. However, researchers (Demir, 2015; Oktiningrum et al., 2016) state that students are not familiar with contextual items. Students can become familiar with the contextual items in mathematical teaching using the items obtained from this study. Thus, the content of mathematical teaching and national assessment and evaluation activities can complement each other.

Demir (2019) revealed the processes and stages of mathematical literacy item writing, usually followed by the item writers. Besides, in another study, Demir (2020) designed a mathematical literacy curriculum for prospective teachers and used specific criteria to determine whether a problem is a mathematical literacy item. In this research, the validity and reliability studies of the obtained items were checked by Demir and Altun (2018) findings, which can be seen as a complement to the studies in the literature.

### **Ethical Text**

The practices carried out in the study were approved by Research and Publication Ethics Committee (at the meeting numbered 2020/01, on 21.04.2020) of a state university. In this article, journal writing rules, publishing principles, research and publishing ethics rules, journal ethics rules are followed. Responsibility belongs to the author for any violations related to the article.

**Author Contribution Rate:** This research was supported by the Scientific Research Projects Unit (BAP) of Dumlupınar University (Project Number: 2018-34). The author's contribution rate is 100%. The author designed the study, developed the theoretical framework, collected and analyzed the data and wrote the manuscript. The author declares that there is no conflict of interest. The author would like to thank Mustafa KÖSE and Efkan ÇAY for their precious contributions on expert reviews. He would like to thank Enis Harun BAŞER, Samet DEMİR and Berna TÜRE KÖSE for their precious contributions in the distribution and collection of test documents to practice schools. In addition, the author would like to thank Dumlupınar University scientific research projects coordinatorship, the governorship of Kütahya, the provincial directorate of national education, the administrators of the practice schools, and all the teachers who earnestly supervised the exams for their precious contributions.

### **REFERENCES**

- Acar, T. (2012). Türkiye'nin PISA 2009 sonuçlarına göre OECD'ye üye ve aday ülkeler arasındaki yeri. *Kuram ve Uygulamada Eğitim Bilimleri*, 12(4), 2561-2572.
- Altun, M. & Akkaya, R. (2014). Matematik öğretmenlerinin PISA matematik soruları ve ülkemiz öğrencilerinin düşük başarı düzeyleri üzerine yorumları. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 29(1), 19-34.
- Altun, M. & Bozkurt, I. (2017). Matematik okuryazarlığı problemleri için yeni bir sınıflama önerisi. *Eğitim ve Bilim*, 42(190), 171-188. <http://dx.doi.org/10.15390/EB.2017.6916>
- Annisah, U. (2011). *Pengembangan soal matematika model PISA pada konten quantity untuk mengukur kemampuan penalaran matematis siswa sekolah menengah pertama*. Prosiding Seminar Nasional Pendidikan 2011. Fakultas Keguruan dan Ilmu Pendidikan, Universitas Sriwijaya. Palembang, 7 Mei 2011.
- Atilgan, H., Kan, A. & Doğan, N. (2017). *Eğitimde ölçme ve değerlendirme* (9. bs.). Anı Yayıncılık.
- Aydın, A., Sarier, Y. & Uysal Ş. (2014). PISA sonuçları bağlamında öğrencilerin akademik başarılarının değerlendirilmesi. *İlköğretim Online*, 13(3), 1065-1074.
- Başol, G. (2019). *Eğitimde ölçme ve değerlendirme* (6. bs.). Pegem Akademi.

- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. ÖSYM Yayınları
- Boaler, J. (1993). The role of contexts in the mathematics classroom: Do they make mathematics more “real”? *For the Learning of Mathematics*, 13(2), 12–17.
- Breakspear, S. (2012). The policy impact of PISA: An exploration of normative effects of international benchmarking in school system performance. *OECD Education Working Papers* (No: 71). OECD Publishing.
- Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. Holt, Rinehart and Winston, 6277 Sea Harbor Drive, Orlando, FL 32887.
- Çakır, İ. (2009). *İlköğretim 5. sınıf ders kitaplarının öğretmen ve öğrenci görüşleri doğrultusunda değerlendirilmesi* (Yayımlanmamış yüksek lisans tezi). Çukurova Üniversitesi.
- Dede, Y. & Yaman, S. (2005). *İlköğretim 6, 7 ve 8. sınıf matematik ve fen bilgisi ders kitaplarının incelenmesi: problem kurma ve çözme etkinlikleri bakımından*. XVI. Ulusal Eğitim Bilimleri Kongresi’nde sunulmuş bildiri, Denizli.
- Demir, F. (2015). *Matematik Okuryazarlığı soru yazma süreç ve becerilerinin gelişimi* (Yayımlanmamış doktora tezi). Uludağ Üniversitesi.
- Demir, F. & Altun, M. (2018). Matematik okuryazarlığı soru yazma süreç ve becerilerinin gelişimi. *Eğitim ve Bilim*, 43(194), 19-41. <http://dx.doi.org/10.15390/EB.2018.7111>
- Demir, F. (2019). Matematik okuryazarlığı soru yazma sürecinde yer alan eylemlerin belirlenmesi ve sıralarının kestirilmesi. *Bayburt Eğitim Fakültesi Dergisi*, 14(28), 372-390. <https://doi.org/10.35675/befdergi.6434>
- Demir, F. (2020). Mathematics literacy curriculum design based on question writing. *International Journal of Education Technology and Scientific Researches*, 5(11), 35-62. <http://dx.doi.org/10.35826/ijetsar.71>
- Ekawati, R., Susanti, S. & Chen, J. C. (2020). Primary students’ mathematical literacy: A case study. *Infinity Journal*, 9(1), 49-58.
- Fan, L. & Zhu, Y. (2000). Problem solving in Singaporean secondary mathematics textbooks. *The Mathematics Educator*, 5(1/2), 117-141.
- İskenderoğlu, T. A. & Baki, A. (2011). İlköğretim 8. sınıf matematik ders kitabındaki soruların PISA matematik yeterlilik düzeylerine göre incelenmesi. *Eğitim ve Bilim*, 36(161), 287-300.
- Jürges, H., Schneider, K., Senkbeil, M. & Carstensen, C. H. (2012). Assessment drives learning: The effect of central exit exams on curricular knowledge and mathematical literacy. *Economics of Education Review*, 31(1), 56-65.
- Kamaliyah, K., Zulkardi, Z. & Darmawijoyo, D. (2013). Developing the sixth level of PISA-like mathematics problems for secondary school students. *Indonesian Mathematical Society Journal on Mathematics Education*, 4(1), 9-28.
- Kohar, A. W., Zulkardi, Z. & Darmawijoyo, D. (2014). Developing PISA-like mathematics tasks to promote students’ mathematical literacy. In *Proceeding in the Second South East Asia Design/Development Research (SEA-DR) Conference* (14-26).

- Leikin, R. & Levav-Waynberg, A. (2007). Exploring mathematics teacher knowledge to explain the gap between theory-based recommendations and school practice in the use of connecting tasks. *Educational Studies in mathematics*, 66(3), 349-371.
- Lestariningsih, L., Nurhayati, E., Susilo, T., Cicinidia, C. & Lutfianto, M. (2020). Development of mathematical literacy problems to empower students' representation. *Journal of Physics: Conference Series*, 1464 012018.
- Lloyd, G. M., Herbel-Eisenmann, B. A. & Star, J. R. (2011). *Developing essential understanding of expressions, equations, and functions for teaching mathematics in grades 6–8*. Reston, VA: National Council of Teachers of Mathematics.
- Mardhiyanti, D., Putri, R. I. I. & Kesumawati, N. (2013). Pengembangan soal matematika model PISA untuk mengukur kemampuan komunikasi matematis siswa sekolah dasar. *Jurnal Pendidikan Matematika UNSRI*, 5(1). 16-29. <https://doi.org/10.22342/jpm.5.1.334>.
- MEB. (2018). *Matematik dersi öğretim programı*. Milli Eğitim Bakanlığı.
- MEB. (2019). *PISA 2018 Türkiye raporu eğitim analiz ve değerlendirme raporları serisi No:10*. Milli Eğitim Bakanlığı.
- Murdiyani, N.M. (2018). Developing non-routine problems for assessing students' mathematical literacy. *Journal of Physics: Conference Series*, 983, 1-6.
- Nasution, R. S., Fauzi, K. M. A. & Syahputra, E. (2019). Developing mathematics problem based on PISA level of space and shape content to measure student's mathematics problem solving ability. *American Journal of Educational Research*, 7(10).
- National Council of Teachers of Mathematics. (2002). *Okul matematiğinin prensipleri ve standartları* (Çev. C.H. Böke). NCTM Publications.
- National Governors Association Center for Best Practices and Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. NGA.
- OECD. (2013a). *PISA 2015 draft mathematics framework*. OECD Publishing.
- OECD. (2013b). *PISA 2012 results: excellence through equity: giving every student the chance to succeed (Volume II), PISA*. OECD Publishing.
- OECD, (2013c). *What Students Know and Can Do: Student Performance in Math, Reading and Science Volume I*. PISA OECD Publishing.
- OECD. (2016). *PISA 2015 assessment and analytical framework. Science, reading, mathematics and financial literacy*. OECD Publishing.
- (OECD. (2019) *PISA 2018 results volume1: What students know and can do*. OECD Publishing.
- Oktiningrum, W., Zulkardi, Z. & Hartono, Y. (2016). Developing PISA-like mathematics task with indonesia natural and cultural heritage as context to assess students' mathematical literacy. *Journal on Mathematics Education*, 7, 1-8. <http://dx.doi.org/10.22342/jme.7.1.2812.1-8>
- Özgen, K. (2013). Problem çözme bağlamında matematiksel ilişkilendirme becerisi: öğretmen adayları örneği. *NWSA-Education Sciences*, 8(3), 323-345.

- Reinke, L. T. (2020). Contextual problems as conceptual anchors: an illustrative case. *Research in Mathematics Education*, 22(1), 3-21. <https://doi.org/10.1080/14794802.2019.1618731>
- Robinson, E. E., Robinson, M. F. & Maceli, J. C. (2000). The impact of standards-based instructional materials in mathematics in the classroom. In M. J. Burke (Ed.), *Learning mathematics for a new century* (Vol. 11, 112–126). National Council of Teachers of Mathematics.
- Silva, E. Y. (2011). *Pengembangan soal matematika model PISA pada konten uncertainty untuk mengukur kemampuan pemecahan masalah matematika siswa sekolah menengah pertama*. Prosiding Seminar Nasional Pendidikan 2011. Fakultas Keguruan dan Ilmu Pendidikan, Universitas Sriwijaya. Palembang, 7 Mei 2011.
- Smith, A. (2004). *Making mathematics count*. The report of Professor Adrian Smith's inquiry into post-14 mathematics education. London.
- Stacey, K. & Vincent, J. (2009). Modes of reasoning in explanations in Australian eighth-grade mathematics textbooks. *Educational Studies in Mathematics*, 72(3), 271.
- Stacey, K. (2013). Answer given by Kaye Stacey, a member of the PISA mathematics expert team, to the question asked by the researcher via e-mail, 07.12.2013.
- Tekin, H. (2019). *Eğitimde ölçme ve değerlendirme* (27. bs.). Yargı Yayınevi.
- Tessmer, M. (1993). *Planning and Conducting Formative Evaluations: Improving the Quality of Education and Training*. Kogan Page.
- Thompson, D. R., Senk, S. L. & Johnson, G. J. (2012). Opportunities to learn reasoning and proof in high school mathematics textbooks. *Journal for Research in Mathematics Education*, 43(3), 253-295. <https://doi.org/10.5951/jresmetheduc.43.3.0253>
- Tomlinson, M. (2004). *14-19 curriculum and qualifications reform*. Final report of the working group on 14-19 reform. London.
- Turgut, M. F. & Baykul, Y. (2010). *Eğitimde ölçme ve değerlendirme* (2. bs.). Pegem Akademi.
- Uysal, E. & Yenilmez, K. (2011). Sekizinci sınıf öğrencilerinin matematik okuryazarlığı düzeyi. *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi*, 12(2), 1-15.
- Wardono, M. S. & Mariani, S. (2014). The Realistic Learning Model With Character Education And PISA Assessment To Improve Mathematics Literacy. *International Journal of Education and Research*, 2(7), 361-372.
- Wong, L. F. & Low, L. (2020). Strategies to foster mathematical reasoning and communication. In N. H. Lee, C. Seto, R. A. Rahim and L. S. Tan (Eds.) *Mathematics Teaching in Singapore*, 107-122. World Scientific.
- Wongpanya, N., Pativisan, S. & Ngamman, J. (2018). *Development of mathematics test items*. 37. Annual Conference of International Association for Educational Assessment. Accessed from <https://iaea.info/documents/development-of-mathematics-test-items/>.
- Yuentrakulchai, T., Kamtet, W. & Dechsri, P. (2018). *Development of mathematics test items*. 37. Annual Conference of International Association for Educational Assessment. Accessed from <https://iaea.info/documents/development-of-science-test-items/>.

**APPENDIX-1 ITEM IDENTIFICATION TABLE**

No	Item ID	Context*	Math. Content**	Math. Process***	$P_j$	$r_{jk}$	$S_j$	$r_j$
1	2	OCP	CR	EMP	0,33	0,46	0,47	0,21
2	3	PER	QNT	EMP	0,33	0,38	0,47	0,18
3	4	OCP	QNT	EMP	0,76	0,41	0,43	0,18
4	5	PER	QNT	EMP	0,52	0,51	0,50	0,25
5	6	PER	SS	EMP	0,19	0,33	0,39	0,13
6	6R1	PER	SS	EMP	0,20	0,32	0,40	0,13
7	7	PER	QNT	FRM	0,26	0,37	0,44	0,16
8	7R1	PER	QNT	FRM	0,33	0,41	0,47	0,19
9	8	PER	CRs	FRM	0,36	0,46	0,48	0,22
10	9	SOC	CR	INT	0,26	0,33	0,44	0,14
11	9R1	SOC	CR	INT	0,26	0,46	0,44	0,20
12	11R1	PER	SS	FRM	0,33	0,44	0,47	0,21
13	12	OCP	SS	EMP	0,63	0,70	0,48	0,34
14	13R1	OCP	CR	FRM	0,24	0,46	0,43	0,19
15	14	OCP	SS	EMP	0,25	0,36	0,43	0,15
16	15A	SOC	QNT	EMP	0,56	0,59	0,50	0,29
17	15B	SOC	QNT	EMP	0,79	0,45	0,41	0,18
18	15C	SOC	QNT	FRM	0,54	0,69	0,50	0,34
19	16	SOC	UD	INT	0,73	0,47	0,45	0,21
20	17	PER	QNT	FRM	0,44	0,64	0,50	0,32
21	18	SCI	CR	FRM	0,28	0,46	0,45	0,21
22	19AR1	PER	QNT	FRM	0,24	0,37	0,43	0,16
23	19BR1	PER	QNT	FRM	0,18	0,35	0,38	0,13
24	19AR2	PER	QNT	FRM	0,54	0,59	0,50	0,30
25	19BR2	PER	QNT	FRM	0,38	0,59	0,48	0,28
26	20	PER	QNT	EMP	0,80	0,47	0,40	0,19
27	21	OCP	CR	INT	0,66	0,71	0,47	0,34
28	22R1	PER	SS	EMP	0,31	0,33	0,46	0,15
29	23AR1	PER	QNT	INT	0,79	0,32	0,41	0,13
30	23B	PER	QNT	INT	0,39	0,47	0,49	0,23
31	24A	OCP	CR	FRM	0,51	0,57	0,50	0,28
32	24B	OCP	CR	FRM	0,18	0,39	0,38	0,15
33	25	PER	QNT	EMP	0,49	0,62	0,50	0,31
34	26	SOC	CR	FRM	0,47	0,60	0,50	0,30
35	27	OCP	CR	INT	0,18	0,30	0,38	0,11
36	27R1	OCP	CR	INT	0,16	0,31	0,36	0,11
37	28A	SCI	QNT	EMP	0,72	0,49	0,45	0,22
38	28B	SCI	QNT	EMP	0,58	0,53	0,49	0,26
39	30	PER	QNT	EMP	0,66	0,50	0,47	0,24
40	32A	PER	QNT	EMP	0,78	0,55	0,42	0,23
41	32B	PER	QNT	EMP	0,41	0,55	0,49	0,27
42	33	OCP	SS	INT	0,33	0,48	0,47	0,23
43	34R2	PER	CR	FRM	0,42	0,58	0,49	0,29
44	34R4	PER	CR	FRM	0,23	0,32	0,42	0,13
45	37	SOC	UD	INT	0,30	0,32	0,46	0,15
46	38	PER	QNT	FRM	0,32	0,52	0,47	0,24
47	40R1	PER	QNT	EMP	0,48	0,45	0,50	0,22
48	42A	PER	QNT	EMP	0,55	0,67	0,50	0,33
49	42B	PER	QNT	EMP	0,66	0,62	0,47	0,29
50	43	PER	QNT	EMP	0,40	0,61	0,49	0,30
51	44A	SCI	QNT	INT	0,69	0,43	0,46	0,20
52	44B	SCI	QNT	INT	0,26	0,34	0,44	0,15
53	44C	SCI	QNT	INT	0,75	0,40	0,43	0,17
54	44D	SCI	QNT	INT	0,31	0,37	0,46	0,17
55	45	PER	QNT	EMP	0,73	0,44	0,44	0,19
56	46	SOC	SS	INT	0,65	0,44	0,48	0,21
57	47	PER	UD	INT	0,44	0,71	0,50	0,35
58	48A	PER	CR	INT	0,66	0,59	0,47	0,28
59	48B	PER	CR	INT	0,45	0,70	0,50	0,35
60	48C	PER	CR	INT	0,31	0,61	0,46	0,28
61	49A	OCP	QNT	FRM	0,36	0,48	0,48	0,23
62	50	OCP	QNT	FRM	0,52	0,64	0,50	0,32
63	51	OCP	UD	INT	0,49	0,41	0,50	0,21
64	51R1	OCP	UD	INT	0,54	0,57	0,50	0,28
65	52	OCP	UD	INT	0,76	0,51	0,43	0,22
66	53R1	SOC	UD	INT	0,34	0,38	0,47	0,18
67	54	OCP	QNT	FRM	0,37	0,59	0,48	0,29
68	55	OCP	QNT	FRM	0,43	0,74	0,49	0,36
69	56	PER	QNT	FRM	0,51	0,58	0,50	0,29
70	57	PER	SS	EMP	0,35	0,46	0,48	0,22
71	61R1	OCP	CR	FRM	0,32	0,30	0,47	0,14
72	62	SOC	CR	INT	0,55	0,54	0,50	0,27
73	64	SCI	CR	INT	0,70	0,51	0,46	0,23
74	65R1	SCI	CR	FRM	0,41	0,31	0,49	0,15



75	66B	SOC	SS	FRM	0,22	0,30	0,41	0,12
76	68	PER	QNT	FRM	0,44	0,37	0,50	0,18
77	69	PER	QNT	FRM	0,56	0,52	0,50	0,26
78	70A	PER	CR	EMP	0,30	0,47	0,46	0,21
79	70AR1	PER	CR	EMP	0,35	0,39	0,48	0,19
80	71A	PER	QNT	FRM	0,59	0,72	0,49	0,35
81	71B	PER	QNT	FRM	0,54	0,69	0,50	0,35
82	72A	PER	QNT	EMP	0,32	0,31	0,47	0,15
83	73	OCP	QNT	EMP	0,51	0,50	0,50	0,25
84	74	PER	CR	INT	0,75	0,49	0,43	0,21
85	75	PER	UD	INT	0,62	0,50	0,49	0,24
86	76A=76AR1	PER	UD	FRM	0,45	0,59	0,50	0,29
87	77	OCP	QNT	EMP	0,61	0,55	0,49	0,27
88	78	OCP	QNT	EMP	0,28	0,47	0,45	0,21
89	79R1	PER	QNT	EMP	0,42	0,33	0,49	0,16
90	80	OCP	CR	FRM	0,27	0,46	0,45	0,20
91	81	PER	QNT	EMP	0,50	0,56	0,50	0,28
92	82A	OCP	QNT	FRM	0,35	0,30	0,48	0,14
93	82AR1	OCP	QNT	FRM	0,34	0,60	0,47	0,28
94	87	SOC	SS	FRM	0,33	0,39	0,47	0,18

\*Contexts: SCI: scientific, PER: personal, OCP: occupational, SOC: societal.

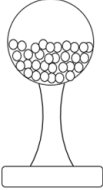
\*\*Mathematical Contents: UD: uncertainty and data, CR: change and relationships, QNT: quantity, SS: space and shape

\*\*\*Mathematical Processes: FRM: formulating situations mathematically, EMP: employing mathematical concepts, facts, procedures, and reasoning, INT: Interpreting, applying and evaluating mathematical outcomes.

APPENDIX-2. SAMPLE ITEMS

**ID\_76A CHEWING GUM MACHINE**

Please answer questions 3 and 4 based on the following information.



The chewing gum machine on the right works with 1 TL\*. For every 1TL, the device gives

- 2 chewing gums for 3 and folds of 3,
- 3 chewing gums for 5 and folds of 5,
- 5 chewing gums for the situation that meet both for having advantageous chewing gums,
- 1 chewing gum in other cases.


For example, when you put 1TL for 3 consecutively, you will have 1 chewing gum from the first one, 1 chewing gum from the second one and 2 chewing gums from the third one. Thus, when you put 3 TL, you will get a total of 4 chewing gums. \*Turkish Monetary Unit

**Question 3:** How many times should Seval put money to have the minimum average cost per chewing gum?

A) 27 times  
B) 28 times  
C) 29 times  
D) 30 times  
E) 31 times

**ID\_66B FOREST**

A road is constructed around the protected forest area given in the figure to intervene in a possible fire and enable people to take a walk. The diameter of the forest area is approximately 400 meters. The fire trucks can intervene at most 50 meters away from their location.



**ITEM 6:** New roads are planned to be added for the fire trucks to intervene in the forest fire. New routes will be added so that the fire trucks can intervene in the entire forest and cut down the minimum number of trees. What is the minimum length of roads to meet the required conditions?

A) 800 meters  
B) 700 meters  
C) 600 meters  
D)  $(300\pi + 200)$  meter  
E)  $(200\pi + 200)$  meter


**ID\_33 PHOTOCOPY**

**Question 7:** An advertisement agency copies an 8 cm x 12 cm rectangular banner in different sizes. In the photocopy process, the 8/12 ratio of the banner is preserved. For this operation, the enlarge/shrink button of the machine is used. For example, if this button is set to 200%, each edge of the rectangular banner will become two times longer than the original length, and if the switch is set to 25%, each side of the rectangular banner becomes 1/4 times of the initial period. What percentage should of the button should be set for a 2400 cm<sup>2</sup> banner?

A) 500  
B) 2500  
C) 250  
D) 50  
E) 25

**ID-80 COPPER CABLES**

**Question 9:**




Copper cables will be hanged between two poles with a 40-meter distance. The copper cables expand in summer and shrink in winter. 1°C change in the temperature leads to a 10 cm change in the 100 m copper cable length. The temperature is 35°C when the copper cables are hanged. It is known that the minimum temperature in this region is -30°C in winter. What should be the minimum cable length between two poles to prevent cables from breaking due to shrinking?

A) 3 840 cm  
B) 4 160 cm  
C) 4 260 cm  
D) 3 740 cm  
E) 4 400 cm

**ID\_62 EVERYONE TO THEIR FIELD!**

**ITEM 9:** Anyone between 5 to 24 years old can join the summer basketball course. Course participants are divided into 5 groups based on their ages and they train on five different basketball fields.

The first group has students between 5-7 years old, the second group has students between 8-10 years old, the third group has students between 11-13 years old, the fourth group has students between 14-16 years old, and the fifth group has students 17 years old or above.



1 Basket Height  
2 Basket Diameter  
3 Diameters of Balls  
4 Weights of Balls

The four properties (Basket Height, Basket Diameter, Diameters of Balls, Weights of Balls) are different for the five basketball fields

The basket diameter for five basketball fields and the widths of the balls are given as follows.

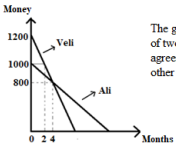
	Basket Diameter:	Ball Diameter:
I. Field	48 cm	19 cm
II. Field	50 cm	19 cm
III. Field	46 cm	22 cm
IV. Field	45 cm	24 cm
V. Field	45 cm	22 cm

Guess the correct group and field match.

