



Development of A 4th Grade Achievement Test About Perimeter Measures: Validity and Reliability Study*

Cumali ÖKSÜZ¹ Galip GENÇ²

Article Info

DOI:

10.29329/jirte.2021.372.1

Article History:

Received: 27/04/2021

Revised: 11/08/2021

Accepted: 15/09/2021

Type: Research

Keywords

Achievement Test

Validity

Reliability

Perimeter Measurement

Abstract

In this study, the aim was to develop an achievement test ensuring valid and reliable measurement of student success related to the topic of measuring perimeter in 4th-grade mathematic lessons. Within the scope of the study, a specification table was prepared according to the acquirements included in the primary school fourth-grade mathematic curriculum and an item pool was created in line with this. After receiving expert opinion, items were included on the test in accordance with difficulty level and the face validity of the test. A duration of 40 minutes was determined to perform the test. In order to determine the measurement power of items for the targeted features and understandability of the test, it was applied to 9 students in the 5th grade who had previously learned the topic with the verbal thinking technique one by one. Necessary corrections were made accordingly. The actual implementation of the test was performed with 164 middle school fifth-grade students. Results obtained from the test were analyzed with the test analysis program (TAP) and jMetric program. From these results, the final 21-item test comprised 18 multiple choice questions, 2 open-ended questions, and 1 short-response question. Reliability analysis of the test found the KR-20 value was 0.81. The response duration for the test was determined as 35 minutes. As a result of analyses of this test prepared about the topic of perimeter measures in primary school 4th grade, it was determined to be a valid and reliable test to determine the success of 4th-grade students.

4. Sınıf Çevre Ölçme Konusunda Başarı Testi Geliştirme: Geçerlik ve Güvenirlik Çalışması

Makale Bilgisi

DOI:

10.29329/jirte.2021.372.1

Makale Geçmişi:

Alındı: 27/04/2021

Düzeltilme: 11/08/2021

Kabul: 15/09/2021

Tür: Araştırma

Anahtar Sözcükler

Başarı testi,

Geçerlik, Güvenirlik,

Çevre Ölçme.

Özet

Bu çalışmada 4. sınıf matematik dersi çevre ölçme konusu ile ilgili öğrenci başarısını geçerli ve güvenilir bir şekilde ölçmeye imkân sağlayacak bir başarı testinin geliştirilmesi amaçlanmıştır. Çalışma kapsamında, ilkokul dördüncü sınıf matematik ders programında yer alan kazanımlara göre belirtke tablosu hazırlanmış ve bu doğrultuda madde havuzu oluşturulmuştur. Uzman görüşü alındıktan sonra maddeler zorluk düzeyine ve testin görünüş geçerliğine uygun olarak teste yerleştirilmiştir. Testin uygulanması için 40 dakikalık bir süre belirlenmiştir. Test, maddelerin amaçladığı özellikleri ölçme gücü ve anlaşılabilirliğinin ortaya konulabilmesi bakımından daha önce konuyu görmüş olan 5. sınıfta öğrenim gören 9 öğrenciye sesli düşünme tekniği ile birebir uygulanmış ve buna bağlı olarak gerekli düzenlemeler yapılmıştır. Testin asıl uygulaması 164 ortaokul 5.sınıf öğrencisi üzerinde yapılmıştır. Testten elde edilen sonuçlar Test Analiz Programı (Test Analysis Program-TAP) ve jMetric programı ile analiz edilmiştir. Elde edilen sonuçlara 18 tane çoktan seçmeli, 2 tane açık uçlu soru ve 1 tane kısa cevaplı sorudan oluşan 21 maddelik nihai test oluşturulmuştur. Testin güvenilirlik analizi sonucunda KR-20 değeri 0,81 olarak bulunmuştur. Testin cevaplama süresi olarak 35 dakika belirlenmiştir. İllkokul 4.sınıf Çevre ölçme konusu ile ilgili hazırlanan bu testin yapılan analizler sonucunda, 4. sınıf öğrencilerinin başarılarını belirlemede geçerli ve güvenilir bir test olduğu ortaya konulmuştur.

Citation: Öksüz, C., & Genç, G. (2021). Development of a 4th grade achievement test about perimeter measures: Validity and reliability study. *Journal of Innovative Research in Teacher Education*, 2(2), 79-114. <https://doi.org/10.29329/jirte.2021.372.1>

¹ Aydın Adnan Menderes University, Primary Education Department, Turkey, cumalioksuz@gmail.com

² Aydın Adnan Menderes University, Primary Education Department, Turkey, galipgencc@gmail.com

* This research was carried out with the approval of Aydın Adnan Menderes University Educational Research Ethics Committee with the decision numbered "2331" in the session dated 09/02/2021.

INTRODUCTION

Education is not a simple process where information is transferred to students. Education should include activities ensuring students internalize information by association with their lives. However, obtaining realistic information about the effect of these activities on student behavior is an important part of education. During the educational process, an attempt is made to determine through measurement the degree to which each student has acquired the desired behavior change, and the level to which the desired skills and abilities are achieved (Atılğan, Kan & Doğan, 2006). Measurement in education is performed with the aim of both making educational decisions and a range of inferences about the development and functioning of students in the teaching process. It is necessary to determine the measurement target according to the suitability of inferences to be made (Fidan, 2013; Turgut & Baykul, 2012). For this, it is necessary to use valid and reliable tests prepared with scientific effort and methodology to measure the success of students in educational environments. At this point, the most frequently used tests are known to be achievement tests.

Achievement tests in education are concrete tools guiding teachers and students in determining the level of the students, observing deficiencies related to the topic, and determining learning-teaching activities related to these. For achievement tests developed in the literature to provide accurate results, each measurement tool must have three important qualities. These qualities can be listed as validity, reliability, and practicality (Turgut & Baykul, 2012). The validity, reliability, and practicality concepts are very frequently used in the measurement. According to Büyükkıdık (2020), reliability is related to the repeatability of research results, while validity is associated more with the accuracy of research results. Ellez (2014) stated validity of a scale tool is the accurate measurement of the targeted feature without confusion with any other feature. Another concept required for scale tools is reliability. According to Karasar (2012), reliability is the stability between independent measurements of any feature. Additionally, Baykul (2000) stated that reliability shows that measurements of any quality made on the same individuals will obtain the same results under similar conditions. Another quality of measurement tool is practicality. According to Özçelik (2013), practicality is defined as the ease of implementation of a measurement tool. For this reason, measurement tools used in education should be valid, reliable, and practical. This situation indicates the high psychometric quality of the measurement tool.

According to Pressley et al. (1997), in Turkey and many other countries, tests comprising multiple-choice questions are used to measure the mental level or mental capacity of students and to allocate students to higher learning institutions (cited in Akbulut & Çepni, 2013). Multiple-choice items are a type of question where the question and prepared distractors are given to students. Multiple-choice tests come to the forefront because they have objective ratings, include the correct response within the item, are easy to implement and grade, and apply to crowded groups and every teaching level (Atılğan et al., 2006). Multiple-choice questions frequently used in large groups occasionally may not be adequate in small groups. This is because measurement questions should have different forms to determine acquisitions in small groups in more detail. For this reason, the tests must contain different types of questions.

Konak (2013) stated the view that the use of only multiple-choice questions may not be sufficient for the measurement of upper-level thinking skills on tests. İn'nami and Koizumi (2009) stated that no question type was full without deficiencies and suitable for every situation (cited in in Koyuncu, 2017). In light of this information, different question items should be used instead of only multiple-choice questions on tests. This will contribute to the test being more valid and reliable. The use of open-ended questions, which allow the opportunity to observe how the students understand the concepts in more depth outside of the choices given, should be included on tests. Hong (1984) stated that open-ended questions left the person free when answering the question and the educators can observe the students' information gaps and errors in a real sense with these questions (cited in Koyuncu, 2017). In this way, students must consider, design, organize and write their answers to the questions for open-ended questions (Koyuncu, 2017; Turgut & Baykul, 2014). For this reason, for achievement tests to be prepared for students at every class level, experts or teachers should pay attention to the necessary validity, reliability, and practicality of measurement tools.

The development process for an accurate achievement test is a systematic process because the features of the test to be developed should be determined beforehand. Stated differently, the preparation of a test with previously determined features makes it necessary to follow a range of systematic process stages (Atilgan et al., 2006). According to Turgut and Baykul (2012) and Atilgan and colleagues (2006), some important steps are recommended when developing tests. Accordingly, firstly it is necessary to determine what aim the test will be used for. Then, the behavior which the test will measure should be determined. Following this, the test items should be written and reviewed. A trial form should be created from the test items and a trial implementation should be completed. After implementing the trial, the answer sheets should be graded to perform item analysis and the selection process for items on the test should be completed. Finally, the final test should be created and statistics should be estimated. In order to determine the features of the achievement test beforehand, it is necessary to know the features of the implementation group well. It is not possible to follow these systematic stages without paying attention to features like the size of the group, age interval, and knowledge level. In light of this information, the expected acquirments of the students in the group may be determined at more realistic levels.

The features to be measured by the achievement test should be the knowledge, skills, and attitudes targeted for acquirement during the education-teaching process. Targets in education, in addition to guiding teaching, also function as a guide to topics and important points encompassed in measurements to be performed (Demirel, 2010; Fidan, 2013). In curricula, the acquirments targeted for acquisition by students at every class level are determined and achievement tests should be developed according to these acquirments. In achievement tests applied at the international level, especially, Turkey appears to remain below average. The TIMMS 2015 exam results are shown in Table 1 (Yıldırım, Özgürlük, Parlak, Gönen & Polat, 2016).

Table 1. TIMMS 2015 4th Grade Mathematics Learning Areas Assessment

TIMMS 2015 exam 4th-grade mathematic learning areas	Mean Success Points
Numbers	489
Data Display	476
Geometric shapes and measures	475

In 2015, 49 countries participated in the international TIMMS exam and Turkey was 36th with 483 points at the 4th-grade level. According to TIMMS 2015 data, the lowest points of 475 were obtained in the geometric shapes and measures field within the scope of learning fields (Yıldırım et al., 2016). Considering this situation, the importance of the "geometric shapes and measures" learning area acquirments emerge at the primary school 4th-grade level.

Geometric shapes and objects comprise the topic of geometry. In the primary school mathematic curriculum, it is very important that children acquire knowledge about features related to shapes and objects, generalizations, classifications, and drawings, and that these reach the level of application (Altun, 2018). In this situation, children must learn to use measures related to geometric shapes. When the primary school mathematic curriculum is investigated, concepts and skills related to the 'measures' area include basic knowledge and skills that will be encountered or required in the students' daily lives (Tan-Şişman & Aksu, 2009). When measures are mentioned in geometry, they mean perimeter, area, and volume (Dağlı & Peker, 2012). For the topic of perimeter measures, the perimeter topic should be taught by showing why perimeter calculations are needed, calculation of perimeter with measurement studies, measurement of perimeter using perimeter formulae, and application of the relationships between shapes and concepts (Altun, 2018). For this reason, to structure the perimeter measures topic better, teachers should not only deal with it as a formula-based topic but as a specific topic that needs to be understood and internalized by students. Research revealed that without understanding concepts like perimeter, area, and volume, especially, formulae are memorized, attempts are made to access results in this way, and errors and misconceptions of students emerge in questions related to the perimeter, area, and volume (Dağlı & Peker, 2012; Emekli, 2001; Gough, 2008; Moreira & Contente, 1997; Tan & Aksu, 2009). For this reason, it is necessary to ensure students have sufficient knowledge levels related to some concepts in their minds to be able to perform visualizations. A test developed in this

way will identify the success levels of students and is a test necessary to determine the degree of effect of the implemented method on success. Though there are studies with separate areas as a basis for the development of achievement tests in the national and international literature (Akgül, 2011; Fidan, 2013; Kurt, 2015; Tutak & Birgin, 2008), there is no research investigating only the topic of the perimeter. The study by Gürkan (2019) developed a achievement test for length and perimeter measures topics comprising 18 items and with a KR-20 test measurement reliability coefficient of 0.87. Six questions on this test were related to perimeter measures. However, the test developed in our research encompasses all acquirments about perimeter measures for the 4th grade and includes more questions related to the topic of only perimeter measures.

On the mathematic curriculum for primary school mathematic lessons updated in 2018, acquirments for students related to perimeter measures in the measurement sub-learning field are included in 3rd and 4th grade. There are 3 acquirments related to this topic in 4th grade (Ministry of National Education [MoNE], 2018). The perimeter topic is included in the 'measurement' topic in 4th grade and comprises calculations of perimeters of squares and rectangles, relationships between edge lengths and perimeter for these shapes, and problem-solving skills related to perimeter measures. For this reason, when literature screening is performed, there do not appear to be many studies related to the 'measurement' field, especially, among test development studies related to 4th-grade mathematic lessons in our country. This study aims to contribute to removing this gap in the literature. In this study, the development process for a achievement test about perimeter measures of geometric objects in the 4th grade within the scope of the 2018 mathematic curriculum is explained in detail. With this aim, the research questions are as follows:

- Is the test prepared to determine the success of 4th-grade students about the 'perimeter measures' topic valid?
- Is the test prepared to determine the success of 4th-grade students about the 'perimeter measures' topic reliable?

METHOD

Study Design

The research is a descriptive study to develop an achievement test prepared to determine success of students related to 4th-grade perimeter measurement. In this method, topics, people, objects, or events included in the research are defined and described as they are within their conditions in the past or present (Karasar, 2012). The research was descriptive considering the present situation of students to measure the success of 4th-grade students. As the research aimed to develop a valid and reliable measurement tool to measure the success of 4th-grade students about perimeter measures, the test development process steps were followed.

Participants

The study included 164 students attending 5th grade in Altıeylül county, Balıkesir province in the 2017-2018 educational year who had completed the 4th-grade 'perimeter measures' topic. The achievement test was created for 4th grade; however, as the 4th-grade classes had not learned the topic at the time of implementation and may leave questions unanswered, the study included 5th-grade students who had completed the topic. As these students were independent of each other, the selection of one student for the sample was not dependent on the selection of any other. For this reason, the students were identified with the simple random sampling method. Of these students, 86 were girls and 78 were boys.

