



(ISSN: 2602-4047)

Çokçalışkan, H. (2024). Development of a spatial ability test for 7<sup>th</sup> grade students, *International Journal of Eurasian Education and Culture*, 9(26), 180-214.

DOI: <http://dx.doi.org/10.35826/ijoecc.1811>

Article Type (Makale Türü): Research Article

---

## DEVELOPMENT OF A SPATIAL ABILITY TEST FOR 7<sup>TH</sup> GRADE STUDENTS<sup>1</sup>

**Hasan ÇOKÇALIŞKAN**

PhD Student, Nevşehir Hacı Bektaş Veli University, Nevşehir, Turkey, [hasancokcaliskan@nevsehir.edu.tr](mailto:hasancokcaliskan@nevsehir.edu.tr)

ORCID: 0000-0002-3269-0144

Received: 23.10.2023

Accepted: 14.05.2024

Published: 01.06.2024

### ABSTRACT

The aim of this study is to develop an achievement test for 7th grade students. The lack of a current achievement test to measure spatial ability at the 7th grade level led to the emergence of this study. For this purpose, firstly, the acquisitions related to the 7th grade "Views of Objects from Different Directions" topic were examined, the literature was reviewed and 84 questions were prepared. After expert opinions were taken, the number of questions was reduced to 64 by making the necessary arrangements. The validity and reliability of the 64-question test were examined, and it was applied to 139 7th grade students in the 2022-2023 academic year as a pilot application. According to the results of the item analysis conducted with the TAP program, the final version of the 25-question Spatial Ability Achievement Test (SATT) was given by taking expert opinions. While the average difficulty index value of the achievement test was 0.647 in the pilot application, it was found to be 0.540 in the final version of the test. The KR-20 value of the reliability analysis of the achievement test was found to be 0.918 in the final version of the test while it was 0.848 in the pilot study. As a result of this study, a 25-question, four-choice multiple-choice achievement test with validity and reliability was developed.

**Keywords:** Achievement test, spatial ability, reliability, validity, test development.

---

<sup>1</sup> This study, was presented as an oral presentation in 2023: NEM 2023 (3rd International Natural Science, Engineering and Material Technologies Conference) Kibris

---

## INTRODUCTION

Measurement is defined as the process of selecting the necessary tools to express certain characteristics of an individual or student numerically in the light of certain methods and rules (Aksu, 2018). Another definition is that measurement is the act of observing the characteristics of people or objects using appropriate tools and expressing the observations with symbols. In summary, determining the extent to which an individual has a desired characteristic for a specific purpose is called measurement (Cizek, 1997). According to Doğan (2019), the reason for measurement is that many properties of the things we observe change or differentiate. In other words, it is stated that change constitutes the basis of measurement. According to Doğan (2019), the existence of differences between individuals triggers the measurement process and change leads to measurement. If there is no change, that is, if everything is constant, there is no need to make measurements, so, there is no need to repeat the process over time after the measurement is made. Because if there is no change, the measurement result will be the same and there will be no need for measurement.

Evaluation is defined as the process of reaching value judgments or conclusions by comparing measurement results with standards in the same field (Yılmaz, 1986). According to Aksu (2018), assessment is the process of collecting all the data necessary to make decisions about the academic development of the individual or student, which is divided into two parts as qualitative and quantitative, and making retrospective or prospective predictions in the light of the data.

It is necessary to use measurement and evaluation processes to determine behavioral change, that is, learning and achievements (Özdemir, 2009). There is no doubt that wherever there is education, there is a need for measurement and evaluation. Student success is tried to be determined through board decisions, teacher opinions, exams, port folios and many other tools (Başol, 2019). In addition, measurement tools such as open-ended questions, multiple-choice questions, true-false questions, fill-in-the-blank questions, and matching questions are widely used to measure student performance in education (Karip, 2012). According to Özçelik (2016), the reasons for making measurements in schools are as follows: to determine the level of readiness for the lesson, to determine whether the actions planned in the lesson have been learned before and to measure each learning unit, to monitor the learning in the unit at the end of each learning unit, at the end of certain periods of the course, to determine the learning levels consistent with the objectives of the course, which occur as a product of the interactions in these periods.

Measurement and evaluation play an important role in students' lifelong mathematics and geometry lessons. The fact that geometry generally involves abstract concepts makes it difficult for most students and causes them to move away from geometry (Gür, Kobak Demir, 2017). Spatial abilities play an important role in geometry lessons, especially in teaching many mathematics topics. Smith (1998) emphasizes the importance of these abilities as follows: "It can be difficult to exist in the world without spatial intelligence. In its absence, we may have difficulty predicting changes in shapes by considering changes in their size and position or expressing

the relationships and position between objects in given directions." Spatial ability is defined as the ability to visualize, move and mentally manipulate the shape of objects (Lohman, 1979; Linn & Petersen, 1985; Carroll, 1993). In another definition, spatial ability is defined as the ability to mentally move or visualize objects consisting of one or more parts and their components in three-dimensional space (Turğut, 2007).

Research measuring spatial ability have been conducted since the 1930s. Guay (1976) stated that tests measuring the rotation of an object around a single axis can be easily solved by simply rotating the test paper. For this reason, Guay (1976) developed a test called the "Purdue Spatial Visualization Test" that requires mental operations related to spatial abilities. Instruments designed to measure spatial ability usually include questions about visualizing 3-D versions of 2-D objects (Kyllonen, Lohman, & Snow, 1984) and 3-D objects rotated about one or more axes (Winter, Lappan, Fitzgerald, & Schroyer, 1989). Eliot and Smith (1983) developed a modified version of the spatial visualization test developed by Guay (1977). This test consists of 24 questions and measures the ability to visualize 3D objects from an imaginary perspective. A review of the literature reveals that tests measuring spatial ability usually include questions that are constructed using unit cubes. Questions requiring the ability to rotate a geometric structure in the mind (MGMP), questions requiring the ability to draw the views of a geometric structure from different directions (Demirkaya, 2017), questions about finding the number of unit cubes in a structure (Battista & Clements, 1996; Ben Chaim et al., 1985; MGMP; Yolcu, 2008), and tests for adding and subtracting unit cubes for certain structures (MGMP) are also found in the literature. Some of the tests for measuring spatial ability encountered in the literature are listed below;

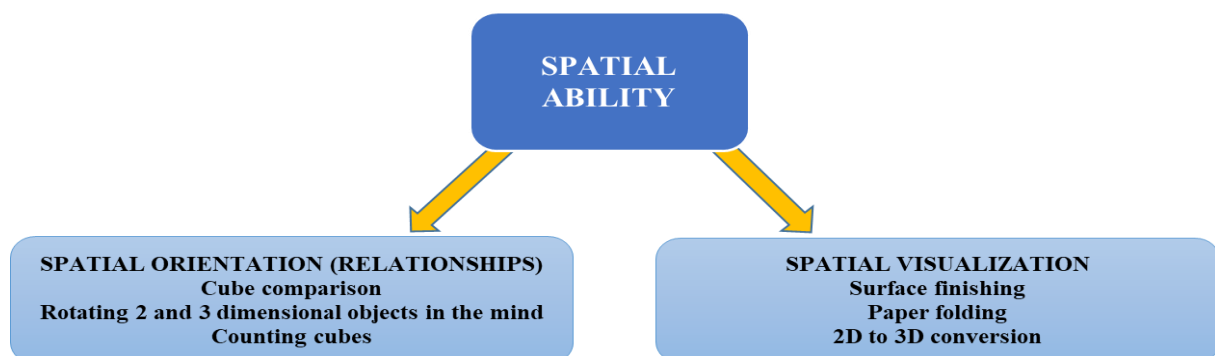
**Table 1.** Tests Developed to Measure Spatial Ability