	I. Field	II. Field	III. Field	IV. Field	V. Field
A) 1. Group	2. Group	3. Group	4. Group	5. Group	
B) 2. Group	1. Group	4. Group	3. Group	5. Group	
C) 2. Group	1. Group	3. Group	5. Group	4. Group	
D) 1. Group	3. Group	4. Group	5. Group	2. Group	
E) 2. Group	1. Group	4. Group	5. Group	3. Group	

**ID\_48 TOUR FRIENDSHIP**

Please answer questions 5, 6, and 7 based on the following information.



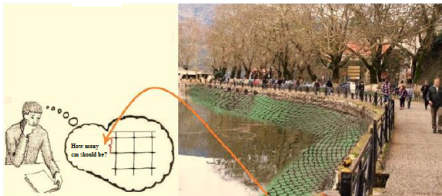
The graphic shows the linear change in the amount of money of two friends on tour across Turkey. According to their agreement before the trip, they will financially support each other till the end of tour.

**Question 5:** ...  
**Question 6:** ...  
**Question 7:** Who and for how many months supported during the tour?

A) Veli to Ali for 2 months.  
B) Veli to Ali for 8 months.  
C) Ali to Veli for 2 months.  
D) Ali to Veli for 8 months.  
E) Ali to Veli for 12 months.

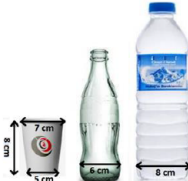
**ID\_87 MUNICIPALITY IS WORKING!**

**Question 6:**



To prevent water pollution, the metropolitan municipality decided to construct a large net consisting of equal squares to cover the Porsuk river's coastline and plans to tie the net, as seen in the picture above. Hence, it will be easier to collect waste, and people's awareness will increase.

Before constructing the net, the waste accumulated in the Porsuk River is collected. Most common waste and measurements of these wastes are given.



Accordingly, which one of the following cannot be the length of one edge of the net's squares?

- I.  $\frac{\pi}{2}$  cm    II.  $\pi$  cm    III.  $2\pi$  cm    IV.  $3\pi$  cm    V.  $4\pi$  cm
- A) Only V  
B) IV and V  
C) III, IV and V  
D) II, III, IV and V  
E) I, II, III, IV and V

APPENDIX-3. SAMPLE REVISION PROCEDURES

**ID\_34 FLOORING**

**QUESTION 6:** Mr. Ali wants to cover the kitchen floor with square tiles. The price offers from three different firms are as follows. In the first offer, 20-cm-square tiles are 2 TL, in the second offer, 30-cm-square tiles are 3 TL, in the third offer, 40-cm-square tiles are 4 TL.

What offer is more profitable for Mr. Ali?

A) If we do not know the kitchen floor's width and length, we cannot determine the most profitable offer.

B) The most profitable offer is the 20-cm-square tiles.

C) The most profitable offer is the 30-cm-square tiles.

D) The most profitable offer is the 40-cm-square tiles.

E) All three offers have the same profitability.

**ID\_34R1 FLOORING**

**QUESTION 6:** Mr. Ali wants to cover the kitchen floor with square tiles. The price offers from three different firms are as follows.

- In the first offer, 20-cm-square tiles are 2 TL,
- In the second offer, 30-cm-square tiles are 3 TL,
- In the third offer, 40-cm-square tiles are 4 TL.

What offer is more profitable for Mr. Ali?

A) If we do not know the kitchen floor's width and length, we cannot determine the most profitable offer.

B) The most profitable offer is the 20-cm-square tiles.

C) The most profitable offer is the 30-cm-square tiles.

D) The most profitable offer is the 40-cm-square tiles.

E) All three offers have the same profitability.

**ID\_34R4 FLOORING**

**QUESTION 10:** Mr. Ali wants to cover the kitchen floor with square tiles. The price offers from three different firms are as follows.

- In the first offer, 20-cm-square tiles are 2 TL,
- In the second offer, 30-cm-square tiles are 3 TL,
- In the third offer, 40-cm-square tiles are 4 TL.

What offer is more profitable for Mr. Ali?

A) The first and second offers are equally profitable, and both are more profitable than the third offer.

B) The most profitable offer is the 20-cm-square tiles.

C) The most profitable offer is the 30-cm-square tiles.

D) The most profitable offer is the 40-cm-square tiles.

E) All three offers have the same profitability.

Item ID	Revise	Group	Option 1 N	Option 1	Option 2 N	Option 2	Option 3 N	Option 3	Option 4 N	Option 4	Option 5 N	Option 5
34		TOTAL	100	0,36	15	0,054	11	0,04	34	0,122	118	0,424
34		High	39	0,293	5	0,038	2	0,015	25	0,188	62	0,466
34		Low	31	0,397	9	0,115	8	0,103	5	0,064	25	0,321
34		Diff	8	-0,104	-4	-0,078	-6	-0,088	20	* 0,124	37	0,146
34		R1 TOTAL	87	0,339	22	0,086	14	0,054	50	0,195	82	0,319
34		R1 High	21	0,304	1	0,014	1	0,014	21	0,304	25	0,362
34		R1 Low	35	0,432	13	0,16	5	0,062	6	0,074	22	0,272
34		R1 Diff	-14	-0,128	-12	-0,146	-4	-0,047	15	* 0,23	3	0,091
34		R4 TOTAL	9	0,03	27	0,089	15	0,049	70	0,23	183	0,602
34		R4 High	0	0	2	0,019	1	0,01	44	0,427	56	0,544
34		R4 Low	6	0,066	15	0,165	8	0,088	10	0,11	52	0,571
34		R4 Diff	-6	-0,066	-13	-0,145	-7	-0,078	34	* 0,317	4	-0,028

**ID\_22 ONE TIME**

**Question 9:** Ayşe, who has bought a new house, will lay up tiles for her bathroom. Everywhere, except the ceiling and door, will be tiled in the bathroom. The bathroom plan and the measurements of the door and one tile are given below.

The tiles are sold with a box with 12 tiles. When the extra and unopened tile boxes are taken back, the price is refunded at the end of the job. Ayşe's house is very far away from the store she bought the tiles, and for delivery, 250 TL\* transportation fee was demanded in addition to the tile fee. Therefore, Ayşe must carry the tiles on her own. Otherwise, it will be expensive to bring back the extra tiles or reorder if there are missing tiles. \*Turkish Monetary Unit

Since everywhere, except the ceiling and door, will be tiled in the bathroom, how many boxes of tiles should Ayşe order?

A) 11  
B) 12  
C) 13  
D) 14  
E) 15

**ID\_22R1 ONE TIME**

**Question 9:** Ayşe, who has bought a new house, will lay up tiles for her bathroom. The bathroom plan and the measurements of the door and one tile are given below.

The tiles are sold with a box with 12 tiles. When the extra and unopened tile boxes are taken back, the price is refunded at the end of the job. Ayşe's house is very far away from the store she bought the tiles, and for delivery, 250 TL\* transportation fee was demanded in addition to the tile fee. Therefore, Ayşe must carry the tiles on her own. Otherwise, it will be expensive to bring back the extra tiles or reorder if there are missing tiles. \*Turkish Monetary Unit

Since everywhere, except the ceiling and door, will be tiled in the bathroom, how many boxes of tiles should Ayşe order?

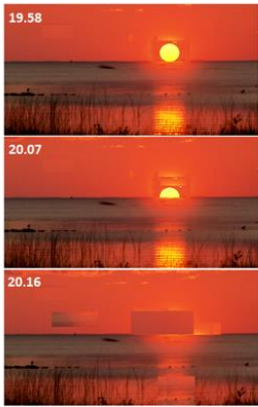
A) 11  
B) 12  
C) 13  
D) 14  
E) 15

Item ID	Revise	Group	Option 1 N	Option 1	Option 2 N	Option 2	Option 3 N	Option 3	Option 4 N	Option 4	Option 5 N	Option 5
22		TOTAL	53	0,175	89	0,294	62	0,205	48	0,158	41	0,135
22		High	23	0,181	54	0,425	20	0,157	16	0,126	11	0,087
22		Low	21	0,183	22	0,191	27	0,235	19	0,165	23	0,2
22		Diff	2	-0,002	32	* 0,234	-7	-0,077	-3	-0,039	-12	-0,113
22		R1 TOTAL	48	0,165	90	0,309	54	0,186	61	0,21	28	0,096
22		R1 High	13	0,157	41	0,494	10	0,12	16	0,193	3	0,036
22		R1 Low	20	0,175	19	0,167	25	0,219	26	0,228	17	0,149
22		R1 Diff	-7	-0,019	22	* 0,327	-15	-0,099	-10	-0,035	-14	-0,113

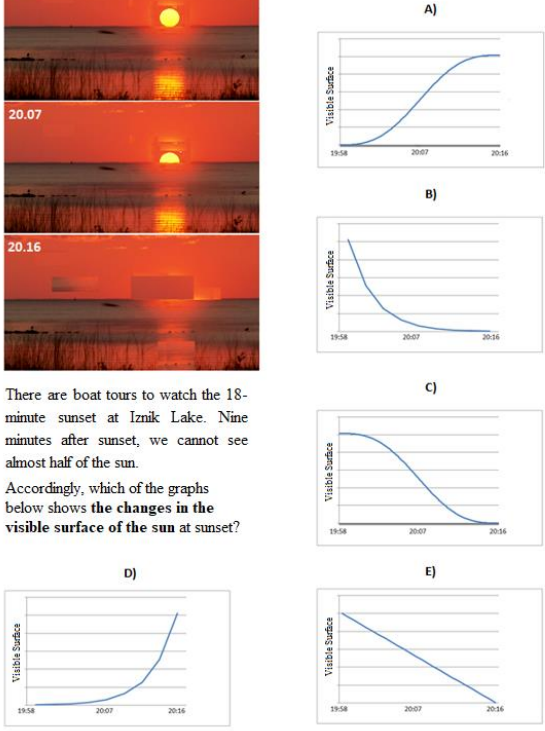
**ID\_65** **SUNSET**

**QUESTION 7**




There are boat tours to watch the 18-minute sunset at Iznik Lake. Nine minutes after sunset, we cannot see almost half of the sun.

Accordingly, which of the graphs below shows the changes in the visible surface of the sun at sunset?



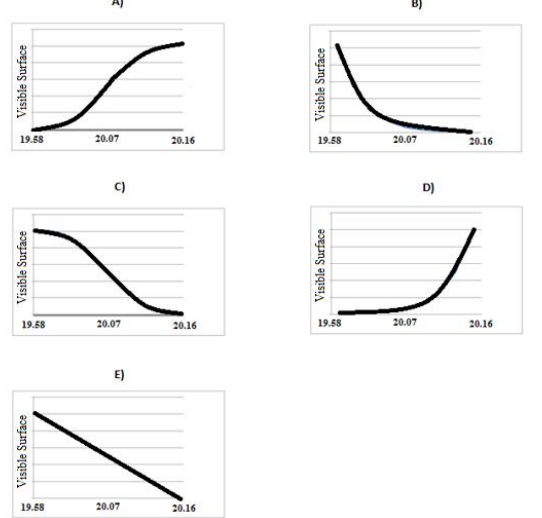
**ID\_65R1** **SUNSET**

**QUESTION 7**



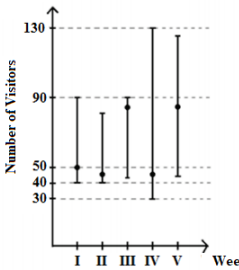
There are boat tours to watch the 18-minute sunset at Iznik Lake. The photographs above were shot during sunset in Iznik, and the time of shoots was written on them. Nine minutes after sunset, we cannot see almost half of the sun.

Accordingly, which of the graphs below shows the changes in the visible surface of the sun at sunset?



Item ID	Revise	Group	Option 1 N	Option 1	Option 2 N	Option 2	Option 3 N	Option 3	Option 4 N	Option 4	Option 5 N	Option 5
65		TOTAL	6	0,019	44	0,143	137	0,445	6	0,019	115	0,373
65		High	0	0	11	0,098	62	0,554	0	0	39	0,348
65		Low	5	0,042	21	0,175	38	0,317	4	0,033	52	0,433
65		Diff	-5	-0,042	-10	-0,077	24	* 0,237	-4	-0,033	-13	-0,085
65	R1	TOTAL	14	0,047	35	0,117	121	0,406	14	0,047	113	0,379
65	R1	High	2	0,019	5	0,049	61	0,592	0	0	35	0,34
65	R1	Low	8	0,064	20	0,16	35	0,28	8	0,064	53	0,424
65	R1	Diff	-6	-0,045	-15	-0,111	26	* 0,312	-8	-0,064	-18	-0,084

**ID\_53** **MUSEUM LOVERS**

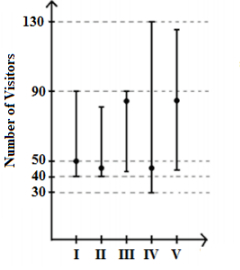


The graphic shows the information on the 5-week number of visitors that passed through the ticket office of a museum open 7 days of a week. Bottom ends of the bars show the number of visitors on the quietest days, upper ends show the number of visitors on the busiest days, dots show the daily average number of visitors in the related week.

Accordingly, which one of the following cannot be the total number of visitors in any week?

A) 273 B) 308 C) 350 D) 420 E) 560

**ID\_53R1** **MUSEUM LOVERS**



The graphic shows the information on the 5-week number of visitors that passed through the ticket office of a museum open 7 days of a week. Bottom ends of the bars show the number of visitors on the quietest days, upper ends show the number of visitors on the busiest days, dots show the daily average number of visitors in the related week.

Accordingly, which one of the following cannot be the total number of visitors in any week?

A) 273  
B) 308  
C) 350  
D) 420  
E) 560

Item ID	Revise	Group	Option 1 N	Option 1	Option 2 N	Option 2	Option 3 N	Option 3	Option 4 N	Option 4	Option 5 N	Option 5
53		TOTAL	85	0,282	37	0,123	38	0,126	35	0,116	97	0,322
53		High	56	0,427	8	0,061	13	0,099	15	0,115	36	0,275
53		Low	12	0,138	18	0,207	19	0,218	7	0,08	30	0,345
53		Diff	44	* 0,29	-10	-0,146	-6	-0,119	8	0,034	6	-0,07
53	R1	TOTAL	92	0,337	38	0,139	37	0,136	25	0,092	74	0,271
53	R1	High	65	0,524	14	0,113	14	0,113	5	0,04	24	0,194
53	R1	Low	11	0,143	18	0,234	12	0,156	16	0,208	18	0,234
53	R1	Diff	54	* 0,381	-4	-0,121	2	-0,043	-11	-0,167	6	-0,04

ID_11	GARDEN SHARE		ID_11R1	GARDEN SHARE																																																																																																																					
<p><b>Question 1:</b> Two siblings will share a 100 m<sup>2</sup> garden. The young one told the old one that "You will have all the land you can cover with this thread; the rest is mine." The thread is 30 meters long.</p> <p>Which shape should the older sibling choose to have the maximum land?</p>																																																																																																																									
<p>A) <b>What is given in the question is not enough to calculate which one is profitable.</b></p> <p>B) Equilateral triangle</p> <p>C) Square</p> <p>D) Rectangle</p> <p>E) Circle</p>		<p>A) Isosceles triangle</p> <p>B) Equilateral triangle</p> <p>C) Square</p> <p>D) Rectangle</p> <p>E) Circle</p>																																																																																																																							
<table border="1" style="width: 100%; border-collapse: collapse; font-size: small;"> <thead> <tr> <th>Item ID</th> <th>Revise</th> <th>Group</th> <th>Option 1 N</th> <th>Option 1</th> <th>Option 2 N</th> <th>Option 2</th> <th>Option 3 N</th> <th>Option 3</th> <th>Option 4 N</th> <th>Option 4</th> <th>Option 5 N</th> <th>Option 5</th> </tr> </thead> <tbody> <tr> <td>11</td> <td></td> <td>TOTAL</td> <td>122</td> <td>0,396</td> <td>36</td> <td>0,117</td> <td>33</td> <td>0,107</td> <td>38</td> <td>0,123</td> <td>78</td> <td>0,253</td> </tr> <tr> <td>11</td> <td></td> <td>High</td> <td>45</td> <td>0,409</td> <td>10</td> <td>0,091</td> <td>9</td> <td>0,082</td> <td>5</td> <td>0,045</td> <td>41</td> <td>0,373</td> </tr> <tr> <td>11</td> <td></td> <td>Low</td> <td>35</td> <td>0,376</td> <td>14</td> <td>0,151</td> <td>14</td> <td>0,151</td> <td>18</td> <td>0,194</td> <td>12</td> <td>0,129</td> </tr> <tr> <td>11</td> <td></td> <td>Diff</td> <td>10</td> <td>0,033</td> <td>-4</td> <td>-0,06</td> <td>-5</td> <td>-0,069</td> <td>-13</td> <td>-0,148</td> <td>29</td> <td>* 0,244</td> </tr> <tr> <td>11</td> <td></td> <td>R1 TOTAL</td> <td>35</td> <td>0,115</td> <td>51</td> <td>0,168</td> <td>51</td> <td>0,168</td> <td>66</td> <td>0,217</td> <td>101</td> <td>0,332</td> </tr> <tr> <td>11</td> <td></td> <td>R1 High</td> <td>8</td> <td>0,078</td> <td>11</td> <td>0,107</td> <td>14</td> <td>0,136</td> <td>12</td> <td>0,117</td> <td>58</td> <td>0,563</td> </tr> <tr> <td>11</td> <td></td> <td>R1 Low</td> <td>13</td> <td>0,143</td> <td>19</td> <td>0,209</td> <td>23</td> <td>0,253</td> <td>25</td> <td>0,275</td> <td>11</td> <td>0,121</td> </tr> <tr> <td>11</td> <td></td> <td>R1 Diff</td> <td>-5</td> <td>-0,065</td> <td>-8</td> <td>-0,102</td> <td>-9</td> <td>-0,117</td> <td>-13</td> <td>-0,158</td> <td>47</td> <td>* 0,442</td> </tr> </tbody> </table>					Item ID	Revise	Group	Option 1 N	Option 1	Option 2 N	Option 2	Option 3 N	Option 3	Option 4 N	Option 4	Option 5 N	Option 5	11		TOTAL	122	0,396	36	0,117	33	0,107	38	0,123	78	0,253	11		High	45	0,409	10	0,091	9	0,082	5	0,045	41	0,373	11		Low	35	0,376	14	0,151	14	0,151	18	0,194	12	0,129	11		Diff	10	0,033	-4	-0,06	-5	-0,069	-13	-0,148	29	* 0,244	11		R1 TOTAL	35	0,115	51	0,168	51	0,168	66	0,217	101	0,332	11		R1 High	8	0,078	11	0,107	14	0,136	12	0,117	58	0,563	11		R1 Low	13	0,143	19	0,209	23	0,253	25	0,275	11	0,121	11		R1 Diff	-5	-0,065	-8	-0,102	-9	-0,117	-13	-0,158	47	* 0,442
Item ID	Revise	Group	Option 1 N	Option 1	Option 2 N	Option 2	Option 3 N	Option 3	Option 4 N	Option 4	Option 5 N	Option 5																																																																																																													
11		TOTAL	122	0,396	36	0,117	33	0,107	38	0,123	78	0,253																																																																																																													
11		High	45	0,409	10	0,091	9	0,082	5	0,045	41	0,373																																																																																																													
11		Low	35	0,376	14	0,151	14	0,151	18	0,194	12	0,129																																																																																																													
11		Diff	10	0,033	-4	-0,06	-5	-0,069	-13	-0,148	29	* 0,244																																																																																																													
11		R1 TOTAL	35	0,115	51	0,168	51	0,168	66	0,217	101	0,332																																																																																																													
11		R1 High	8	0,078	11	0,107	14	0,136	12	0,117	58	0,563																																																																																																													
11		R1 Low	13	0,143	19	0,209	23	0,253	25	0,275	11	0,121																																																																																																													
11		R1 Diff	-5	-0,065	-8	-0,102	-9	-0,117	-13	-0,158	47	* 0,442																																																																																																													

ID_40	TOASTING BREAD		ID_40R1	TOASTING BREAD																																																																																																																					
<p><b>Question 8:</b> Ece will toast one slice of bread for her son, husband, and her. But she can only fit two slices in the pan at one time. One side of the bread is toasted in 5 minutes, what is the minimum time for Ece to toast both sides of 3 slices of bread?</p>																																																																																																																									
<p>A) <b>20</b></p> <p>B) 17,5</p> <p>C) 15</p> <p>D) 12,5</p> <p>E) 10</p>		<p>A) 17,5</p> <p>B) 15</p> <p>C) 12,5</p> <p>D) 10</p> <p>E) <b>7,5</b></p>																																																																																																																							
<table border="1" style="width: 100%; border-collapse: collapse; font-size: small;"> <thead> <tr> <th>Item ID</th> <th>Revise</th> <th>Group</th> <th>Option 1 N</th> <th>Option 1</th> <th>Option 2 N</th> <th>Option 2</th> <th>Option 3 N</th> <th>Option 3</th> <th>Option 4 N</th> <th>Option 4</th> <th>Option 5 N</th> <th>Option 5</th> </tr> </thead> <tbody> <tr> <td>40</td> <td></td> <td>TOTAL</td> <td>210</td> <td>0,675</td> <td>12</td> <td>0,039</td> <td>68</td> <td>0,219</td> <td>5</td> <td>0,016</td> <td>15</td> <td>0,048</td> </tr> <tr> <td>40</td> <td></td> <td>High</td> <td>95</td> <td>0,669</td> <td>2</td> <td>0,014</td> <td>40</td> <td>0,282</td> <td>2</td> <td>0,014</td> <td>3</td> <td>0,021</td> </tr> <tr> <td>40</td> <td></td> <td>Low</td> <td>66</td> <td>0,695</td> <td>8</td> <td>0,084</td> <td>10</td> <td>0,105</td> <td>2</td> <td>0,021</td> <td>9</td> <td>0,095</td> </tr> <tr> <td>40</td> <td></td> <td>Diff</td> <td>29</td> <td>-0,026</td> <td>-6</td> <td>-0,07</td> <td>30</td> <td>* 0,176</td> <td>0</td> <td>-0,007</td> <td>-6</td> <td>-0,074</td> </tr> <tr> <td>40</td> <td></td> <td>R1 TOTAL</td> <td>76</td> <td>0,26</td> <td>139</td> <td>0,476</td> <td>20</td> <td>0,068</td> <td>43</td> <td>0,147</td> <td>11</td> <td>0,038</td> </tr> <tr> <td>40</td> <td></td> <td>R1 High</td> <td>13</td> <td>0,146</td> <td>65</td> <td>0,73</td> <td>2</td> <td>0,022</td> <td>7</td> <td>0,079</td> <td>1</td> <td>0,011</td> </tr> <tr> <td>40</td> <td></td> <td>R1 Low</td> <td>26</td> <td>0,317</td> <td>23</td> <td>0,28</td> <td>9</td> <td>0,11</td> <td>18</td> <td>0,22</td> <td>6</td> <td>0,073</td> </tr> <tr> <td>40</td> <td></td> <td>R1 Diff</td> <td>-13</td> <td>-0,171</td> <td>42</td> <td>* 0,45</td> <td>-7</td> <td>-0,087</td> <td>-11</td> <td>-0,141</td> <td>-5</td> <td>-0,062</td> </tr> </tbody> </table>					Item ID	Revise	Group	Option 1 N	Option 1	Option 2 N	Option 2	Option 3 N	Option 3	Option 4 N	Option 4	Option 5 N	Option 5	40		TOTAL	210	0,675	12	0,039	68	0,219	5	0,016	15	0,048	40		High	95	0,669	2	0,014	40	0,282	2	0,014	3	0,021	40		Low	66	0,695	8	0,084	10	0,105	2	0,021	9	0,095	40		Diff	29	-0,026	-6	-0,07	30	* 0,176	0	-0,007	-6	-0,074	40		R1 TOTAL	76	0,26	139	0,476	20	0,068	43	0,147	11	0,038	40		R1 High	13	0,146	65	0,73	2	0,022	7	0,079	1	0,011	40		R1 Low	26	0,317	23	0,28	9	0,11	18	0,22	6	0,073	40		R1 Diff	-13	-0,171	42	* 0,45	-7	-0,087	-11	-0,141	-5	-0,062
Item ID	Revise	Group	Option 1 N	Option 1	Option 2 N	Option 2	Option 3 N	Option 3	Option 4 N	Option 4	Option 5 N	Option 5																																																																																																													
40		TOTAL	210	0,675	12	0,039	68	0,219	5	0,016	15	0,048																																																																																																													
40		High	95	0,669	2	0,014	40	0,282	2	0,014	3	0,021																																																																																																													
40		Low	66	0,695	8	0,084	10	0,105	2	0,021	9	0,095																																																																																																													
40		Diff	29	-0,026	-6	-0,07	30	* 0,176	0	-0,007	-6	-0,074																																																																																																													
40		R1 TOTAL	76	0,26	139	0,476	20	0,068	43	0,147	11	0,038																																																																																																													
40		R1 High	13	0,146	65	0,73	2	0,022	7	0,079	1	0,011																																																																																																													
40		R1 Low	26	0,317	23	0,28	9	0,11	18	0,22	6	0,073																																																																																																													
40		R1 Diff	-13	-0,171	42	* 0,45	-7	-0,087	-11	-0,141	-5	-0,062																																																																																																													