The Development Process of The Data Collection Tool

In the first step, the achievement test for the 'perimeter measures' topic in the 'measures' learning area of the primary school 4th-grade mathematic lessons linked to the mathematic curriculum (2018) was developed. In the stage of developing the test, the acquirments related to the 'perimeter measures' topic in the primary school 4th-grade mathematic curriculum (2018) were listed. A specification table was created in Table 2 using the Bloom taxonomy in line with these acquirments. Later the researchers wrote trial items according to the specification table. The trial items were sent to experts (Turkish, mathematics, and measurement assessment experts) and feedback was obtained about validity. The test form was revised according to feedback from the

experts. In the next stage, the test form was prepared and this test was applied individually to 9 students in 5th grade who had completed the topic. After this pilot study, questions that were not understood and erroneous or deficient were revised. After this stage, the test was implemented for 164 students. Finally, item analysis was performed and the test was given its final form.

The research determined acquirments about the 'perimeter measures' topic according to the primary school 4th-grade mathematic curriculum updated in 2018. These acquirments were used to create the specification table below by the researchers according to the content on the curriculum and Bloom taxonomy levels. The cognitive areas of the Bloom taxonomy containing the acquirments about perimeter measures are shown in Table 2.

Table 2. Specification Table for achievement Test

Learning Area (Perimeter)	Cognitive Field (Bloom)						Total	Percentage
Acquirments	Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation		
1. Explain the relationship between perimeter and edge length for squares and rectangles	3 (Q 4, 8 and 9)	1 (Q 2)	3 (Q 11, 12 and 13)	2 (Q 5 and 16)	2 (Q 24 and 27)	3 (Q 19, 21 and 23)	14	51.8%
2. Create different geometric shapes with the same perimeter			1 (Q7)	2 (Q 25 and 27)			3	11.1%
3. Solve problems related to calculating the perimeters of shapes	1 (Q 3)	1 (Q 1)	4 (Q 6,10,14 and 15)	4 (Q 17, 18, 20, and 26)			10	37.1%
Total	4	2	8	8	2	3	27	100%
Percentage	14.8%	7.4%	29.6%	29.6%	7.4%	11.2%	10%	

An item pool containing a total of 27 items was created in accordance with acquirments. Of these questions, 20 items were developed by the researchers, and 7 items were derived from achievement tests in the studies given in Table 3.

Table 3. Test Questions, Acquirments, and References Used

Acquirments	Question number	References used
Explain the relationship between perimeter and edge length for squares and rectangles	4, 5, 9, 12, 13, 14, 18, 19, 20, 21, 22, 24, 25, 26	İnan (2016) Küçükaydın, (2017)
Create different geometric shapes with the same perimeter	3, 6, 27	Karadağ, Balcı, Abdik & Demiralp, (2017)
Solve problems related to calculating the perimeters of shapes	1, 2, 7, 8, 10, 11, 15, 16, 17, 23	

The test items created by the researchers and using the above studies were included in a trial item text including multiple-choice questions in addition to open-ended and practice questions to be able to observe the decision-making and upper-level skill development of students and reveal the difficulty of the mathematical expressions. According to Turgut and Baykul (2014), open-ended questions ask students to think for themselves, design, organize and write the desired answers. This construct was considered for this test.

Data Analysis

After writing the trial test items, they should be reviewed from multiple aspects and deficient aspects and deficiencies of the items should be corrected (Atılğan et al., 2006). The trial items to be used in the achievement test in the research were sent to experts in the mathematical field, Turkish field, and measurement-assessment field for checks. In line with these opinions, suitability criteria were revised again in terms of validity, scientific accuracy, linguistics, and technical aspects. After the necessary revisions and corrections, the prepared items were distributed according to the degree of difficulty and acquirement hierarchy so the test form comprised 20% difficult, 60% moderate, and 20% easy items. Appropriate directions for the mental level of the students were added. The researchers determined the font size of 14 was appropriate for the development level of primary school 4th-grade students. In the stage of implementing the test, firstly, a duration of 40 minutes was determined for the students to complete the test. Later, this test was administered to nine 5th grade students one-on-one in a pilot study, and misunderstood and erroneous questions were revised. Students used the oral thinking technique to explain their thoughts about the items while solving the problems and gestures and facial expressions were examined to observe their reactions to the items. In this context, analysis was performed and the items were given their final form. After the one-to-one implementation, 164 students attending 5th grade in Altieylül county, Balıkesir province were determined. These students were given the achievement test comprising a total of 27 items, 24 multiple-choice and 3 open-ended, within the duration.

Within the scope of the validity and reliability study for the achievement test, the test analysis program (TAP) was used for data analysis based on the classic test concept. The responses of students to each question were entered into the TAP program and thus, selection analysis was performed for each item. For the total of 24 multiple-choice questions the four choices were investigated, while for open-ended questions, two choices of correct-wrong (1-0) were used to analyze the distribution of responses in the TAP program. In this context, the result values for the wrong answers of the students to the open-ended questions were taken into consideration and these values were entered into the system as an option and analyzed.

For example, for the question "Draw a shape with perimeter equal to a rectangle with 20 cm perimeter on the dotted paper below", choice A was entered as the correct choice for squares and rectangles with 20 cm perimeter, while shapes with 14 cm perimeter were assigned to choice B, shapes with 22 cm perimeter were assigned to choice C and shapes with perimeter 18 cm were assigned to choice D and entered into the program. The selections created according to the distribution of responses to these open-ended questions were also stated in the analysis for the test. Situations, where different test selections may develop in different implementations, require the attention of researchers. Additionally, questions left blank by students were entered into the TAP program with a ". (full stop)". The number of wrong answers did not affect the number of correct answers when evaluating the test results. Accordingly, each correct response was given "1" point and each wrong answer was given "0" points. The highest points that can be obtained on the test are "27" with the lowest points of "0". Reliability and validity analyses for the measurement tool within the scope of the study were completed with the jMetric 4.1.1 program based on the item response concept. For assessment of results obtained from the program, unweighted and standardized fit statistics by Linacre (2002) were used and acceptable intervals for the criteria are shown in Table 4.

Table 4. Acceptable Intervals for Fit Statistics

Criterion	Interval	Interpretation
WMS and UMS	> .20	Deviates from topic features to be measured and lowers the quality of the measure
	1.5 – 2.0	Inefficient for measurement process but not too bad
	0.5 – 1.5	Suitable for measurement
	< .50	Less efficient for the measurement process, but not very bad. May produce reliable and distinctive values at high levels.
Std. WMS and Std. UMS	≥ 3	Data do not fit the model. Maybe trialed with a larger sample.
	2.0 – 2.9	Data cannot be estimated at a certain level
	-1.9 – 1.9	Data are predictable at reasonable levels
	≤ -2.0	Data is very predictable. Perhaps answer patterns in other dimensions may be limited.

When Table 4 is investigated, the ideal interval for WMS and UMS values was determined to be between 0.50 and 1.50. Similarly, the ideal interval for standardized WMS and UMS values was between -1.90 and 1.90. The WMS and Std. WMS columns show the weighted and standardized square mean incompatibility statistics. After these, the UMS and Std. UMS column shows the statistics outside the unweighted and standardized square mean criteria (Güzeller, Aksu & Eser, 2019).

FINDINGS

In this stage of the study, the item difficulty, item discrimination index, and reliability results from the TAP program are explained. Item analyses with the TAP program are shown in Table 5.

Table 5. Item Difficulty Index and Item Discrimination Index

Questions	Item Difficulty	Item Discrimination Index	Difficulty of items	Distribution according to answer selections (correct responses)				KR-20
				A	B	C	D	
1	0.21	0.24	Very difficult	28*	63	30	42	
2	0.65	0.39	Easy	10	18	25	106*	
3	0.39	0.19	Moderate	39	64*	8	24	
4	0.96	0.07	Very easy	1	3	158*	1	
5	0.11	-0.06	Very difficult	97	15	16	18*	
6	0.29	0.38	Difficult	48*	23	5	51	
7	0.46	0.80	Moderate	11	75*	38	34	
8	0.61	0.59	Moderate	11	17	100*	16	
9	0.34	0.55	Difficult	11	55*	56	8	
10	0.89	0.23	Very Easy	4	10	146*	1	
11	0.77	0.45	Easy	6	126*	6	17	
12	0.50	0.77	Moderate	17	82*	11	40	
13	0.40	0.82	Moderate	7	65*	13	71	
14	0.52	0.70	Moderate	17	8	85*	37	0.81
15	0.60	0.49	Moderate	13	98*	11	23	
16	0.50	0.62	Moderate	9	82*	8	43	
17	0.22	0.36	Difficult	11	80	10	36*	
18	0.23	0.23	Difficult	9	37*	28	23	
19	0.36	0.61	Moderate	34	20	17	59*	
20	0.33	0.14	Difficult	22	39	5	54*	
21	0.41	0.57	Moderate	27	17	13	68*	
22	0.23	0.09	Difficult	20	28	25	38*	
23	0.21	0.49	Difficult	13	35*	23	72	
24	0.33	0.48	Difficult	25	58	6	54*	
25	0.30	0.36	Difficult	18	48	13	49*	
26	0.34	0.56	Difficult	56*	11	20	18	
27	0.37	0.66	Moderate	61*	19	32	15	

*: Shows correct responses.

When the item analyses included in Table 5 are assessed, for items to be very good the item discrimination index value should be 0.40 and for items to be good and remain in the scale without any correction it should be from 0.30 to 0.39, according to Büyüköztürk, Kılıç Çakmak, Akgün, Karadeniz, & Demirel (2014). Additionally, for items to be revised and developed, it was stated the discrimination index value should be from 0.20 to 0.29. For an item to be removed from the test or revised completely, the item discrimination index value should be lower than 0.20.

Accordingly, item analyses with the TAP program found questions 22, 5, 3, 20, and 4 had discrimination index values lower than 0.20 and were removed from the test. Questions 1 and 22 with item discrimination index values from 0.20 to 0.29 were revised and remained in the test. To find the reliability values, each item was individually removed from the test and the final value was calculated according to the variation in reliability coefficients. As a result of the reliability analysis of the test, the KR-20 value was 0.84. Reliability coefficients of 0.70 and above are accepted as adequate for reliability calculations (Büyüköztürk et al., 2014). In this sense, the reliability coefficient obtained from the test is within this interval showing the measurement tool has adequate reliability.

After the necessary revisions were made to the achievement test, the total points students gained on the achievement test were the number of correct answers; each wrong answer on the achievement test was given "0" points and each correct answer was given "1" point. Table 6 shows the item analysis values for the final form of the achievement test according to the three acquisitions and lists the questions from easy to difficult for each acquisition.

Table 6. Item Difficulty Index and Item Discrimination Index for Perimeter Measures achievement Test

Questions	Item Difficulty	Item Discrimination Index	Distribution according to answer selections (correct responses)				KR-20
			A	B	C	D	
1	0.89	0.23	4	10	146*	1	0.837
2	0.77	0.46	6	126*	6	17	
3	0.50	0.61	9	82*	8	43	
4	0.50	0.82	17	82*	11	40	
5	0.40	0.83	7	65*	13	71	
6	0.52	0.72	17	8	85*	37	
7	0.36	0.52	34	20	17	59*	
8	0.33	0.46	25	58	6	54*	
9	0.34	0.52	11	55*	56	8	
10	0.22	0.34	11	80	10	36*	
11	0.23	0.23	9	37*	28	23	
12	0.50	0.82	17	82*	11	40	
13	0.60	0.53	13	98*	11	23	
14	0.34	0.56	56*	11	20	18	
15	0.46	0.81	11	75*	38	34	
16	0.61	0.64	11	17	100*	16	
17	0.41	0.51	27	17	13	68*	
18	0.29	0.32	48*	23	5	51	
19	0.21	0.52	13	35*	23	72	
20	0.30	0.37	18	48	13	49*	
21	0.34	0.56	56*	11	20	18	
22	0.21	0.24	28*	63	30	42	

According to Table 6, the final form of the test included 10 questions about the first acquisition related to perimeter measures, 3 questions related to the second acquisition, and 9 questions related to the third acquisition. The final form of the test included a clear and open statement of test directions for students and the duration of the test was determined as 35 minutes. For responses to multiple-choice questions not to have the quality of a code (like A-B-C-A-B-C), selections were revised again to create the test form. After this procedure, analyses were completed with the Rasch model based on the item response concept (IRC). The difficulty related to items, discrimination values, and fit index values for the 22 items included on the test are shown in Table 7.

Table 7. Rasch Analysis Results for Items on the Scale

Item no.	Difficulty	Std. Error	WMS	Std. WMS	UMS	Std. UMS
1	1.74	0.23	1.20	1.43	1.52	1.64
2	-1.15	0.19	1.16	1.93	1.77	3.84
6	0.83	0.20	1.17	1.68	1.27	1.49
7	-0.13	0.18	0.77	-3.13	0.68	-2.98
8	-0.95	0.18	0.92	-1.12	0.80	-1.40
9	0.60	0.19	0.99	-0.02	1.33	1.99
10	-3.90	0.37	0.96	-0.03	0.57	-0.47
11	-1.90	0.20	0.90	-0.99	0.82	-0.69
12	-0.36	0.18	0.76	-3.30	0.67	-3.05
13	0.21	0.19	0.77	-2.85	0.68	-2.77
14	-0.46	0.18	0.89	-1.52	0.87	-1.02
15	-0.88	0.18	1.04	0.52	1.04	0.31
16	-0.36	0.18	0.92	-1.07	0.94	-0.48
17	1.29	0.21	1.07	0.63	1.53	2.11
18	1.29	0.21	1.18	1.47	1.28	1.22
19	0.42	0.19	1.07	0.74	1.05	0.36
21	0.10	0.18	1.07	0.87	1.07	0.58
23	1.34	0.22	0.90	-0.80	0.78	-0.94
24	0.64	0.19	1.13	1.34	1.24	1.49
25	0.75	0.20	1.23	2.17	1.38	2.11
26	0.56	0.19	0.99	-0.11	0.96	-0.18
27	0.31	0.19	0.90	-1.17	0.79	-1.63

The fit statistics related to the items shown in Table 7 include the Unweighted Mean Square (UMS) and Weighted Mean Square (**WMS**) fit statistics. Of these values, WMS is accepted as the infit criterion, while UMS is accepted as the outfit criterion. When Table 5 is investigated, it appears that the measurement process performed according to WMS and UMS values for the tool comprising 22 items are very suitable. Though item number 2 had UMS value above 1.50, which is accepted as critical, it was concluded the fit statistics for the item are very good as 1.50-2.00 was stated to be the acceptable interval. According to standardized WMS and UMS values, items are between -2.00 and 2.00, and the model-data fit was ensured. However, when the item difficulty values are investigated, item number 10 had a difficulty value outside the acceptable margins of -3.00 and 3.00. According to analysis results obtained in a later stage, the decision will be made about whether to include this item in the test or not. Analyses were completed with the differential item functioning (DIF) to determine whether the items included on the tool are biased according to the gender variable. The results obtained based on the common odds ratio with the Mantel-Haenszel method are shown in Table 8.