Name of Test	Researcher(s) Who Developed The Test
Mental Cutting Test	CEEB (1939)
Paper Folding Test	French et al.(1963)
Spatial Visualization Test	Lappan et al.( 1983)
Minnesota Paper Covering Test	Likert (1970)
Purdue Spatial Visualization Test	Guay (1977)
Kesti Surface Test	Cohen & Hegarty (2012)
Spatial Visualization Test	Yüksel & Bülbül (2014)
Card Flip and Cube Comparison Test	Ekstrom et al.( 1976)
Mental Rotation Test	Vandenberg & Kuse (1978)
Wheatley Spatial Ability Test	Wheatley (1978)
Guildford-Zimmerman Spatial Orientation Test	Guildford-Zimmerman (1948)
Santa Barbara Direction Perception Scale	Hegarty et al.( 2002)
Revised Purdue Spatial Visualization Test	Yoon Yoon (2011)

There is no consensus on the sub-dimensions of spatial ability in tests designed to measure spatial ability. McGee (1979), Lohman (1979), Clements (1998) and Kayhan (2005) define the sub-dimensions of spatial ability as spatial visualization and spatial orientation; Linn and Petersen (1985) and Okagaki and Frensch (1996) define spatial visualization, spatial perception, and mental rotation; Olkun and Altun (2003) and Turğut (2007) as spatial orientation (relations) and spatial visualization; Contero (2005) as spatial visualization, spatial

orientation and spatial relations. In the literature, the commonly accepted dimensions of spatial ability are spatial relations and spatial visualization. Although spatial abilities have sub-dimensions with different names, the operations covered by spatial abilities generally include rotating objects in space and visualizing them in two or three dimensions. In addition, it is also important to perform these operations quickly (Olkun, 2003). Therefore, tests measuring spatial ability have time limitations.

According to Olkun and Altun (2003), spatial relations are the ability of the student to mentally rotate two-dimensional and three-dimensional geometric shapes as a whole and perceive them in different positions. In addition, these tests are defined as tests that require fast and accurate decision making. Spatial visualization is defined as the ability to visualize new situations that arise because of two- and three-dimensional movement of objects in three-dimensional space, consisting of one or more parts and images of these parts. When the items in the standard tests measuring this skill are examined, it is seen that they require mental actions such as complex shapes consisting of moving parts and/or mental folding or mental integration to transform from 2-D to 3-D (Pellegrino et al., 1984 as cited in Olkun & Altun, 2003:2). Unlike spatial relations, these tests focus on the accuracy of increasingly complex objects more than speed. The contents of the sub-dimensions of spatial ability accepted in the literature are shown in Figure 1.



**Figure 1.** Spatial Ability Sub-dimensions

Measuring student performance requires a well-crafted measurement tool that covers the whole student performance. Students' cognitive levels can be measured with tools developed by using Bloom's Taxonomy, developed by Benjamin Bloom (1956) and named after him. Bloom's taxonomy predicts that questions can be prepared to measure students' thinking skills and that this can be done not only in the cognitive domain but also in the affective and psychomotor domains (Çepni, 2003). The categories of cognitive domains start with the recall of dry and memorized knowledge and extend to comprehension, application, analysis, synthesis and evaluation. As we move from the knowledge stage to the evaluation stage, behaviors become more complex and difficult to learn and develop (Tekin, 1991). According to Bloom's proposed taxonomy, students' cognitive levels can be measured as follows: "knowledge", "comprehension" and "application", whereas higher cognitive levels (HLC) include "analysis", "synthesis" and "evaluation" (Colletta & Chiappetta, 1989). In the revised

Bloom's Taxonomy, these steps were changed and updated as "Remembering, Understanding, Applying, Analyzing, Evaluating and Creating".

Renewed in 2018 and still valid and implemented, the new curriculum is based on the constructivist approach and accepts that knowledge is constructed by students. The change in the curriculum also brought about changes in course content, teaching methods, tools and materials used, and measurement and evaluation methods (Gelbal & Kelecioğlu, 2007). The unit "Views of Objects from Different Directions", which is directly related to spatial ability, is included in the curriculum as one of the last topics of the mathematics course at the 7th grade level. The fact that the subject is included in the last units, that the written exams of the students are completed before this subject, and that the mock exams held in schools are also completed before this unit showed that it is very difficult or even impossible to evaluate this unit seriously in the interviews with the course teachers. The literature reviews also revealed that there is a need for a new and up-to-date spatial ability assessment tool at the secondary school level because of the constantly renewed and developing education programs. Therefore, the aim of this study is to develop a spatial ability achievement test with proven validity and reliability at the secondary school level to meet the needs of researchers and teachers who conduct academic studies.

## **METHOD**

### **Research Pattern**

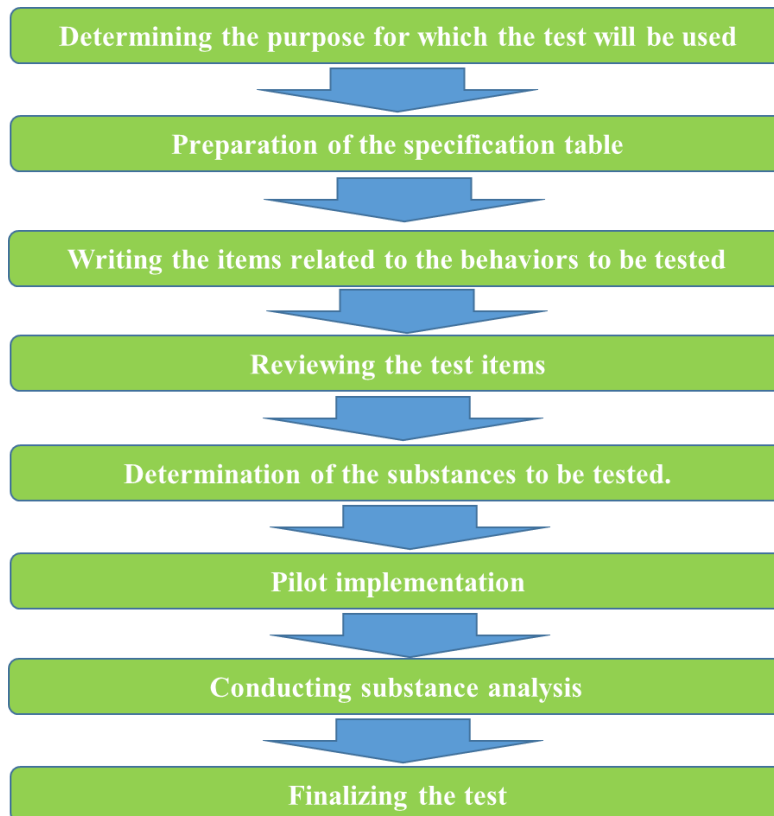
This study aimed to develop a "Spatial Ability Test" (SAT) at the secondary school level and was conducted using the survey model, which is one of the quantitative research methods. This model aims to describe a past or present situation as it exists, and to reveal the opinions, abilities and skills of individuals regarding a subject or event (Fraenkel & Wallen, 1996; Karasar, 2011). This model was preferred since the "spatial ability levels" of the students were to be determined through multiple-choice questions. In the study, a reliability analysis was conducted to determine how sensitive the test can measure the feature it measures and how clean the results are from errors, and a validity analysis was conducted to determine the level of the test's ability to measure the feature to be measured without confusing it with other features (Küçükahmet, 2002).