ID_79	RENT A CAR		ID_79R1	RENT A CAR																																																																																																																					
<p><b>QUESTION 3:</b></p> <table border="1" style="width: 100%; border-collapse: collapse; font-size: small;"> <thead> <tr> <th>Brand</th> <th>Fuel Type</th> <th>Fuel per 100 km (lt)</th> <th>Daily Rental Fee</th> </tr> </thead> <tbody> <tr> <td>Alfa</td> <td>Gasoline</td> <td>5</td> <td>160 TL</td> </tr> <tr> <td>Beta</td> <td>Diesel</td> <td>3,5</td> <td>180 TL</td> </tr> <tr> <td>Gama</td> <td>Diesel</td> <td>3</td> <td>200 TL</td> </tr> <tr> <td>Tetra</td> <td>Gasoline</td> <td>5,5</td> <td>150 TL</td> </tr> </tbody> </table> <p><b>Diesel Price: 6TL/lt and Gasoline Price: 6,5 TL/lt</b></p> <p>Mr. Osman will rent a car for three days to drive from Kütahya to Adana. The distance between Kütahya and Adana is 700 kilometers. Which of the vehicles on the table would be the most economical for Mr. Osman?</p>		Brand	Fuel Type	Fuel per 100 km (lt)	Daily Rental Fee	Alfa	Gasoline	5	160 TL	Beta	Diesel	3,5	180 TL	Gama	Diesel	3	200 TL	Tetra	Gasoline	5,5	150 TL																																																																																																				
Brand	Fuel Type	Fuel per 100 km (lt)	Daily Rental Fee																																																																																																																						
Alfa	Gasoline	5	160 TL																																																																																																																						
Beta	Diesel	3,5	180 TL																																																																																																																						
Gama	Diesel	3	200 TL																																																																																																																						
Tetra	Gasoline	5,5	150 TL																																																																																																																						
<p>A) Alfa</p> <p>B) Beta</p> <p>C) Gama</p> <p>D) Tetra</p> <p>E) There is no difference between vehicles.</p>		<p><b>QUESTION 3:</b></p> <table border="1" style="width: 100%; border-collapse: collapse; font-size: small;"> <thead> <tr> <th>Brand</th> <th>Fuel Type</th> <th>Fuel Fee per 100 km</th> <th>Daily Rental Fee</th> </tr> </thead> <tbody> <tr> <td>Alfa</td> <td>Gasoline</td> <td>32,5 TL</td> <td>160 TL</td> </tr> <tr> <td>Beta</td> <td>Diesel</td> <td>21 TL</td> <td>180 TL</td> </tr> <tr> <td>Gama</td> <td>Diesel</td> <td>18 TL</td> <td>200 TL</td> </tr> <tr> <td>Tetra</td> <td>Gasoline</td> <td>35,75 TL</td> <td>150 TL</td> </tr> </tbody> </table> <p>Mr. Osman will rent a car for three days to drive from Kütahya to Adana. The distance between Kütahya and Adana is 700 kilometers. Which of the vehicles on the table would be the most economical for Mr. Osman?</p>			Brand	Fuel Type	Fuel Fee per 100 km	Daily Rental Fee	Alfa	Gasoline	32,5 TL	160 TL	Beta	Diesel	21 TL	180 TL	Gama	Diesel	18 TL	200 TL	Tetra	Gasoline	35,75 TL	150 TL																																																																																																	
Brand	Fuel Type	Fuel Fee per 100 km	Daily Rental Fee																																																																																																																						
Alfa	Gasoline	32,5 TL	160 TL																																																																																																																						
Beta	Diesel	21 TL	180 TL																																																																																																																						
Gama	Diesel	18 TL	200 TL																																																																																																																						
Tetra	Gasoline	35,75 TL	150 TL																																																																																																																						
<p>A) Alfa</p> <p>B) Beta</p> <p>C) Gama</p> <p>D) Tetra</p> <p>E) There is no difference between vehicles.</p>		<p>A) Alfa</p> <p>B) Beta</p> <p>C) Gama</p> <p>D) Tetra</p> <p>E) There is no difference between vehicles.</p>																																																																																																																							
<table border="1" style="width: 100%; border-collapse: collapse; font-size: small;"> <thead> <tr> <th>Item ID</th> <th>Revise</th> <th>Group</th> <th>Option 1 N</th> <th>Option 1</th> <th>Option 2 N</th> <th>Option 2</th> <th>Option 3 N</th> <th>Option 3</th> <th>Option 4 N</th> <th>Option 4</th> <th>Option 5 N</th> <th>Option 5</th> </tr> </thead> <tbody> <tr> <td>79</td> <td></td> <td>TOTAL</td> <td>30</td> <td>0,103</td> <td>83</td> <td>0,285</td> <td>84</td> <td>0,289</td> <td>83</td> <td>0,285</td> <td>10</td> <td>0,034</td> </tr> <tr> <td>79</td> <td></td> <td>High</td> <td>8</td> <td>0,073</td> <td>32</td> <td>0,291</td> <td>36</td> <td>0,327</td> <td>27</td> <td>0,245</td> <td>7</td> <td>0,064</td> </tr> <tr> <td>79</td> <td></td> <td>Low</td> <td>14</td> <td>0,116</td> <td>37</td> <td>0,306</td> <td>30</td> <td>0,248</td> <td>38</td> <td>0,314</td> <td>2</td> <td>0,017</td> </tr> <tr> <td>79</td> <td></td> <td>Diff</td> <td>-6</td> <td>-0,043</td> <td>-5</td> <td>-0,015</td> <td>6</td> <td>* 0,079</td> <td>-11</td> <td>-0,069</td> <td>5</td> <td>0,047</td> </tr> <tr> <td>79</td> <td></td> <td>R1 TOTAL</td> <td>18</td> <td>0,064</td> <td>118</td> <td>0,418</td> <td>92</td> <td>0,326</td> <td>43</td> <td>0,152</td> <td>11</td> <td>0,039</td> </tr> <tr> <td>79</td> <td></td> <td>R1 High</td> <td>1</td> <td>0,01</td> <td>60</td> <td>0,588</td> <td>25</td> <td>0,245</td> <td>15</td> <td>0,147</td> <td>1</td> <td>0,01</td> </tr> <tr> <td>79</td> <td></td> <td>R1 Low</td> <td>14</td> <td>0,118</td> <td>31</td> <td>0,261</td> <td>51</td> <td>0,429</td> <td>17</td> <td>0,143</td> <td>6</td> <td>0,05</td> </tr> <tr> <td>79</td> <td></td> <td>R1 Diff</td> <td>-13</td> <td>-0,108</td> <td>29</td> <td>* 0,328</td> <td>-26</td> <td>-0,183</td> <td>-2</td> <td>0,004</td> <td>-5</td> <td>-0,041</td> </tr> </tbody> </table>					Item ID	Revise	Group	Option 1 N	Option 1	Option 2 N	Option 2	Option 3 N	Option 3	Option 4 N	Option 4	Option 5 N	Option 5	79		TOTAL	30	0,103	83	0,285	84	0,289	83	0,285	10	0,034	79		High	8	0,073	32	0,291	36	0,327	27	0,245	7	0,064	79		Low	14	0,116	37	0,306	30	0,248	38	0,314	2	0,017	79		Diff	-6	-0,043	-5	-0,015	6	* 0,079	-11	-0,069	5	0,047	79		R1 TOTAL	18	0,064	118	0,418	92	0,326	43	0,152	11	0,039	79		R1 High	1	0,01	60	0,588	25	0,245	15	0,147	1	0,01	79		R1 Low	14	0,118	31	0,261	51	0,429	17	0,143	6	0,05	79		R1 Diff	-13	-0,108	29	* 0,328	-26	-0,183	-2	0,004	-5	-0,041
Item ID	Revise	Group	Option 1 N	Option 1	Option 2 N	Option 2	Option 3 N	Option 3	Option 4 N	Option 4	Option 5 N	Option 5																																																																																																													
79		TOTAL	30	0,103	83	0,285	84	0,289	83	0,285	10	0,034																																																																																																													
79		High	8	0,073	32	0,291	36	0,327	27	0,245	7	0,064																																																																																																													
79		Low	14	0,116	37	0,306	30	0,248	38	0,314	2	0,017																																																																																																													
79		Diff	-6	-0,043	-5	-0,015	6	* 0,079	-11	-0,069	5	0,047																																																																																																													
79		R1 TOTAL	18	0,064	118	0,418	92	0,326	43	0,152	11	0,039																																																																																																													
79		R1 High	1	0,01	60	0,588	25	0,245	15	0,147	1	0,01																																																																																																													
79		R1 Low	14	0,118	31	0,261	51	0,429	17	0,143	6	0,05																																																																																																													
79		R1 Diff	-13	-0,108	29	* 0,328	-26	-0,183	-2	0,004	-5	-0,041																																																																																																													

## MATEMATİK OKURYAZARLIĞI SORU HAVUZU GELİŞTİRME ÇALIŞMASI

### ÖZ

Bu çalışmanın amacı alan yazında matematik okuryazarlığı soru yazma çalışmaları kapsamında üretilmiş soruları dil, yapı ve içerik açısından incelemek, revize etmek, sahada denemek, değerlendirmek ve nihayetinde geçerli, güvenilir ve uygulanabilir maddeleri içeren bir soru havuzu elde etmektir. Çalışmanın yönteminde geliştirme araştırmasının biçimlendirici değerlendirme türü esas alınmıştır. Buna göre küçük grup testleri 8. sınıf, saha testleri 9 ve 10. sınıf öğrencileri ile gerçekleştirilmiştir. Küçük grup testlerine 20-30, birinci ve ikinci saha testlerine sırasıyla 3053 ve 2918 öğrenci katılmıştır. Toplam 102 adet soru denenmiştir. Küçük grup testlerinin ardından öğrenci çözümleri analiz edilmiş ve buna göre revize işlemleri gerçekleştirilmiştir. Birinci ve ikinci saha testlerinin sonunda % 27'lik üst ve alt gruplar yöntemi kullanılmış ve soruların madde analizleri yapılmıştır. Birinci saha testi sonuçlarına göre soruların madde ayırt edicilik değerleri ile üst ve alt gruptaki öğrencilerin çeldiricilere yığılma oranları dikkate alınarak 33 adet soru revize edilmiş ve ikinci saha testinde tekrar denenmiştir. Revize edilen soruların ikinci saha testi analizlerine göre sorularda geçerliği etkileyen yapısal ve bilişsel etkenlere sonuçlarda yer verilmiştir. Bu etkenler revize metodolojisi alanında katkı sağlayabilir. Çalışma sonunda madde ayırt edicilik değeri  $r_{jk} \geq 0,30$  şartını sağlayan 94 adet soru elde edilmiştir. Ayrıca bu soruların her birinin ölçtüğü matematik okuryazarlığı alt boyutları bu çalışma kapsamında belirlenmiştir. Sorulara, soruların eşleştirildiği matematik okuryazarlığı alt boyutlarına ve madde istatistiklerine eklerde yer verilmiştir. Bu çalışma kapsamında elde edilen sorulara matematik eğitimi araştırmalarında, ders içi etkinliklerde ve ölçme değerlendirme faaliyetlerinde yer verilmesinin öğrencilerin matematik okuryazarlığının geliştirilmesi açısından yararlı olacağı düşünülmüştür.

**Anahtar kelimeler:** Madde analizi, matematik okuryazarlığı, PISA, revize metodolojisi, soru havuzu.

---

Bu araştırma, Dumlupınar Üniversitesi Bilimsel Araştırma Projeleri Birimi (BAP) tarafından desteklenmiştir (Proje No: 2018-34). Bu çalışma iki bölüme ayrılmış, 2020 ve 2021 yıllarında sırasıyla şu kongrelerde sunulmuştur: Uluslararası Avrasya Sosyal Bilimler Kongresi-4, Bodrum-Muğla, Türkiye ve 14. Ulusal Fen ve Matematik Eğitimi Kongresi, Burdur, Türkiye.

## GİRİŞ

Uluslararası Öğrenci Değerlendirme Programının [PISA] sonuçları birçok ülkede ulusal bir karne niteliğinde değerlendirilmektedir. Ekonomik İşbirliği ve Kalkınma Örgütü [OECD] tarafından yayınlanan PISA sonuçlarına ek olarak her ülke kendi ulusal raporlarını hazırlamakta ve bu sonuçlara odaklanan birçok ulusal araştırma yürütmektedir. Bunun nedeni, sonuçların ülkelere sadece sıralamaya yönelik bilgiler vermekle kalmayıp ölçüte dayalı birçok değerlendirme yapma fırsatı sunmasıdır. Bu ölçütlerden biri matematik okuryazarlığıdır. Matematik okuryazarlığı, bireyin matematiği yaşamsal ortamlarda çeşitli şekillerde formüle etme, uygulama ve yorumlamada kullanma kapasitesi olarak tanımlanmaktadır (OECD, 2013a; OECD, 2016).

PISA, matematik okuryazarlığı alanındaki sorularını içerdiği (i) bağlamlara, çözümünde gerekli (ii) matematiksel içerik bilgisine ve (iii) matematiksel süreçlere göre sınıflandırmaktadır. Böylece öğrencilere bir bağlam içinde sunulan sorunun çözümünde öğrencilerin matematiksel içerik bilgisini kullanarak çözüm için gerekli matematiksel süreçleri ne düzeyde yürütebildiğini belirlemeyi amaçlamaktadır. Bu üç boyuta ek olarak “matematik okuryazarlığı yeterlik düzeyleri ölçeği” aracılığıyla ülkelere beceri temelli ve yeterlik odaklı yargısal çıkarımlar yapma fırsatı sunmaktadır. Zira 6 düzeyden oluşan bu ölçeğin (OECD, 2019) temelini her bir düzey için anlamlı olan sekiz karakteristik matematiksel beceri teşkil etmektedir. Öğrencilerin performansları arasındaki farklılıklara sözlü bir açıklama vermek amacıyla oluşturulan yeterlik düzeylerinin kapsamı şu yöntemle belirlenmektedir: “Soruların belirlenen örneklem tarafından çözülmesinin ardından, bir zorluk bandına düşen tüm sorulara ve temel matematiksel becerilerin bu soruların çözümünde nasıl gerekli olduğuna bakılmaktadır.” (Stacey, 2013). Ardından öğrencilerin düzeyleri, söz konusu matematiksel becerilere dayalı olarak geliştirilmiş olan matematik okuryazarlığı yeterlik düzeyleri ölçeği üzerine yerleştirilmektedir. Sonuç olarak katılan tüm öğrencilerin sahip olduğu ve kullanabildiği becerilerinden yola çıkarak matematik okuryazarlığı alanındaki yeterlikleri özetlenmiş olur. Böylece katılımcı ülkelerin 15 yaş grubu bireylerinin bu alandaki yeterlikleri hakkında genel bir fikir sahibi olmaları sağlanır. Bu özellikler, PISA uygulamalarından elde edilen puanları daha anlamlı ve daha işlevsel kılmaktadır.

PISA'nın matematik okuryazarlığı sonuçları birçok ülkenin eğitim programlarını yapılandırmasında referans alınmaktadır (Breakspear, 2012). Ayrıca son 30 yılda, araştırmacıların (Lloyd, Herbel-Eisenmann ve Star, 2011) önerileriyle uyumlu olarak İngiltere ve Amerika Birleşik Devletleri'nin de [ABD] içinde bulunduğu bazı ülkeler öğretim programlarında, bağlam içinde matematiksel fikirler geliştirmeye yönelmiştir (Boaler, 1993; Robinson, Robinson ve Maceli, 2000). Tomlinson (2004) ve Smith (2004) tarafından hazırlanan raporların ardından İngiltere, matematik okuryazarlığını okullarında temel bir hedef hâline getirmiştir. ABD çekirdek standartlarında, devlet okullarında öğretilen tüm temel içerik alanlarında (İngilizce, matematik, fen, sosyal) öğrencilerin okuryazarlık becerilerini geliştirmek öncelikli bir hedef olarak ele alınmış ve okuryazarlığa dayalı öğretim teşvik edilmiştir (National Governors Association Center for Best Practices, 2010). Türkiye’de 2018 yılında yürürlüğe giren ortaokul matematik öğretim programının genel amaçlar bölümünde “matematiksel okuryazarlık becerilerini geliştirebilecek ve etkin bir şekilde kullanabilecektir.” (Milli Eğitim Bakanlığı [MEB], 2018) ifadesine, ilk sırada yer

verilmiştir. Öğrencilerin matematik okuryazarlığını geliştirmeye yönelik, ülkelerin bir takım teorik tedbirler aldığı anlaşılmaktadır. Pratikte ise bireylerin matematik okuryazarlığının tespit edilmesine ve geliştirilmesine ilişkin çabalar aşağıda verilen çalışmalar üzerinden değerlendirilebilir.

Ulusal sınavlarda okuryazarlık problemlerinin sayısının artırılması tavsiye edilmiştir (Jürges, Schneider, Senkbeil ve Carstensen, 2012; Uysal ve Yenilmez, 2011). Bu tavsiye olumlu görülmeyle birlikte direk olarak ölçme değerlendirmeye yöneliktir ve kapsamının eğitim öğretim sürecini de içine alacak şekilde genişletilmesi gerekir. Zira öğretim sürecinde verilen temel matematik eğitiminin hedeflerine ne ölçüde ulaşılabildiğinin değerlendirilmesi eğitimin tamamlayıcı bir unsuru olarak görülmelidir (National Council of Teacher of Mathematics, 2002). O hâlde kapsamın anılan şekilde genişletilebilmesi için ders içi etkinliklere kaynaklık edecek matematik okuryazarlığı sorularına ihtiyaç vardır. Bu çalışma bu ihtiyacın karşılanmasına katkı sağlayacak soruları alan yazına kazandırmayı hedeflemektedir. Böyle bir girişim matematik öğretim sürecinin, ölçme değerlendirme ile ilişkisini güçlendirebilir. Ayrıca matematik okuryazarlığının tanımı gereği matematik öğretiminin yaşamla ilişkilendirilmesine katkı sağlayabilir.

Özgen (2013), matematik öğretmen adaylarının matematiği yaşamla ilişkilendirmede güçlük çektiklerini ifade etmiştir. Leikin ve Levav-Waynberg de (2007) ortaokul öğretmenlerinin matematik konularını gerçek yaşamla ilişkilendirmede güçlük çektiklerini, bu nedenle soyut örnekler vermeyi tercih ettiklerini belirtmiştir. Bunun, öğretimin yaşamla ilişkilendirilmekten yoksun olarak yürütülmesine (Altun ve Akkaya, 2014) ve sonuç olarak öğrencilerin matematik okuryazarlık becerilerinin geliştirilememesine neden olduğu söylenebilir. Bu durumda gösterilebilecek çabalardan biri, öğretmenleri ve öğrencileri, PISA tabanlı (bağlamsal) bir öğrenme planlamaya ve PISA'nın özelliklerine göre öğretim değerlendirmeleri tasarlamaya teşvik etmektir (Kohar, Zulkardi ve Darmawijoyo, 2014).

Reinke (2020) standart bir sınıf ortamında bir öğrencinin sonraki görevleri anlamlandırmak için başlangıçta bağlamsal problemlerle uğraşarak geliştirdiği görüntüleri kullandığı bir durum sunarak matematik öğretiminde bağlamların kullanımına ilişkin alanyazına katkıda bulunmuştur. Uysal ve Yenilmez (2011) öğrencilerin matematik okuryazarlık becerilerini geliştirmek için öğrenme ortamlarında PISA'da yer alan türden problem durumları ile sıkça karşı karşıya getirilmesi gerektiğini belirtmiştir. Wardono ve Mariani'ye (2014) göre matematik okuryazarlığının geliştirilmesini desteklemeye yönelik öğrenme araçlarının ve öğrenci etkinliklerinin bulunmaması, öğrencilerin matematiksel okuryazarlık becerilerinin geliştirilememesinin nedenlerinden biridir. Altun ve Bozkurt (2017) öğrenme ortamlarında matematik içeriği olan yaşamsal metinlerin tartışılmasının, öğrencilerde bu tür metinlere aşinalık kazandırılmasının bir ihtiyaç olduğunu belirtmiştir. Matematik öğretiminde bağlamsal sorular ve matematiksel içeriğe sahip metinler eklemek ve bunlarla ilgili sorular sormak, tartışmalar açmak suretiyle bu sorunun giderilebileceğini belirtmiştir. Ekawati, Susanti ve Chen (2020) ortaokul öğrencilerine matematik okuryazarlığı problemlerini öğrenme ve çözme konusunda daha fazla fırsat tanıyacak öğrenme etkinlikleri tasarlamayı önermektedir. Alan yazındaki bu öneriler geçerli, güvenilir ve uygulanabilir matematik



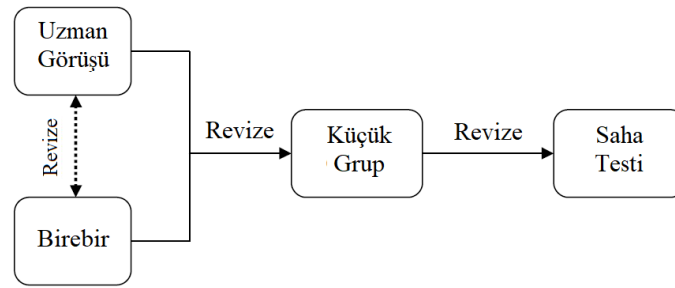
okuryazarlığı soruları-etkinlikleri geliştirmenin bir ihtiyaç olduğunu ortaya koymaktadır. Bu ihtiyaç, araştırmacıları bu tip soruları-etkinlikleri üretmeye yönelik çalışmalara yönlentmiştir.

Annisah (2011) öğrencilerin matematiksel akıl yürütme becerilerini ölçmek için, Mardhiyanti, Putri ve Kesumawati (2013) öğrencilerin matematiksel iletişim becerilerini ölçmek için, Oktiningrum, Zulkardi ve Hartono (2016) öğrencilerin matematik okuryazarlığını ölçmek için, Nasution, Fauzi ve Syahputra (2019) öğrencilerin problem çözme yeteneğini ölçmek için PISA benzeri sorular geliştirmeye yönelik çalışmalar yürütmüştür. Lestariningsih, Nurhayati, Susilo, Cicinidia ve Lutfianto (2020) öğrencilerin temsil (representation) becerileri üzerinde potansiyel etkisi olan matematik okuryazarlığı soruları geliştirmek için çalışmıştır. Bunlara ek olarak Silva (2011) belirsizlik (uncertainty) kategorisinde sorular geliştirip bunları matematik okuryazarlığı yeterlik kümelerine göre kategorize etmiştir. Kamaliyah, Zulkardi ve Darmawijoyo (2013) PISA benzeri altıncı düzey matematik okuryazarlığı soruları geliştirmiştir. Son on yılda bu alanda yukarıda sıralanan uluslararası birçok çalışmaya rastlanmasına, Türkiye'nin PISA'da elde ettiği matematik okuryazarlığı sonuçlarının düşük olmasına (Acar, 2012; Altun ve Akkaya, 2014; Aydın, Sarier ve Uysal, 2014; OECD, 2013b), başarı düşüklüğüne matematik okuryazarlığının geliştirilmesini desteklemeye yönelik soru ve etkinliklerin bulunmamasının bir sebep olarak gösterilmesine (Wardono ve Mariani, 2014) rağmen, matematik okuryazarlığı soruları-etkinlikleri üretmeye yönelik az sayıda ulusal çalışmaya rastlanmıştır.

Yukarıda anılan araştırmaların sonuç ve önerileri bir bütün halinde değerlendirildiğinde şu yargıya ulaşılabilir: Matematik öğretiminin yaşamla ilişkilendirilmesine katkı sağlayacak şekilde tasarlanan ders içi etkinliklere kaynaklık edecek bağlamsal soruları içeren bir soru havuzu güncel bir ihtiyaçtır. Bu nedenle bu çalışmada, geçerli, güvenilir ve uygulanabilir matematik okuryazarlığı soruları üretmek amaçlanmıştır. Bu amaç doğrultusunda, Demir ve Altun'un (2018) ve Demir'in (2019; 2020) öğretmen adayları ile yürüttükleri çalışmalarından elde ettikleri matematik okuryazarlığı soruları dil, yapı ve içerik açısından değerlendirilmiş, geliştirilmiş, uygulanmış ve madde analizine tabi tutulmuştur.