Table 8. DIF Analysis Results

Item No	Chi-square	p-value	Valid N	Effect size	Class
1	0.79	0.38	135.00	0.67	A
2	1.09	0.30	141.00	0.67	A
6	0.83	0.36	137.00	0.71	A
7	0.37	0.54	58.00	0.51	A
8	0.16	0.69	119.00	1.24	A
9	0.82	0.37	154.00	1.46	A
10	0.56	0.45	96.00	0.70	A
11	1.55	0.21	131.00	0.63	A
12	0.00	0.98	151.00	0.99	A
13	0.15	0.70	69.00	1.32	A

14	0.09	0.76	119.00	0.88	A
15	0.51	0.47	116.00	0.74	A
16	4.77	0.03	113.00	2.67	B-
17	0.15	0.70	131.00	1.15	A
18	0.22	0.64	137.00	1.20	A
19	0.18	0.67	139.00	1.19	A
21	0.00	0.97	113.00	1.02	A
23	1.24	0.27	129.00	1.64	A
24	1.52	0.22	133.00	1.73	A
25	0.70	0.40	152.00	1.32	A
26	1.61	0.20	139.00	1.68	A
27	2.50	0.11	132.00	1.97	A

When Table 8 is investigated, only items number 13 and 26 have moderate level (B+/B-) differential item functioning; in other words, it was determined that only two items may have moderate levels of bias. The item characteristics curves (ICC) showing the probability of correct responses from male students, determined to be the reference group, with girls, determined to be the focus group, at all ability levels for the two items with moderate levels of bias are shown in Figure 1.

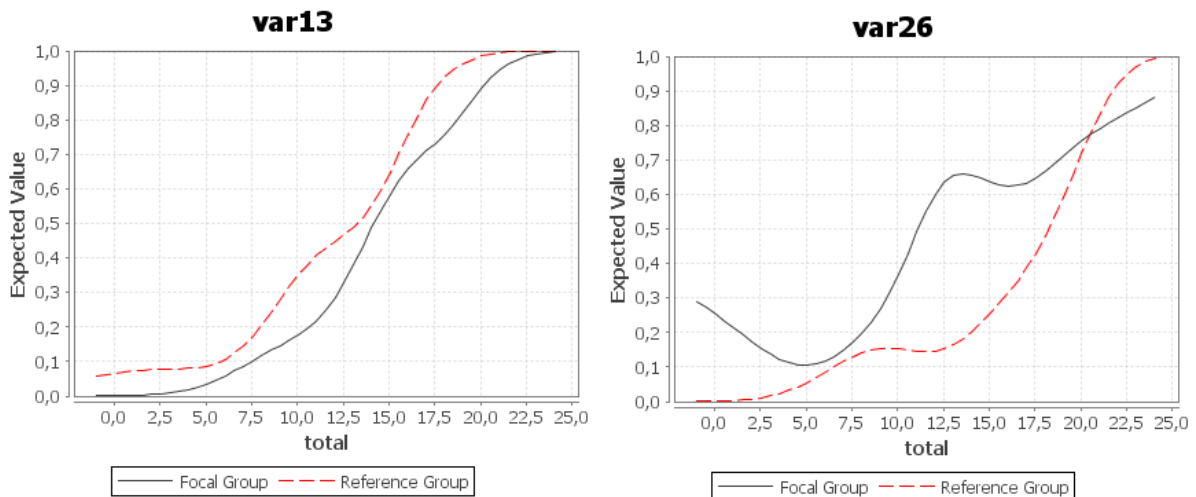


Figure 1. ICC Related to Items 13 and 26 on the test

As seen in Figure 1, the correct answers for students accepted as having the lowest ability level and without any correct answers in total and for students accepted as having the highest ability level with a total of 22 correct answers are shown on the horizontal axis and the probability of correct responses are shown on the horizontal axis. When the item characteristic curves are investigated, it appears both items are not biased. Additionally, as DIF must be at C level at least for one item to be biased, it was concluded that all 22 items on the tool are not biased (Koyuncu, Aksu & Kelecioğlu, 2018). Though item number 10 was determined not to be biased, item characteristic curves were used to determine the probability of correct responses at all ability levels related to the item, and results are shown in Figure 2.

When Figure 2 is investigated, the probability of correct responses to this item from female students with no correct answers on the test was above 60%, while the probability of correct responses from male students appeared to be above 90%. Additionally, these rates were determined as 100% for students with 10 or more correct answers on the test. This result shows the item does not serve its purpose; in other words, it does not discriminate between those who know and don't know the topic. For this reason, the decision was made to remove item number 10 with low item validity from the test. As a result of this process, the test statistics for the final 21-item form of the perimeter measures achievement test are shown in Table 9.

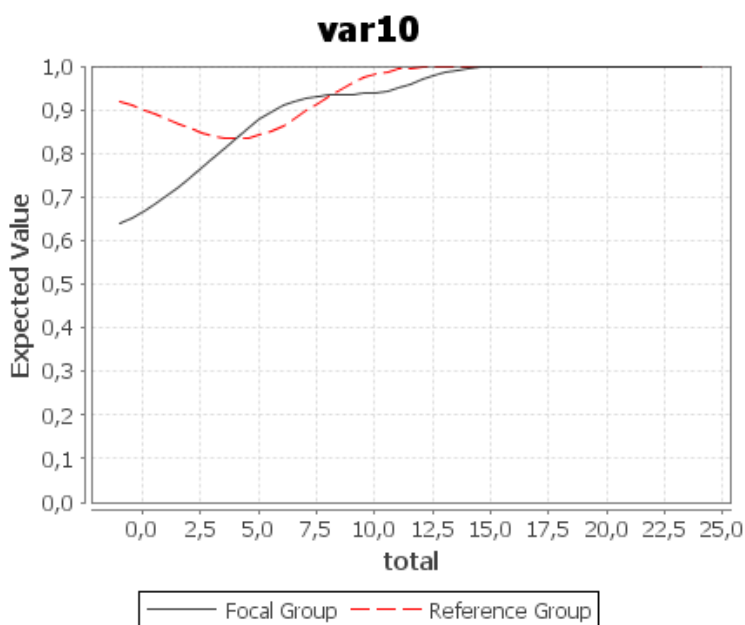


Figure 2. ICC related to Item 10 on the test

Table 9. Test Statistics

Features	Value
Item number	21
Individual number	164
Minimum	1.00
Maximum	21.00
Mean	8.57
Median	8.00
Standard deviation	4.74
Quarter shift	7.00
Skewness	0.52
Kurtosis	-0.79
Mean difficulty	0.41
Mean discrimination	0.40
KR- 20	0.81

When Table 9 is investigated, the lowest points that were obtained from the test were 1.00 and the highest points were 21.00. The mean points were 8.57 and the standard deviation was 4.74, with skewness and kurtosis determined as 0.52 and -0.79, respectively. Additionally, the mean difficulty value for items included on the test was 0.41, and the mean discrimination index value was calculated as 0.40. The reliability coefficient for the achievement test about perimeter measures comprising a total of 21 items was determined as 0.81. Accordingly, when all results are assessed, the developed measurement tool was determined to be a valid and reliable measurement tool.

DISCUSSION, CONCLUSION, AND RECOMMENDATIONS

In this study, an achievement test related to the topic of perimeter measures in primary school 4th-grade mathematics lessons were developed and the validity and reliability studies for the test were completed. The teaching of the measurement topic has an important place in showing the use of mathematics in daily life and in ensuring the development of many mathematical concepts and skills (Tan & Aksu, 2009). For this reason, measurement is an important topic in primary school mathematic programs. Additionally, memorizing the

perimeter calculation formula for geometric shapes like squares and rectangles makes it difficult for students to conceptualize this topic. As a result, it is necessary to configure the perimeter measures topic in a concrete way. This test may measure the efficacy of teaching methods by determining the success levels of students. The mean difficulty of this achievement test was found to be 0.43. According to Karip (2008) and Ayvacı and Durmuş (2016), mean difficulty of an achievement test being about 0.50 is a desired situation. For this reason, this test developed related to the perimeter measures topic appears to have moderate difficulty, close to being difficult. The total discrimination of the test being 0.50 shows that items have good discrimination features.

Analyses of the test found the KR-20 value was 0.81 for reliability analysis. This reliability value is calculated by removing each question item individually. When each item is removed, the final value is calculated according to the variation in the reliability coefficient. According to Özdamar (1999) and Tavşancıl (2006), a test being very reliable is linked to the reliability coefficient value being between 0.60 and 0.80. In this sense, the reliability coefficient obtained for the test (KR-20=0.81) is above this interval and the measurement tool appears to have adequate reliability.

In line with the results obtained in the research, it is possible to make a range of recommendations for studies in the future. Applications may be made by increasing the number of individuals in the sample for validity and reliability studies to achieve more realistic results for achievement tests. Thus, measurement of the audience targeted by the achievement test results may reach more realistic levels. Additionally, more open-ended questions may be asked with the aim of measuring more upper-level cognitive skills of students. Achievement tests may also be developed about new areas due to the different learning and sub-learning fields on the revised Primary School Mathematic Program.

Statement of Researchers

Researchers contribution rate statement: The authors contributed equally to the article.

Conflict statement: The authors declared no potential conflicts of interest regarding the research, authorship, or publication of this article.

Support and thanks: The researcher did not receive support from any institution or organization in the study.

TÜRKÇE SÜRÜM

GİRİŞ

Eğitim öğrencilere bilginin aktarıldığı yalın bir süreç değildir. Çünkü eğitim, öğrencilerin bilgileri hayat ile ilişkilendirerek içselleştirmelerini sağlayacak etkinlikleri de içermelidir. Ancak bu etkinliklerin öğrenci davranışlarına etkisi ile ilgili gerçekçi bilgilerin elde edilmesi de eğitimin diğer önemli parçasıdır. Eğitim sürecinde, öğrencilerin her birinin, kazandırılmak istenilen davranış değişikliğini ne derece gerçekleştirdiği, istenilen beceri ve yeteneğe ne düzeyde ulaşmış olduğu ölçme ile belirlenmeye çalışılır (Atılğan, Kan ve Doğan, 2006). Eğitimde ölçme hem eğitimsel kararlar verilmesinde hem de öğrencilerin öğretim sürecindeki gelişimi ve işleyişi hakkında elde edilecek birtakım çıkarımlar amaçlanarak yapılmaktadır. Yapılacak çıkarımın uygunluğuna göre ölçme amacının belirlenmesi gerekmektedir (Fidan, 2013; Turgut ve Baykul, 2012). Bunun için de eğitim ortamlarında öğrencilerin başarılarının ölçülmesi için bilimsel bir çaba ve metodoloji ile hazırlanan geçerli ve güvenilir testler kullanılması gerekmektedir. Bu aşamada en sık kullanılan testlerin ise başarı testleri olduğu bilinmektedir.

Eğitimde başarı testleri, öğrencilerin hangi seviyede olduğunun belirlenmesinde, konuya ilişkin eksikliklerinin görülmesinde ve buna ilişkin öğrenme-öğretme etkinliklerinin belirlenmesinde öğretmen ve öğrenciye yol gösterici somut araçlardır. Literatürde geliştirilen başarı testlerinin doğru sonuçlara ulaşabilmesi için, her ölçme aracında olması gereken üç önemli niteliğe sahip olması gerekir. Bu nitelikler de geçerlik, güvenilirlik ve kullanılabilirlik olarak sıralanmaktadır (Turgut ve Baykul, 2012). Bu nitelikler alanyazında şu şekilde açıklanmaktadır: Büyükkıdık'a (2020) göre güvenilirlik araştırma sonuçlarının tekrar edilebilirliği ile ilişkili geçerlik daha çok araştırma sonuçlarının doğruluğu ile ilişkilendirilmektedir. Ellez (2014) ise geçerliği bir ölçme aracının amaçlanan özelliğe herhangi bir şeyin karıştırılmadan doğru şekilde ölçülmesi olarak açıklamaktadır. Ölçme araçlarında bulunması gereken bir diğer kavram ise güvenilirliktir. Karasar'a (2012) göre güvenilirlik, herhangi bir şeyin bağımsız yapılan ölçümleri arasındaki kararlılıktır. Bunun yanı sıra Baykul (2000) da güvenilirliği herhangi bir niteliğe ait aynı bireyler üzerinde yapılan ölçmelerin, benzer şartlarda aynı sonuçların elde edilebilir olmasını vurgulamaktadır. Ölçme araçlarındaki diğer nitelik ise kullanılabilirliktir. Özçelik'e (2013) göre kullanılabilirlik, bir ölçme aracının uygulanmasındaki kolaylık olarak tanımlanmaktadır. Bu sebeple eğitimde ölçme araçları geçerli, güvenilir ve kullanılabilir olmalıdır. Bu durum ölçme araçlarının psikometrik kalitesinin yüksekliğine işaret eder.

Pressley vd. (1997) göre ülkemiz ve birçok ülkede çoktan seçmeli sorulardan oluşan testler, öğrencilerin zihinsel seviyelerini ya da zihinsel kapasitelerini ölçmek ve öğrencileri bir üst öğrenim kurumuna yerleştirmek amacı ile kullanılmaktadır (akt. Akbulut ve Çepni, 2013). Çoktan seçmeli madde, öğrenciye bir sorunun hazırlanmış çeldiriciler ile verildiği soru türüdür. Puanlamanın objektif olması, doğru yanıtın maddenin içinde yer alması, uygulanması ve puanlanmasının kolay olması, kalabalık gruplara ve her öğretim kademesine uygulanabilirliği çoktan seçmeli testleri daha fazla ön plana çıkarmaktadır (Atılğan vd., 2006). Büyük gruplarda sıklıkla kullanılan çoktan seçmeli sorular küçük gruplarda zaman zaman yeterli olmayabilir. Çünkü küçük gruplardaki kazanımların daha ayrıntılı belirlenebilmesi ölçme sorularının farklı formlarını gerekli kılmaktadır. Bu sebeple testlerde farklı tür soru türlerinin olması önem arz etmektedir.