### **Population and Sample**

The population of this study consists of 7th grade students studying in the central district of Nevşehir province in the 2022-2023 academic year. While determining the sample in this study, 139 students studying in the 7th grade in a public secondary school in the central district of Nevşehir province were studied with the convenience sampling method. The convenience sampling method was preferred because it is both fast and easy (Büyükoztürk et al., 2012).

### Test Development Process

The steps applied in the development process of SAT in this study are given in Figure 2 (Turgut & Baykul, 2014).



**Figure 2.** SAT Development Steps (Turgut & Baykul, 2014).

First, the purpose of the test was determined. With the development of the test in this study, it was aimed to determine the spatial ability levels of the students before the unit was taught, to determine the spatial ability levels and learning deficiencies of the students after the unit was taught, to enable teachers and researchers to benefit from the test, to benefit the researchers who will work in this field and to contribute to the literature.

After the purpose of the test was determined, a specification table was prepared. In this study, it was tried to prepare questions with the data obtained within the literature research in accordance with the spatial ability sub-dimensions in a way to cover the acquisitions of the 7th grade mathematics course "Views of Objects from Different Directions" unit. In 2009, the unit was included in the 6th, 7th, and 8th grade curricula in the MoNE curriculum, but in the curriculum renewed in 2018 and still used, it is only included as a 7th grade unit. Table 2 below shows the learning outcomes of the unit in 2009 and Table 3 below shows the objectives of the unit in 2018.

**Table 2.** Learning Outcomes and Grade Levels Related to the Views of Objects from Different Directions in the 2009 MoNE Curriculum

GRADE	2009 MoNE Curriculum Learning Outcomes
6. GRADE	Draws the views of the structures formed with equal cubes from different directions.
7. GRADE	Constructs the structures whose drawings are given for the views of their faces from different directions with unit cubes and draws them on isometric paper.
8. GRADE	Constructs the given structures with multiple cubes and draws the views of the structures constructed with multiple cubes.

**Table 3.** Learning Outcomes Related to the Views of Objects from Different Directions in the 2018 MoNE Curriculum and Grade Level

GRADE	2018 MoNE Curriculum Learning Outcomes
<b>M.7.3.4. Views of Objects from Different Directions</b>	
7. GRADE	M.7.3.4.1. Draws two-dimensional views of three-dimensional objects from different directions.
	a) Structures made of equal cubes and known geometric objects are used. Suitable squared papers are used for drawing. Students are asked to associate the views of the structures from different directions (such as symmetrical front-back and right-left views).
	b) Interactive studies can be included with appropriate information and communication technologies.
	M.7.3.4.2. Constructs structures given drawings of their views from different directions.
a) Structures made of equal cubes and known geometric objects are used. Isometric paper can be used to draw structures made of equal cubes.	
b) Interactive activities with appropriate information and communication technologies can be included.	

In the 2018 curriculum, the subject of the appearance of objects from different directions is taught in 5 lesson hours with 2 learning outcomes. The items prepared for the sub-dimensions of spatial ability to cover these acquisitions were classified according to the "Revised Bloom's Taxonomy" and a specification table was created. The specification table of the final version of the SAT is presented in Table 4.

**Table 4.** Specification Table for the Final Test

TOPICS	COGNITIVE DOMAIN			
	Applying	Analyzing	Evaluating	Creating
Counting cubes	4, 14, 15, 16, 17, 21	1, 4, 5, 10, 11, 12, 13, 23	1, 2, 3, 5, 6, 7, 8, 9, 10, 12, 18, 19, 20	22, 24, 25
Cube Comparison				22, 24, 25
2D to 3D and 3D to 2D Visualization	14, 15, 16, 17	13	18, 19, 20	
Rotating 2D and 3D objects in the mind	4, 14, 15, 16, 17	1, 4, 5, 10, 12, 23	1, 2, 5, 6, 7, 8, 9, 10, 12, 18, 19, 20	22, 24, 25
Surface Completion (Mental Integration)	4	4, 11	3	24, 25
Parsing Shape in Mind	14, 15, 16, 17, 21	1, 5, 10, 11, 12, 13	1, 2, 3, 5, 6, 7, 8, 9, 10, 12, 18, 19, 20	22, 24, 25

When Table 4 is examined, it is seen that the questions were not included in the recall and comprehension levels at all, and one question was included in more than one sub-dimension and at many levels. During the writing of the items related to the behaviors to be tested; before the questions were prepared, the learning outcomes of the unit "Views of Objects from Different Directions" in the mathematics curriculum were examined and analyzed. Then, the 7th grade mathematics textbook, auxiliary course resources and previous spatial ability tests in the literature were examined. As a result of the examinations, no question was taken in the same way and all of the questions were prepared by the researcher. An item pool consisting of 84 questions with each sub-dimension and appropriate to the learning outcomes was created.

At the stage of reviewing the items, i.e. the validity of the test and determining the items to be included in the test, a specification table was prepared to check the content validity and expert opinion was taken. For the questions in the item pool, the opinions of 3 faculty members, 1 research assistant, 1 language expert, 1 educational sciences expert and 2 mathematics teachers working in schools affiliated to National Education were taken. After the expert opinions, it was decided to reduce the number of questions in the item pool to 64. Table 5 below was prepared for the expert opinions and the experts were asked to evaluate 84 questions.

**Table 5.** Expert Question Evaluation Table

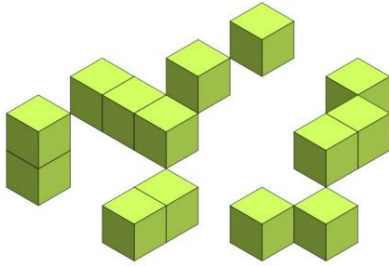
---

Item Number	Suitable	Not Suitable	Correction Required	Correction Suggestions
-------------	----------	--------------	---------------------	------------------------

---

As a result of the expert opinions, questions with the same characteristics or aimed at measuring the same outcome were removed from the item pool and a 64-question test was obtained. 20 questions were removed from the pool and some questions were corrected. Experts gave their opinions on whether the prepared questions were appropriate for the relevant learning outcomes, their suitability to the specified cognitive domain level, the suitability and comprehensibility of the questions to the student level, their suitability to the grammar rules, etc. For example, the statement in the question item given in Figure 3 was revised by adding the phrase "You are asked to perform operations by dividing the shapes into unit cubes" in parentheses as expert feedback.