## **YÖNTEM**

Bu çalışmanın amacı alan yazında matematik okuryazarlığı soru yazma çalışmaları kapsamında üretilmiş soruları dil, yapı ve içerik açısından incelemek, revize etmek, sahada denemek, değerlendirmek ve nihayetinde geçerli, güvenilir ve uygulanabilir maddeleri içeren bir soru havuzu elde etmektir. Bu çalışmada yöntem olarak, birbir öğrenci görüşmelerini, uzman incelemesini, küçük grup ve saha testlerini içeren geliştirme araştırmasının (development research) biçimlendirici değerlendirme türü esas (Tessmer, 1993) alınmıştır. Şekil 1'de biçimlendirici değerlendirmenin genel sırası gösterilmiştir.

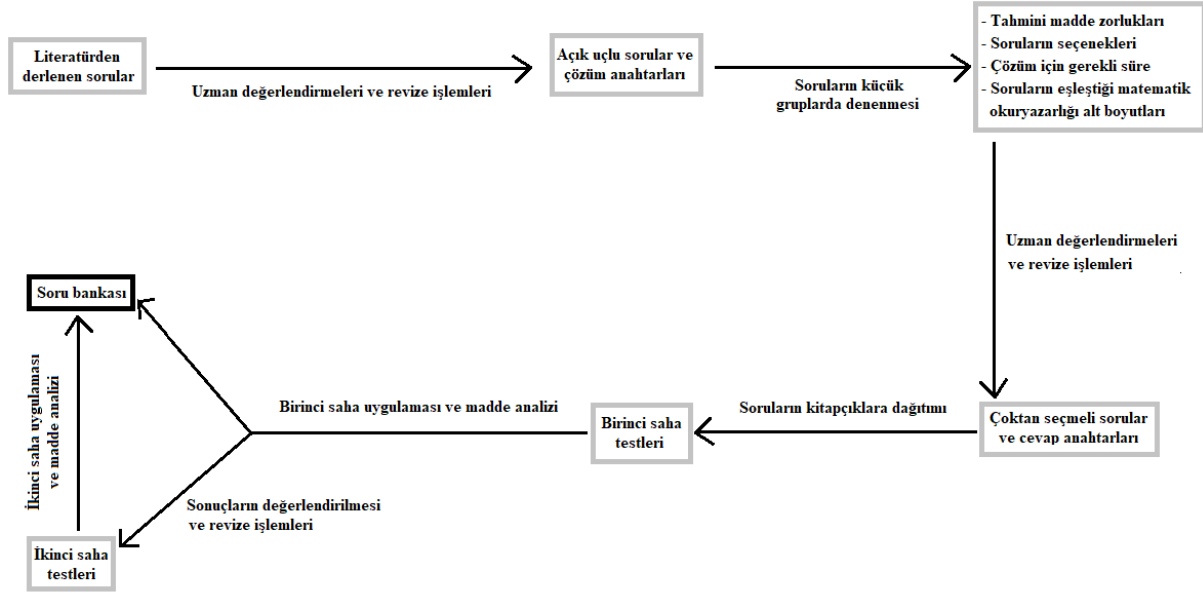


Şekil 1. Biçimlendirici Değerlendirmenin Genel Sırası (Tessmer, 1993)

Tessmer'in (1993) yöntemi birçok soru geliştirme çalışmasında kullanılmıştır (örn. Oktiningrum ve ark., 2016; Nasution ve ark., 2019) ve bu çalışma ile de (çalışmanın amacı itibarıyla) uyumludur. Ancak "birebir" aşamasındaki öğrenci görüşmeleri genellikle geliştirilmesi hedeflenen soru sayısının sadece birkaç tane ile sınırlı olduğu çalışmalar için uygulanabilir görünmektedir. Bu çalışmada ise 102 adet sorunun öğrencilerle deneneceği göz önünde bulundurulmuş ve yöntemin her bir soru için birkaç öğrenci ile görüşme yapmayı içeren "birebir" basamağına yer verilmemiştir. Bu durum, geliştirilmesi hedeflenen soru sayısının fazla olmasından dolayı bu çalışmanın bir sınırlılığı olarak nitelenebilir. Bu sınırlılığın etkisini azaltacak bir önlem alınmış ve bu önlem aşağıda yöntemin dördüncü sürecinde açıklanmıştır. Çalışmanın yönteminde izlenen süreçler detaylandırılarak 10 madde halinde aşağıda sunulmuştur.

- 1) Matematik okuryazarlığı soruları alan yazından derlenmiştir.
- 2) Derlenen soruların dil, yapı ve içerik açısından uzman değerlendirmeleri yapılmış, uzmanların yorumları ve önerileri alınmıştır. Buna göre revize işlemleri gerçekleştirilmiştir.
- 3) Açık uçlu soru havuzu ve çözüm anahtarları oluşturulmuştur.
- 4) Açık uçlu sorular küçük gruplarla denenmiştir. Burada soruların açık uçlu sorulmasının bir nedeni, Tessmer'in (1993) yönteminin "birebir" basamağında sorulara ilişkin topladığı verileri bu çalışmada küçük gruplarla yapılan denemelerde toplamaktır. Öğrencilerin açık uçlu çözümlerinin, görüşmede elde edilebilecek verilerin bir kısmını sağlayabileceği varsayılmıştır. Bu şekilde çalışmanın sınırlılığını azaltmak hedeflenmiştir.
- 5) Küçük gruplardan elde edilen verilere göre sorulara seçenekler yazılmış, çözüm için gerekli süre yaklaşık olarak belirlenmiştir. Bu sürecin sonunda artık çoktan seçmeli tipe dönüşen maddelerin dil, yapı ve içerik açısından uzman değerlendirmeleri yapılmış, uzmanların yorumları ve önerileri alınmıştır. Bunlara göre revize işlemleri yapılmıştır.
- 6) Soruların eşleştiği ağırlıklı matematik okuryazarlığı alt boyutları (bağlam, matematiksel içerik, matematiksel süreç) ve küçük gruplardan elde edilen verilere göre soruların tahmini madde zorlukları belirlenmiştir. Bu şekilde her bir soru alt boyutları ve tahmini madde zorluğu verileri ile etiketlenmiştir.
- 7) Çoktan seçmeli taslak soru havuzu ve cevap anahtarları oluşturulmuştur.
- 8) Sorular, tahmini madde zorlukları ve eşleştiği matematik okuryazarlığı boyutları göz önünde bulundurularak kitapçıklara paylaştırılmıştır.
- 9) 1. saha çalışması yapılmıştır. Bunun verilerine göre madde analizi ve revize çalışmaları yapılmıştır.
- 10) 2. saha çalışması yapılmıştır. Bunun verilerine göre madde analizi yapılmıştır.

Yukarıda verilen süreçlerden ilk 7'si alan yazından yeni soru eklendikçe sürdürülmüştür. Başka bir ifadeyle yeni sorular geldikçe ilk 7 süreçten geçirilmiştir. İlk 7 süreci tamamlayan sorular saha çalışmaları için gerekli izinler alınana kadar biriktirilmiştir. Yöntemin her aşamasında etik kurallara uyulmuştur.



Şekil 2. Yöntemde İzlenen Süreçlerin Modeli

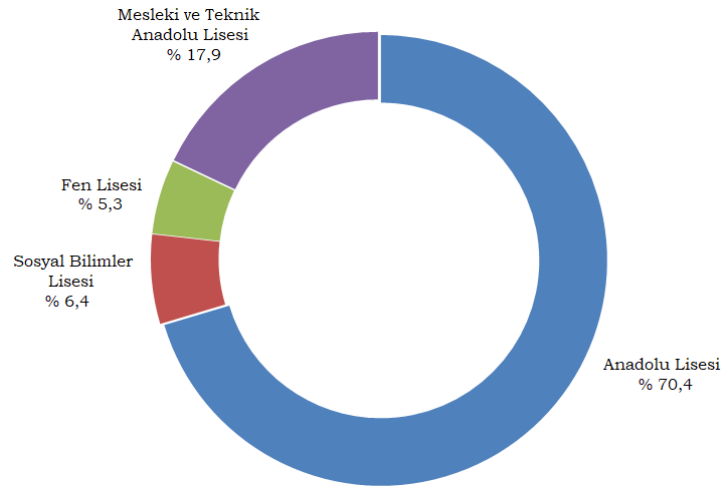
Çalışmanın yönteminde izlenen süreçler Şekil 2'de modellenmiştir. Alan yazında yer alan soru havuzu/bankası geliştirme çalışmalarında da (örn. Wongpanya, Pativisan ve Ngamman, 2018; Yuentrakulchai, Kamtet ve Dechsri, 2018) yöntemle ilişkin benzer modeller kullanılmıştır.

### Çalışma Grubu

Bu araştırmada 2 çalışma grubu yer almaktadır. Birinci çalışma grubu, yöntemin 4. sürecinde yer alan ve uygulamaların yapıldığı günlerde öğrenci sayılarının 20-30 aralığında değiştiği iki adet 8. sınıf şubesidir. Sınıflar Akdeniz bölgesinde bir ilçe merkezinde yer alan bir devlet okulundan seçilmiştir. İlçenin nüfusu yaklaşık 160 000'dir. Öğrencilerin sosyoekonomik durumu orta düzeyde değerlendirilebilir. Bu iki sınıfın seçilmesinin nedeni şu şekilde açıklanabilir: Yöntemin 4. sürecinin en önemli amacı, soruların netliğini (soruda anlatılmak istenen ile, öğrencinin sorudan anladığının örtüşüp örtüşmediğini) test etmektir. Bu nedenle sorular açık uçlu olarak yöneltilmiş ve öğrenciler çözüm için teşvik edilmiştir. Soruların netliğine ilişkin geçerli ve güvenilir veriler elde etmek, öğrencilerin çözüme odaklanmalarına bağlıdır. Öğrencilerin çözüme odaklanmaları için teşvik edilmesinde kendi öğretmenlerinin rolü oldukça büyüktür. Bu iki sınıfın öğrencilerinin yöntemde küçük gruplarla yürütülen açık uçlu soru çözümleri için seçilmesindeki en büyük etken, öğretmenin bu konuda araştırmaya destek sağlamayı kabul etmesidir.

İkinci çalışma grubu, yöntemin 9. ve 10. süreçlerinde yer alan 1. ve 2. saha çalışmalarına katılan öğrencilerdir. İkinci çalışma grubu Ege bölgesinde yer alan bir il merkezinde 12 adet lisenin tüm 9 ve 10. sınıf öğrencileridir. 110

adet sınıftan oluşmaktadır. Bu 110 sınıfta yer alan öğrenciler her iki saha testine de katılmıştır. Saha testlerine sırasıyla 3053 ve 2918 öğrenci katılmıştır. Aradaki fark, uygulamanın yapıldığı günlerde okula gelmeyen öğrencilerden kaynaklanmaktadır. Katılımcıların % 55'i 9. sınıf, % 45'i 10. sınıf öğrencisidir. Başka sınıf düzeylerine göre 9 ve 10. sınıfların tercih edilmesinin nedeni, PISA 2018 Türkiye örnekleminin % 96,5'inin (MEB, 2019) 9 ve 10. sınıf öğrencilerinden oluşmasıdır.



Şekil 3. Saha Testlerine Katılan Öğrencilerin Okul Türlerine Göre Dağılımı

Çalışmaya 8 Anadolu, 2 Mesleki Teknik Anadolu, 1 Fen, 1 Sosyal Bilimler lisesi katılmıştır. Saha testlerine katılan öğrenci sayılarının okul türlerine göre yüzdeleri dağılımı Şekil 3'te verilmiştir. PISA 2018 Türkiye örnekleminin % 81,4'ü (MEB, 2019) bu 4 okul türünün öğrencilerinden oluşmaktadır.

### Veri Toplama Araçları, Verilerin Toplanması ve Analizi

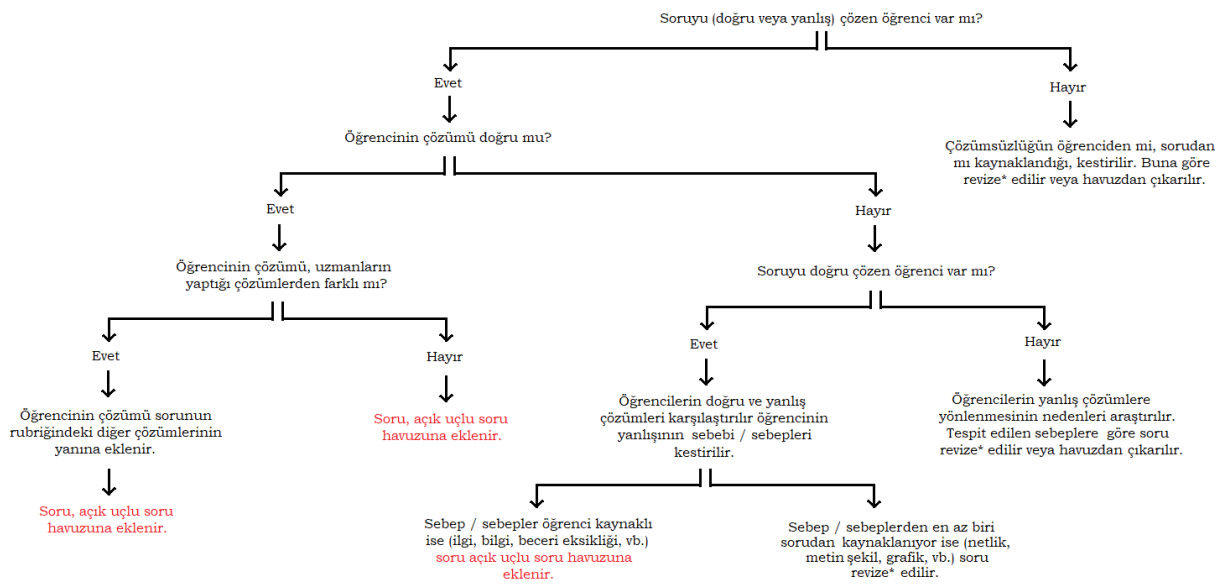
Araştırmanın yöntemi 10 süreç halinde açıklanmıştır. Birinci süreç alan yazından (Demir ve Altun, 2018; Demir 2019; Demir, 2020) matematik okuryazarlığı sorularının derlenmesini içermektedir. Bu sürecin sonunda toplam 102 adet açık uçlu soru derlenmiştir.

İkinci süreç, derlenen soruların dil, yapı ve içerik açısından uzman değerlendirmelerini içermektedir. Bu süreçte değerlendirmeler, iki Türkçe eğitimi uzmanı, bir matematik öğretmeni (küçük grupları oluşturan iki adet sekizinci sınıf şubesinin öğretmeni), bir matematik eğitimi (araştırmacı) ve bir ölçme değerlendirme uzmanı tarafından yapılmıştır. Türkçe eğitimi uzmanları çalışmanın yapıldığı yıl itibarıyla doktora tez aşamasında iki araştırma görevlisidir. Matematik öğretmeni, 8 yıllık tecrübeye sahiptir. Ölçme değerlendirme uzmanı, aynı zamanda matematik öğretmenliği lisansına sahiptir, PISA ve matematik okuryazarlığı alanlarında yayınlanmış akademik çalışmaları mevcuttur. Türkçe eğitimi uzmanları soruları dil açısından değerlendirmiştir. Diğer uzmanlardan soruları çözmeleri, 15 yaş grubu öğrencilerin seviyelerine uygunluğunu değerlendirmeleri, metin, şekil ve netlik açısından yorum ve önerilerini ifade etmeleri istenmiştir. Uzmanların yorum ve önerilerine göre sorular revize edilmiştir.

Üçüncü süreç, her bir sorunun çözüm anahtarının (rubrik) hazırlanmasını içermektedir. Bu sürecin sonunda, sorulardan ve uzmanların yaptığı çözümlerden oluşan açık uçlu soru havuzu ve çözüm anahtarları (rubrikler) elde edilmiştir.

Dördüncü süreç, açık uçlu soru havuzundaki maddeleri 20-30 kişilik küçük gruplarda denemeyi içermektedir. Geçerlik ve güvenilirliği artırmak için bu süreçte öğretmen çeşitli argümanlarla (teşvik edici ifadeler kullanmak, kanaat notu kullanmak, sorulara karşı merak uyandırmak, vb.) öğrencileri çözüme odaklanmaya teşvik etmiştir. Öğrencilerin çözüme odaklanmalarının, sorular hakkında geçerli ve güvenilir sonuçların ortaya çıkma olasılığını artıracığı düşünülmüştür. Bu süreçte her bir soruyu en az 7-8 öğrencinin çözmesi sağlanmıştır. Sürecin sonunda, öğrencilerin çözümleri toplanmıştır.

Beşinci süreçte, öğrenci çözümleri analiz edilmiştir. Bu çözümlerin analizinde takip edilen algoritma Şekil 4'te sunulmuştur. Bu analiz sonucunda açık uçlu soru havuzuna eklenen sorulara, küçük grup uygulamalarında elde edilen öğrenci çözümlerine göre seçenekler yazılmıştır.



\* Revize edilmesi kararlaştırılan maddeler yöntemin ikinci sürecine geri gönderilir.

Şekil 4. Açık Uçlu Soruların Öğrenci Çözümlerinin Analizinde Takip Edilen Algoritma

Beşinci sürecin sonunda çoktan seçmeli tipe dönüşen maddelerin dil, yapı ve içerik açısından uzman değerlendirmeleri yapılmış, uzmanların yorumları ve önerileri alınmıştır. Bunlara göre revize işlemleri yapılmıştır. Ayrıca 10-11 sorunun çözümü için öğrencilere 45 dakikalık süre verilmesinin uygun olduğu belirlenmiştir.

Altıncı süreç soruların, küçük gruplarda elde edilen verilere göre tahmini madde zorluklarını ve uzman değerlendirmelerine göre eşleştiği ağırlıklı matematik okuryazarlığı alt boyutlarını (bağlam, matematiksel içerik, matematiksel süreç) belirlemeyi içermektedir. Tahmini madde zorluğu her bir soru için 0-10 aralığında bir tam sayı olarak belirlenmiştir. Madde zorlaştıkça 10'a daha yakın değerler ile temsil edilmiştir. Soruları matematik okuryazarlığı alt boyutları ile eşleştirmede, PISA raporlarında alt boyutlar için verilen açıklamalar ve yayınlanmış

PISA sorularının sınıflandırmaları ölçüt olarak kabul edilmiştir. Buna göre eşleştirme işlemleri araştırmacı ve ölçme değerlendirme uzmanı tarafından ayrı ayrı yapılmıştır. Bu süreçte farklılık gösteren eşleştirmeler üzerine yukarıda anılan PISA argümanlarına dayandırılan tartışmalar sonucunda madde hakkında karar verilmiştir. Her madde bu özellikleri ile etiketlenmiştir. Yedinci süreç gereği, yukarıdaki 6 süreçten geçen toplam 102 soru cevap anahtarları ile birlikte çoktan seçmeli taslak soru havuzuna eklenmiştir.

Sekizinci süreç, tahmini madde zorluklarını ve eşleştiği matematik okuryazarlığı boyutlarını göz önünde bulundurarak soruları kitapçıklara paylaşımını içermektedir. Burada kitapçıkların zorluk düzeylerinin birbirine yakın olmasını ve her kitapçıkta matematik okuryazarlığının her alt boyutundan yaklaşık olarak eşit sayıda soru bulunmasını sağlamak hedeflenmiştir. Amaç denk kitapçıklar oluşturmak ve böylece saha çalışmalarından elde edilecek verilerin geçerliğini ve güvenilirliğini sağlamaktır. Bu bağlamda birinci uygulama için 98 soru, her biri 10'ar sorudan oluşan 10 farklı kitapçığa paylaştırılmıştır. 4 soru henüz yukarıdaki süreçleri tamamlamadığından ikinci saha çalışmasına bırakılmıştır. Bir kitapçığa düşen soru sayısının (10) belirlenmesinde küçük gruplarla yapılan denemelerde öğrencilerin 45 dakikada çözebildikleri matematik okuryazarlığı soru sayısı dikkate alınmıştır.

Dokuzuncu süreç 1. saha çalışmasını, verilerinin analizini ve revize işlemlerini içermektedir. Bu süreçte ilgili kurumlarla görüşmeler yapılmış, birinci ve ikinci uygulamayı kapsayacak şekilde gerekli resmi izinler alınmıştır. Ardından iznin kapsadığı tüm okullardan dokuzuncu ve onuncu sınıfa devam eden öğrenci sayıları alınmıştır. Birinci uygulama, 6 Mart 2019 Çarşamba günü yapılmıştır.

#### ***Birinci Saha Çalışmasında Geçerlik ve Güvenirliğin Sağlanması İçin Yapılan Çalışmalar***

1. Her bir sınıfın kitapçıkları ayrı zarflara yerleştirilmiştir. Her sınıfın zarfında, 10 kitapçık türünden her birine istisnasız yer verilmiştir. Örneğin 32 kişilik bir sınıfın zarfına her bir kitapçık türünden en az 3'er tane yerleştirilmiştir. Bu işlem, çalışmanın geçerlik ve güvenilirliğine iki açıdan katkı sağlamıştır. Birincisi, kitapçıklardaki sorular farklı olduğundan sınav anında soruların cevaplarına ilişkin öğrencilerin birbiri ile iletişimi önlenmiştir. İkincisi, kitapçıkların tüm örnekleme homojen bir şekilde dağılması ve her bir sorunun maksimum sayıda öğrenci tarafından çözülmesi sağlanmıştır.

2. Sınav evrakları (sınav zarfları, kitapçık ve optik formlar) üzerine öğrenciler ve öğretmenler tarafından dikkate alınması gereken hususlar ayrıntılı bir şekilde yazılmıştır.

3. Sınav evraklarının dağıtımı esnasında okulda sorumlu yöneticiye, zarfların üzerindeki sınav kurallarının gözetmenlik yapacak öğretmenler tarafından mutlaka okunması gerektiği ifade edilmiştir.

4. Sınavların 6 Mart 2019 Çarşamba günü 3. ders saatinde, dersine rastlayan öğretmenin gözetmenliğinde tüm okullarda ve sınıflarda eş zamanlı olarak yapılması sağlanmıştır.

5. Sınav için toplam 50 dakikalık süre verilmiştir. Öğrencilere, kendilerine rastlayan kitapçıkta yer alan 10 adet matematik okuryazarlığı sorusunu çözmeleri için 45 dakikalık süre verilmiştir. Kalan 5 dakika öğrencilerin sınav düzenini alması, kitapçıkların ve optik formların dağıtımı ve toplanması için verilmiştir.

6. Sınavın bitmesi ile birlikte gözetmen öğretmenler tarafından okul idarecilerine teslim edilen sınav evrakları araştırmacı tarafından teslim alınmıştır.

### **Birinci Saha Çalışmasında Toplanan Verilerin Analizi**

Optik forma hatalı kodlama yapmış olmalarından veya optik formda kitapçık türünü kodlamamış olmalarından dolayı 49 öğrencinin formları analize dâhil edilmemiştir. “72B” ID numaralı, “H10” (H: Kitapçık türü, 10: Sorunun sırası) konumlu soru, seçeneklerin hatalı yazılmış olmasından dolayı iptal edilmiş ve analize dâhil edilmemiştir. Böylece kayıp verilerin elenmesinin ardından 3004 öğrenciye ve 97 matematik okuryazarlığı sorusuna ait veriler “% 27’lik üst ve alt gruplar yöntemi” (Baykul, 2000) ile “Test Analysis Program” (TAP) kullanılarak analiz edilmiştir. Analiz sonucunda her bir sorunun madde güçlüğüne ( $p_j$ ), ayırt edicilik gücüne ( $r_{jx}$ ) ve sorulardaki seçeneklerin çeldiriciliğine ilişkin madde istatistikleri elde edilmiştir. Madde güçlük değeri soruya doğru yanıt verenlerin tüm yanıtlayıcı sayısına oranıdır (Atılğan, Kan ve Doğan, 2017). Bu nedenle 0 ve 1 ( $0 \leq p_j \leq 1$ ) aralığında bir değer alır. Bu değer, soru zorlaştıkça 0’a, soru kolaylaştıkça 1’e yaklaşır. Madde geçerlik katsayısı olarak da adlandırılan (Atılğan ve ark., 2017) madde ayırt edicilik değeri, üst grupta soruyu doğru yanıtlayan öğrenci oranından, alt grupta soruyu doğru yanıtlayan öğrenci oranının çıkarılması ile bulunur (Başol, 2019). Bu nedenle -1 ve 1 ( $-1 \leq r_{jx} \leq 1$ ) arasında bir değer alır.