Konak (2013) testlerde üst düzey düşünme becerileri için sadece çoktan seçmeli soru türünün kullanılmasının ölçmede yeterli olamayacağı görüşünü belirtmiştir. İn'nami ve Koizumi (2009) de hiçbir soru türünün tam anlamıyla eksiksiz ve her durum için uygun olmadığını belirtmişlerdir (akt. Koyuncu, 2017). Bu bilgiler ışığında testlerde sadece çoktan seçmeli sorular yerine farklı soru maddeleri de kullanılmalıdır. Bu durum testin daha geçerli ve güvenilirlik olmasına katkı sağlamaktadır. Özellikle öğrencilerin seçeneklerden hariç, kavramları nasıl anladığı hakkında daha derin gözlemlenebilirlik fırsatına sahip olmaya olanak sağladığı açık uçlu soruların kullanımı testlerde yer alabilir. Hong (1984), açık uçlu soruların, kişilerin bu soruyu cevaplarırken özgür olduklarını ve eğitimcilerin de bu sorularla öğrencilerin bilgi kesikliği ve hatalarını gerçek anlamda gözleyebildiklerini belirtmiştir (akt. Koyuncu, 2017). Bu neticede açık uçlu sorular sayesinde öğrenciler istenen cevapları kendileri düşünecek, tasarlayacak, düzenleyecek ve yazacaktır (Koyuncu, 2017; Turgut ve Baykul, 2014). Bu sebeple her

sınıf düzeyindeki öğrenciler için hazırlanacak başarı testlerinde uzmanlar ya da öğretmenler ölçme araçlarında gerekli olan geçerlilik, güvenirlik ve kullanılabilirlik özelliklerine dikkat etmelidirler.

Doğru bir başarı testinin geliştirilmesi, sistematik bir süreçtir. Çünkü geliştirilmek istenen testin özelliklerinin önceden belirlenmiş olması söz konusudur. Başka bir ifade ile özellikleri önceden belirlenmiş bir testin hazırlanması, birtakım sistematik işlem basamaklarının takip edilmesini gerekli kılar (Atılğan vd., 2006). Turgut ve Baykul (2012) ile Atılğan vd. (2006), test geliştirmede bazı önemli adımların izlenmesini önermektedir. Buna göre öncelikle testin hangi amaçla kullanılacağına belirlenmesi gerekir. Ardından bu testle hangi davranışların ölçüleceği de belirlenmelidir. Bunu takiben, test maddelerinin yazımı ve gözden geçirilmesi sağlanmalıdır. Yazılan test maddelerinden deneme formları oluşturulmalı ve deneme uygulaması gerçekleştirilmelidir. Deneme uygulaması yapıldıktan sonra cevap kâğıtları puanlanarak madde analizleri yapılmalı ve test için maddelerin seçim işlemi tamamlanmalıdır. Son olarak da nihai testin oluşturulması ve istatistiklerinin kestirilmesi gerçekleştirilmelidir. Başarı testinin özelliklerinin önceden belirlenebilmesi de uygulanması planlanan grubun özelliklerinin iyi bilinmesini gerektirir. Grubun büyüklüğü, yaş aralığı, bilgi seviyesi gibi özellikleri dikkate alınmadan bu sistematik basamakların takibi mümkün değildir. Bu bilgiler ışığında gruptaki öğrencilerin sahip olması beklenen kazanımlar daha gerçekçi düzeyde belirlenebilir.

Başarı testleri ile ölçülmeye çalışılan özellikler, eğitim-öğretim sürecinde kazanılması hedeflenen bilgi, beceri ve tutumlar olmalıdır. Eğitimde hedefler, öğretimi yönlendirmenin yanında yapılacak ölçmelere de kapsanacak konular ve önem verilecek noktalar hakkında rehberlik etme işlevine sahiptir (Demirel, 2010; Fidan, 2013). Öğretim programlarında her sınıf düzeyi için öğrencilerin edinmeleri hedeflenen kazanımlar belirlenmiştir ve bu kazanımlara göre başarı testleri geliştirilebilir. Özellikle uluslararası düzeyde uygulanan başarı testlerinde ülkemizin ortalamasının altında kaldığı görülmektedir. Özellikle son zamanlarda yapılan TIMMS 2015 sınav sonuçları Tablo 1’de gösterilmiştir (Yıldırım, Özgürlük, Parlak, Gönen ve Polat, 2016).

Tablo 1. TIMMS 2015 2015 4.Sınıf Matematik Öğrenme Alanları Değerlendirmesi

TIMMS 2015 Sınavındaki 4.Sınıf Matematik Öğrenme Alanları	Ortalama Başarı Puanı
Sayılar	489
Veri Gösterimi	476
Geometrik şekiller ve ölçüler	475

Uluslararası TIMMS sınavına 2015 yılında 49 ülke katılmış ve Türkiye 4.sınıf düzeyinde 483 puanla 36. olmuştur. TIMMS 2015 verilerine göre öğrenme alanları kapsamında en düşük puanı 475 puanla Geometrik Şekiller ve Ölçüler alanında alınmıştır (Yıldırım vd., 2016). Bu durum göz önüne alındığında ilkökul 4.sınıf düzeyinde “Geometrik Şekiller ve Ölçme” öğrenme alanının kazanımlarının önemli olduğu öne çıkmaktadır.

Geometrik şekiller ve cisimler geometrinin konusunu oluşturmaktadır. İlkokul matematik programında çocukların özellikle şekil ve cisimlerle ilgili özellikler, genellemeler, sınıflandırma ve çizim bilgisi kazanmaları ve bunların uygulanabilir düzeye gelmesi de çok önemlidir (Altun, 2018). Bu durumda çocukların geometrik şekillerle ilgili ölçü kullanmayı öğrenmeleri de gerekmektedir. İlkokul matematik programı incelendiğinde ‘ölçme’ alanına ait kavram ve beceriler, öğrencilerin günlük hayatta sıklıkla karşılaşacağı ya da ihtiyaç duyacağı temel bilgi ve becerileri içermektedir (Tan-Şişman ve Aksu, 2009). Geometride ölçmeden bahsedildiğinde çevre uzunluğu, alan ve hacim gelmektedir (Dağlı ve Peker, 2012). Çevre ölçme konusunda çevre hesaplamaya neden ihtiyaç duyulduğu, ölçme çalışmaları ile çevre hesaplama, çevre formüllerini kullanarak çevre ölçme ve şekillerin, kavramların bağlantılarının uygulamalarının yapılması ile çevre ölçme konusu öğretilmelidir (Altun, 2018). Bu yüzden çevre ölçme konusunun daha iyi yapılandırılabilmesi için öğretmenin sadece formüle dayalı bir konu olarak ele almaması, öğrencilerin içlerinde anlamlandırılması gereken özellikli bir konudur. Yapılan araştırmalarda özellikle çevre, alan, hacim gibi kavramların anlamlarını bilmeden, formüllerinin ezberletildiği, bu şekilde sonuca ulaşılmaya çalışıldığı ve çevre, alan, hacim ile ilgili sorularda öğrencilerin hata ve kavram yanılgılarına sahip oldukları ortaya çıkmıştır (Dağlı ve Peker, 2011; Emekli, 2011; Gough, 2008; Moreira ve Contente, 1997; Tan-Şişman ve Aksu, 2009). Dolayısıyla öğrencilerin zihinlerinde bazı kavramlara ilişkin yeterli düzeyde bilgiler ile canlandırmalar yapmalarının sağlanması gerekmektedir. Bu durumda geliştirilen test,

öğrencilerin başarı düzeyleri tespit etmek ve uygulanan yöntemlerin başarıya etkisinin ne derecede olduğunu belirlemek için gereklidir. Ulusal ve uluslararası alan yazında başarı testi geliştirme konularını ayrı ayrı temele alan çalışmalar olmasına rağmen, (Akgül, 2011; Fidan, 2013; Kurt, 2015; Tutak ve Birgin, 2008) sadece çevre ölçme konusunu inceleyen herhangi bir araştırmaya rastlanmamıştır. Sadece Gürkan (2019) çalışmasında uzunluk ve çevre ölçme konuları için 18 maddeden oluşan ve KR-20 test ölçüm güvenirlik katsayısı 0.87 olan bir başarı testi geliştirmiştir. Bu testte yer alan 6 soru çevre ölçme ile ilgilidir. Fakat araştırmamızda geliştirilen test 4.sınıf çevre ölçme konusundaki tüm kazanımları kapsamakta ve daha fazla soru içererek sadece çevre ölçme konusu ile ilgilidir.

İlkokul Matematik dersi içinde yenilenen 2018 matematik öğretim programında öğrencilerin ölçme alt öğrenme alanı, çevre ölçme ile ilgili kazanımları 3. ve 4. sınıftadır. Dördüncü sınıfta bu konu alanına ilişkin 3 kazanım bulunmaktadır (MEB, 2018). Çevre ölçme konusu 4.sınıfta "Ölçme" konusunda sonra yer almakta; özellikle kare ve dikdörtgenin çevre uzunluklarını hesaplama, bu şekiller arasında kenar uzunlukları ve çevreleri arasındaki ilişki ve çevre ölçme ile ilgili problem çözme becerilerinden oluşmaktadır. Bu sebeple literatür taraması yapıldığında ülkemizde dördüncü sınıf matematik dersi ile ilgili test geliştirme çalışmalarının özellikle "Ölçme" alanı ile ilgili çalışmaların çok fazla olmadığı görülmektedir. Yapılan bu çalışma ile ilgili alandaki boşluğun giderilmesine katkı sunacağı umulmaktadır. Bu çalışmada, 2018 matematik programı kapsamında 4. sınıflar için geometrik cisimlerin çevresini ölçme konusunda başarı testinin geliştirilme süreci detaylı olarak anlatılmıştır. Belirlenen bu amaçla birlikte araştırma soruları şu şekildedir:

- 4. sınıf öğrencilerinin "Çevre Ölçme" konusundaki başarılarını belirlemek için hazırlanan test geçerli midir?
- 4. sınıf öğrencilerinin "Çevre Ölçme" konusundaki başarılarını belirlemek için hazırlanan test güvenilir midir?

YÖNTEM

Çalışmanın Deseni

Araştırma, öğrencilerin 4.sınıf Çevre Ölçme ile ilgili başarılarını belirlemek için hazırlanan bir başarı testi geliştirme çalışması olduğu için betimsel bir araştırmadır. Bu yöntemde, araştırmaya dahil olan konular, kişiler, nesnelere ya da olaylar olduğu gibi, geçmişteki ya da hali hazırdaki kendi koşulları içinde var olduğu gibi tanımlanır ve tasvir edilir (Karasar, 2012). Araştırmada 4.sınıf öğrencilerinin başarılarını ölçebilmek için öğrencilerin var olan durumları ele alındığı için betimsel olarak çalışma ortaya konmuştur. Araştırmada 4. sınıf öğrencilerinin çevre ölçme konusundaki başarılarını ölçmek amacıyla geçerli ve güvenilir bir ölçme aracı geliştirilmesi amaçlandığından test geliştirme süreci adımları takip edilmiştir.

Katılımcılar

Bu çalışmanın katılımcılarını 2017-2018 eğitim-öğretim yılında ilkökul 4.sınıf "çevre ölçme" konusunu tamamlamış Balıkesir ili Altıeylül ilçesine bağlı 5. sınıfta okuyan 164 öğrenci oluşturmaktadır. Başarı testi 4.sınıf için belirlenmiş fakat uygulamanın yapıldığı dönemde 4.sınıflar konuyu görmediği için ve soruları boş bırakma ihtimalleri olduğu için konuyu hali hazırda görmüş olan 5.sınıf öğrencileri oluşturmaktadır. Bu öğrenciler birbirinden bağımsız olduğu için, öğrencilerden birinin örneklem için seçilmesi diğerinin seçilmesine bağlı olmamaktadır. Bu yüzden öğrenciler basit seçkisiz örnekleme yöntemi ile belirlenmiştir. Bu öğrencilerden 86'sı kız öğrenci, 78'i erkek öğrencidir.

Veri Toplama Aracının Geliştirilme Süreci

İlk adımda testin amacı, Matematik programına bağlı olarak (2018) ilkökul 4. sınıf matematik dersi "Ölçme" öğrenme alanındaki "Çevre Ölçme" konusuna yönelik bir başarı testi geliştirmektir. Testin geliştirilmesi aşamasında ilkökul 4. Sınıf Matematik Programı'nda (2018) "Çevre Ölçme" konusu ile ilgili kazanımlar listelenmiştir. Bu kazanımlar doğrultusunda Bloom taksonomisine göre Tablo 2'de bir belirtke tablosu oluşturulmuştur. Daha sonra araştırmacılar tarafından belirtke tablosuna göre deneme maddeleri yazılmıştır. Deneme maddeleri uzmanlara (Türkçe, Matematik ve Ölçme Değerlendirme uzmanı) gönderilerek geçerlilik konusunda dönütler alınmıştır ve uzmanlardan gelen dönütlere göre test formu tekrar düzenlenmiştir. Daha sonraki aşamada, test formu hazırlanmış; test bu konuyu görmüş olan 5. Sınıftan 9 öğrenciye birebir uygulanmıştır. Bu uygulama sonrasında anlaşılammış ve hatalı, eksik olan sorular tekrar gözden geçirilmiştir.

Bu aşamadan sonra test 164 öğrenciye uygulanmıştır. En son olarak da madde analizi yapılarak teste son hali verilmiştir.

Araştırmada 2018 yılında yenilenen ilkököl 4. sınıf matematik programına göre "Çevre Ölçme" konusuna yönelik kazanımlar belirlenmiştir. Bu kazanımların programdaki içeriğine göre ve Bloom taksonomisindeki düzeylerine göre, araştırmacılar tarafından aşağıdaki gibi bir belirtke tablosu oluşturulmuştur. Çevre ölçme konusundaki kazanımların Bloom taksonomisindeki bilişsel alanın hangi basamağında yer aldığı Tablo 2'de gösterilmiştir.