If it is desired to obtain a cube with the largest volume by using only the unit cubes given in the next figure, how many unit cubes will increase? (You are asked to perform the operation by dividing the shapes into unit cubes).

- A) 12
- B) 10
- C) 8
- D) 6

**Figure 3.** Pilot study sample revised test item.

In data collection, students were asked to answer the questions voluntarily without concern for grades. Thus, it was assumed that subject bias was prevented. Students were given 40+40 minutes, i.e. 2 class hours, to answer the test. The test was administered to the students in 2 sections. According to expert opinions, a time limit of 80 minutes was set for the test applied to measure spatial ability so that the solution could not be reached by using cognitive strategies (Linn & Petersen, 1985).

In the step of conducting the pilot application, the spatial ability test was applied to 9 seventh grade students as a pre-pilot application to test whether the questions were understandable and whether the time given for answering the questions was sufficient. In line with the feedback received, it was seen that the questions were understood by the students and the time given was sufficient to answer the test. Then, the pilot application, the SAT, consisting of 64 multiple-choice questions, was applied to 139 seventh grade students studying in the central district of Nevşehir province in the 2022-2023 academic year. For the application, permission was obtained from Nevşehir Hacı Bektaş Veli University Rectorate Scientific Research and Publication Ethics Committee dated 27.03.2023 and decision number 2023.03.84. The necessary permissions were also obtained by applying to the Provincial Directorate of National Education.

At the stage of item analysis and finalizing the test, the reliability analysis of the test was calculated by looking at the KR-20 internal consistency coefficient. When scoring the 64-question SAT, each correct answer was given a score of "1" and each incorrect answer was given a score of "0" and the evaluations were made over a total of 64 points. Since the total sample size was 139, no grouping was made as lower group or upper group, and all students were evaluated over the exam results.

Bu bölüm içerisinde araştırmanın modeli, evren-örneklem, çalışma grubu, veri toplama araçları, geçerlik-güvenirlik, verilerin analizi gibi alt başlıklar yer almalıdır. Araştırmanın deseni bu kısımda ayrıntılı olarak açıklanmalıdır. Yöntemin kuramsal tanımının verilmesi yerine süreç detaylı şekilde baştan aşağı anlatılmalıdır. Etik kurul onayı yöntem kısmında detaylandırılmalıdır.

**FINDINGS**

**Findings Related to Item Analysis of SAT**

Item difficulty and item discrimination analyses of the 64-question pilot version of the test were conducted in TAP Item and Test Analysis Program. The results of the item analysis of the pilot test are given in Table 6.

**Table 6.** Item Analysis Values Obtained as a Result of the Pilot Application of the Spatial Ability Test

Item	Item Difficulty Index (Pj)	Item Discrimination Index (rjx)	Item	Item Difficulty Index (Pj)	Item Discrimination Index (rjx)
Item 1*	0.87	0.28	Item 33*	0.87	0.41
Item 2*	0.98	0.05	Item 34*	0.91	0.28
Item 3*	0.93	0.13	Item 35	0.77	0.42
Item 4*	0.86	0.21	Item 36*	0.78	0.42
Item 5*	0.88	0.13	Item 37*	0.81	0.29
Item 6*	0.93	0.03	Item 38*	0.90	0.33
Item 7*	0.81	0.41	Item 39*	0.86	0.34
Item 8*	0.10	0.12	Item 40*	0.91	0.26
Item 9	0.53	0.41	Item 41	0.52	0.55
Item 10*	0.74	0.44	Item 42*	0.76	0.64
Item 11	0.38	0.49	Item 43	0.61	0.60
Item 12*	0.47	0.46	Item 44*	0.60	0.60
Item 13*	0.41	0.41	Item 45	0.62	0.50
Item 14	0.42	0.53	Item 46*	0.76	0.37
Item 15*	0.81	0.23	Item 47	0.58	0.53
Item 16*	0.88	0.11	Item 48*	0.63	0.52
Item 17	0.63	0.57	Item 49	0.45	0.61
Item 18*	0.60	0.30	Item 50	0.53	0.68
Item 19*	0.42	0.34	Item 51*	0.50	0.34
Item 20	0.49	0.60	Item 52	0.37	0.44
Item 21*	0.61	0.57	Item 53*	0.80	0.51
Item 22	0.54	0.60	Item 54	0.67	0.42
Item 23*	0.63	0.80	Item 55*	0.83	0.46
Item 24	0.55	0.68	Item 56*	0.81	0.49
Item 25	0.42	0.52	Item 57	0.61	0.82
Item 26*	0.68	0.45	Item 58*	0.71	0.32
Item 27*	0.71	0.49	Item 59*	0.34	0.18
Item 28	0.65	0.50	Item 60	0.64	0.40
Item 29	0.65	0.62	Item 61	0.62	0.55
Item 30*	0.65	0.57	Item 62*	0.59	0.18
Item 31	0.38	0.47	Item 63	0.53	0.58
Item 32	0.34	0.37	Item 64*	0.58	0.55

\*Marked Items were excluded from the test.

Item difficulty analysis performed on the test items provides information about how easy or difficult the item is for the group to which it is applied. As this value approaches 1.00, it means that the number of correct answers increases and the Item becomes easier, and as it approaches 0.00, the number of incorrect answers increases, and the Item becomes more difficult (Kan, 2011). According to Table 6, it is seen that the difficulty indices of the items in the SAT obtained as a result are between 0.34 and 0.77 and the average difficulty index of the test is 0.647. According to this information, it can be said that there are questions from all levels of difficult and easy Items in the test and the average difficulty of the test is at the medium level.

The Item discrimination value on the test items determines the extent to which the items measure the trait to be measured. As this value approaches +1, it is stated that the Item better measures the characteristic that the test aims to measure, and as it approaches 0, it is stated that the Item is insufficient in measuring the characteristic that the test aims to measure. If this value is negative, it can be considered that the Item measures a characteristic other than the one intended (Kan, 2011). According to Table 6, it is seen that the discrimination indices of the Items in the SAT obtained as a result took values between 0.33 and 0.82 and the average discrimination index of the test was 0.429. According to this result, it can be said that the discrimination of the test is at a sufficient level. The results of the TAP analysis of the pilot and final version of the SAT are given in Table 7.

**Table 7.** TAP Analysis Results for the Pilot and Final Implementations of the Spatial Ability Test

Statistics	Pilot Test	Final Test
Item number	64	25
Number of People Implemented	139	139
Average	41.432	13.489
Standard Deviation	11.396	5.638
Variance	129.871	31.789
Minimum	11	2
Maximum	60	24
Skewness	-0.447	-0.043
Kurtosis	-0.434	-0.906
Average Item difficulty	0.647	0.540
Average Item discrimination	0.429	0.527

Table 7 shows that the mean of the pilot version of the test was 41.432, the standard deviation was 11.396, the variance was 129.871, the minimum score obtained from the test consisting of 64 questions was 11, the maximum score was 60, the skewness value was -0.447, the kurtosis value was -0.434, the average difficulty index of the test was 0.647, and the discrimination index of the test was 0.429. The final version of the test had a mean of 13.489, a standard deviation of 5.638, a variance of 31.789, a minimum score of 2, a maximum score of 24, a skewness value of -0.043, a kurtosis value of -0.906, an average difficulty index of 0.540, and a discrimination index of 0.527.