Madde hakkında karar verilmesi: Madde güçlüğü ( $p_j$  değeri), oluşturulmak istenen testin güçlüğüne göre değişiklik gösterebilir. Crocker ve Algina’ya (1986) göre ayırt edicilik gücü ( $r_{jx}$  değeri),

- 0,3’ten büyük veya 0,3’e eşit ( $r_{jx} \geq 0,3$ ) olan maddeler (sorular) iyi işlemektedir, teste olduğu gibi kullanılabilir.
- $0,20 \leq r_{jx} < 0,29$  ise soru, düzeltilmek suretiyle kullanılabilir.
- $r_{jx} < 0,20$  ise soru, kullanılmamalı veya tamamen düzeltilerek kullanılmalıdır.

Analiz, maddelerin sadece ayırt edicilik gücü ( $r_{jx}$ ) ile sınırlandırılmamıştır. Bunun yanı sıra her bir maddenin seçeneklerine ilişkin veriler de tek tek incelenmiştir. Üst ve alt gruplara dayalı madde analizlerinde, herhangi bir çeldiricinin üst gruptakilerden çok, alt gruptaki öğrencileri kendine çekmesi beklenir (Tekin, 2019). Tüm bunlar bir arada değerlendirilerek aşağıda sıralanan iki şarttan herhangi birini taşıyan maddeler revize edilmiştir.

- i.  $r_{jx} < 0,3$  olmasıdır (29 madde).
- ii.  $r_{jx} \geq 0,3$  olmasına rağmen alt gruba göre, üst gruptan daha yüksek oranda öğrencinin yanlış seçeneğe/seçeneklere yığılmış olmasıdır (4 madde).

Tablo 1’de bu iki şarta göre revize edilen 33 adet sorunun konuları, ID numaraları ve revize nedenleri verilmiştir.

Tablo 1. Maddeler ve Revize Nedenleri

Konum	Soru ID	Revize Nedeni	Konum	Soru ID	Revize Nedeni
A1	1	i	F2	60	i
D4	6	ii	D8	61	i
D2	7	ii	F7	65	i
I1	9	ii	D3	79	i
A2	10	i	F10	83	i
E1	11	i	J5	84	i
I2	13	i	B1	19A	i
C9	22	i	B2	19B	i
B6	27	i	C2	23A	i
J3	29	i	H3	49B	i
C4	31	i	H5	66A	i
I6	34	i	H6	66B	i
B7	39	i	G10	70B	i
B8	40	i	I4	76B	i
H1	51	ii	I7	82A	i
H7	53	i	I8	82B	i
A10	58	i			

Tablo 1’de verilen 33 sorunun her biri arařtırmacı ve matematik öğretmenliđi lisansına sahip bir ölçme deđerlendirme uzmanı tarafından birinci uygulamanın verilerine göre Őekil, metin, ierik ve seenekler aısından gözden geirilmiş ve revize edilmiştir. Bazı soruların birden fazla revize versiyonu türetilmiş ve toplamda 44 adet revize versiyon soru elde edilmiştir. Böylece revize sonucunda, kullanılabilir sorular ortaya koyma olasılıđı artırılmıştır. Bu revize versiyonların her biri, sorunun ID numarasının ardına “R1”, “R2”, ... yazılarak isimlendirilmiştir. Örneđin “60” ID numaralı sorunun ikinci saha alışmasında kullanılmak üzere türetilen 3 farklı revize versiyonu “60R1”, “60R2” ve “60R3” Őeklinde isimlendirilmiştir.

Tablo 1’de verilen 33 soruya ikinci saha alışmasının kitapıklarında yer verilmemiştir. Bu 33 sorudan türetilen 44 soru (revize versiyonlar) ikinci saha alışmasının kitapıklarına yerleřtirilmiştir. Ayrıca birinci uygulamada denenmemiş 4 yeni soru da yöntemin önceki süreçlerini tamamladıđı ve oktan semeli taslak soru havuzuna eklendiđi için kitapıklara yerleřtirilmiştir. Soruların kitapıklara dađıtımı, yöntemin 8. sürecinde verilen ölçütlere göre yapılmıştır. Bununla birlikte aynı soruya ait farklı revize versiyonlar ikinci saha alışmasında farklı kitapık türlerine dađıtılmıştır. Böylece ikinci saha uygulamasında kullanılacak kitapıklar hazır hale getirilmiştir.

Yöntemin onuncu süreci 2. saha alışmasını ve verilerinin analizini içermektedir. Sınavın geerliđi ve güvenilirliđi için birinci uygulamada alınan tüm önlemler ikinci uygulamada da alınmıştır. İkinci uygulama, 3 Mayıs 2019 Cuma günü birinci uygulamanın yapıldıđı okullarda yine aynı sınıfların katılımı ile geekleřtirilmiştir.

### **İkinci Saha alışmasında Toplanan Verilerin Analizi**

Optik forma hatalı kodlama yapmış olmalarından veya optik formda kitapık türünü kodlamamış olmalarından dolayı 72 öğrencinin formları analize dâhil edilmemiştir. Böylece kayıp verilerin elenmesinin ardından 2846 öğrenciye ait veriler “% 27 ‘lik üst ve alt gruplar yöntemi” (Baykul, 2000) ile “Test Analysis Program” (TAP) kullanılarak analiz edilmiştir. Analiz sonucunda her bir sorunun madde güçlüđüne ( $p_j$ ), ayırt edicilik gücüne ( $r_{jx}$ ) ve sorulardaki seeneklerin eldiriciliđine iliřkin madde istatistikleri elde edilmiştir.



Birinci ve ikinci uygulamadan elde edilen  $p_j$  ve  $r_{jx}$  değerlerine bağlı olarak (Başol, 2019) tüm maddelerin standart sapmaları ( $s_j = \sqrt{p_j \cdot (1 - p_j)}$ ) ve güvenilirlik katsayıları ( $r_j = r_{jx} \cdot s_j$ ) hesaplanmıştır. Bu değerlerin tümü çalışmanın bulgularında (Ek-1'de) sunulmuştur.

Tıpkı birinci uygulama verilerinin değerlendirilmesinde olduğu gibi, burada da Crocker ve Algina'nın (1986) ölçütlerine göre maddeler "olduğu gibi kullanılabilir", "düzeltilmek suretiyle kullanılabilir" ve "kullanılmamalı veya tamamen düzeltilmelidir" şeklinde üç kategoriye ayrılmıştır. Veri analizi işlemlerinin sonunda,  $0,29 < r_{jx} < 0,295$  aralığındaki madde ayırt edicilik gücü değerleri  $r_{jx} = 0,29$  olacak şekilde,  $0,295 \leq r_{jx} < 0,3$  aralığındaki madde ayırt edicilik gücü değerleri  $r_{jx} = 0,3$  olacak şekilde yuvarlanmıştır. Yuvarlama işlemi, üç maddenin ( $r_{jx}$  değeri 0,296 olan "27",  $r_{jx}$  değeri 0,298 olan "66B", ve  $r_{jx}$  değeri 0,296 olan "82A" ID numaralı soruların) kategori değiştirmesine neden olmuştur. Bu üç madde "olduğu gibi kullanılabilir" kategorisine dahil edilmiştir.

## BULGULAR

Yöntemin birinci, ikinci ve üçüncü süreçlerinin sonunda 102 adet açık uçlu soru elde edilmiştir. Bu sorulardan biri (Kargo Şirketi) yöntemin dördüncü ve beşinci sürecinde elde edilen bulguları da içerdiği haliyle Şekil 5'te sunulmuştur. Diğer bir ifadeyle sorunun küçük grupla denenen açık uçlu haline, örnek bir öğrenci çözümüne ve yapılan revize işlemleri sonunda elde edilen çoktan seçmeli haline Şekil 5'te yer verilmiştir.

**KARGO ŞİRKETİ**

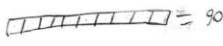
SSDT\_MDF\_7\_51

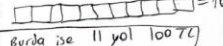
Bir kargo şirketi sahip olduğu araçlara yakıt sağlayacak bir petrol şirketiyle anlaşmak istiyor. İki petrol şirketi de buna karşılık kargo şirketine özel olarak yapabileceği kampanyaları şu şekilde açıklıyorlar:

- A petrol şirketi, normal yakıtın satış fiyatı üzerinden % 10 indirim yapacağını belirtiyor.
- B petrol şirketi, fiyatta indirim yapmayacağını ancak özel olarak üreteceği bir yakıt vereceğini ve bununla araçların normal yakıtta göre % 10 daha fazla yol alabileceğini belirtiyor.

Kargo şirketi için hangi petrol şirketi ile çalışmak daha karlı olacaktır? İşleminizi gösteriniz.

*Bence A şirketi mesela yakıtı 100 TL değil A şirketinde %10 indirimle 90 TL'ye gelecek. B'de ise 100 TL olacak ama biraz daha fazla gidilecek.*

A  = 90 TL  
Burda 10 yol 90 TL oluyor  
11 yol 99 TL oluyor

B  = 100 TL  
Burda ise 11 yol 100 TL  
yani A daha karlı

Dördüncü sürecin sonunda elde edilen bir örnek öğrenci çözümü

**KARGO ŞİRKETİ**

SSDT\_MDF\_7\_61

**SORU 8:** Bir kargo şirketi sahip olduğu araçlara yakıt sağlayacak bir petrol şirketiyle anlaşmak istiyor. Üç petrol şirketi de buna karşılık kargo şirketine özel olarak yapabileceği kampanyaları şu şekilde açıklıyorlar:

- K petrol, normal yakıtın satış fiyatı üzerinden % 10 oranında indirim yapacağını belirtiyor.
- L petrol, fiyatta indirim yapmayacağını ancak özel olarak üreteceği bir yakıt vereceğini ve bununla araçların normal yakıtta göre % 10 daha fazla yol alabileceğini belirtiyor.
- M petrol, fiyatta indirim yapmayacağını ancak % 10 oranında bedava yakıt verebileceğini belirtiyor.

Bu verilere göre kargo şirketi için aşağıdakilerden hangisi doğrudur?

A) Üç şirketin de teklifi aynı düzeyde karlılık sağlar.  
B) En karlı olan K, sonra L sonra M petroldür.  
C) En karlı olan M, sonra K sonra L petroldür.  
D) En karlı olan K petroldür. L ve M petrol ise birbiriyle aynı düzeyde karlıdır.  
E) En karlı olan M petroldür. K ve L petrol ise birbiriyle aynı düzeyde karlıdır.

Beşinci sürecin sonunda sorunun çoktan seçmeli tipe dönüştürülmüş hali

**Şekil 5.** Örnek Sorunun Üçüncü, Dördüncü ve Beşinci Süreç Sonundaki Durumları

Yukarıda (solda) öğrenci çözümüne bakıldığında soruda anlatılmak istenen ile, öğrencinin sorudan anladığının örtüştüğü görülmektedir. Bu durumda soru 4. süreci başarı ile geçmiştir. Sorunun açık uçlu halinde tercih edilebilecek sadece 2 şirket sunulmuştu. Bu durum sorunun çoktan seçmeli hâli için 5 seçenek yazılmasını zorlaştırmıştır. Bu nedenle sorunun çoktan seçmeli hâle revize edilmesi sürecinde (5. süreçte) üçüncü bir şirketin daha eklenmesi konusunda uzmanlar ile görüş birliğine varılmıştır. Bu sorunun bulgularında örnek olarak seçilmesinin nedeni diğer sorulara göre daha kapsamlı bir revize işlemi görmüş olmasıdır. Beşinci sürecin sonunda sorunun nihai durumu Şekil 5'te (sağda) sunulmuştur.

Yöntemin altıncı sürecinden soruların kestirilen madde zorluk düzeylerine ve soruların ölçtüğü matematik okuryazarlığı alt boyutlarına ilişkin bulgular elde edilmiştir. Her bir zorluk düzeyinde yer alan soru sayıları Tablo 2’de sunulmuştur. Değerler 0-10 aralığına göre verilmiştir. Soru, zorlaştıkça daha büyük bir değerle temsil edilmiştir.

**Tablo 2.** Soruların Zorluk Düzeylerine Dağılımı

Zorluk Düzeyi	1	2	3	4	5	6	7	8	9	10	Toplam
Soru Sayısı	0	0	0	17	19	26	20	12	2	6	102

Tablo 2’de soruların kestirilen zorluk düzeylerinin 4-10 aralığında değiştiği görülmektedir. Orta zorluk düzeyindeki (4, 5, 6.) sorular çoğunluktadır. Bu bulgular çoktan seçmeli soru havuzundaki maddelerin kestirilen zorluk düzeylerinin ortalamasının 6,21 olduğunu göstermektedir.

Matematik okuryazarlığı alt boyutlarının her biri ile eşleşen soru sayıları Tablo 3’te verilmiştir. Buna göre, kişisel bağlamda ve nicelik içerik bilgisini ölçmeye yönelik soruların diğerlerine göre daha fazla olduğu görülmektedir. Soruların alt boyutlarla birebir eşleştirildiği tabloya Ek-1’de yer verilmiştir.

**Tablo 3.** Alt Boyutlara Düşen Soru Sayıları

Boyut	Bağlam*				Matematiksel İçerik**				Matematiksel Süreç***		
Alt Boyut	BİL	KS	ML	TPL	BV	Dİ	NC	UZ	FRM	KAY	MCY
Soru Sayısı	11	48	28	15	10	27	49	16	41	36	25
Toplam	102				102				102		

\* Bağlam: BİL: bilimsel, KS: kişisel, ML: mesleki, TPL: toplumsal.

\*\* Matematiksel İçerik: BV: belirsizlik ve veri, Dİ: değişim ve ilişkiler, NC: nicelik, UZ: uzay ve şekil.

\*\*\* Matematiksel Süreç: FRM: Durumları, problemleri matematiksel olarak formüle etme, KAY: Matematiksel kavramları, gerçekleri, yöntemleri kullanma ve akıl yürütme, MCY: Matematiksel çıktıları yorumlama, uygulama ve değerlendirme

Örnek olarak Şekil 5’te verilen “Kargo Şirketi” sorusunun tahmini madde zorluğu 10 üzerinden “7”, bağlamı “meslekî”, matematiksel içeriği “değişim ve ilişkiler” matematiksel süreci “durumları, problemleri matematiksel olarak formüle etme” olarak belirlenmiştir. Soru bu özellikleri ile etiketlenmiş ve “D” seçeneğinde verilen “En kârlı olan K petroldür. L ve M petrol ise birbiriyle aynı düzeyde kârlıdır.” doğru cevabı ile çoktan seçmeli soru havuzuna eklenmiştir. Yöntemin yedinci sürecinde tüm sorular bu örnek soruda verilen özellikleri ile etiketlenerek çoktan seçmeli taslak soru havuzu oluşturulmuştur. Sekizinci süreçte soruların kitapçıklara dağıtımı, Tablo 2 ve Tablo 3’te yer alan bu bulgulara ve Ek-1’de yer alan birebir “madde-ölçtüğü matematik okuryazarlığı alt boyutları” eşleştirmelerine göre yapılmıştır.

Yöntemin dokuzuncu ve onuncu süreçlerinden sırasıyla birinci ve ikinci saha çalışmalarının bulguları elde edilmiştir. Bu bulgular soruların olduğu gibi kullanılmasına/revize edilmesine karar vermek üzere değerlendirilmiştir. Bu konuda karar vermede soruların saha çalışmalarından elde edilen madde ayırt edicilik değerleri kullanılmıştır. Birinci saha çalışmasında 98 adet matematik okuryazarlığı sorusu denenmiştir. Bu 98 sorudan 33 tanesi revize edilmiştir. Revize edilen 33 sorudan (bazıları için birden fazla revize versiyon türetildiğinden) 44 soru elde edilmiştir. Revize edilen bu maddelere ilişkin bulgular, revize edilme nedenlerine

göre iki kategoride incelenmiştir. *Birinci kategori* 29, *ikinci kategori* 4 soru içermektedir. Kategorilere ilişkin açıklamalar aşağıda sunulmuştur.

*Birinci kategori*, birinci uygulamanın verilerine göre  $r_{jk} < 0,3$  olduğu için revize edilen maddeleri içermektedir. Bu kategoride yer alan 29 maddenin birinci uygulama sonundaki  $r_{jk}$  değerleri ile bunların revize edilmesi ile elde edilen 40 maddenin ikinci uygulama sonundaki  $r_{jk}$  değerleri Tablo 4'te verilmiştir.

**Tablo 4.** Birinci Kategoride Yer Alan Soruların Madde Ayırt Edicilik Değerlerinin Karşılaştırılması

Soru ID	$r_{jk}$ (1. Uygulama)	Soru ID (revize versiyon)	$r_{jk}$ (2. Uygulama)
1	0,127	1R1	0,255
10	0,087	10R1	0,107
10	0,087	10R2	0,118
11	0,244	11R1	0,442
13	0,293	13R1	0,458
22	0,234	22R1	0,327
27	0,296	27R1	0,313
29	0,158	29R1	0,245
31	0,154	31R1	0,244
34	0,124	34R1	0,23
34	0,124	34R2	0,584
34	0,124	34R3	0,225
34	0,124	34R4	0,317
39	0,124	39R1	0,216
40	0,176	40R1	0,45
53	0,29	53R1	0,381
58	0,277	58R1	0,269
58	0,277	58R2	0,118
60	0,257	60R2	0,255
60	0,257	60R1	0,262
60	0,257	60R3	0,277
61	0,228	61R1	0,301
65	0,237	65R1	0,312
79	0,079	79R1	0,328
83	0,202	83R1	0,145
84	0,195	84R2	0,163
84	0,195	84R3	0,13
84	0,195	84R1	0,092
19A	0,247	19AR1	0,373
19A	0,247	19AR2	0,593
19B	0,19	19BR1	0,345
19B	0,19	19BR2	0,588
23A	0,281	23AR1	0,315
49B	0,225	49BR1	0,15
66A	0,152	66AR1	0,28
66B	0,298	66BR1	0,225
70B	0,238	70BR1	0,26
76B	0,196	76BR1	0,142
82A	0,296	82AR1	0,597
82B	0,258	82BR1	0,293

Tablo 4'e bakıldığında, birinci kategoride yer alan 29 maddeden 14 tanesine (11, 13, 22, 27, 34, 40, 53, 61, 65, 79, 19A, 19B, 23A, 82A) ait revize versiyonların  $r_{jk} \geq 0,3$  şartını sağladığı görülmektedir. Bu 14 adet maddeden 3 tanesinin (34, 19A, 19B) ikişer revize versiyonu  $r_{jk} \geq 0,3$  şartını sağlamıştır. Böylece revize sonunda  $r_{jk} \geq 0,3$  şartını sağlayan 17 madde elde edilmiştir.

Yapılan revize işlemlerini örneklendirmek için birinci kategoride yer alan 19A ve 19B maddeleri 1. ve 2. uygulamada kullanıldığı haliyle aşağıda verilmiştir. Bu soruların bulgularda örnek olarak seçilmesinin nedeni, revize işleminin en etkili istatistiksel sonuçları verdiği maddelerden biri olmasıdır. Değişiklik yapılan yerler, madde üzerinde taranmıştır. Ayrıca maddelerin revize öncesinde ve sonrasındaki güçlük ve ayırt edicilik değerleri ile güvenilirlik katsayıları Şekil 6’da sunulmuştur.

1. UYGULAMADA:		2. UYGULAMADA:	
SSDT_RNK_7_19		SSDT_RNK_7_19R2	
OTOBÜS BEKLEME		OTOBÜS BEKLEME	
1 ve 2. soruları aşağıdaki bilgilere göre cevaplayınız.		1 ve 2. soruları aşağıdaki bilgilere göre cevaplayınız.	
Durakta otobüs bekleyen Ahmet şunları bilmektedir. Bekledikleri otobüs 30 dakikada bir geçmektedir. Bu otobüs trafiğin yoğunluğuna göre ya normal saatinden en fazla 8 dakika geç veya en fazla 6 dakika erken gelebilmektedir.		Durakta otobüs bekleyen Ahmet şunları bilmektedir. • Bekledikleri otobüs normalde 30 dakikada bir geçmektedir. • Bu otobüs trafiğin yoğunluğuna göre; o Ya en fazla 8 dakika geç, o Ya da en fazla 6 dakika erken gelebilmektedir.	
SORU 1: Otobüsü son anda kaçan Ahmet bir sonraki için en fazla kaç dakika bekler?		SORU 1: Buna göre, otobüsü son anda kaçan Ahmet bir sonraki için en fazla kaç dakika bekler?	
A) 44		A) 44	
B) 38		B) 33	
C) 37	Madde güçlüğü: 0,174	C) 35	Madde güçlüğü: 0,543
D) 14	Madde Ayırt Ediciliği: 0,247	D) 14	Madde Ayırt Ediciliği: 0,593
E) 6	Güvenirlik Katsayısı: 0,094	E) 6	Güvenirlik Katsayısı: 0,295
SORU 2: Otobüsü son anda kaçan Ahmet bir sonraki için en az kaç dakika bekler?		SORU 2: Buna göre, otobüsü son anda kaçan Ahmet bir sonraki için en az kaç dakika bekler?	
A) 24		A) 20	
B) 14		B) 14	
C) 16	Madde güçlüğü: 0,116	C) 16	Madde güçlüğü: 0,375
D) 18	Madde Ayırt Ediciliği: 0,190	D) 18	Madde Ayırt Ediciliği: 0,588
E) 22	Güvenirlik Katsayısı: 0,061	E) 21	Güvenirlik Katsayısı: 0,284

**Şekil 6.** 19A ve 19B ID Numaralı Sorularda Yapılan Revize İşlemleri ve Madde İstatistikleri

19A ve 19B maddelerinin Şekil 6’da verilen istatistiklerine bakıldığında yapılan revize işleminin maddelerin ayırt ediciliklerini önemli ölçüde artırdığı ve  $r_{jx} \geq 0,3$  şartını sağladığı görülmektedir. Madde güçlüklerine bakıldığında, her iki maddenin bu işlem sonucunda kolaylaştığı ve önceki değerine göre 0,5 düzeyine yaklaştığı görülmektedir. Maddenin  $r_{jx}$  ve  $p_j$  değerlerindeki bu değişimlerin her ikisi de güvenilirlik katsayısını olumlu yönde etkilemiştir. Bu bulgular, revize işleminin maddenin geçerlik ve güvenilirliğini sağlamada başarılı olduğunu istatistiksel olarak kanıtlamaktadır.

*İkinci kategori*, birinci uygulamanın verilerine göre  $r_{jx} \geq 0,3$  olmasına rağmen alt gruba göre üst gruptan daha fazla oranda öğrencinin yanlış seçeneğe/seçeneklere yığıldığı maddelerdir. Bu kapsamda 4 adet madde yer almaktadır. Bu kategoride yer alan maddelerin birinci ve ikinci uygulama sonundaki  $r_{jx}$  değerleri Tablo 5’te verilmiştir.

**Tablo 5.** İkinci Kategoride Yer Alan Soruların Madde Ayırt Edicilik Değerlerinin Karşılaştırılması

Soru ID	1. Uygulama $r_{jx}$ değeri	Soru ID (revize versiyon)	2. Uygulama $r_{jx}$ değeri
6	0,325	6R1	0,315
7	0,372	7R1	0,409
9	0,326	9R1	0,458
51	0,411	51R1	0,57

Tablo 5’e bakıldığında ikinci kategoride yer alan 4 maddeden 7, 9 ve 51 ID numaralı maddelerin ayırt edicilik değerlerinin revize işlemi sonucunda arttığı görülmektedir. Örnek olması bakımından bu kategorideki 9 ID

numaralı ("Baraj" adlı) sorunun ve revize edilmiş versiyonunun (9R1) madde istatistikleri Tablo 6'da sunulmuştur. Ayrıca bu istatistiklere göre yapılan çıkarımlara istinaden yürütülen revize işlemleri, Şekil 7'de soru üzerinde gösterilmiştir.