Tablo 2. Başarı Testi Belirtke Tablosu

Öğrenme Alanı (Çevre)	Bilişsel Alan (Bloom)						Toplam	Yüzde
	Bilgi	Kavrama	Uygulama	Analiz	Sentez	Değerlendirme		
1.Kare ve dikdörtgenin çevre uzunlukları ile kenar uzunlukları arasındaki ilişkiyi açıklar.	3 (4., 8. ve 9.soru)	1 (2. soru)	3 (11., 12. ve 13.soru)	2 (5. ve 16.soru)	2 (24. ve 27.soru)	3 (19., 21. ve 23.soru)	14	%51,8
2.Aynı çevre uzunluğuna sahip farklı geometrik şekiller oluşturur.			1 (7. soru)	2 (25. ve 27. soru)			3	%11,1
3.Şekillerin çevre uzunluklarını hesaplamayla ilgili problemleri çözer.	1 (3.soru)	1 (1.soru)	4 (6., 10., 14. ve 15.soru)	4 (17., 18., 20. ve 26. soru)			10	%37,1
Toplam	4	2	8	8	2	3	27	%100
Yüzde	%14,8	%7,4	%29,6	%29,6	%7,4	%11,2	%100	

Kazanımlara uygun olarak toplam 27 soruluk madde havuzu oluşturulmuştur. Bu sorulardan 20 madde araştırmacılar tarafından geliştirilmiş, 7 madde ise Tablo 3'te verilen çalışmalardaki başarı testlerindeki maddelerden türetilmiştir.

Tablo 3. Test Soruları, Kazanımları ve Taranan Kaynaklar

Kazanımlar	Soru No	Faydalanılan kaynaklar
Kare ve dikdörtgenin çevre uzunlukları ile kenar uzunlukları arasındaki ilişkiyi açıklar.	4, 5, 9, 12, 13, 14, 18, 19, 20, 21, 22, 24, 25, 26	İnan (2016)
Aynı çevre uzunluğuna sahip farklı geometrik şekiller oluşturur.	3, 6, 27	Küçükaydın, (2017)
Şekillerin çevre uzunluklarını hesaplamayla ilgili problemleri çözer.	1, 2, 7, 8, 10, 11, 15, 16, 17, 23	Karadağ, Balcı, Abdik ve Demiralp, (2017)

Araştırmacılar ve yukarıda belirtilen çalışmalardan yararlanılarak düzenlenen test maddeleri, denemelik madde yazımında öğrencilerin karar verme, üst düzey becerileri gelişimlerinin görülebilmesi, matematiksel ifade güçlerinin ortaya konulabilmesi amacıyla çoktan seçmeli sorulara ek olarak, açık uçlu ve alıştırma tarzı sorulara da yer verilmiştir. Turgut ve Baykul'a (2014) göre açık uçlu sorularla öğrenciden istenen cevapları kendisinin düşünmesi, tasarlaması, düzenlemesi ve yazması istenmektedir. Bu testte de bu yapı göz önünde bulundurulmuştur.

Verilerin Analizi

Denemelik test maddeleri yazıldıktan sonra birçok yönden gözden geçirilerek maddelerin eksik yönleri ve eksiklikleri düzeltilmelidir (Atılğan vd., 2006). Araştırmada kullanılacak başarı testinin denemelik madde kontrolü Matematik alan uzmanı, Türkçe alan uzmanı ve Ölçme Değerlendirme uzmanlarına gönderilmiş ve onların görüşleri doğrultusunda geçerlilik, bilimsel doğruluk, dilbilgisi ve teknik açıdan uygunluk kriterleri tekrar gözden geçirilmiştir. Gerekli düzenlemeler ve düzeltmeler yapıldıktan sonra, hazırlanan maddelerin zorluk (güçlük) derecesine ve kazanım hiyerarşisine göre test formunda %20 zor, %60 orta ve %20 zor olacak şekilde dağılım yapılmıştır. Teste öğrencilerin zihinsel seviyelerine uygun yönergeler eklenmiştir. Ardından araştırmacılar tarafından ilkokul 4.sınıf öğrencilerinin gelişim seviyesine göre testin punto büyüklüğü 14 punto olarak belirlenmiştir. Testin uygulanması aşamasında öncelikli olarak testin tamamlanması için öğrencilere verilecek süre 40 dakika olarak belirlenmiştir. Daha sonra bu test ön deneme uygulaması olarak 5. sınıftan dokuz öğrenciye birebir uygulanarak anlaşılmayan ve hatalı olan sorular tekrar değerlendirilmiştir. Öğrencilerin soruları çözerken maddelere yönelik düşünceleri sesli düşünme tekniği ile alınmış, jest ve mimiklerine bakılarak maddelere verdikleri tepkiler izlenmiş ve bu bağlamda çözümlenmeler yapılarak maddelerin son hali oluşturulmuştur. Birebir uygulamadan sonra Balıkesir ili Altıeylül ilçesindeki 5. sınıfta okuyan 164 öğrenci belirlenmiştir. Bu öğrencilere 24 çoktan seçmeli test sorusu ve 3 açık uçlu olmak üzere toplam 27 sorudan oluşan başarı testi, verilen süre çerçevesinde uygulanmıştır.

Başarı testinin geçerlik ve güvenilirlik çalışması kapsamında Klasik Test Kuramına dayalı veri analizi için Test Analiz Programı (Test Analysis Program-TAP) kullanılmıştır. Öğrencilerin her soruya hangi cevabı verdiği TAP test analiz programına girilmiş ve böylece her maddenin seçenek analizi de yapılmıştır. Toplam 24 çoktan seçmeli sorunun seçenekleri dört seçenek olarak incelenmiş, açık uçlu soruların seçenekleri verilen cevapların yığılmalarına göre TAP programında analiz edilmek doğru-yanlış (1-0) olacak şekilde iki seçenekli hale getirilmiştir. Bu bağlamda, açık uçlu sorularda öğrencilerin hatalı cevaplarındaki sonuç değerleri ele alınarak bu değerler seçenek olarak sisteme girilmiş ve analize tabii tutulmuştur.

Örneğin; "Çevre uzunluğu 20 cm olan bir dikdörtgenin çevre uzunluğuna eşit bir şekli aşağıdaki noktalı kâğıda çiziniz" Sorusu için çevresi 20 cm olarak çizilen kare ve dikdörtgenler için A doğru seçenek olarak girilmiş, çevre uzunluğu 14 cm olarak çizilen şekiller için B seçeneği, çevre uzunluğu 22 olarak çizilen şekiller için C seçeneği ve çevre uzunluğu 18 olarak çizilen şekiller için D seçeneği programa girilmiştir. Açık uçlu bu tür soruların cevaplardaki yığılmalarına göre oluşturulan seçenekler bu test için analizde ayrıca belirtilmiştir. Farklı uygulamalarda farklı test seçeneklerinin oluşabileceği durumlar için araştırmacıların dikkatli olmaları gerekmektedir. Ayrıca öğrencilerin boş bıraktığı sorular TAP programında ". (nokta)" girilerek doldurulmuştur. Testin sonuçlarının değerlendirilmesinde yanlış cevaplar doğru sayısını etkilememektedir. Buna göre her doğru cevaba "1" puan ve her yanlış cevaba "0" puan verilmiştir. Oluşturulan testten alınabilecek en yüksek puan "27" ve en düşük puan "0" olarak belirlenmiştir. Çalışma kapsamında kullanılan ölçme aracının güvenilirlik ve geçerlik analizleri Madde Tepki Kuramına dayalı olarak jMetrik 4.1.1 programı ile gerçekleştirilmiştir. Programdan elde edilen sonuçların değerlendirilmesinde Linacre (2002) tarafından yapılan ağırlandırılmamış ve standardize edilmiş uyum istatistikleri için kullanılacak olan ölçütler için kabul edilebilir aralıklar Tablo 4'te gösterilmiştir.

Tablo 4. Uyum İstatistiklerinin Kabul Edilebilir Aralıkları

Ölçüt	Aralık	Yorum
WMS ve UMS	> .20	Ölçmeye konu özelliği saptırmaktadır ve ölçmenin kalitesini düşürür.
	1.5 – 2.0	Ölçme işlemi için verimsiz ama çok da kötümser değil.
	0.5 – 1.5	Ölçme için gayet uygun.
	< .50	Ölçme işlemi için daha az verimli ancak çok kötü değil. Yüksek düzeyde güvenilirlik ve ayırt edicilik değeri üretebilir.
Std.WMS ve Std.UMS	≥ 3	Veriler model ile uyum göstermiyor. Daha büyük örneklem ile çalışılabilir.
	2.0 – 2.9	Veriler belirgin ölçüde tahmin edilemez.
	-1.9 – 1.9	Veriler makul düzeyde tahmin edilebilirliğe sahiptir.
	≤ -2.0	Veriler oldukça tahmin edilebilir. Belki diğer boyutlar yanıt örüntülerini sınırlandırabilir.

Tablo 4 incelendiğinde WMS ve UMS değerleri için en ideal aralığın 0,50 ile 1,50 arasında olduğu belirlenmiştir. Benzer şekilde Standardize edilmiş WMS ve UMS değerleri için en ideal aralığın -1,90 ile 1,90 arasında olduğu söylenebilir. *WMS* ve *Std.WMS* isimli sütunlar sırasıyla ağırlıklandırılmış ve standardize edilmiş kareli ortalama uyumsuzluk istatistiklerini göstermektedir. Bunları takiben *UMS* ve *Std.UMS* isimli sütunlar ağırlıklandırılmamış ve standardize edilmiş kareli ortalama ölçüt dışı olma istatistiklerini göstermektedir (Güzeller, Eser ve Aksu, 2018).

BULGULAR

Çalışmanın bu aşamasında TAP programında yapılan madde güçlüğü ve madde ayırt edicilik indeksi ve güvenilirlik sonuçları açıklanmıştır. Başarı testinin TAP programındaki madde analizleri Tablo 5'te gösterilmiştir.

Tablo 5. Madde Güçlük İndeksi ve Madde Ayırt Edicilik İndeksi

Sorular	Madde Güçlüğü	Madde Ayırt edicilik İndeksi	Maddelerin Zorluğu	Cevapların Seçeneklere Göre Dağılım (Doğru Cevaplayanlar)				KR-20
				A	B	C	D	
1	0.21	0.24	Çok zor	28*	63	30	42	0.81
2	0.65	0.39	Kolay	10	18	25	106*	
3	0.39	0.19	Orta	39	64*	8	24	
4	0.96	0.07	Çok kolay	1	3	158*	1	
5	0.11	-0.06	Çok zor	97	15	16	18*	
6	0.29	0.38	Zor	48*	23	5	51	
7	0.46	0.80	Orta	11	75*	38	34	
8	0.61	0.59	Orta	11	17	100*	16	
9	0.34	0.55	Zor	11	55*	56	8	
10	0.89	0.23	Çok kolay	4	10	146*	1	
11	0.77	0.45	Kolay	6	126*	6	17	
12	0.50	0.77	Orta	17	82*	11	40	
13	0.40	0.82	Orta	7	65*	13	71	
14	0.52	0.70	Orta	17	8	85*	37	
15	0.60	0.49	Orta	13	98*	11	23	
16	0.50	0.62	Orta	9	82*	8	43	
17	0.22	0.36	Zor	11	80	10	36*	
18	0.23	0.23	Zor	9	37*	28	23	
19	0.36	0.61	Orta	34	20	17	59*	
20	0.33	0.14	Zor	22	39	5	54*	
21	0.41	0.57	Orta	27	17	13	68*	
22	0.23	0.09	Zor	20	28	25	38*	
23	0.21	0.49	Zor	13	35*	23	72	
24	0.33	0.48	Zor	25	58	6	54*	
25	0.30	0.36	Zor	18	48	13	49*	
26	0.34	0.56	Zor	56*	11	20	18	
27	0.37	0.66	Orta	61*	19	32	15	

*: Doğru yanıtı göstermektedir.

Tablo 5'te yer alan madde analizleri değerlendirildiğinde; Büyüköztürk vd. (2014), maddenin çok iyi madde olmasının madde ayırt edicilik indeksi değerinin 0.40 olması; maddenin iyi madde olmasının ve herhangi bir düzeltme yapmadan ölçekte tutulabilmesi için 0.30 ile 0,39 arasında olması gerektiğini belirtmektedir. Bunun yanı sıra maddenin düzeltilerek geliştirilmesi için ayırt edicilik indeksi değeri 0.20 ile 0.29 arasında olması gerektiği belirtilmiştir. Aynı zamanda maddenin testten çıkarılması veya bütünüyle gözden geçirilmesi için madde ayırt edicilik indeksi değeri 0.20'den küçük olmalıdır.

Buna göre TAP test analiz programında yapılan madde analizlerinde 22-5-3-20 ve 4. sorunun ayırt edicilik indeks değeri 0,20'den düşük olması nedeniyle testten çıkarılmış ve madde ayırt edicilik indeks değeri 0,20-0.29 arası olan 1. ve 22. maddeler düzeltilerek testte tutulmuştur. Bu güvenilirlik değerinin bulunmasında her bir soru maddesi sırayla; tek tek çıkarılarak güvenilirlik katsayısının değişimine göre son değer hesaplanmıştır. Testin yapılan güvenilirlik analizi sonucunda KR-20 değeri 0,84 olarak bulunmuştur. Güvenirlik katsayısının 0.70 ve üzerinde bir değer olması güvenilirlik hesaplaması için yeterli olarak kabul görmektedir (Büyüköztürk vd., 2014). Bu anlamda testten elde edilen güvenilirlik katsayısının bu aralıkta olması ölçme aracının güvenilirliği için yeterli görülmüştür.

Başarı testi ile ilgili gerekli düzenlemeler yapıldıktan sonra başarı testinde öğrencilerin toplam puanları, yaptıkları doğru sayısı kadar olup; başarı testinin sonuçları her yanlış cevap için "0" puan ve her doğru cevap için "1" puan verilerek değerlendirilmektedir. Başarı testinin son hali üç kazanıma göre ve her kazanımda sorular kolaydan zora doğru sıralanarak Tablo 6'da madde analiz değerleri ile birlikte gösterilmiştir.