After the analysis and evaluations, the Item discrimination index. 30 and the average difficulty index. 20 and greater than 0.80 were removed from the test. To ensure the content validity of the tests, expert opinions were taken both in the determination of the critical gains to be measured by the tests and in the writing of the test Items. In the content validity analysis of the tests, specification tables showing the outcomes measured by each test and the cognitive levels of these outcomes in Bloom's taxonomy were prepared. Questions measuring similar dimensions were also removed in line with the expert opinions and the final SAT consisting of 25 questions was obtained.

### Findings Regarding SAT Reliability

Reliability is defined as the relationship between individuals' responses to test items (Büyüköztürk, 2018). According to Turgut and Baykul (2014), the Kuder-Richardson-20 (KR-20) test, which reveals the consistency of the items in the measurement tool with each other, can be applied as a method of determining the reliability of a measurement tool with a single application. In many studies where achievement tests are developed, it is seen that the reliability is reported by calculating the KR-20 reliability coefficient. Büyüköztürk (2017) stated that if the answers to the test items have two options such as true/false, yes/no, it would be appropriate to use the KR-20 coefficient and that this coefficient of .70 and above would be sufficient for the reliability of the test scores. Kehoe (1995) states that the KR-20 value should be 0.50 for multiple-choice tests consisting of 10-15 items and at least 0.80 for tests consisting of 50 or more items (as cited in Tan, Kayabaşı, & Erdoğan, 2003). The KR-20 values for the pilot and final measurements of the achievement test are given in Table 8.

**Table 8.** Spatial Ability Achievement Test Reliability Analysis Results

TEST TYPE	NUMBER OF QUESTIONS	KR-20 COEFFICIENT
PILOT TEST	64	.848
FINAL TEST	25	.918

As a result of the pilot study conducted with 139 participants, the KR-20 reliability coefficient of the 64-question preliminary version of the SAT was found to be .848. The KR-20 reliability coefficient of the final version of the 25-question test obtained by removing 39 questions was calculated as .918. According to these results, it can be said that the reliability of the developed SAT is at a sufficient level.

### SAT Construct Validity

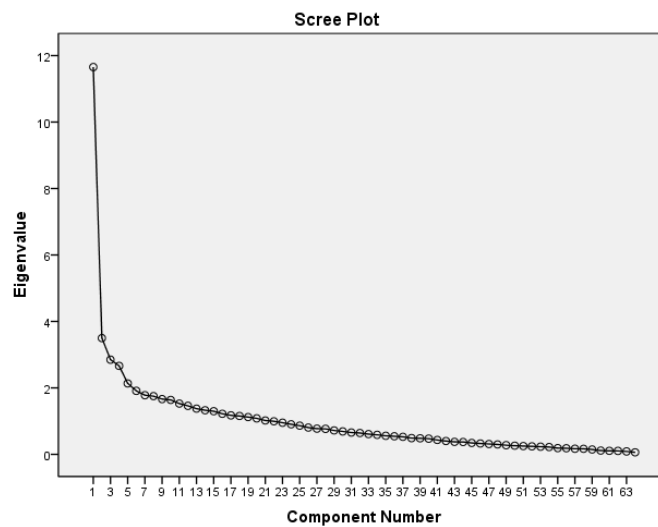
Kaiser-Meyer-Olkin (KMO) value and Barlett's test results are important for conducting exploratory factor analyses of measurement tools. A KMO value of 0.7 and above means a good sample and a KMO value between 0.5-0.7 means a sample that will provide sufficient relationship. Significant Barlett's test results provide information about the normal distribution of the sample (Seçer, 2013). In the light of these values, it can be said that the sample size of 139 people is sufficient for the 64-question test. As a result of the analysis, KMO value was found to be .706 and Barlett's test was found to be significant as .000 ( $p < .05$ ). The results of KMO and Barlett's test show that the distribution is normal and allow the continuation of factor analysis.

According to Seçer (2013), to determine the number of factors as a result of the test, the eigenvalue of each factor should be above "1" and explain at least 5% of the total variance. Table 9 shows the total variance values of the test.

**Table 9.** Total Variance Values of Spatial Ability Test

Factor	Core Value	% of variance	Total %
1	11.655	18.211	18.211
2	3.497	5.464	23.675
3	2.850	4.452	28.128
4	2.662	4.160	32.287
5	2.134	3.335	35.622
6	1.910	2.984	38.607
7	1.779	2.780	41.387
...	...	...	...
...	...	...	...
63	.087	.136	99.904
64	.061	.096	100.000

When Table 9 is examined, it is seen that there are two factors with eigenvalues above "1" and explaining 5% of the total variance, but in the development of measurement tools, only looking at the total variance table is not enough to decide how many factors the tool has. The "ScreePlot" graph also provides information about the factor structure of the instrument (Seçer, 2013). Figure 4 shows the ScreePlot graph of the test.



**Figure 4.** Spatial Ability Test "ScreePlot" Graph

In the "ScreePlot" graph, between two points indicates a factor. As seen in the graph, the test is one-factor. Accordingly, it can be said that the test is one-factor. Following the tetrachoric correlation analysis, test having one factor and consisting of 25 items was developed for use in the study. Since there is no factor, there is no need to look at the factor loadings of the test.

### **Interpretation of SAT Scores**

The developed SAT consists of 25 questions at the levels of application, analysis, evaluation, and creation. The test includes questions from 6 sub-dimensions: counting cubes, comparing cubes, visualizing from 2-D to 3-D and from 3-D to 2D, rotating 2 and 3 dimensional objects in the mind, completing surfaces and decomposing shapes in the mind. The test is multiple-choice and has 4 options in accordance with the 7th grade level. In the evaluation of the SAT, incorrect and blank answers were given "0" points and correct answers were given "1" point. The lowest score that can be obtained from the test is 0 and the highest score is 25. The increase in the score obtained from the test means the strength of the spatial ability and the decrease means the weakness of the spatial ability.