**Tablo 6.** Baraj Adlı Sorunun Birinci ve İkinci Uygulama Sonrasında Madde İstatistikleri

Seçenekler											
Soru ID	Grup	*N <sub>A</sub>	A**	N <sub>B</sub>	B	N <sub>C</sub>	C***	N <sub>D</sub>	D	N <sub>E</sub>	E
9	Total	42	0,151	76	0,273	73	0,263	27	0,097	57	0,205
9	Üst	17	0,128	27	0,203	57	0,429	8	0,06	21	0,158
9	Alt	12	0,154	34	0,436	8	0,103	8	0,103	16	0,205
9	Fark	5	-0,026	-7	-0,233	49	0,326	0	-0,042	5	-0,047
9R1	Total	37	0,144	76	0,296	66	0,257	37	0,144	38	0,148
9R1	Üst	10	0,145	9	0,13	35	0,507	5	0,072	9	0,13
9R1	Alt	15	0,185	32	0,395	4	0,049	16	0,198	13	0,16
9R1	Fark	-5	-0,04	-23	-0,265	31	0,458	-11	-0,125	-4	-0,03

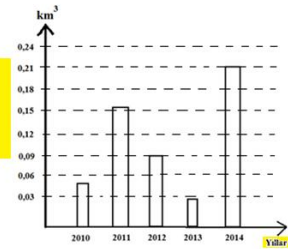
\*Seçeneği işaretleyen öğrenci sayısı

\*\*Seçeneği işaretleyen öğrenci sayısının ilgili gruptaki öğrenci sayısına oranı

\*\*\*Doğru seçenek

Tablo 6 incelendiğinde sorunun birinci uygulama verilerine göre A, D ve E seçeneklerini üst ve alt gruptan işaretleyen öğrenci oranlarının birbirine yakın olduğu görülmüştür. Bu veriler bu 3 seçeneğin sorunlu olabileceğini düşündürmüştür ancak bununla birlikte sorunun içeriği itibarıyla türetilebilecek seçenek sayısı sınırlıdır ve sadece benzer seçenekler türetilir. Buna göre sorunun, seçeneklerin kendisinden değil, seçeneklerdeki ifadelerin iç içe yazılmış olmasından kaynaklanabileceği düşünülmüştür. Bu nedenle seçenekleri değiştirmek yerine, sorunun seçeneklerinde yer alan ifadelerin daha ayrık (bir tablo gibi) yazılmasına karar verilmiştir. Sorunun revize öncesi ve sonrasındaki hali Şekil 7'de sunulmuştur. Bu sorunun bulgularda örnek olarak seçilmesinin nedeni, kategorisindeki diğer sorulara göre ağırlıklı olarak görsel açıdan revize edilmiş olmasıdır. Değişiklik yapılan yerler madde üzerinde taranmış ya da çerçeve içine alınmıştır.

## 1. Uygulama

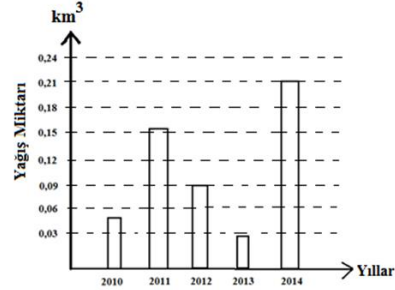
SSDT\_TNY\_7\_9  
SORU 1:

- A) 2010, 2012 ve 2013'te kıtlık, sadece 2014'te taşkın  
 B) 2010 ve 2013'te kıtlık, 2011 ve 2014'te taşkın  
 C) 2010 ve 2013'te kıtlık, sadece 2014'te taşkın  
 D) Sadece 2013'te kıtlık, 2011 ve 2014'te taşkın  
 E) Sadece 2013'te kıtlık ve sadece 2014'te taşkın

BARAJ

Kapasitesi 200 milyon m<sup>3</sup> olan baraja 2010-2014 yılları arasında düşen yağış miktarı grafikte verilmiştir. Yakın çevrede su kıtlığı yaşanmaması için barajın doluluk oranının % 30'un altına düşmemesi gerekmektedir. Buna göre hangi yıllarda su kıtlığı, hangi yıllarda taşkın tehlikesi ortaya çıkmış olabilir?

## 2. Uygulama

SSDT\_TNY\_7\_9R1  
SORU 1:

Kapasitesi 200 milyon metreküp olan baraja 2010-2014 yılları arasında düşen yağış miktarı grafikte verilmiştir. Yakın çevrede su kıtlığı yaşanmaması için barajın doluluk oranının % 30'un altına düşmemesi gerekmektedir. Buna göre hangi yıllarda su kıtlığı, hangi yıllarda taşkın tehlikesi yaşanmış olabilir?

- | Kıtlık                   | Taşkın          |
|--------------------------|-----------------|
| A) 2010, 2012 ve 2013'te | Sadece 2014'te  |
| B) 2010 ve 2013'te       | 2011 ve 2014'te |
| C) 2010 ve 2013'te       | Sadece 2014'te  |
| D) Sadece 2013'te        | 2011 ve 2014'te |
| E) Sadece 2013'te        | Sadece 2014'te  |

Şekil 7. 9 ID Numaralı Soruda Yapılan Revize İşlemleri

Şekil 7'de görüldüğü gibi, seçeneklerin yanı sıra soru kökünde yer alan grafik ve metin de sayfada daha ayrıntılı görünecek şekilde konumlandırılmıştır. Bu revize işlemlerinin ardından sorunun ikinci uygulamadan elde edilen verilerine göre (bkz. tablo 6'da 9R1 verileri), madde güçlük değeri neredeyse aynı düzeyde kalmış (0,263 ve 0,257), ayırt edicilik değeri 0,326'dan 0,458'e, güvenilirlik katsayısı 0,143'ten 0,2'ye yükselmiştir.

Bulgularda revize işlemlerini örneklendirmek için sunulan sorular çeşitliliği sağlayacak şekilde seçilmiştir: Buna göre en kapsamlı revize işlemi (Şekil 5), istatistiksel olarak en anlamlı farkı ortaya çıkaran revize işlemlerinden biri (Şekil 6) ve ağırlıklı olarak görsel öğeler üzerinde yapılan bir revize işlemi (Şekil 7) bulgularda sunulmuştur. Yukarıda örnek sorular üzerinden açıklanan revize işlemleri benzer şekilde 33 sorunun her biri için gerçekleştirilmiştir. Soruların geçerliğine etkileri istatistiksel olarak kanıtlanan ve sonuçlar bölümünde yedi madde halinde sunulan etkenleri örnekleyen diğer revize işlemlerine Ek-3'te yer verilmiştir.

## TARTIŞMA, SONUÇ ve ÖNERİLER

Çalışma sonucunda madde ayırt edicilik değerlerine göre 94 adet soru "Olduğu gibi kullanılabilir" kategorisinde değerlendirilmiştir. Bunların içinden 10 tanesi aynı soruya ait revize versiyonlardır. Başka bir ifadeyle, bu 94 maddeden 10 tanesi aslında aynı maddenin "Olduğu gibi kullanılabilir" kategorisinde değerlendirilen iki farklı (6-6R1, 7-7R1, 9-9R1, 19AR1-19AR2, 19BR1-19BR2, 27-27R1, 34R1-34RR2, 51-51R1, 70A-70AR1, 82A-82AR1) versiyonudur. Net olarak 84 tane farklı madde elde edilmiştir. Bu maddelerin güçlük değerleri 0,158 ve 0,845, güvenilirlik katsayıları 0,113 ve 0,364 arasında değişmektedir. Bu kategoride yer alan maddeler bu çalışmanın

ürünü olan çoktan seçmeli madde havuzunda (soru bankası) yerini almıştır. Tüm soruların madde istatistikleri Ek-1'de sunulmuştur. Örnek olması bakımından 7 adet soruya Ek-2'de yer verilmiştir.

Madde ayırt edicilik değeri 0,3 ile 0,2 arasında ( $0,20 \leq r_{jk} < 0,30$ ) olan sorular "Revize edilmeli" kategorisinde değerlendirilmiştir. Madde ayırt edicilik değeri 0,2'den küçük olan ( $r_{jk} < 0,2$ ) sorular "Kullanılmamalı veya tamamen revize edilmeli" kategorisinde değerlendirilmiştir. Bu maddeler, istatistiksel veriler ışığında revize edilerek yeni uygulamalarda tekrar denenebilir.

Birinci saha testlerinin verilerine göre yapılan ve başarılı olan revize işlemlerinin kapsamaları genel olarak dört başlık altında toplanabilir. Bunlar sorunun (i) metninde, (ii) görselinde, (iii) seçeneklerinde ve (iv) verilerinde yapılan revize işlemlerdir. Bu çalışma kapsamındaki herhangi bir soru için yapılan ve başarılı olan revize işlemlerinin bu başlıklardan sadece birini ya da herhangi ikisini içerdiği görülmüştür, bu başlıklardan üç ya da dördünü birlikte içeren revize işlemine rastlanmamıştır. Sadece metninde veya sadece görselinde veya sadece seçeneklerinde değişiklik yapılarak kurtarılabilen sorular olmasına rağmen sadece verilerinde değişiklik yapılarak kurtarılabilen soruya rastlanmamıştır. Bu revize işlemleri teknik olarak detaylandırıldığında çoktan seçmeli soruların geçerliğini artıran ya da azaltan durumlara ilişkin aşağıdaki sonuçlar elde edilmiştir.

1. Soru metninde farklı koşullarda farklı durumların ortaya çıktığı ifade ediliyorsa, her bir koşulun ve sonucunda ortaya çıkan durumun maddeler halinde (karşılaştırma için bkz. ek-3 soru ID: 34→34R1) sunulması geçerliği artırmaktadır.
2. Soruda şeklin/tablonun hem üstünde hem altında açıklamalara (metne) yer verilmişse, vurgulanmak istenen ifadeye soru metninin şekilden/tablodan sonraki kısmında (karşılaştırma için bkz. ek-3, soru ID: 22→22R1) yer verilmelidir. Diğer bir ifadeyle bu tip sorularda vurgulanmak istenen ifadenin soru metninin sonuna (bitişine doğru) yerleştirilmesi geçerliği artırmaktadır.
3. Bir soruda birbiri ile ilişkili birden fazla görsel içerik (nesne, fotoğraf, şekil, resim) bulunabilir. Bu tip sorularda (bkz. ek-3, soru ID: 65→65R1) görsel içeriklerin aralarındaki ilişkiyi (süreç, akış, vb.) yansıtacak şekilde sıralanması ve soru metninde bu sıralamaya ilişkin ifadelere (ölçüte veya sağdan sola veya yukarıdan aşağıya, vb.) yer verilmesi geçerliği artırmaktadır. Ayrıca görsel içerikler üzerinde yazı ya da sayılar varsa bunların neyi temsil ettiğine ilişkin ifadeler soru metninde yer verilmesi geçerliği artırmaktadır.
4. Vurgulanmak istenen kelimelerin (sadece bir ya da iki kelime için) karartılması (kalın harflerin kullanılması) ve seçeneklerin yan yana yerine alt alta dizilmesi (karşılaştırma için bkz. ek-3, soru ID: 53→53R1) bu çalışmada geçerliği artırdığı sonucuna ulaşılan diğer etkenlerdir.
5. "Soruda verilenler sonuca ulaşmak için yeterli değildir." veya bununla aynı anlamı ifade eden cümlelerin seçeneklerden çıkarılması (bkz. ek-3, soru ID: 11→11R1 ve 34R1→34R4) geçerliği artırmaktadır.
6. En basit bakış açısıyla soruya yaklaşıldığında işaretlenmesi beklenen sonucun seçeneklerden çıkarılması (karşılaştırma için bkz. ek-3, soru ID: 40→40R1) geçerliği artırmaktadır.
7. Verilen koşulların her biri için aralarında herhangi bir örüntü bulunmaksızın salt ayrı ayrı işlemler yapmayı gerektiren durumları içeren sorularda, işlem sayısını azaltmaya yönelik revizeler (karşılaştırma için bkz. ek-3, soru ID: 79→79R1) geçerliği artırmaktadır.

Yukarıda verilen sonuçlardan ilk dört tanesi çoktan seçmeli test maddelerinin yazımı ile ilgili alan yazında (Turgut ve Baykul, 2010; Atılğan ve ark., 2017; Tekin, 2019) bulunan soru metni ve görselleri hakkındaki genel hususlarla (netlik, metin ve görsel arasındaki tutarlılık) örtüşmektedir. Ancak bu çalışmanın bulguları ve sonuçları, teknik açıdan daha özel detayları ve soru yazarlarını net bir şekilde yönlendirebilecek ifadeleri içermektedir.

Yukarıdaki sonuçlardan beş, altı ve yedincisi, soruların bilişsel özellikleri ile ilgilidir. Beşinci sonuç, *sorularda verilenlerin çözüm için yetersiz olduğunu ifade eden bir seçeneğin bulunması durumunda*, üst gruptan beklenmedik sayıda ( $r_{jx}$  değerini 0,30'un altına düşürecek kadar) öğrencinin bu seçeneğe yönelmesinden kaynaklanmaktadır. Buna göre söz konusu seçenek eğer doğru cevap değilse sorunun geçerliğini önemli düzeyde ve olumsuz bir şekilde etkilemektedir. Bu seçeneğin (ifadesinin) uzunluk bakımından diğer seçeneklerden farklı olması (Turgut ve Baykul, 2010) öğrencileri kısmen bu seçeneğe yönelten bir (bilişsel olmayan-yapısal) etkenin varlığını düşündürülebilir. Ancak 34R1 ve 34R4 ID numaralı soruların her ikisinde de "A" seçeneği diğer seçeneklere göre daha uzun olmasına rağmen 34R4 ID numaralı soruda üst gruptan hiçbir öğrencinin "A" seçeneğini işaretlememesi, böyle bir yapısal etkenin burada etkisinin olmadığını göstermektedir.

Altıncı sonuç, *en basit bakış açısıyla soruya yaklaşıldığında işaretlenmesi beklenen sonuca seçeneklerde yer verilmesi durumunda*, üst gruptan beklenmedik sayıda ( $r_{jx}$  değerini 0,30'un altına düşürecek kadar) öğrencinin bu seçeneğe yönelmesinden kaynaklanmaktadır. Buna göre söz konusu seçenek eğer doğru cevap değilse sorunun geçerliğini önemli düzeyde ve olumsuz bir şekilde etkilemektedir. Beşinci ve altıncı sonuçlar, taranan alan yazında rastlanmayan, bu çalışmaya özgün sonuçlardır. Bu iki sonuç çoktan seçmeli matematik sorularını hazırlamaya ya da revize etmeye ilişkin alan yazına metodolojik açıdan katkı sunabilir. Soru ile ilgili tespit edilmiş başka olumsuz etkenler yoksa, bu iki durumun her ikisinde de sadece bu tip seçeneğin çıkarılıp yerine sıradan bir seçenek yazılması ile sorunun kullanılabilir düzeyde geçerliğe erişmesi sağlanabilmektedir.

Yedinci sonuç ilgili sorunun çözümünde, çok sayıda işlemin her bir koşul için tekrarlanmasını gerektirmesinden kaynaklanmaktadır. Bu durum ilgili sorunun zorluk düzeyini artırmış ve geçerliğini de ( $r_{jx}$  değerini 0,30'un altına düşürecek kadar) önemli düzeyde azaltmıştır. Bu sonuç Başol'un (2019) zorluk düzeyiyle öğrencileri yıldırان sorulardan kaçınılması gerektiği ifadesini doğrulamakta ve şu detayı eklemektedir: Buna göre çözümünde, aralarında bir örüntü bulunmayan koşulların her biri için salt ayrı ayrı işlemler yapmayı gerektiren sorular geçerliğini yitirmektedir.

Matematik okuryazarlığı sorusu geliştirmeye ilişkin alan yazında "PISA benzeri sorular" terimine ve özellikle PISA'nın yayınladığı soruları esas alan çalışmalara (Kamaliyah vd., 2013; Oktiningrum vd., 2016; Murdiyani, 2018; Nasution vd., 2019) sıkça rastlanmaktadır. Bu durumun, soru yazma çalışmalarında özgün matematik okuryazarlığı soruları üretme şansını ve üretilen soruların zengin bir bağlam çeşitliliği sunma olasılığını azaltabileceği düşünülmektedir. Bu sınırlandırmanın yerine matematik okuryazarlığının tanımını, amaçlarını ve kapsamını bir bütün olarak ele alan ölçütleri esas almak, üretilen soruların özgünlüğünü, çeşitliliğini ve dolayısıyla geçerliğini artırabilir. Çünkü çeşitli bağlamlarda yazılmış soruları içeren testler öğrencilerin matematik



okuryazarlığını daha kapsamlı bir şekilde ölçebilir. Bu bağlamda, bu çalışmaya kaynaklık eden sorular için ölçüt PISA benzeri olmaları değil, matematik okuryazarlığının tanımı, amaçları ve kapsamına uygun olmalarıdır. Bu farklılık, bu çalışmada ortaya konan soruları, alan yazında yer alan söz konusu çalışmalardan elde edilen sorulardan ayıran en önemli özelliğdir. Bunun yanı sıra çok sayıda sorunun denenmiş olması çeşitliliği ve dolayısıyla geçerliliği artıran diğer bir değişkendir. Böylece matematik okuryazarlığı genel değerlendirme çerçevesinde (OECD, 2013c) yer alan her bir alt boyutta öğrencileri değerlendirebilecek sorular ortaya konmuştur.

Ek-1’de, madde istatistiklerinin yanı sıra soruların ölçtüğü matematik okuryazarlığı alt boyutlarına yer verilmiştir. Çoktan seçmeli soru havuzunda, her bir alt boyutun ölçülmesine yönelik en az 9 madde yer almaktadır. Bu çalışmadan elde edilen soruların farklı kombinasyonlarla bir araya getirilmesi ile, ölçülmek istenen alt boyutlarda (bağlam, matematiksel içerik, matematiksel süreç) ve düzeyde (kolay, orta, zor) matematik okuryazarlığı testleri elde edilebilir. Bu alanda test geliştirme çalışmaları yürütülebilir.

Derlenen 102 adet sorunun içinde, kişisel bağlamda ve nicelik matematiksel içerik bilgisini ölçmeye yönelik soruların diğerlerine göre sayıca daha fazla olduğu görülmektedir (bkz. tablo 3). Demir ve Altun’a (2018) göre “çağrışım” (günlük hayatta karşılaşılan bir olaydan/durumdan yola çıkarak soru yazma) matematik okuryazarlığı soru yazarları tarafından fırsat olarak değerlendirilmektedir. Buna göre günlük hayatta soru yazarlarının, kişisel bağlamda ve çözümünde nicelik matematiksel içerik bilgisini gerektiren sorunlarla diğer durumlara göre daha sık karşılaştığı söylenebilir. Bu çıkarım şimdilik soruların derlendiği çalışmaların soru yazarları için sınırlandırılmalıdır. Ancak “Genel olarak günlük hayatta karşılaşılan ve çözümünde matematik bilgiye ihtiyaç duyulan durumlar bu alt boyutlara nasıl dağılmaktadır?” sorusuna cevap arayan araştırmalarla bir genellemeye varmak mümkün olabilir. Böyle bir girişimle genel bir matematik okuryazarlığı testinde geçerlik için soruların alt boyutlara nasıl dağıtılması gerektiğine ilişkin veriler elde edilebilir.

Çoktan seçmeli sorular büyük örneklemelerin ölçülmesinde daha elverişli olmasına karşın öğrencilerin matematik okuryazarlığı yeterliklerini (akıl yürütme, iletişim, temsil etme, vb.) yansıtma ve araştırmacılara bu konuda kanıt sağlama konularında açık uçlu sorulara göre daha az fırsat sunmaktadır. Çalışmanın yönteminin üçüncü sürecinde oluşturulan açık uçlu soru havuzundaki maddeler bu konuda değerlendirilebilir. Her bir sorunun çözümde gerektirdiği yeterlikler belirlenebilir. Böylece sorular, öğrencilerin matematik okuryazarlığı yeterliklerinin ölçülmesini içeren çalışmalarda araç olarak kullanılabilir.

Ders içi etkinlikler ve ders kitapları, bu çalışmada elde edilen sorular için önemli birer kullanım alanıdır. Ders kitabı araştırmaları matematik okuryazarlığına ilişkin yeterliklerden akıl yürütme ve iletişim kurma için genellikle yeterli fırsatlar sunmaz, çünkü ders kitabı araştırmalarının çoğu rutin yöntemlerin tatbikatını ve uygulanmasını hedefler (Fan ve Zhu, 2000; Stacey ve Vincent, 2009; Thompson, Senk ve Johnson, 2012). Ulusal alan yazında da derslerde ve ders kitaplarında matematik okuryazarlığı sorularına yeterince yer verilmediğine (Altun ve Akkaya, 2014; Çakır, 2009; Dede ve Yaman, 2005; İskenderoğlu ve Baki, 2011) ve matematik öğretiminin gerçek dünyadan kopuk bir süreç gibi yürütüldüğüne ilişkin çalışmalar yer almaktadır. Begel’in (1973; aktaran Wong ve Low, 2020) belirttiği

gibi, ders kitabının içeriği “bir yandan manipüle edilebilen ve diğer yandan öğrenci öğrenmesini etkileyebilen” bir değişkendir. Bu bağlamda, bu araştırmadan elde edilen sorular, matematik okuryazarlığı alanında yapılacak soru-etkinlik geliştirme çalışmalarında araç olarak kullanılabilir. Öğretmenlerin sorulara matematik öğretim sürecinde ve bu sürecin tamamlayıcısı olan ölçme değerlendirme faaliyetlerinde yer vermesi önerilebilir. Soruların ders içi etkinliklerde kullanımında öğrencilerin matematik okuryazarlığının önemli yeterliklerinden olan matematiksel akıl yürütme ve iletişim becerilerini geliştirmek için Wong ve Low’un (2020) “öğretmen rolüne” ve “sınıfın doğasına” ilişkin önerileri dikkate alınabilir.

Türkiye’de 2018 yılında uygulamaya konan ortaokul matematik öğretim programının genel amaçlarında “matematiksel okuryazarlık becerilerini geliştirebilecek ve etkin bir şekilde kullanabilecektir.” ifadesi yer almaktadır (MEB, 2018). Ulusal sınavlarda okuryazarlık problemlerinin sayısının artırılması tavsiye edilmektedir (Jürges ve ark., 2012; Uysal ve Yenilmez, 2011). Burada öğretimin amacı ile ulusal ölçme değerlendirme faaliyetlerine ilişkin tavsiyeleri içeren söz konusu çalışmaların birbirini tamamladığı görülmektedir. Ancak araştırmacılar (Demir, 2015; Oktiningrum vd., 2016), öğrencilerin bağlamsal sorulara aşına olmadıklarını ifade etmektedir. Bu çalışmadan elde edilen soruların matematik öğretim sürecinde kullanılması ile öğrencilerin bağlamsal sorulara aşinalık kazanmasına katkı sağlanabilir. Böylece matematik öğretiminin içeriği ile ulusal ölçme değerlendirme faaliyetleri için tavsiye edilen içeriğin birbirini tamamlayıcı bir nitelik kazanmasına katkı sağlanabilir.

Demir (2019) çalışmasında matematik okuryazarlığı soru yazma sürecinde yer alan eylemleri ve bu eylemlerin soru yazarları tarafından genellikle hangi sıra ile takip edildiğini ortaya koymuştur. Ayrıca diğer bir çalışmada Demir (2020) öğretmen adaylarına yönelik bir matematik okuryazarlığı öğretim programı tasarlamış ve bir sorunun, matematik okuryazarlığı sorusu olma niteliği taşıyıp taşımadığına ilişkin bir takım ölçütler kullanmıştır. Bu araştırmada, sözü edilen çalışmalar ile Demir ve Altun’un (2018) çalışmasından elde edilen soruların geçerlik ve güvenilirlik çalışmaları yapılmıştır. Bu yönüyle bu araştırma alan yazındaki söz konusu çalışmaların bir tamamlayıcısı olarak görülebilir.

#### **Etik Metni**

Çalışmada yürütülen uygulamalar, bir devlet üniversitesinin Bilimsel Araştırma ve Yayın Etik Kurulu (21.04.2020 tarih, 2020/01 sayılı toplantı) tarafından onaylanmıştır. Bu makalede dergi yazım kurallarına, yayın ilkelerine, araştırma ve yayın etiği kurallarına, dergi etik kurallarına uyulmuştur. Makale ile ilgili doğabilecek her türlü ihlallerde sorumluluk yazara aittir.