Tablo 6. Çevre Ölçme Başarı Testi'nin Madde Güçlük İndeksi ve Madde Ayırtedicilik İndeksi

Sorular	Madde güçlüğü	Madde ayırt edicilik indeksi	Cevapların seçeneklere göre dağılım (doğru cevaplayanlar)				KR-20
			A	B	C	D	
1	0,89	0,23	4	10	146*	1	0,837
2	0,77	0,46	6	126*	6	17	
3	0,50	0,61	9	82*	8	43	
4	0,50	0,82	17	82*	11	40	
5	0,40	0,83	7	65*	13	71	
6	0,52	0,72	17	8	85*	37	
7	0,36	0,52	34	20	17	59*	
8	0,33	0,46	25	58	6	54*	
9	0,34	0,52	11	55*	56	8	
10	0,22	0,34	11	80	10	36*	
11	0,23	0,23	9	37*	28	23	
12	0,50	0,82	17	82*	11	40	
13	0,60	0,53	13	98*	11	23	
14	0,34	0,56	56*	11	20	18	
15	0,46	0,81	11	75*	38	34	
16	0,61	0,64	11	17	100*	16	
17	0,41	0,51	27	17	13	68*	
18	0,29	0,32	48*	23	5	51	
19	0,21	0,52	13	35*	23	72	
20	0,30	0,37	18	48	13	49*	
21	0,34	0,56	56*	11	20	18	
22	0,21	0,24	28*	63	30	42	

Tablo 6'ya göre çevre ölçme ile ilgili ilk kazanıma ait 10 soru, ikinci kazanıma ait 3 soru ve üçüncü kazanıma ait 9 soru testin son halinde yer almaktadır. Testin son halinde öğrenciler için test yönergesi açık ve net bir biçimde ifade edilmiş; sınav uygulama süresi de 35 dakika olarak belirlenmiştir. Çoktan seçmeli soruların cevaplarında belirli bir şifre niteliği (A-B-C-A-B-C gibi) taşımaması adına seçenekler tekrar gözden geçirilerek test formu oluşturulmuştur. Bu işlemin ardından Madde Tepki Kuramına (MTK) dayalı Rasch modeli ile analizler gerçekleştirilmiştir. Testte yer alan 22 madde için maddelere ilişkin güçlük ve ayırt edicilik değerleri ile uyum indeksleri değerleri Tablo 7'de gösterilmiştir.

Tablo 7. Ölçekteki Maddelere İlişkin Rasch Analizi Sonuçları

Madde No	Güçlük	Std. Hata	WMS	Std. WMS	UMS	Std. UMS
1	1,74	0,23	1,20	1,43	1,52	1,64
2	-1,15	0,19	1,16	1,93	1,77	3,84
6	0,83	0,20	1,17	1,68	1,27	1,49
7	-0,13	0,18	0,77	-3,13	0,68	-2,98
8	-0,95	0,18	0,92	-1,12	0,80	-1,40
9	0,60	0,19	0,99	-0,02	1,33	1,99
10	-3,90	0,37	0,96	-0,03	0,57	-0,47
11	-1,90	0,20	0,90	-0,99	0,82	-0,69
12	-0,36	0,18	0,76	-3,30	0,67	-3,05
13	0,21	0,19	0,77	-2,85	0,68	-2,77
14	-0,46	0,18	0,89	-1,52	0,87	-1,02
15	-0,88	0,18	1,04	0,52	1,04	0,31
16	-0,36	0,18	0,92	-1,07	0,94	-0,48
17	1,29	0,21	1,07	0,63	1,53	2,11
18	1,29	0,21	1,18	1,47	1,28	1,22
19	0,42	0,19	1,07	0,74	1,05	0,36
21	0,10	0,18	1,07	0,87	1,07	0,58
23	1,34	0,22	0,90	-0,80	0,78	-0,94
24	0,64	0,19	1,13	1,34	1,24	1,49
25	0,75	0,20	1,23	2,17	1,38	2,11
26	0,56	0,19	0,99	-0,11	0,96	-0,18
27	0,31	0,19	0,90	-1,17	0,79	-1,63

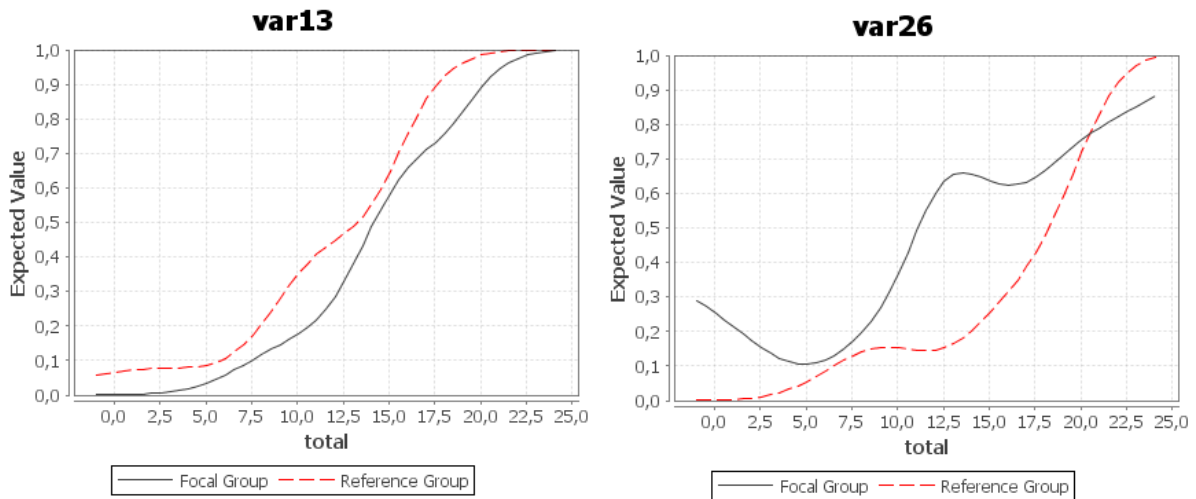
Tablo 7’de gösterilen Unweighted Mean Square (UMS) ve Weighted Mean Square (WMS) fit statistics maddelere ilişkin uyum istatistikleridir. Bu değerlerden WMS infit ölçütü, UMS ise outfit ölçütü olarak kabul edilmektedir. Tablo 5 incelendiğinde 22 maddeden oluşan ölçeğin WMS ve UMS değerlerine göre yapılacak ölçme işlemi için gayet uygun oldukları görülmektedir. Her ne kadar 2 numaralı madden UMS değerleri kritik olarak kabul edilen 1,50’nin üzerinde olsa da 1,50-2,00 aralığının kabul edilebilir olarak belirtilmesi sebebiyle maddelerin uyum istatistiklerinin oldukça iyi olduğu sonucuna ulaşılmıştır. Standardize edilmiş WMS ve UMS değerlerine göre maddelerinin -2,00 ile 2,00 aralığında olması sebebiyle model veri uyumunu sağlandığı belirlenmiştir. Ancak madde güçlük değerleri incelendiğinde 10 numaralı maddenin kabul edilebilir sınırlar olarak belirtilen -3,00 ve +3,00 aralığının dışında bir güçlük değerine sahip olduğu görülmektedir. Bu maddenin bir sonraki aşamada elde edilen analiz sonuçlarına göre testte yer almayacağına karar verilecektir. Ölçekte yer alan maddelerin cinsiyet değişkenine göre yanlı olup olmadığını belirlemek amacıyla Değişen Madde Fonksiyonu (DMF) yardımıyla analizler gerçekleştirilmiştir. Mantel-Haenszel yöntemiyle ortak olabilirlik oranları (*common odds ratio*) değerleri esas alınarak elde edilen sonuçlar Tablo 8’de gösterilmiştir.

Tablo 8. DMF Analizi Sonuçları

Madde No	Kikare	p-değeri	N	Etki büyüklüğü	Sınıf
1	0,79	0,38	135,00	0,67	A
2	1,09	0,30	141,00	0,67	A
6	0,83	0,36	137,00	0,71	A
7	0,37	0,54	58,00	0,51	A
8	0,16	0,69	119,00	1,24	A
9	0,82	0,37	154,00	1,46	A
10	0,56	0,45	96,00	0,70	A
11	1,55	0,21	131,00	0,63	A
12	0,00	0,98	151,00	0,99	A
13	0,15	0,70	69,00	1,32	A
14	0,09	0,76	119,00	0,88	A

15	0,51	0,47	116,00	0,74	A
16	4,77	0,03	113,00	2,67	B-
17	0,15	0,70	131,00	1,15	A
18	0,22	0,64	137,00	1,20	A
19	0,18	0,67	139,00	1,19	A
21	0,00	0,97	113,00	1,02	A
23	1,24	0,27	129,00	1,64	A
24	1,52	0,22	133,00	1,73	A
25	0,70	0,40	152,00	1,32	A
26	1,61	0,20	139,00	1,68	A
27	2,50	0,11	132,00	1,97	A

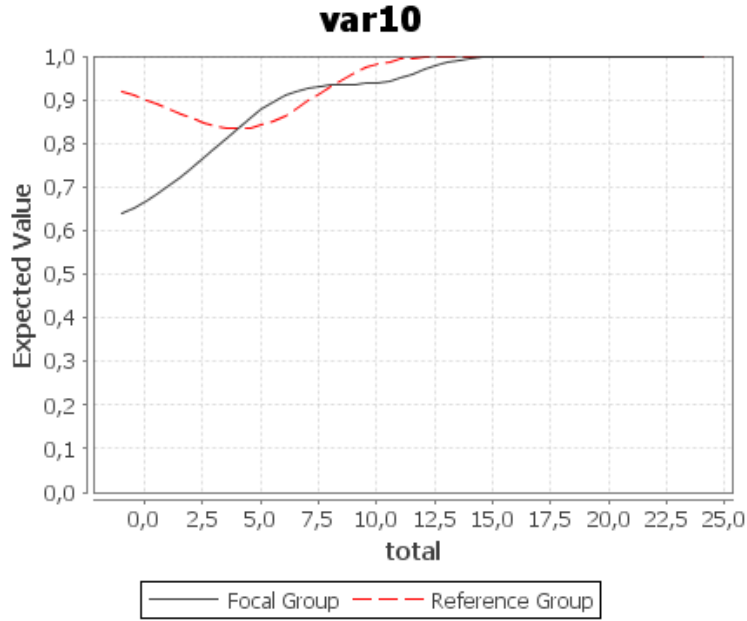
Tablo 8 incelendiğinde sadece 13 ve 26 numaralı maddelerin orta düzeyde (B+/B-) değişen madde fonksiyonuna sahip olduğu başka bir ifadeyle sadece iki maddede orta düzeyde yanlı olabileceği belirlenmiştir. Orta düzeyde yanlı olduğu belirlenen iki maddenin tüm yetenek düzeylerinde odak grup olarak belirlenen kızlar ile referans grup olarak belirlenen erkek öğrenciler tarafından doğru yanıtlanma olasılıklarını gösteren Madde Karakteristik Eğrileri (MKE) Şekil 1’de gösterilmiştir.



Şekil 1. Testte Yer Alan 13 ve 26 numaralı Maddelere İlişkin MKE

Şekil 1’de görüldüğü üzere en düşük yetenek düzeyi olarak kabul edilen ve toplamda hiç doğrusu olmayan öğrenciler ile en yüksek yetenek düzeyi olarak kabul edilen ve toplamda 22 doğrusu olan öğrencilere ilişkin doğru sayıların yatay ekseninde ve doğru yanıtlanma olasılıklarının dikey ekseninde gösterildiği madde karakteristik eğrileri incelendiğinde her iki maddenin de yanlı olmadığı görülmektedir. Bunun yanında bir maddenin yanlı olarak belirtilebilmesi için en az C düzeyinde DMF göstermesi gerektiğinden ölçekte yer alan 22 maddenin tamamı için yanlı olmadığı sonucuna ulaşılmıştır (Koyuncu, Aksu ve Kelecioğlu, 2018). Her ne kadar 10 numaralı maddenin yanlı olmadığı belirlense de ilgili maddenin tüm yetenek düzeylerinde doğru cevaplanma olasılığını belirlemek amacıyla madde karakteristik eğrisi elde edilmiş ve sonuçlar Şekil 2’de gösterilmiştir.

Şekil 2 incelendiğinde ilgili maddenin testte hiç doğrusu olmayan kız öğrenciler tarafından doğru yanıtlanma olasılığı %60’ın üzerinde iken erkek öğrenciler tarafından doğru yanıtlanma olasılığı %90’ın üzerinde olduğu görülmektedir. Bunun yanında testte toplam 10 ve daha fazla doğrusu olan öğrenciler için bu oran %100 olarak belirlenmiştir. Elde edilen bu sonuç maddenin amaca hizmet etmediğini başka bir ifadeyle bilenle bilmeyeni ayırt etmediğini göstermektedir. Bu nedenle madde geçerliği düşük olan 10 numaralı maddenin de testten çıkarılmasına karar verilmiştir. Bu işlemler sonucunda Çevre ölçme başarı testinin 21 maddelik son haline ilişkin test istatistikleri Tablo 9’da gösterilmiştir.



Şekil 2. Testte Yer Alan 10 numaralı Maddelere İlişkin MKE

Tablo 9. Test İstatistikleri

Özellikler	Değer
Madde Sayısı	21
Birey sayısı	164
Minimum	1,00
Maksimum	21,00
Ortalama	8,57
Medyan	8,00
Standart sapma	4,74
Çeyrek kayma	7,00
Çarpıklık	0,52
Basıklık	-0,79
Ortalama güçlük	0,41
Ortalama ayırt edicilik	0,40
KR- 20	0,81

Tablo 9 incelendiğinde testten elde edilecek en düşük puanın 1,00 ve en yüksek puanın 21,00 olduğu belirlenmiştir. Testten elde edilen puanların ortalaması 8,57 ve standart sapması 4,74 iken çarpıklık ve basıklık katsayıları sırasıyla 0,52 ve -0,79 olarak belirlenmiştir. Bunların yanında testte yer alan maddelere ilişkin ortalama güçlük değeri 0,41 ve ortalama ayırt edicilik değeri 0,40 olarak hesaplanmıştır. Toplam 21 maddeden oluşan çevre ölçme konusu başarı testinin güvenilirlik katsayısı 0,81 olduğu belirlenmiştir. Buna göre sonuçlar bir bütün olarak değerlendirildiğinde geliştirilen ölçme aracının geçerli ve güvenilir bir ölçme aracı olduğu belirlenmiştir.