### **CONCLUSION and DISCUSSION**

The SAT is a test consisting of 25 multiple-choice items with four options to determine spatial ability, including counting cubes, comparing cubes, visualization from 2D to 3D and from 3D to 2D, mental rotation of 2D and 3D objects, surface completion (mental integration), and mental decomposition of shape. While developing the test, the acquisitions of the current mathematics course 7th grade "Views of Objects from Different Directions" unit were taken into consideration. When middle school level spatial ability tests are examined in the literature, we come across the MGMP Spatial Visualization Test developed by the instructors of the mathematics department of Michigan State University for second level students. This test was translated into Turkish by Turğut (2007) and an adaptation study was conducted. While the original test consists of 10 different types of 32 questions with 5 choices, the new test adapted to Turkish consists of 6 types of 31 questions. The pilot application of the test was conducted in 4 primary schools with a total of 382 students (128 sixth graders, 150 seventh graders and 104 eighth graders). As a result of the analysis of the collected data, the reliability coefficient was found to be 0.830. The test measures two subcomponents of spatial ability and the name of the Turkish version of the test was changed to MGMP Spatial Ability Test. The Spatial Visualization Test, which is also included in the literature and frequently used in the studies, was prepared for the project called "Middle Grades Mathematics Project" carried out in the United States for the second level of primary education, and used by Yıldız and Tüzün (2011) in their study. The test consists of 15 questions and each question has 5 answer choices. The test questions include questions about isometric views of structures made of unit cubes and their views from the right, left, front and back. The reliability of the test was found to be 0.971 (N=108) when calculated according to post-test scores. The "Mental Rotation Test" developed by Peters et al. (1995), which Yıldız and Tüzün (2011) used in their study, consists of a total of 24 questions with the same nature of each question. It contains questions about finding the new form of a shape made of unit cubes when it is rotated in different directions and at different angles. Each question has 4 choices. The reliability of the test was found to be 0.661 (N=108) when calculated according to the post-test scores. In their study, Dokumacı Sütçü and Oral (2019) developed a medium difficulty test consisting of 29 Items and high discrimination Items. The KR-20 internal consistency coefficient of the test consisting of two factors was found to be .77 for the first

factor consisting of 14 Items and .78 for the second factor consisting of 15 Items. For the overall test, this value was found to be .78. The pilot application of the test was conducted in five different secondary schools in the Central Districts of Diyarbakır Province in the 2016-2017 academic year. The application was carried out with 301 seventh grade students. The test was introduced to the literature as the "Spatial Visualization Test". The reliability coefficients of the Paper Folding and Surface Development Tests developed by Ekstrom, French, Harmon, and Derman (1976) and translated into Turkish by Delialioğlu (1996) were .84 and .82, respectively, while the reliability coefficients of the Spatial Visualization Test in Two-Dimensional Geometry developed by Olkun (2003) were .77. Unlike the other tests, the spatial visualization test developed by Yüksel and Bülbül (2014) for university students underwent Item analysis, validity-reliability analysis, and was found to have sufficient single-factor fit indices. The reliability coefficient was calculated as .84.

Although there are different definitions of spatial ability and its components and different measurement tools have been developed, the common point of these studies is the importance of spatial ability. The fact that the student groups studied in the studies were evaluated with the same test despite having different characteristics reveals this importance. Although this study was prepared by considering the 7th grade unit outcomes, the lack of a recent and comprehensive spatial ability test that can be used at the middle school level in the literature formed the basis of this study. The reliability coefficient of the SAT was calculated as .918.

#### **SUGGESTIONS**

The test developed in this study was designed to assess seventh grade students' spatial abilities using mathematical context. Researchers can develop new and updated tests for different grade levels. Researchers are also recommended to develop tests that address spatial abilities at different grade levels in different subjects in other courses. For teachers to benefit from achievement tests whose reliability and validity have been studied, these tests can be included in textbooks and published on the website of the Ministry of National Education General Directorate of Measurement, Evaluation and Measurement Services. Performance tests can include multiple-choice questions as well as open-ended questions depending on the performance and grade level. It is suggested that achievement tests consisting of open-ended questions can also be developed. Research can also be conducted to determine teacher and student knowledge about the achievement test development process.

#### **ETHICAL TEXT**

The practices carried out in the study were approved by Nevşehir Hacı Bektaş Veli University Rectorate Scientific Research and Publication Ethics Committee dated 27.03.2023 and decision number 2023.03.84. The necessary permissions were also obtained by applying to the Provincial Directorate of National Education. In this article, the journal writing rules, publication principles, research and publication ethics, journal ethical rules were followed. The responsibility belongs to the author for any violations that may arise regarding.

**Author(s) Contribution Rate:** The author's contribution to this article is 100%.

## REFERENCES

- Aksu, B. (2018). *Ölçme ve Değerlendirme*. Cinius Yayınları, 1. Baskı.
- Alias, M., Black, T. R., & Gray, D. E. (2002). Effect of instructions on spatial visualization ability in civil engineering students. *International Education Journal*, 3(1), 1-12.
- Başol, G. (2019). *Eğitimde Ölçme ve Değerlendirme*. Pegem Akademi, 6. Baskı.
- Burnett, S., Lane, D. (1980). Effects of Academic Instruction on Spatial Visualization, *Intelligence*, Sayı 4, 233-342.
- Büyüköztürk, Ş., (2018). *Sosyal bilimler için veri analizi el kitabı istatistik, araştırma deseni SPSS uygulamaları ve yorum*. Pegem A. yayıncılık.
- Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş. & Demirel, F. (2012). *Bilimsel araştırma yöntemleri*. Pegem Akademi, 15. Baskı.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Cizek, G. J. (1997). *Learning, Achievement and Assessment: Constructs at a Crossroads*. In G. D. Phye (Ed.), *Handbook of Classroom Assessment. Learning, Achievement and Adjustment*. California: Academic Press.
- Clements, D.H. (1998). *Geometric and spatial thinking in young children*. (ERIC Servis No. ED436232)
- Colletta, A.T. ve Chiappetta, E.L. (1989). *Science Introduction in The Middle and Secondary Schools* (2nd edn.). Ohio, USA: Merrill Publishing Company
- Contero, M., Naya, F., Company, P., Saorin, J.K., & Conesa, J.(2005). Improving visualization skills in engineering education. *IEEE Computer Graphics and Applications* (5), 24-31.
- Çepni, S. (2003). An Analysis of University Science Instructors' Examination Questions According to The Cognitive Levels. *Educational Sciences: Theory&Practice*, 3(1), 65-84.
- Delialioğlu, Ö. (1996). *Contribution of students' logical thinking ability, mathematical skills and spatial ability on achievement in secondary school physics*. Yayımlanmamış yüksek lisans tezi. Orta Doğu Teknik Üniversitesi.
- Doğan, N. (2019). *Eğitimde Ölçme ve Değerlendirme*. Pegem Akademi, 1. Baskı.
- Dokumacı Sütçü, N. & Oral, B. (2019). Uzamsal Görselleştirme Testinin Geliştirilmesi: Geçerlik ve Güvenirlik Çalışmaları. *Kastamonu Eğitim Dergisi*, 27 (3), 1179-1195. DOI: 10.24106/kefdergi.2826.
- Ekstrom, R. B., Dermen, D., & Harman, H. H. (1976). *Manual for kit of factor-referenced cognitive tests (Vol. 102)*. Princeton, NJ: Educational Testing Service.
- Fraenkel, J. K., & Wallen, N. E. (1996). *How to design and evaluate research in education (3. edition)*. New York: McGraw-Hill, Inc.
- Gelbal, S. ve Kelecioğlu, H. (2007). *H. Ü. Eğitim Fakültesi Dergisi (H. U. Journal of Education)*, 33, 135-145.
- Guay, R. B. (1976). *Purdue Spatial Visualization Test*. West Lafayette, Purdue Research Foundation.