**Yazarın Katkı Oranı Beyanı:** Bu araştırma, Dumlupınar Üniversitesi Bilimsel Araştırma Projeleri Birimi (BAP) tarafından desteklenmiştir (Proje No: 2018-34). Yazarın katkı oranı %100’dür. Yazar çalışmayı tasarlayan, teorik çerçeveyi geliştiren, verileri toplayan, analiz eden ve makaleyi yazan kişi olarak, herhangi bir çıkar çatışması olmadığını beyan etmektedir. Uzman görüşleri konusundaki değerli katkılarından dolayı araştırma görevlisi Mustafa KÖSE ve matematik öğretmeni Efan ÇAY’a; testlerin uygulama okullarına dağıtılması ve

toplanmasındaki değerli katkılarından dolayı araştırma görevlileri Enis Harun BAŞER, Samet DEMİR ve Berna TÜRE KÖSE'ye teşekkür ederim. Ayrıca, Dumlupınar Üniversitesi Bilimsel Araştırma Projeleri Koordinatörlüğü'ne, Kütahya Valiliği'ne, İl Milli Eğitim Müdürlüğü'ne, uygulama okullarının yöneticilerine ve ciddiyetle gözetmenlik yapan tüm öğretmenlerine değerli katkılarından dolayı teşekkür ederim.

#### KAYNAKÇA

- Acar, T. (2012). Türkiye'nin PISA 2009 sonuçlarına göre OECD'ye üye ve aday ülkeler arasındaki yeri. *Kuram ve Uygulamada Eğitim Bilimleri*, 12(4), 2561-2572.
- Altun, M. & Akkaya, R. (2014). Matematik öğretmenlerinin PISA matematik soruları ve ülkemiz öğrencilerinin düşük başarı düzeyleri üzerine yorumları. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 29(1), 19-34.
- Altun, M. & Bozkurt, I. (2017). Matematik okuryazarlığı problemleri için yeni bir sınıflama önerisi. *Eğitim ve Bilim*, 42(190), 171-188. <http://dx.doi.org/10.15390/EB.2017.6916>
- Annisah, U. (2011). *Pengembangan soal matematika model PISA pada konten quantity untuk mengukur kemampuan penalaran matematis siswa sekolah menengah pertama*. Prosiding Seminar Nasional Pendidikan 2011. Fakultas Keguruan dan Ilmu Pendidikan, Universitas Sriwijaya. Palembang, 7 Mei 2011.
- Atılğan, H., Kan, A. & Doğan, N. (2017). *Eğitimde ölçme ve değerlendirme* (9. bs.). Anı Yayıncılık.
- Aydın, A., Sarier, Y. & Uysal Ş. (2014). PISA sonuçları bağlamında öğrencilerin akademik başarılarının değerlendirilmesi. *İlköğretim Online*, 13(3), 1065-1074.
- Başol, G. (2019). *Eğitimde ölçme ve değerlendirme* (6. bs.). Pegem Akademi.
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. ÖSYM Yayınları
- Boaler, J. (1993). The role of contexts in the mathematics classroom: Do they make mathematics more "real"? *For the Learning of Mathematics*, 13(2), 12-17.
- Breakspear, S. (2012). The policy impact of PISA: An exploration of normative effects of international benchmarking in school system performance. *OECD Education Working Papers* (No: 71). OECD Publishing.
- Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. Holt, Rinehart and Winston, 6277 Sea Harbor Drive, Orlando, FL 32887.
- Çakır, İ. (2009). *İlköğretim 5. sınıf ders kitaplarının öğretmen ve öğrenci görüşleri doğrultusunda değerlendirilmesi* (Yayımlanmamış yüksek lisans tezi). Çukurova Üniversitesi.
- Dede, Y. & Yaman, S. (2005). *İlköğretim 6, 7 ve 8. sınıf matematik ve fen bilgisi ders kitaplarının incelenmesi: problem kurma ve çözme etkinlikleri bakımından*. XVI. Ulusal Eğitim Bilimleri Kongresi'nde sunulmuş bildiri, Denizli.
- Demir, F. (2015). *Matematik Okuryazarlığı soru yazma süreç ve becerilerinin gelişimi* (Yayımlanmamış doktora tezi). Uludağ Üniversitesi.
- Demir, F. & Altun, M. (2018). Matematik okuryazarlığı soru yazma süreç ve becerilerinin gelişimi. *Eğitim ve Bilim*, 43(194), 19-41. <http://dx.doi.org/10.15390/EB.2018.7111>

- Demir, F. (2019). Matematik okuryazarlığı soru yazma sürecinde yer alan eylemlerin belirlenmesi ve sıralarının kestirilmesi. *Bayburt Eğitim Fakültesi Dergisi*, 14(28), 372-390. <https://doi.org/10.35675/befdergi.6434>
- Demir, F. (2020). Mathematics literacy curriculum design based on question writing. *International Journal of Education Technology and Scientific Researches*, 5(11), 35-62. <http://dx.doi.org/10.35826/ijetsar.71>
- Ekawati, R., Susanti, S. & Chen, J. C. (2020). Primary students' mathematical literacy: A case study. *Infinity Journal*, 9(1), 49-58.
- Fan, L. & Zhu, Y. (2000). Problem solving in Singaporean secondary mathematics textbooks. *The Mathematics Educator*, 5(1/2), 117-141.
- İskenderoğlu, T. A. & Baki, A. (2011). İlköğretim 8. sınıf matematik ders kitabındaki soruların PISA matematik yeterlilik düzeylerine göre incelenmesi. *Eğitim ve Bilim*, 36(161), 287-300.
- Jürges, H., Schneider, K., Senkbeil, M. & Carstensen, C. H. (2012). Assessment drives learning: The effect of central exit exams on curricular knowledge and mathematical literacy. *Economics of Education Review*, 31(1), 56-65.
- Kamaliyah, K., Zulkardi, Z. & Darmawijoyo, D. (2013). Developing the sixth level of PISA-like mathematics problems for secondary school students. *Indonesian Mathematical Society Journal on Mathematics Education*, 4(1), 9-28.
- Kohar, A. W., Zulkardi, Z. & Darmawijoyo, D. (2014). Developing PISA-like mathematics tasks to promote students' mathematical literacy. In *Proceeding in the Second South East Asia Design/Development Research (SEA-DR) Conference* (14-26).
- Leikin, R. & Levav-Waynberg, A. (2007). Exploring mathematics teacher knowledge to explain the gap between theory-based recommendations and school practice in the use of connecting tasks. *Educational Studies in mathematics*, 66(3), 349-371.
- Lestariningsih, L., Nurhayati, E., Susilo, T., Cicinidia, C. & Lutfianto, M. (2020). Development of mathematical literacy problems to empower students' representation. *Journal of Physics: Conference Series*, 1464 012018.
- Lloyd, G. M., Herbel-Eisenmann, B. A. & Star, J. R. (2011). *Developing essential understanding of expressions, equations, and functions for teaching mathematics in grades 6–8*. Reston, VA: National Council of Teachers of Mathematics.
- Mardhiyanti, D., Putri, R. I. I. & Kesumawati, N. (2013). Pengembangan soal matematika model PISA untuk mengukur kemampuan komunikasi matematis siswa sekolah dasar. *Jurnal Pendidikan Matematika UNSRI*, 5(1). 16-29. <https://doi.org/10.22342/jpm.5.1.334>.
- MEB. (2018). *Matematik dersi öğretim programı*. Milli Eğitim Bakanlığı.
- MEB. (2019). *PISA 2018 Türkiye raporu eğitim analiz ve değerlendirme raporları serisi No:10*. Milli Eğitim Bakanlığı.
- Murdiyani, N.M. (2018). Developing non-routine problems for assessing students' mathematical literacy. *Journal of Physics: Conference Series*, 983, 1-6.

- Nasution, R. S., Fauzi, K. M. A. & Syahputra, E. (2019). Developing mathematics problem based on PISA level of space and shape content to measure student's mathematics problem solving ability. *American Journal of Educational Research*, 7(10).
- National Council of Teachers of Mathematics. (2002). *Okul matematiğinin prensipleri ve standartları* (Çev. C.H. Böke). NCTM Publications.
- National Governors Association Center for Best Practices and Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. NGA.
- OECD. (2013a). *PISA 2015 draft mathematics framework*. OECD Publishing.
- OECD. (2013b). *PISA 2012 results: excellence through equity: giving every student the chance to succeed (Volume II), PISA*. OECD Publishing.
- OECD, (2013c). *What Students Know and Can Do: Student Performance in Math, Reading and Science Volume I*. PISA OECD Publishing.
- OECD. (2016). *PISA 2015 assessment and analytical framework. Science, reading, mathematic and financial literacy*. OECD Publishing.
- (OECD. (2019) *PISA 2018 results volumel: What students know and can do*. OECD Publishing.
- Oktiningrum, W., Zulkardi, Z. & Hartono, Y. (2016). Developing PISA-like mathematics task with indonesia natural and cultural heritage as context to assess students' mathematical literacy. *Journal on Mathematics Education*, 7, 1-8. <http://dx.doi.org/10.22342/jme.7.1.2812.1-8>
- Özgen, K. (2013). Problem çözme bağlamında matematiksel ilişkilendirme becerisi: öğretmen adayları örneği. *NWSA-Education Sciences*, 8(3), 323-345.
- Reinke, L. T. (2020). Contextual problems as conceptual anchors: an illustrative case. *Research in Mathematics Education*, 22(1), 3-21. <https://doi.org/10.1080/14794802.2019.1618731>
- Robinson, E. E., Robinson, M. F. & Maceli, J. C. (2000). The impact of standards-based instructional materials in mathematics in the classroom. In M. J. Burke (Ed.), *Learning mathematics for a new century* (Vol. 11, 112–126). National Council of Teachers of Mathematics.
- Silva, E. Y. (2011). *Pengembangan soal matematika model PISA pada konten uncertainty untuk mengukur kemampuan pemecahan masalah matematika siswa sekolah menengah pertama*. Prosiding Seminar Nasional Pendidikan 2011. Fakultas Keguruan dan Ilmu Pendidikan, Universitas Sriwijaya. Palembang, 7 Mei 2011.
- Smith, A. (2004). *Making mathematics count*. The report of Professor Adrian Smith's inquiry into post-14 mathematics education. London.
- Stacey, K. & Vincent, J. (2009). Modes of reasoning in explanations in Australian eighth-grade mathematics textbooks. *Educational Studies in Mathematics*, 72(3), 271.
- Stacey, K. (2013). E-posta aracılığıyla araştırmacı tarafından sorulan soruya PISA matematik uzman ekibinde yer alan Kaye Stacey tarafından verilen cevap, 07.12.2013.
- Tekin, H. (2019). *Eğitimde ölçme ve değerlendirme* (27. bs.). Yargı Yayınevi.

- Tessmer, M. (1993). *Planning and Conducting Formative Evaluations: Improving the Quality of Education and Training*. Kogan Page.
- Thompson, D. R., Senk, S. L. & Johnson, G. J. (2012). Opportunities to learn reasoning and proof in high school mathematics textbooks. *Journal for Research in Mathematics Education*, 43(3), 253-295. <https://doi.org/10.5951/jresemetheduc.43.3.0253>
- Tomlinson, M. (2004). *14-19 curriculum and qualifications reform*. Final report of the working group on 14-19 reform. London.
- Turgut, M. F. & Baykul, Y. (2010). *Eğitimde ölçme ve değerlendirme* (2. bs.). Pegem Akademi.
- Uysal, E. & Yenilmez, K. (2011). Sekizinci sınıf öğrencilerinin matematik okuryazarlığı düzeyi. *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi*, 12(2), 1-15.
- Wardono, M. S. & Mariani, S. (2014). The Realistic Learning Model With Character Education And PISA Assessment To Improve Mathematics Literacy. *International Journal of Education and Research*, 2(7), 361-372.
- Wong, L. F. & Low, L. (2020). Strategies to foster mathematical reasoning and communication. In N. H. Lee, C. Seto, R. A. Rahim and L. S. Tan (Eds.) *Mathematics Teaching in Singapore*, 107-122. World Scientific.
- Wongpanya, N., Pativisan, S. & Ngamman, J. (2018). *Development of mathematics test items*. 37. Annual Conference of International Association for Educational Assessment. Accessed from <https://iaea.info/documents/development-of-mathematics-test-items/>.
- Yuentrakulchai, T., Kamtet, W. & Dechsri, P. (2018). *Development of mathematics test items*. 37. Annual Conference of International Association for Educational Assessment. Accessed from <https://iaea.info/documents/development-of-science-test-items/>.

## EK 1. MADDE TANIMLAMA TABLOSU

Sıra No	Soru ID	Bağlam*	Mat. İçerik**	Mat. Süreç***	P <sub>j</sub>	r <sub>jk</sub>	S <sub>j</sub>	r <sub>j</sub>
1	2	ML	Di	KAY	0,33	0,46	0,47	0,21
2	3	KS	NC	KAY	0,33	0,38	0,47	0,18
3	4	ML	NC	KAY	0,76	0,41	0,43	0,18
4	5	KS	NC	KAY	0,52	0,51	0,50	0,25
5	6	KS	UZ	KAY	0,19	0,33	0,39	0,13
6	6R1	KS	UZ	KAY	0,20	0,32	0,40	0,13
7	7	KS	NC	FRM	0,26	0,37	0,44	0,16
8	7R1	KS	NC	FRM	0,33	0,41	0,47	0,19
9	8	KS	Di	FRM	0,36	0,46	0,48	0,22
10	9	TPL	Di	MCY	0,26	0,33	0,44	0,14
11	9R1	TPL	Di	MCY	0,26	0,46	0,44	0,20
12	11R1	KS	UZ	FRM	0,33	0,44	0,47	0,21
13	12	ML	UZ	KAY	0,63	0,70	0,48	0,34
14	13R1	ML	Di	FRM	0,24	0,46	0,43	0,19
15	14	ML	UZ	KAY	0,25	0,36	0,43	0,15
16	15A	TPL	NC	KAY	0,56	0,59	0,50	0,29
17	15B	TPL	NC	KAY	0,79	0,45	0,41	0,18
18	15C	TPL	NC	FRM	0,54	0,69	0,50	0,34
19	16	TPL	BV	MCY	0,73	0,47	0,45	0,21
20	17	KS	NC	FRM	0,44	0,64	0,50	0,32
21	18	BİL	Di	FRM	0,28	0,46	0,45	0,21
22	19AR1	KS	NC	FRM	0,24	0,37	0,43	0,16
23	19BR1	KS	NC	FRM	0,18	0,35	0,38	0,13
24	19AR2	KS	NC	FRM	0,54	0,59	0,50	0,30
25	19BR2	KS	NC	FRM	0,38	0,59	0,48	0,28
26	20	KS	NC	KAY	0,80	0,47	0,40	0,19
27	21	ML	Di	MCY	0,66	0,71	0,47	0,34
28	22R1	KS	UZ	KAY	0,31	0,33	0,46	0,15
29	23AR1	KS	NC	MCY	0,79	0,32	0,41	0,13
30	23B	KS	NC	MCY	0,39	0,47	0,49	0,23
31	24A	ML	Di	FRM	0,51	0,57	0,50	0,28
32	24B	ML	Di	FRM	0,18	0,39	0,38	0,15
33	25	KS	NC	KAY	0,49	0,62	0,50	0,31
34	26	TPL	Di	FRM	0,47	0,60	0,50	0,30
35	27	ML	Di	MCY	0,18	0,30	0,38	0,11
36	27R1	ML	Di	MCY	0,16	0,31	0,36	0,11
37	28A	BİL	NC	KAY	0,72	0,49	0,45	0,22
38	28B	BİL	NC	KAY	0,58	0,53	0,49	0,26
39	30	KS	NC	KAY	0,66	0,50	0,47	0,24
40	32A	KS	NC	KAY	0,78	0,55	0,42	0,23
41	32B	KS	NC	KAY	0,41	0,55	0,49	0,27
42	33	ML	UZ	MCY	0,33	0,48	0,47	0,23
43	34R2	KS	Di	FRM	0,42	0,58	0,49	0,29
44	34R4	KS	Di	FRM	0,23	0,32	0,42	0,13
45	37	TPL	BV	MCY	0,30	0,32	0,46	0,15
46	38	KS	NC	FRM	0,32	0,52	0,47	0,24
47	40R1	KS	NC	KAY	0,48	0,45	0,50	0,22
48	42A	KS	NC	KAY	0,55	0,67	0,50	0,33
49	42B	KS	NC	KAY	0,66	0,62	0,47	0,29
50	43	KS	NC	KAY	0,40	0,61	0,49	0,30
51	44A	BİL	NC	MCY	0,69	0,43	0,46	0,20
52	44B	BİL	NC	MCY	0,26	0,34	0,44	0,15
53	44C	BİL	NC	MCY	0,75	0,40	0,43	0,17
54	44D	BİL	NC	MCY	0,31	0,37	0,46	0,17
55	45	KS	NC	KAY	0,73	0,44	0,44	0,19
56	46	TPL	UZ	MCY	0,65	0,44	0,48	0,21
57	47	KS	BV	MCY	0,44	0,71	0,50	0,35
58	48A	KS	Di	MCY	0,66	0,59	0,47	0,28
59	48B	KS	Di	MCY	0,45	0,70	0,50	0,35
60	48C	KS	Di	MCY	0,31	0,61	0,46	0,28
61	49A	ML	NC	FRM	0,36	0,48	0,48	0,23
62	50	ML	NC	FRM	0,52	0,64	0,50	0,32
63	51	ML	BV	MCY	0,49	0,41	0,50	0,21
64	51R1	ML	BV	MCY	0,54	0,57	0,50	0,28
65	52	ML	BV	MCY	0,76	0,51	0,43	0,22
66	53R1	TPL	BV	MCY	0,34	0,38	0,47	0,18
67	54	ML	NC	FRM	0,37	0,59	0,48	0,29
68	55	ML	NC	FRM	0,43	0,74	0,49	0,36
69	56	KS	NC	FRM	0,51	0,58	0,50	0,29
70	57	KS	UZ	KAY	0,35	0,46	0,48	0,22
71	61R1	ML	Di	FRM	0,32	0,30	0,47	0,14
72	62	TPL	Di	MCY	0,55	0,54	0,50	0,27
73	64	BİL	Di	MCY	0,70	0,51	0,46	0,23
74	65R1	BİL	Di	FRM	0,41	0,31	0,49	0,15

---

75	66B	TPL	UZ	FRM	0,22	0,30	0,41	0,12
76	68	KS	NC	FRM	0,44	0,37	0,50	0,18
77	69	KS	NC	FRM	0,56	0,52	0,50	0,26
78	70A	KS	Dİ	KAY	0,30	0,47	0,46	0,21
79	70AR1	KS	Dİ	KAY	0,35	0,39	0,48	0,19
80	71A	KS	NC	FRM	0,59	0,72	0,49	0,35
81	71B	KS	NC	FRM	0,54	0,69	0,50	0,35
82	72A	KS	NC	KAY	0,32	0,31	0,47	0,15
83	73	ML	NC	KAY	0,51	0,50	0,50	0,25
84	74	KS	Dİ	MCY	0,75	0,49	0,43	0,21
85	75	KS	BV	MCY	0,62	0,50	0,49	0,24
86	76A=76AR1	KS	BV	FRM	0,45	0,59	0,50	0,29
87	77	ML	NC	KAY	0,61	0,55	0,49	0,27
88	78	ML	NC	KAY	0,28	0,47	0,45	0,21
89	79R1	KS	NC	KAY	0,42	0,33	0,49	0,16
90	80	ML	Dİ	FRM	0,27	0,46	0,45	0,20
91	81	KS	NC	KAY	0,50	0,56	0,50	0,28
92	82A	ML	NC	FRM	0,35	0,30	0,48	0,14
93	82AR1	ML	NC	FRM	0,34	0,60	0,47	0,28
94	87	TPL	UZ	FRM	0,33	0,39	0,47	0,18

---

\* Bağlılar: BİL: bilimsel, KS: kişisel, ML: mesleki, TPL: toplumsal.

\*\* Matematiksel İçerikler: BV: belirsizlik ve veri, Dİ: değişim ve ilişkiler, NC: nicelik, UZ: uzay ve şekil.

\*\*\* Matematiksel Süreçler: FRM: Durumları, problemleri matematiksel olarak formüle etme, KAY: Matematiksel kavramları, gerçekleri, yöntemleri kullanma ve akıl yürütme, MCY: Matematiksel çıktıları yorumlama, uygulama ve değerlendirme.

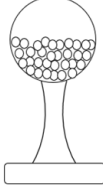


## EK 2. ÖRNEK SORULAR

ID\_76A

SAKIZ MAKİNESİ

3 ve 4. soruları aşağıdaki bilgilere göre cevaplayınız.



Yandaki paralı sakız makinesi 1 TL ile çalışmaktadır. Atılan her 1 TL'ye karşılık,

- Sırası 3 ve 3'ün katlarına rast gelen atışlarda 2 sakız,
- Sırası 5 ve 5'in katlarına rast gelen atışlarda 3 sakız,
- Avantajlı sakız alma durumlarını ikisini birden sağlayan sıraya rast gelen atışlarda 5 sakız,
- Diğer durumlarda 1 sakız vermektedir.

Örneğin 3 defa peşe 1 TL attığımızda; birinci atıştan 1, ikinci atıştan 1, üçüncü atıştan 2 sakız alırız. Böylece 3 TL attığımızda toplam 4 sakız almış oluruz.

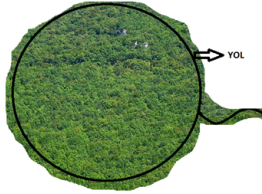
**SORU 3:** Buna göre Şevval aşağıdaki sayılardan hangisi kadar atış yaparsa kendisi için bir sakızın ortalama maliyeti en düşük olur?

- A) 27 atış  
B) 28 atış  
C) 29 atış  
D) 30 atış  
E) 31 atış

ID\_66B

ORMAN

Şekildeki koruma altına alınmış olan ormanlık alanın etrafına hem insanların gezebilmesi amacıyla hem de her hangi bir yangın çıkması durumunda müdahale edilebilmesi amacıyla yol yapılmıştır. Ormanlık alanın çapı yaklaşık 400 metredir. Yangına müdahale eden araçlar ise buldukları yerden en fazla 50 metre uzaklıktaki mesafeye kadar olan bölgeye müdahale edebilmektedir.



**SORU 6:** Araçların ormanın içlerinde çıkabilecek yangınlara da müdahale edebilmesi amacıyla yeni yollar eklenmesi düşünülmektedir. Yeni eklenecek yolların hem araçların ormanın tümüne müdahale edileceği şekilde oluşturulması hem de en az sayıda ağaç kesilerek yapılması planlanmaktadır. Bu plana uygun olarak ormanın içine yapılacak ve gerekli şartları sağlayacak yolların uzunluğu en az kaç metre olur?

- A) 800 metre  
B) 700 metre  
C) 600 metre  
D)  $(300\pi + 200)$  metre  
E)  $(200\pi + 200)$  metre

ID\_33

FOTOKOPI

**SORU 7:** Bir reklam firması boyutları 8 cm x 12 cm olan dikdörtgen şeklindeki bir afiş fotoğraf makinesiyle farklı boyutlarda çoğaltmaktadır. Çoğaltma işleminde afişin ebatları arasındaki 8/12 oranı korunmaktadır. Bu işlem için makinenin büyütme/küçültme tuşu kullanılmaktadır. Örneğin bu tuş % 200 e ayarlanırsa dikdörtgen afişin her bir kenarı orijinal halinin iki katı, % 25 e ayarlanırsa dikdörtgen afişin her bir kenarı orijinal halinin 1/4 katı uzunluğa olmaktadır.

Alanı 2400 cm<sup>2</sup> bir afiş elde etmek için bu tuş % kaçta ayarlanmalıdır?

- A) 500  
B) 2500  
C) 250  
D) 50  
E) 25

ID\_80

BAKIR KABLOLAR

SORU 9:



İki elektrik direği arasındaki mesafe 40 metre olacak şekilde, direkler arasında bakır kablolar asılacaktır. Bakır teller özellikleri gereği yazın genişlemekte, kışın büzümektedir. Sıcaklıktaki 1°C'lik değişim, 100 m bakır telin boyunda 10 cm değişime neden olmaktadır. Kabloların asıldığı gün sıcaklık 35°C'dir. Bu yörede kışın en düşük sıcaklığın -30°C olduğu bilinmektedir. Kışın bu kabloların büzülerek kopmasını önlemek için iki direk arasında asılan kabloların uzunluğu en az kaç cm olarak ayarlanmalıdır?

- A) 3 840  
B) 4 160  
C) 4 260  
D) 3 740  
E) 4 400

ID\_62

HERKES KENDİ SAHASINA!

**SORU 9:** Yaz basketbol kursuna 5-24 yaş arası herkes katılabilmektedir. Kurs katılanlar yaşlarına göre 5 gruba ayrılarak beş farklı basketbol sahasında çalışmaktadır.