SONUÇ, TARTIŞMA VE ÖNERİLER

Bu çalışmada ilkökul 4. sınıf matematik dersi çevre ölçme konusu ile ilgili başarı testi geliştirilmiş ve testin geçerlilik ve güvenilirlik çalışmaları açıklanmıştır. Ölçme konusunun öğretimi öğrencilere hem matematiğin günlük hayatta kullanımını göstermede, hem de birçok matematiksel kavram ve becerinin geliştirilmesini sağlamada önemli bir yer tutmaktadır (Tan vd., 2009). Bu nedenledir ki ölçme, ilkökul matematik programlarının önemli bir konusudur. Ayrıca kare ve dikdörtgen gibi geometrik şekillerin çevre hesaplama formüllerinin ezberlenmesi bu konuyu öğrencilerin yapılandırılmalarını zorlaştırmaktadır. Bu nedenle çevre ölçme konusunun somut bir şekilde yapılandırılması gerekmektedir. Geliştirilen bu test ile öğrencilerin başarı düzeyleri belirlenerek öğretim yöntemlerinin etkililiği de ölçülebilecektir. Geliştirilen başarı testinin ortalama güçlüğü 0,43 olarak bulunmuştur. Karip (2008) ve Ayvaci ve Durmuş'a (2016) göre başarı testinin ortalama güçlülüğünün 0,50

civarında olması istenen bir durum olarak görülmektedir. Bu sebeple çevre ölçme konusu ile ilgili geliştirilen bu testin zora yakın olmakla birlikte orta güçlükte olduğunu göstermektedir. Testin toplam ayırt ediciliğinin de 0,50 olması maddeler iyi ayırt edici özellikte olduğunu göstermektedir.

Test için yapılan analizlerde güvenilirlik analizi KR-20 değeri 0,81 olarak bulunmuştur. Bu güvenilirlik değerinin bulunmasında soru maddelerinin teker teker çıkarılarak hesaplanması söz konusudur. Her madde çıkarıldığında güvenilirlik katsayısının değişimine göre son değer hesaplanmıştır. Özdamar (1999) ve Tavşancıl'a göre (2006) bir testin oldukça güvenilir olması güvenilirlik katsayısı değerinin 0.60 ile 0.80 arasında olmasına bağlıdır. Bu anlamda testten elde edilen güvenilirlik katsayısının (KR-20=0,81) bu aralıkta olması ölçme aracının güvenilirliği için yeterli görülmüştür.

Araştırmada elde edilen sonuçlar doğrultusunda geleceğe yönelik çalışmalara ilişkin birtakım öneriler getirmek mümkündür. Başarı testlerinin daha gerçek sonuçlara ulaşması adına yapılan geçerlik ve güvenilirlik çalışmaları için belirlenen örneklemdaki birey sayısı daha da arttırılarak uygulama yapılabilir. Böylece oluşturulan başarı testi sonuçları hedeflediği kitlenin ölçülmesini çok daha gerçek seviyede sağlayabilir. Buna ek olarak öğrencilerin daha üst düzey bilişsel becerilerini ölçebilmek amacıyla daha fazla açık uçlu sorular sorulabilir. Ayrıca yenilenen İlkokul Matematik Programı daha farklı öğrenme ve alt öğrenme alanlarının oluşturulmasını sağladığından bu alanlara yönelik başarı testleri geliştirilebilir.

Araştırmacıların Beyanı

Araştırmacıların katkı oranı beyanı:

Yazarlar makaleye eşit oranda katkı sağlamıştır.

Çatışma beyanı:

Yazarlar bu makalenin araştırılması, yazarlığı veya yayımlanmasına ilişkin herhangi bir potansiyel çıkar çatışması beyan etmemiştir.

Destek ve teşekkür:

Araştırmacılar çalışmada herhangi bir kurum veya kuruluştan destek almamıştır.

REFERENCES

- Altun, M. (2018). İlkokullarda matematik öğretimi [Teaching mathematics in primary schools]. Bursa: Aktüel.
- Akbulut, H. İ. & Çepni, S. (2013). Bir üniteye yönelik başarı testi nasıl geliştirilir? ilköğretim 7. sınıf kuvvet ve hareket ünitesi [How to develop an achievement test for a unit?: A study for grade 7 force and motion unit]. *Amasya Üniversitesi Eğitim Fakültesi Dergisi*, 2(1), 18-44.
- Akgül, E. (2011). İlköğretim 4. sınıf öğrencilerine matematik dersinde "açılar" konusunun öğretilmesinde proje tabanlı öğretim yönteminin etkisinin incelenmesi [Analizing the effect of project based learning method in mathematics on teaching primary school fourth grade students the angles subject]. Unpublished master thesis, Marmara University, İstanbul.
- Atılğan, H., Kan, A. & Doğan, N. (2006). Test geliştirme [Test Development]. In Atılğan, H. (Ed.), *Eğitimde ölçme ve değerlendirme* [Measurement and evaluation in education] (pp. 378-393). Ankara: Anı.
- Ayvacı, Ş.H. & Durmuş, A. (2016). Bir başarı testi geliştirme çalışması: Isı ve sıcaklık başarı testi geçerlik ve güvenilirlik araştırması, [Developing an achievement test: Heat and temperature achievement test validity and reliability research]. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi/OMU J. Fac. Educ.*, 35(1), 787-103.
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme; klasik test teorisi ve uygulanması* [Measurement in education and psychology: Classical test theory and its application]. Ankara: ÖSYM.
- Büyükkıdık, S. (2020). Geçerlik ve güvenilirlik [Validity and reliability]. In Oğuz, E. (Ed.), *Eğitimde araştırma yöntemleri* [Research methods in education] (pp. 169-201). Ankara: Eğiten Kitap.
- Büyükoztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş. & Demirel, F. (2014). *Bilimsel araştırma yöntemleri* [Scientific research methods]. (16th ed.), Ankara: Pegem.
- Dağlı, H. & Peker, M. (2012). İlköğretim 5. sınıf öğrencileri geometrik şekillerin çevre uzunluğunu hesaplamaya ilişkin ne biliyor? [What do 5th grades know related the geometric shapes' perimeter?]. *The Journal of Theoretical Educational Science*, 5(3), 330-351.
- Demirel, Ö. (2010). *Kuramdan uygulamaya eğitimde program geliştirme*[Curriculum development in education from theory to practise]. Ankara: Pegem.
- Ellez, M. (2014). Ölçme araçlarında bulunması gereken özellikler [Features that must be present in measuring tools]. In Tanrıoğen, A. (Ed.), *Bilimsel araştırma yöntemleri* [Scientific research methods] (pp. 167-190). Ankara: Anı.
- Emekli, A. (2001). Ölçüler konusunun öğretiminde yanlışların teşhisi ve alınması gereken tedbirler [Diagnosing the misconceptions in teaching measurement and required measures]. Unpublished master thesis, Selçuk University, Konya.
- Fidan, E. (2013). İlkokul öğrencileri için matematik dersi sayılar öğrenme alanında başarı testi geliştirilmesi [Development of achievement tests in the number domain of mathematics course for elementary school students]. Unpublished master thesis, Ankara University, Ankara.
- Gough, J. (2008). Fixing misconceptions: Length, area and volume. *Australian Mathematics Teacher*, 64(2), 34-35. <https://files.eric.ed.gov/fulltext/EJ802697.pdf>
- Gürkan, E. (2019). Sınıf dışı öğrenme ortamlarında farklılaştırılmış öğretime yönelik bir eylem araştırması [An action research on differentiated learning activities in outdoor learning]. Unpublished master thesis, Ege University, İzmir.
- Güzeller, C. O., Aksu, G. & Eser, M. T. (2019). Klasik test kuramı ve madde tepki kuramına ilişkin veri analizi yazılımı: Jmetrik [Data on classical test theory and item response theory analysis software: Jmetric]. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 10(2), 165-178.

- Hong, L. K. (1984). List processing free responses: analysis of open-ended questions with word processor. *Qualitative Sociology*, 7, 98-109. <https://doi.org/10.1007/BF00987110>
- İnan, Ş. (2016). *4. sınıf ilkokul matematik not defteri* [4th grade primary school math notebook]. Ankara: Mutlu.
- Innami, Y., & Koizumi, R. (2009). A meta-analysis of test format effects on reading and listening test performance: Focus on multiple-choice and open-ended formats. *Language Testing*, 26(2), 219-244. <https://journals.sagepub.com/doi/10.1177/0265532208101006>
- Karadağ, S., Balcı, M., Abdik, E. & Demiralp, A. (2017). *İlkokul 4 matematik öğretmen kılavuz kitabı* [Primary school 4 math teacher guide book]. Ankara: Yakın Çağ.
- Karasar, N. (2012). *Bilimsel araştırma yöntemi* [Scientific research method] (24th Edition). Ankara: Nobel.
- Karip, E. (2008). *Ölçme ve değerlendirme* [Measurement and evaluation]. Ankara: Pegem.
- Konak, Ö. A., (2013). Öğretim programlarının güncellenmesi üzerine, *CITO Eğitim: Kuram ve uygulama*, 20, 4-8.
- Koyuncu, E. (2017). Temel eğitimden ortaöğretime geçiş (TEOG) sınavında açık uçlu soruların kullanılmasına ilişkin öğrenci görüşlerinin incelenmesi [Determination of student views on the use of open-ended questions on transition from primary to secondary education exam (TEOG)]. Unpublished master thesis, Gaziantep University, Gaziantep.
- Koyuncu, İ., Aksu, G., & Kelecioğlu, H. (2018). Mantel-Haenszel, Lojistik regresyon ve olabilirlik oranı değişen madde fonksiyonu inceleme yöntemlerinin farklı yazılımlar kullanılarak karşılaştırılması [Comparison of Mantel-Haenszel, logistic regression and likelihood ratio methods to evaluate differential item functioning by using different computer software]. *Elementary Online*, 17(2), 909-925.
- Kurt, E. S. (2015). Gerçekçi matematik eğitiminin uzunluk ölçme konusunda başarı ve kalıcılığa etkisi [The effect of realistic mathematics education on achievement and recall in length measurement]. Unpublished master thesis, Ondokuz Mayıs University, Samsun.
- Küçükaydın, A. (2017). *4. sınıf matematik* [4th grade mathematics]. Ankara: Mavi Deniz.
- Linacre, J. M. (2002). What do Infit and Outfit, mean-square and standardized mean?. *Rasch Measurement Transactions*, 16, 878.
- Ministry of National Education [NOME]. (2018). *Matematik dersi öğretim programı (ilkokul ve ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8. sınıflar)* [Mathematics curriculum primary and secondary school grades 1, 2, 3, 4, 5, 6, 7 and 8]. Ankara: Ministry of National Education.
- Moreira, C. Q. & Contente, M. do R. (1997). The role of writing to foster pupil's learning about area. In E. Pehkonen (Ed.), *Proceedings of the 21st PME International Conference*, 3, (pp. 256-263). Lanti, Finland.
- Özdamar, K. (1999). *Paket programlar ile istatistiksel veri analizi* [Statistical data analysis with package programs]. Eskişehir: Kaan.
- Özçelik, D.A. (2013). *Test hazırlama kılavuzu*, [Test preparation guide]. Ankara: Pegem.
- Tan-Şişman, G., & Aksu, M. (2009). Yedinci sınıf öğrencilerinin alan ve çevre konularındaki başarıları [Seventh grade students' success on the topics of area and perimeter]. *Elementary Online*, 8(1), 243-253.
- Tavşancıl, E. (2006). *Tutumların ölçülmesi ve SPSS ile veri analizi* [Measuring attitudes and data analysis with SPSS]. Ankara: Nobel.
- Turgut, M. F., & Baykul, Y. (2012). *Eğitimde ölçme ve değerlendirme* [Measurement and evaluation in education]. Ankara: Pegem.
- Turgut, M. F., & Baykul, Y. (2014). *Eğitimde ölçme ve değerlendirme* [Measurement and evaluation in education]. Ankara: Pegem.

Tutak, T., & Birgin, O. (2008). Geometri öğretiminde bilgisayar destekli öğretimin öğrenci başarısına etkisi [The effects of computer assisted instruction on the students' achievement in geometry]. In IETC 2008 Proceedings Book, 06-08 May, (pp. 1062-1065). Eskişehir, Turkey.

Yıldırım, A., Özgürlük, B., Parlak, B., Gönen, E., & Polat, M. (2016). TIMSS 2015 ulusal matematik ve fen bilimleri ön raporu 4. ve 8. sınıflar [TIMSS 2015 national math and science preliminary report] Erişim adresi: http://timss.meb.gov.tr/wp-content/uploads/TIMSS_2015_Ulusal_Rapor.pdf

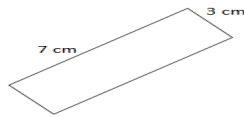
Appendix

4th Class Primary School Perimeter Measures Achievement Test

Explanation

Dear Students,

1. This test includes 18 multiple-choice questions, 2 open-ended questions and 1 short-response question.
2. There is only one answer for multiple-choice questions. You can use the spaces to solve the problems. For open-ended questions, answer by using the space given below the question. For short-response questions, answer using the space beside the question.
3. The time to complete the test is 35 minutes.
4. During assessment, only correct answers will be noted, wrong answers will not be noted.



1. How many cm is the perimeter of the rectangle above?

- A) 10 B) 15 C) 20 D) 25

2. What is the edge length in cm of a square with 68 cm perimeter?

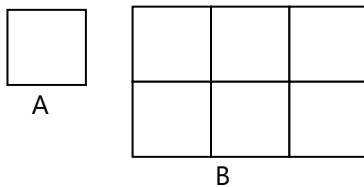
- A) 12 B) 17 C) 27 D) 34

3. What is the long edge in cm for a rectangle with perimeter 92 cm and short edge length 18 cm?

- A) 18 B) 56 C) 36 D) 28

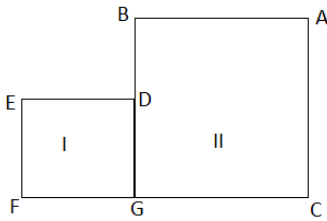
4. A rectangle has width of 6 cm and length of 10 cm. A square has perimeter which is equal to the perimeter of this rectangle. What is the length of the edges of this square in cm?

- A) 4 B) 8 C) 10 D) 16



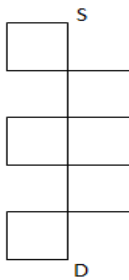
5. If the perimeter of square A shown above is 20 cm, what is the perimeter of rectangle B made of these squares in cm?