- Gür, H. & Kobak Demir, M. (2017). Pergel Cetvel Kullanarak Temel Geometrik Çizimlerin Öğretmen Adaylarının Geometrik Düşünme Düzeylerine Ve Tutumlarına Etkisi. *Eğitimde Kuram ve Uygulama*, 13(1), 88-110.
- Kan, A. (2011). *Ölçme aracı geliştirme*. S. Tekindal (Ed.), Eğitimde ölçme ve değerlendirme (ss. 249-257). Pegem Akademi.
- Karasar, N. (2011). *Bilimsel araştırma yöntemleri*. Nobel Yayınları, 3. Baskı.
- Karip, E. (Ed.) (2012). *Ölçme ve değerlendirme*. Pegem Akademi, 5. Baskı.
- Kayhan, E.B. (2005). *Investigation of high school students' spatial ability*. Yayımlanmamış doktora tezi, ODTÜ, Ankara.
- Kehoe, J. (1995). Basic Item Analysis for Multiple-Choice Tests. (ERIC Servis No. ED398237)
- Kyllonen, P. C., Lohman, D. F., & Snow, R. E. (1984). Effects of aptitudes, strategy training, and task facets on spatial task performance. *Journal of Educational Psychology*, 76(1), 130.
- Küçükahmet, L. (2002). *Öğretimde planlama ve değerlendirme*. Nobel Yayın Dağıtım, 13. Baskı.
- Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of gender differences in spatial abilities: A meta- analysis. *Child Development*, 56, 1479-1498.
- Lohman, D. F. (1979). *Spatial ability: Individual differences in speed and level*. Aptitude Research Project, School of Education, Stanford University.
- McGee, M.G. (1979). *Human spatial abilities: Sources of sex differences*. Praeger.
- Odell, R.L. (1993). *Relationship Among Three Dimensional Laboratory Models, Spatial Visualization Ability, Gender and Earth Science Achievement*. Yayımlanmamış Doktora Tezi School of Education, Indiana University, Indiana.
- Okagaki, L.R., & Frensch, P.A. (1996). *Effects of video game playing on measures of spatial performance: Gender effects in late adolescents*. In P. Greenfield & R. Cocking (Eds.), *Interacting with video* (pp. 115-140) Norwood, NJ: Ablex Corporation.
- Olkun S. (2003). Making connections: Improving spatial abilities with engineering drawing activities. *International Journal of Mathematics Teaching and Learning*, 3(1), 1-10. <https://pdfs.semanticscholar.org/0328/1231a0411b83f44268d5c1145940067e8556.pdf> adresinden 13/09/2023 tarihinde erişilmiştir.
- Olkun, S., Altun, A. (2003). İlköğretim Öğrencilerinin Bilgisayar Deneyimleri ile Uzamsal Düşünme ve Geometri Başarıları Arasındaki ilişki, *Turkish Journal of Educational Technology*, 2 (4), Article 13.
- Özçelik, D. A. (2016). *Ölçme ve Değerlendirme*. Pegem Akademi, 5. Baskı.
- Özdemir, S. M., (2009). Eğitimde program değerlendirme ve Türkiye'de eğitim programlarını değerlendirme çalışmalarının incelenmesi. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 6(2), 126-149.
- Pellegrino, J.W., Alderton, D.L., Shute, V.J. (1984). Understanding Spatial Ability, *Educational Psychologist*, 19, 239-253.
- Seçer, İ. (2013). *SPSS ve LISREL ile pratik veri analizi: Analiz ve raporlaştırma*. Anı Yayıncılık.
- Tan, Ş., & Kayabaşı, Y. (2003). *Öğretimi Planlama ve Değerlendirme*. Anı Yayıncılık, 4. Baskı.
- Tekin, H. (1991). *Eğitimde Ölçme ve Değerlendirme*. Yargı Yayınları.

- Turğut, M. (2007). *İlköğretim II. kademedeki öğrencilerin uzamsal yeteneklerinin incelenmesi*. Yayınlanmamış yüksek lisans tezi, Dokuz Eylül Üniversitesi, İzmir.
- Turğut, M. & Yılmaz, S. (2012). İlköğretim 7. Ve 8. Sınıf Öğrencilerinin Uzamsal Yeteneklerinin İncelenmesi. Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi , (19), 69-79. <https://dergipark.org.tr/en/pub/zgefd/issue/47945/606597> adresinden 18/09/2023 tarihinde erişilmiştir.
- Turgut, M. F., & Baykul, Y. (2014). *Eğitimde ölçme ve değerlendirme*. Pegem Akademi, 6. Baskı.
- Winter, J. W., Lappan, G., Fitzgerald, W. & Shroyer, J. (1989). *Middle Grades Mathematics Project: Spatial Visualization*. NY: Addison-Wesley.
- Yıldız, B. ve Tüzün, H. (2011). Üç Boyutlu Sanal Ortam Ve Somut Materyal Kullanımının Uzamsal Yeteneğe Etkileri. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi (H. U. Journal of Education)* 41: 498-508.
- Yılmaz, H. (1998). *Eğitimde Ölçme ve Değerlendirme*. Mikro Yayınları.
- Yolcu, B. & Kurtuluş, A. (2010). 6. Sınıf Öğrencilerinin Uzamsal Görselleştirme Yeteneklerini Geliştirme Üzerine bir Çalışma. *İlköğretim Online*, 9 (1) , 256-274. <https://dergipark.org.tr/en/pub/ilkonline/issue/8596/106967> adresinden 18/09/2023 tarihinde erişilmiştir.
- Yüksel, N. S., & Bülbül, A. (2014). Uzamsal görselleştirme üzerine test geliştirme çalışması. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 8(2), 124-142.

## 7. SINIF ÖĞRENCİLERİNE YÖNELİK UZAMSAL YETENEK TESTİNİN GELİŞTİRİLMESİ

### Öz

Bu araştırmanın amacı, 7. sınıf öğrencilerine yönelik bir başarı testi geliştirmektir. Alan yazın incelendiğinde 7. sınıf düzeyinde uzamsal yeteneği ölçecek güncel bir başarı testinin bulunmaması bu çalışmanın ortaya çıkmasını sağlamıştır. Bu amaçla ilk olarak 7. sınıf "Cisimlerin Farklı Yönlerden Görünümleri" konusu ile ilgili kazanımlar incelenmiş, alan yazın taranmış ve 84 soru hazırlanmıştır. Uzman görüşleri alınmış ardından gerekli düzenlemeler yapılarak soru sayısı 64'e düşürülmüştür. Hazırlanan 64 soruluk testin geçerliği ve güvenilirliği incelenmiş, pilot uygulama için 2022-2023 eğitim-öğretim yılında 139 kişilik 7. sınıf öğrencilerine uygulanmıştır. TAP programı ile yapılan madde analizi sonucuna göre uzman görüşleri de alınarak Uzamsal Yetenek Başarı Testinin (UYBT) 25 soruluk son hali verilmiştir. Başarı testinin ortalama güçlük indeksi değeri pilot uygulamada 0.647 iken testin son halinde 0.540 bulunmuştur. Başarı testinin güvenilirlik analizi KR-20 değeri pilot uygulamada 0.848 iken testin son halinde 0.918 olarak bulunmuştur. Bu çalışma sonucunda 25 soruluk dört seçenekten oluşan çoktan seçmeli geçerliği ve güvenilirliği sağlanmış bir başarı testi geliştirilmiştir.