Birinci grup 5-7 yaş

İkinci grup 8-10 yaş

Üçüncü grup 11-13 yaş

Dördüncü grup 14-16 yaş aralığında,

Beşinci grup 17 ve üzeri yaş

öğrencilerden oluşturulmaktadır.

- 1 Potaların Yükseklikleri  
2 Potaların Çapları  
3 Kullanılan Topların Çapları  
4 Kullanılan Topların Ağırıkları



Bu beş basketbol sahasında, yanda belirtilen dört özellik farklılık göstermektedir.

Beş sahaya ait potaların çapları ve sahalarda kullanılan topaların çapları aşağıda verilmiştir.

	Pota Çapı:	Top Çapı:
I. Saha	48 cm	19 cm
II. Saha	50 cm	19 cm
III. Saha	46 cm	22 cm
IV. Saha	45 cm	24 cm
V. Saha	45 cm	22 cm

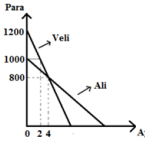
Her bir sahanın kaçınıcı gruba ait olabileceğini tahmin ediniz.

I. Saha	II. Saha	III. Saha	IV. Saha	V. Saha
A) 1. Grup	2. Grup	3. Grup	4. Grup	5. Grup
B) 2. Grup	1. Grup	4. Grup	3. Grup	5. Grup
C) 2. Grup	1. Grup	3. Grup	5. Grup	4. Grup
D) 1. Grup	3. Grup	4. Grup	5. Grup	2. Grup
E) 2. Grup	1. Grup	4. Grup	5. Grup	3. Grup

ID\_48

TUR ARKADAŞLIĞI

5, 6 ve 7. soruları aşağıdaki bilgilere göre cevaplayınız.



Grafik bisikletle Türkiye turuna çıkan iki arkadaşın cepelerindeki paranın zamana göre değişimini doğrusal olarak göstermektedir. Tura çıkmadan önce yaptıkları anlaşmaya göre paraları olan, paraları bitene tur tamamlanana kadar destek olacaktır.

SORU 5: ...

SORU 6: ...

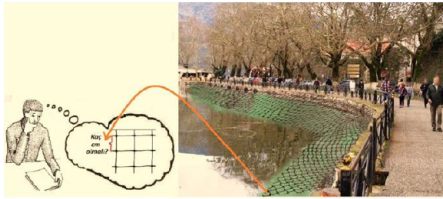
SORU 7: Tur boyunca kim kime kaç ay destek olmuştur? .

- A) Veli, Ali'ye 2 ay destek olmuştur.  
B) Veli, Ali'ye 8 ay destek olmuştur.  
C) Ali, Veli'ye 2 ay destek olmuştur.  
D) Ali, Veli'ye 8 ay destek olmuştur.  
E) Ali, Veli'ye 12 ay destek olmuştur.

ID-87

BÜYÜKŞEHİR ÇALIŞIYOR!

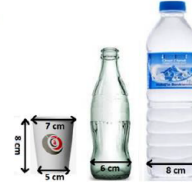
SORU 6:



Büyükşehir Belediyesi su kirliliğini önlemek amacıyla

Persuk Çayının kıyı perdesinin bir kısmını boydan boya kaplayacak şekilde es karelerden oluşan büyük bir ağ oducağı ve ağ yüzeyindeki resimde görüldüğü gibi bağlamayı planlanmaktadır. Böylece hem anıkarların toplanması kolaylaşacak hem de insanların farkındalığı artırılmış olacaktır.

Ağ örneğine işleme geçmeden önce Persuk Çayında biriken atıklar toplanmıştır. Bu atıkların içinde en çok rastlanılan ve ölçüleri yanda verilmiştir. Buna göre aşağıdaki resim hangileri örülecek ağdaki karelerin bir kenarının uzunluğu olmaz?



- I.  $\frac{\pi}{2}$  cm II.  $\pi$  cm III.  $2\pi$  cm IV.  $3\pi$  cm V.  $4\pi$  cm

- A) Yalnız V  
B) IV ve V  
C) III, IV ve V  
D) II, III, IV ve V  
E) I, II, III, IV ve V

## EK 3. ÖRNEK REVİZE İŞLEMLERİ

**ID\_34 ZEMİN KAPLAMA**

**SORU 6:** Ali Bey mutfağının zeminini kare fayanslarla kaplamak istiyor. Üç farklı şirketten aldığı fiyat teklifleri şu şekildedir. Birinci teklifte bir kenarı 20 cm olan kare fayansların tanesi 2 TL, ikinci teklifte ise bir kenarı 30 cm olan kare fayansların tanesi 3 TL, üçüncü teklifte bir kenarı 40 cm olan kare fayansların tanesi 4 TL'dir. Hangi teklifi kabul ettiğinde Ali Bey daha kârlı olur?

A) Mutfağın zemininin enini ve boyunu bilmeden hangisinin kârlı olduğunu hesaplayamayız.  
B) En kârlı tercih kenarı 20 cm olmalıdır.  
C) En kârlı tercih kenarı 30 cm olmalıdır.  
D) En kârlı tercih kenarı 40 cm olmalıdır.  
E) Üç teklif de kârlılık açısından aynıdır.

**ID\_34R1 ZEMİN KAPLAMA**

**SORU 6:** Ali Bey mutfağının zeminini kare fayanslarla kaplatacağıdır. Üç farklı şirketten aldığı fiyat teklifleri şu şekildedir.

- Birinci teklifte bir kenarı 20 cm olan kare fayansların tanesi 2 TL,
- İkinci teklifte ise bir kenarı 30 cm olan kare fayansların tanesi 3 TL,
- Üçüncü teklifte bir kenarı 40 cm olan kare fayansların tanesi 4 TL'dir.

Buna göre Ali Bey hangi teklifi kabul ettiğinde daha kârlı olur?

A) Mutfağın zemininin enini ve boyunu bilmeden hangisinin kârlı olduğunu hesaplayamayız.  
B) En kârlı teklif kenarı 20 cm olmalıdır.  
C) En kârlı teklif kenarı 30 cm olmalıdır.  
D) En kârlı teklif kenarı 40 cm olmalıdır.  
E) Üç teklif de kârlılık açısından aynıdır.

**ID\_34R4 ZEMİN KAPLAMA**

**SORU 10:** Ali Bey mutfağının zeminini kare fayanslarla kaplatacağıdır. Üç farklı şirketten aldığı fiyat teklifleri şu şekildedir.

- Birinci teklifte bir kenarı 20 cm olan kare fayansların tanesi 2 TL,
- İkinci teklifte ise bir kenarı 30 cm olan kare fayansların tanesi 3 TL,
- Üçüncü teklifte bir kenarı 40 cm olan kare fayansların tanesi 4 TL'dir.

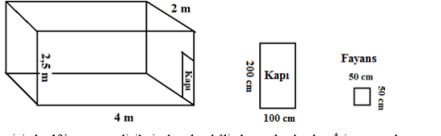
Buna göre Ali Bey hangi teklifi kabul ettiğinde daha kârlı olur?

A) Birinci ve ikinci teklif aynı düzeyde kârlıdır, her ikisi de üçüncü teklife göre daha kârlıdır.  
B) İkinci ve üçüncü teklif aynı düzeyde kârlıdır, her ikisi de birinci teklife göre daha kârlıdır.  
C) En kârlı teklif kenarı 20 cm olmalıdır.  
D) En kârlı teklif kenarı 30 cm olmalıdır.  
E) Üç teklif de kârlılık açısından aynıdır.

Item ID	Revize	Group	Option 1 N	Option 1	Option 2 N	Option 2	Option 3 N	Option 3	Option 4 N	Option 4	Option 5 N	Option 5
34		TOTAL	100	0,36	15	0,054	11	0,04	34	0,122	118	0,424
34		High	39	0,293	5	0,038	2	0,015	25	0,188	62	0,466
34		Low	31	0,397	9	0,115	8	0,103	5	0,064	25	0,321
34		Diff	8	-0,104	-4	-0,078	-6	-0,088	20	* 0,124	37	0,146
34		R1 TOTAL	87	0,339	22	0,086	14	0,054	50	0,195	82	0,319
34		R1 High	21	0,304	1	0,014	1	0,014	21	0,304	25	0,362
34		R1 Low	35	0,432	13	0,16	5	0,062	6	0,074	22	0,272
34		R1 Diff	-14	-0,128	-12	-0,146	-4	-0,047	15	* 0,23	3	0,091
34		R4 TOTAL	9	0,03	27	0,089	15	0,049	70	0,23	183	0,602
34		R4 High	0	0	2	0,019	1	0,01	44	0,427	56	0,544
34		R4 Low	6	0,066	15	0,165	8	0,088	10	0,11	52	0,571
34		R4 Diff	-6	-0,066	-13	-0,145	-7	-0,078	34	* 0,317	4	-0,028

**ID\_22 TEK SEFER**

**SORU 9:** Yeni ev alan Ayşe banyosuna fayans döşetecektir. Tavan ve kapı dışındaki her yer fayansla kaplanacaktır. Banyonun krokisi, kapısının ölçüleri ve bir fayansın ölçüleri aşağıda verilmiştir.



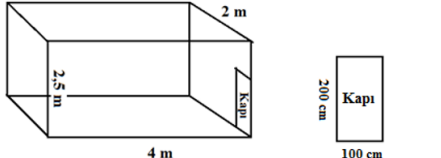
Fayanslar, içinde 12'şer tane dizilmiş kutular halinde satılmaktadır. İşin sonunda, artan açılmamış fayans kutuları geri götürüldüğünde bedelleri iade edilmektedir. Ayşe'nin evi fayansları satın alacağı yere oldukça uzaktır ve teslimat için fayansların fiyatının üzerine 250 TL taşıma ücreti vermesi istenmiştir. Bu nedenle Ayşe tek taşımada işini halletmelidir. Artanları geri getirmek ya da eksik kalırsa tekrar sipariş vermek ona pahalıya mâl olacaktır.

Bu durumda Ayşe kaç kutu fayans sipariş etmelidir?

A) 11  
B) 12  
C) 13  
D) 14  
E) 15

**ID\_22R1 TEK SEFER**

**SORU 9:** Yeni ev alan Ayşe banyosuna fayans döşetecektir. Banyonun krokisi, kapısının ölçüleri ve bir fayansın ölçüleri aşağıda verilmiştir.



Fayanslar, içinde 12'şer tane dizilmiş kutular halinde satılmaktadır. İşin sonunda, artan açılmamış fayans kutuları geri götürüldüğünde bedelleri iade edilmektedir. Ayşe'nin evi fayansları satın alacağı yere oldukça uzaktır ve teslimat için fayansların fiyatının üzerine 250 TL taşıma ücreti vermesi istenmiştir. Bu nedenle Ayşe tek taşımada işini halletmelidir. Artanları geri getirmek ya da eksik kalırsa tekrar sipariş vermek ona pahalıya mâl olacaktır.

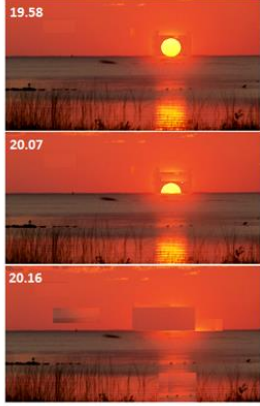
Banyoda tavan ve kapı hariç, her yer fayansla kaplanacağına göre, bu durumda Ayşe kaç kutu fayans sipariş etmelidir?

A) 11  
B) 12  
C) 13  
D) 14  
E) 15

Item ID	Revize	Group	Option 1 N	Option 1	Option 2 N	Option 2	Option 3 N	Option 3	Option 4 N	Option 4	Option 5 N	Option 5
22		TOTAL	53	0,175	89	0,294	62	0,205	48	0,158	41	0,135
22		High	23	0,181	54	0,425	20	0,157	16	0,126	11	0,087
22		Low	21	0,183	22	0,191	27	0,235	19	0,165	23	0,2
22		Diff	2	-0,002	32	* 0,234	-7	-0,077	-3	-0,039	-12	-0,113
22		R1 TOTAL	48	0,165	90	0,309	54	0,186	61	0,21	28	0,096
22		R1 High	13	0,157	41	0,494	10	0,12	16	0,193	3	0,036
22		R1 Low	20	0,175	19	0,167	25	0,219	26	0,228	17	0,149
22		R1 Diff	-7	-0,019	22	* 0,327	-15	-0,099	-10	-0,035	-14	-0,113

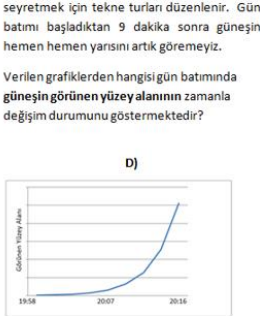
**ID\_65** **GÜN BATIMI**

**SORU 7:**




İznik gölünde 18 dakika süren gün batımını seyretmek için tekne turları düzenlenir. Gün batımı başladıktan 9 dakika sonra güneşin hemen hemen yarısını artık göremeyiz.

Verilen grafiklerden hangisi gün batımında güneşin görünen yüzey alanının zamanla değişim durumunu göstermektedir?



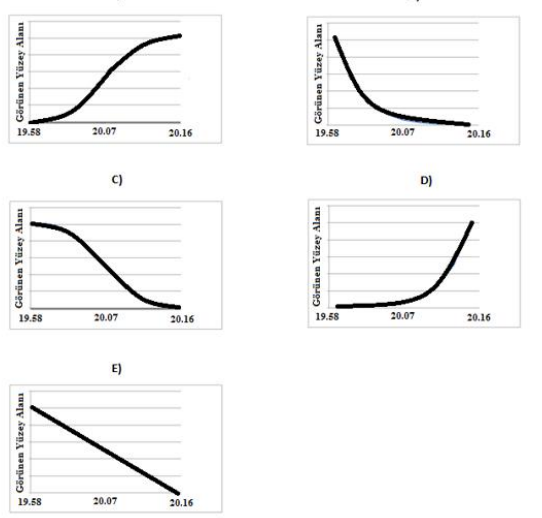
**ID\_65R1** **GÜN BATIMI**

**SORU 7:**



İznik gölünde 18 dakika süren gün batımını seyretmek için tekne turları düzenlenir. Yukarıdaki fotoğraflar İznik'te bir gün batımında çekilmiş ve çekildiği saatler üstlerine yazılmıştır. Gün batımı başladıktan 9 dakika sonra güneşin hemen hemen yarısını artık göremeyiz.

Buna göre, aşağıda verilen grafiklerden hangisi gün batımında güneşin görünen yüzeyinin alanının zamanla değişim durumunu göstermektedir?



Item ID	Revize	Group	Option 1 N	Option 1	Option 2 N	Option 2	Option 3 N	Option 3	Option 4 N	Option 4	Option 5 N	Option 5
65		TOTAL	6	0,019	44	0,143	137	0,445	6	0,019	115	0,373
65		High	0	0	11	0,098	62	0,554	0	0	39	0,348
65		Low	5	0,042	21	0,175	38	0,317	4	0,033	52	0,433
65		Diff	-5	-0,042	-10	-0,077	24	* 0,237	-4	-0,033	-13	-0,085
65		R1 TOTAL	14	0,047	35	0,117	121	0,406	14	0,047	113	0,379
65		R1 High	2	0,019	5	0,049	61	0,592	0	0	35	0,34
65		R1 Low	8	0,064	20	0,16	35	0,28	8	0,064	53	0,424
65		R1 Diff	-6	-0,045	-15	-0,111	26	* 0,312	-8	-0,064	-18	-0,084

**ID\_11** **BAHÇE PAYLAŞIMI**

**SORU 1:** İki kardeş 100 m<sup>2</sup> bahçeyi paylaşacaklardır. Küçük kardeş abisine "Bu makaraya sarılı olan ip ile etrafını çevreleyebildiğin topraklar senin, gerisi benim" demmiştir. Makaraya sarılı ipin uzunluğu 30 metredir.

Büyük kardeş toprağın etrafını çevrelerken ipe aşağıdaki şekillerden hangisini verirse daha kârlı olur?

A) Soruda verilenler kârlı olanı hesaplamak için yeterli değildir.

B) Eşkenar Üçgen

C) Kare

D) Dikdörtgen

E) Çember

**ID\_11R1** **BAHÇE PAYLAŞIMI**

**SORU 1:** İki kardeş 100 m<sup>2</sup> bahçeyi paylaşacaklardır. Küçük kardeş abisine "Bu makaraya sarılı olan ip ile etrafını çevreleyebildiğin topraklar senin, gerisi benim" demmiştir. Makaraya sarılı ipin uzunluğu 30 metredir.

Büyük kardeş toprağın etrafını çevrelerken ipe aşağıdaki şekillerden hangisini verirse daha kârlı olur?

A) İkizkenar Üçgen

B) Eşkenar Üçgen

C) Kare

D) Dikdörtgen

E) Çember

Item ID	Revize	Group	Option 1 N	Option 1	Option 2 N	Option 2	Option 3 N	Option 3	Option 4 N	Option 4	Option 5 N	Option 5
11		TOTAL	122	0,396	36	0,117	33	0,107	38	0,123	78	0,253
11		High	45	0,409	10	0,091	9	0,082	5	0,045	41	0,373
11		Low	35	0,376	14	0,151	14	0,151	18	0,194	12	0,129
11		Diff	10	0,033	-4	-0,06	-5	-0,069	-13	-0,148	29	* 0,244
11		R1 TOTAL	35	0,115	51	0,168	51	0,168	66	0,217	101	0,332
11		R1 High	8	0,078	11	0,107	14	0,136	12	0,117	58	0,563
11		R1 Low	13	0,143	19	0,209	23	0,253	25	0,275	11	0,121
11		R1 Diff	-5	-0,065	-8	-0,102	-9	-0,117	-13	-0,158	47	* 0,442

**ID\_53 MÜZESEVERLER**

**SORU 7:**

Grafikte, haftanın 7 günü açık olan bir müzenin girişinden 5 hafta boyunca geçiş yapan ziyaretçi sayıları ile ilgili bilgiler verilmiştir. Çubukların alt uçları en sakin günlerdeki ziyaretçi sayılarını, üst uçları en yoğun günlerdeki ziyaretçi sayılarını, noktalar ilgili haftanın **günlük ortalama** ziyaretçi sayılarını göstermektedir.

Buna göre aşağıdakilerden hangisi her hangi bir haftanın toplam ziyaretçi sayısı olamaz?

A) 273 B) 308 C) 350 D) 420 E) 560

**ID\_53R1 MÜZESEVERLER**

**SORU 7:**

Yandaki grafikte, haftanın 7 günü açık olan bir müzenin girişinden 5 hafta boyunca geçiş yapan ziyaretçi sayıları ile ilgili bilgiler verilmiştir. Çubukların alt uçları en sakin günlerdeki ziyaretçi sayılarını, üst uçları en yoğun günlerdeki ziyaretçi sayılarını, noktalar ise ilgili haftanın **günlük ortalama** ziyaretçi sayılarını göstermektedir.

Buna göre aşağıdakilerden hangisi her hangi bir haftanın ziyaretçi sayısı olamaz?

A) 273  
B) 308  
C) 350  
D) 420  
E) 560

Item ID	Revize	Group	Option 1 N	Option 1	Option 2 N	Option 2	Option 3 N	Option 3	Option 4 N	Option 4	Option 5 N	Option 5
53		TOTAL	85	0,282	37	0,123	38	0,126	35	0,116	97	0,322
53		High	56	0,427	8	0,061	13	0,099	15	0,115	36	0,275
53		Low	12	0,138	18	0,207	19	0,218	7	0,08	30	0,345
53		Diff	44	* 0,29	-10	-0,146	-6	-0,119	8	0,034	6	-0,07
53		R1 TOTAL	92	0,337	38	0,139	37	0,136	25	0,092	74	0,271
53		R1 High	65	0,524	14	0,113	14	0,113	5	0,04	24	0,194
53		R1 Low	11	0,143	18	0,234	12	0,156	16	0,208	18	0,234
53		R1 Diff	54	* 0,381	-4	-0,121	2	-0,043	-11	-0,167	6	-0,04

**ID\_40 EKMEK KIZARTMA**

**SORU 8:** Ece sabah kahvaltısı için eşine, oğluna ve kendisine birer dilim ekme kızırtacaktır. Fakat tavaya **en fazla 2 dilim** sığmaktadır. Tavada **ekmeğin bir yüzü 5 dakikada kızartırsa**, Ece 3 dilim ekmeğin her iki yüzünü de **en az kaç** dakikada kızartabilir?

A) 20  
B) 17,5  
C) 15  
D) 12,5  
E) 10

**ID\_40R1 EKMEK KIZARTMA**

**SORU 8:** Ece sabah kahvaltısı için eşine, oğluna ve kendisine birer dilim ekme kızırtacaktır. Fakat tavaya **en fazla 2 dilim** sığmaktadır. Tavada **ekmeğin bir yüzü 5 dakikada kızartmaktadır**.

Buna göre Ece 3 dilim ekmeğin her iki yüzünü de **en az kaç** dakikada kızartabilir?

A) 17,5  
B) 15  
C) 12,5  
D) 10  
E) 7,5

Item ID	Revize	Group	Option 1 N	Option 1	Option 2 N	Option 2	Option 3 N	Option 3	Option 4 N	Option 4	Option 5 N	Option 5
40		TOTAL	210	0,675	12	0,039	68	0,219	5	0,016	15	0,048
40		High	95	0,669	2	0,014	40	0,282	2	0,014	3	0,021
40		Low	66	0,695	8	0,084	10	0,105	2	0,021	9	0,095
40		Diff	29	-0,026	-6	-0,07	30	* 0,176	0	-0,007	-6	-0,074
40		R1 TOTAL	76	0,26	139	0,476	20	0,068	43	0,147	11	0,038
40		R1 High	13	0,146	65	0,73	2	0,022	7	0,079	1	0,011
40		R1 Low	26	0,317	23	0,28	9	0,11	18	0,22	6	0,073
40		R1 Diff	-13	-0,171	42	* 0,45	-7	-0,087	-11	-0,141	-5	-0,062

**ID\_79 ARAÇ KİRALAMA**

**SORU 3:**

Marka	Yakıt Türü	100 km'de Tüketilen Yakıt Miktarı (lt)	Günlük Kiralama Bedeli
Alfa	Benzin	5	160 TL
Beta	Mazot	3,5	180 TL
Gama	Mazot	3	200 TL
Tetra	Benzin	3,5	150 TL

Mazot litre fiyatı: 6 TL ve Benzin litre fiyatı: 6,5 TL

Osman Bey, Kütahya'dan Adana'ya gidip gelmek için 3 günlüğüne araç kiralayacaktır. Kütahya-Adana arası 700 kilometredir. Tabloda bilgileri verilen araçlardan hangisini seçmesi, Osman Bey için en ekonomik olur?

A) Alfa  
B) Beta  
C) Gama  
D) Tetra  
E) Araçlar arasında fark yok.

**ID\_79R1 ARAÇ KİRALAMA**

**SORU 3:**

Marka	Yakıt Türü	100 km'de Tüketilen Yakıt Bedeli	Günlük Kiralama Bedeli
Alfa	Benzin	32,5 TL	160 TL
Beta	Mazot	21 TL	180 TL
Gama	Mazot	18 TL	200 TL
Tetra	Benzin	35,75 TL	150 TL

Osman Bey, Kütahya'dan Adana'ya gidip gelmek için 3 günlüğüne araç kiralayacaktır. Kütahya - Adana arası 700 kilometredir. Tabloda bilgileri verilen araçlardan hangisini seçmesi, Osman Bey için en ekonomik olur?

A) Alfa  
B) Beta  
C) Gama  
D) Tetra  
E) Araçlar arasında fark yoktur.

Item ID	Revize	Group	Option 1 N	Option 1	Option 2 N	Option 2	Option 3 N	Option 3	Option 4 N	Option 4	Option 5 N	Option 5
79		TOTAL	30	0,103	83	0,285	84	0,289	83	0,285	10	0,034
79		High	8	0,073	32	0,291	36	0,327	27	0,245	7	0,064
79		Low	14	0,116	37	0,306	30	0,248	38	0,314	2	0,017
79		Diff	-6	-0,043	-5	-0,015	6	* 0,079	-11	-0,069	5	0,047
79		R1 TOTAL	18	0,064	118	0,418	92	0,326	43	0,152	11	0,039
79		R1 High	1	0,01	60	0,588	25	0,245	15	0,147	1	0,01
79		R1 Low	14	0,118	31	0,261	51	0,429	17	0,143	6	0,05
79		R1 Diff	-13	-0,108	29	* 0,328	-26	-0,183	-2	0,004	-5	-0,041