- A) 50 B) 40 C) 60 D) 120



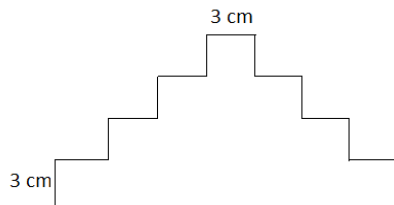
6. In the figure above, if the square region I has perimeter 24 cm and square region II has perimeter 44 cm, what length is FC in cm?

A) 11 B) 15 C) 17 D) 20



7. In the figure $SD = 35$ cm. What is the perimeter of this shape made of equal squares in cm?

A) 70 B) 80 C) 110 D) 140



8. The short edges in the figure above have equal lengths. If the length of the short edges is 3 cm, what is the perimeter of the shape in cm?

A) 21 B) 48 C) 56 D) 66

9. A square has perimeter 44 cm. What is the perimeter in cm of a rectangle created by combining four of these squares side by side?

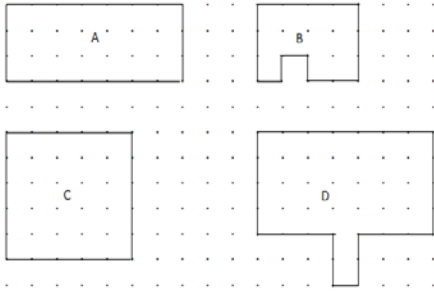
A) 121 B) 110 C) 132 D) 55

10. A rectangle has short edge length that is $\frac{1}{4}$ the long edge length. If the perimeter of this rectangle is 140 cm, what is the short edge length in cm?

A) 5 B) 28 C) 14 D) 56

11. A square has perimeter 44 cm. What is the perimeter in cm of a rectangle created by combining four of these squares side by side?

A) 121 B) 110 C) 132 D) 55



12. Which of the shapes above have equal perimeters?

- A) A and B B) A and C C) B and C D) C and D

13. Write a perimeter problem using at least two of the following; square, triangle and rectangle.

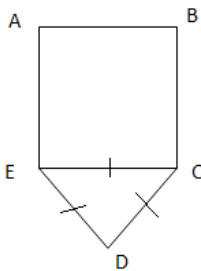
.....

14. The perimeter of an equilateral triangle with edge length 8 cm is equal to the perimeter of a square. Accordingly, what is the edge length of the square in cm?

- A) 2 B) 6 C) 8 D) 32

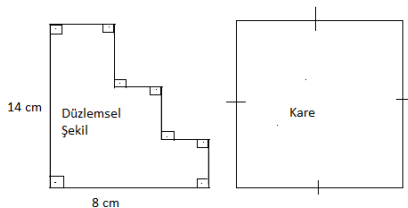
15. What is the length in cm of the unequal side of an isosceles triangle with one edge of equal length 21 cm and perimeter 70 cm?

- A) 16 B) 21 C) 28 D) 42



16. In the shape above, AECB is a square. EDC is an equilateral triangle. Accordingly, if the perimeter of the triangle is 39 cm, what is the perimeter of the square in cm?

- A) 13 B) 29 C) 39 D) 52

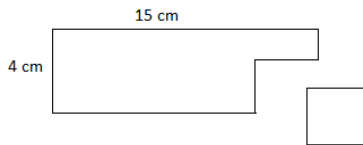
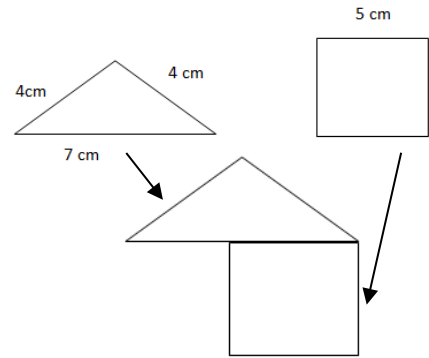


17. If the planar shape and the square above have equal perimeters, what is the edge length of the square in cm?

- A) 11 B) 16 C) 18 D) 22

18. The triangle and square with edge lengths given are combined as shown. What is the perimeter of this new shape in cm?

- A) 20 B) 25 C) 30 D) 35



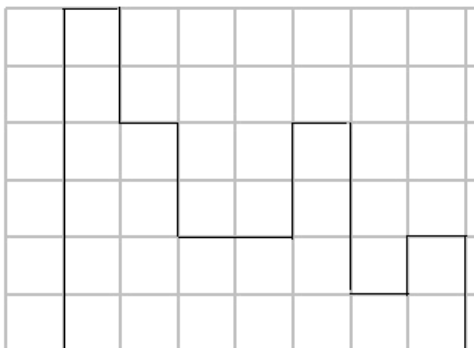
19. A square with edge length 3 cm is cut from the rectangle above. After cutting, what is the perimeter of the remaining shape in cm?

- A) 19 B) 26 C) 35 D) 38

20. Write a perimeter problem using at least two of the following; square, triangle and rectangle.

.....

21. Write the perimeter of the shape below shown on graph paper in units. (The length of the edge of each square is one unit.)



.....unit

Answer Key

1	C	7	D	13	-	19	D
2	B	8	D	14	B	20	-
3	D	9	B	15	C	21	(32 unit)
4	B	10	C	16	D		
5	A	11	B	17	A		
6	C	12	B	18	B		

Note: Q 13, 20 and 21 are open-ended.

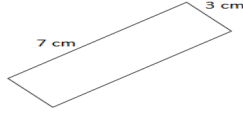
EK

İlkokul 4. Sınıf Çevre Ölçme Başarı Testi

Açıklama

Sevgili Öğrenciler;

5. Bu testte 18 çoktan seçmeli, 2 açık uçlu ve 1 kısa cevaplı soru bulunmaktadır.
6. Çoktan seçmeli soruların sadece bir cevabı bulunmaktadır. Soruların çözümü için boşlukları kullanabilirsiniz. Açık uçlu sorular için de sorunun altında verilen boşluğu kullanarak cevaplayabilirsiniz. Kısa cevaplı soru için sorunun yanında verilen boşluğu kullanarak cevaplayabilirsiniz.
7. Testin cevaplama süresi yaklaşık 35 dakikadır.
8. Değerlendirme sırasında sadece doğru cevaplar dikkate alınacak, yanlış cevaplar dikkate alınmayacaktır.



1. Yukarıdaki dikdörtgenin çevre uzunluğu kaç cm'dir?

- B) 10 B) 15 C) 20 D) 25

2. Çevre uzunluğu 68 cm olan bir karenin bir kenar uzunluğu kaç cm'dir?

- B) 12 B) 17 C) 27 D) 34

3. Çevre uzunluğu 92 cm ve kısa kenar uzunluğu 18 cm olan bir dikdörtgenin uzun kenarı kaç cm'dir?

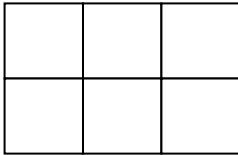
- B) 18 B) 56 C) 36 D) 28

4. Bir dikdörtgenin eninin uzunluğu 6 cm, boyunun uzunluğu 10 cm'dir. Bir karenin çevre uzunluğu da bu dikdörtgenin çevre uzunluğuna eşittir. Bu karenin bir kenarının uzunluğu kaç cm'dir?

- B) 4 B) 8 C) 10 D) 16



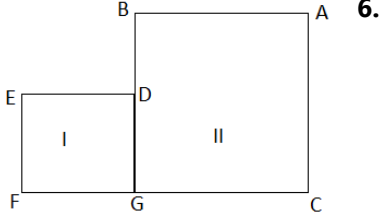
A



B

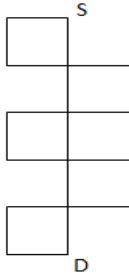
5. Yukarıdaki A karesinin çevre uzunluğu 20 cm ise bu karelerden oluşan B dikdörtgeninin çevresi kaç cm'dir?

- B) 50 B) 40 C) 60 D) 120



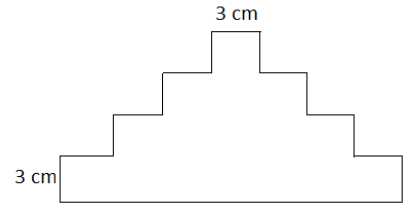
6. Yukarıdaki şekilde I. Karesel bölgenin çevresi 24 cm ve II. Karesel bölgenin çevresi 44 cm olduğuna göre, FC uzunluğu kaç cm'dir?

- A) 11 B) 15 C) 17 D) 20



7. Şekilde $SD = 35$ cm'dir. Eş karelerden oluşan bu şeklin çevresi kaç cm'dir?

- B) 70 B) 80 C) 110 D) 140



8. Yukarıdaki şekilde kısa kenarların uzunlukları birbirine eşittir. Kısa kenarın kenar uzunluğu 3 cm ise şeklin çevresi kaç cm'dir?

- A) 21 B) 48 C) 56 D) 66

9. Bir karenin çevre uzunluğu 44 cm'dir. Bu karelerin dört tanesinin yan yana birleşmesiyle oluşan dikdörtgenin çevre uzunluğu kaç cm'dir?

- A) 121 B) 110 C) 132 D) 55

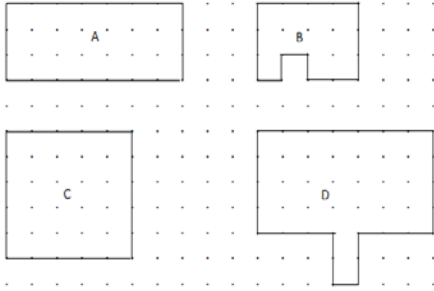
10. Bir dikdörtgenin kısa kenar uzunluğu uzun kenar uzunluğunun $\frac{1}{4}$ 'üdür. Bu dikdörtgenin çevre uzunluğu 140 cm ise kısa kenarı kaç cm'dir?

B) 5 B) 28 C) 14 D) 56

11. Bir karenin çevre uzunluğu 44 cm'dir. Bu karelerin dört tanesinin yan yana birleşmesiyle oluşan dikdörtgenin çevre uzunluğu kaç cm'dir?

A) 121 B) 110 C) 132 D) 55

12.



Yukarıdaki şekillerden hangilerinin çevre uzunlukları birbirine eşittir?

A) A ve B B) A ve C C) B ve C D) C ve D

13. Kare, üçgen ve dikdörtgenden en az iki tanesini kullanarak bir çevre problemi kurunuz.

.....

.....

.....

.....

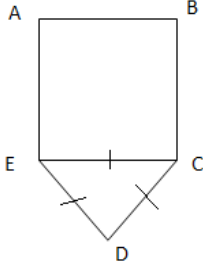
14. Bir kenar uzunluğu 8 cm olan eşkenar üçgenin çevre uzunluğu, bir karenin çevre uzunluğuna eşittir. Buna göre karenin bir kenar uzunluğu kaç cm'dir?

B) 2 B) 6 C) 8 D) 32

15. Eşit kenarlardan biri 21 cm ve çevre uzunluğu 70 cm olan bir ikizkenar üçgenin eş olmayan kenarı kaç cm'dir?

B) 16 B) 21 C) 28 D) 42

16.

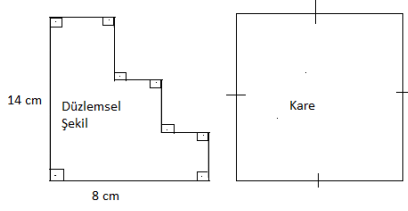


Yukarıdaki şekilde AECB bir karedir. EDC bir eşkenar üçgendir.

Buna göre eşkenar üçgenin çevre uzunluğu 39 cm ise; karenin çevre uzunluğu kaç cm'dir?

- B) 13 B) 29 C) 39 D) 52

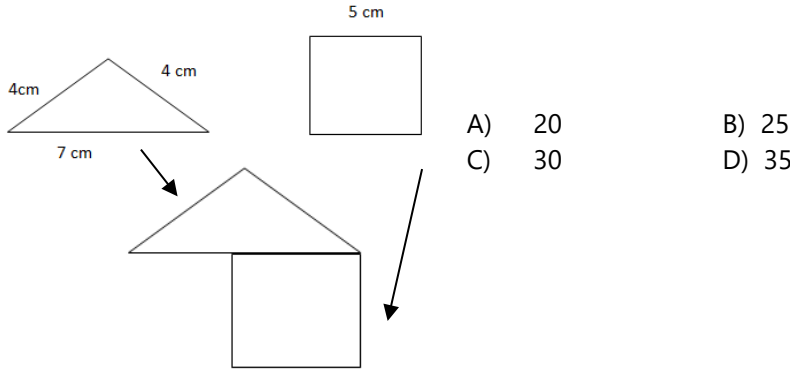
17.



Yukarıdaki düzlemsel şekil ve karenin çevre uzunlukları birbirine eşitse; karenin bir kenar uzunluğu kaç cm'dir?

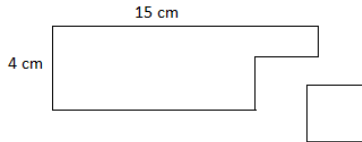
- B) 11 B) 16 C) 18 D) 22

18. Kenar uzunlukları verilen üçgen ve kare şeklindeki gibi birleştiriliyor. Oluşturulan bu şeklin çevresi kaç cm'dir?



- A) 20 B) 25
C) 30 D) 35

19.



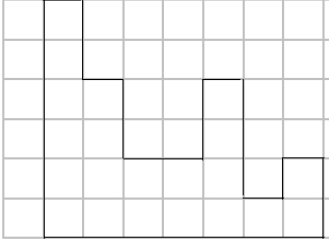
Yukarıdaki dikdörtgenden bir kenar uzunluğu 3 cm olan kare kesiliyor. Kesildikten sonra oluşan şeklin çevre uzunluğu kaç cm'dir?

- A) 19 B) 26 C) 35 D) 38

20. Kare, üçgen ve dikdörtgenden en az iki tanesini kullanarak bir çevre problemi kurunuz.

.....
.....
.....
.....

21. Birim kareli düzlem üzerinde verilen aşağıdaki şeklin çevre uzunluğu kaç birim olduğunu aşağıdaki boşluğa yazınız. (Karelerin bir kenar uzunluğu bir birimdir.)



.....birim

Cevap Anahtarı

1	C	7	D	13	-	19	D
2	B	8	D	14	B	20	-
3	D	9	B	15	C	21	(32birim)
4	B	10	C	16	D		
5	A	11	B	17	A		
6	C	12	B	18	B		

Not:13., 20. ve 21. ve sorular açık uçludur.