**Anahtar kelimeler:** Başarı testi, uzamsal yetenek, güvenilirlik, geçerlik, test geliştirme.

## GİRİŞ

Ölçme, bir bireyin veya öğrencinin belirli özelliklerini belirli yöntem ve kurallar ışığında sayısal olarak ifade etmek için gerekli araçların seçilmesi süreci olarak tanımlanmaktadır (Aksu, 2018). Diğer bir tanım ise ölçmenin, kişilerin veya nesnelerin özelliklerinin uygun araçlar kullanılarak gözlemlenmesi ve gözlemlerin sembollerle ifade edilmesi eylemi olduğudur. Özetle, bireyin belirli bir amaç için arzu edilen bir özelliğe ne ölçüde sahip olduğunun belirlenmesine ölçme denilmektedir (Cizek, 1997). Doğan' a (2019) göre ölçüm yapılmasının nedeni gözlemlendiğimiz şeylerin birçok özelliğinin değişmesi veya farklılaşmasıdır. Bir başka ifadeyle ölçme yapılmasının temelini değişimin oluşturduğunu ifade edilmektedir. Doğan'a (2019) göre bireyler arasındaki farklılıkların varlığı ölçüm sürecini tetiklemekte ve değişim ölçüme yol açmaktadır. Eğer değişim yoksa yani her şey sabitse ölçüm yapmaya yani ölçüm yapıldıktan sonra işlemin zaman içinde tekrarlanmasına gerek yoktur. Çünkü değişiklik olmadığı takdirde ölçüm sonucu aynı olacak ve ölçüm işlemine gerek kalmayacaktır.

Değerlendirme, ölçüm sonuçlarını aynı alandaki standartlarla karşılaştırarak değer yargılarına veya sonuçlara ulaşma süreci olarak tanımlanmaktadır (Yılmaz, 1986). Aksu'ya (2018) göre değerlendirme, nitel ve nicel olarak iki kısma ayrılan ve bireyin veya öğrencinin akademik gelişimine ilişkin karar vermek için gerekli tüm verileri toplama ve veriler ışığında geçmişe veya ileriye dönük tahminlerde bulunma sürecidir.

Davranış değişikliğinin yani öğrenmenin, kazanımların belirlenmesi için ölçme ve değerlendirme süreçlerinden yararlanmak gerekmektedir (Özdemir, 2009). Hiç şüphe yok ki eğitimin olduğu yerde ölçme ve değerlendirmeye de ihtiyaç duyulmaktadır. Kurul kararları, öğretmen görüşleri, sınavlar, port folyolar ve daha birçok araç aracılığı ile öğrenci başarısı belirlenmeye çalışılmaktadır (Başol, 2019). Ayrıca eğitimde öğrenci performansını ölçmek amacıyla açık uçlu sorular, çoktan seçmeli sorular, doğru-yanlış soruları, boşluk doldurma soruları ve eşleştirme soruları gibi ölçme araçları yaygın olarak kullanılmaktadır (Karip, 2012). Özçelik'e (2016) göre, okullarda ölçüm yapılmasının nedenleri şunlardır: derse hazır bulunuşluk düzeyini belirlemek, derste planlanan eylemlerin daha önce öğrenilip öğrenilmediğini belirlemek ve her öğrenme birimini ölçmek, her bir öğrenme ünitesi sonunda üniteye öğrenmeleri izleme, dersin belli dönemlerinin sonunda, bu dönemlerdeki etkileşimlerin ürünü olarak meydana gelen, dersin hedefleriyle tutarlı öğrenme düzeylerini belirlemektir.

Ölçme ve değerlendirme öğrencilerin yaşam boyu matematik ve geometri derslerinde de önemli bir rol oynamaktadır. Geometrinin genel olarak soyut kavramları içermesi çoğu öğrenciyi zora sokmakta ve geometriden uzaklaşmalarına neden olmaktadır (Gür, Kobak Demir, 2017). Uzamsal yetenekler geometri derslerinde, özellikle de birçok matematik konusunun öğretiminde önemli bir rol oynamaktadır. Smith (1998), bu yeteneklerin önemini “ Uzamsal zekâ olmadan dünyada var olmak zor olabilir. Bunun eksikliğinde şekillerin boyut ve konumlarındaki değişiklikleri göz önünde tutarak değişimlerini tahmin etmede veya verilen yönleriyle nesnelere arasındaki ilişkileri ve konumu ifade ederken zorlanabiliriz.” şeklinde vurgulamıştır. Uzamsal yetenek, nesnelere zihinde görselleştirme, hareket ettirme ve nesnelere şeklini zihinsel olarak manipüle etme yeteneği olarak tanımlanmaktadır (Lohman, 1979; Linn ve Petersen, 1985; Carroll, 1993). Başka bir tanımda ise uzamsal

yetenek, bir veya daha fazla parça ve bunların bileşenlerinden oluşan nesnelere üç boyutlu uzayda zihinsel olarak hareket ettirme veya görselleştirme yeteneği olarak tanımlanmaktadır. (Turğut, 2007).

1930' lardan günümüze uzamsal yeteneği ölçen araştırmalar yapılmaktadır. Guay (1976), bir nesnenin tek bir eksen etrafındaki dönüşünü ölçen testlerin, test kâğıdının basitçe döndürülmesiyle kolayca çözülebileceğini ifade etmiştir. Bu nedenle Guay (1976) "Purdue Uzamsal Görselleştirme Testi" adı verilen ve uzamsal yeteneklerle ilgili zihinsel işlemleri gerektiren bir test geliştirmiştir. Uzamsal yeteneği ölçmek için tasarlanan araçlar genellikle 2 boyutlu nesnelere 3 boyutlu versiyonlarını görselleştirmeyi (Kyllonen, Lohman ve Snow, 1984) ve bir veya daha fazla eksen etrafında döndürülmüş 3 boyutlu nesnelere görselleştirmeyi içeren sorular içermektedir (Winter, Lappan, Fitzgerald ve Schroyer, 1989). Eliot ve Smith (1983), Guay (1977) tarafından geliştirilen uzamsal görselleştirme testinin değiştirilmiş bir versiyonunu geliştirmiştir. Bu test 24 sorudan oluşmakta ve 3 boyutlu nesnelere hayali bir perspektiften görselleştirme yeteneğini ölçmektedir. Literatür incelendiğinde, uzamsal yeteneği ölçen testlerin genellikle birim küpler kullanılarak oluşturulan soruları içerdiği ortaya çıkmaktadır. Geometrik bir yapıyı zihinde döndürme becerisi gerektiren sorular (MGMP), bir geometrik yapının farklı yönlerden görünüşlerini çizibilme becerisi gerektiren sorular (Demirkaya, 2017), yapı içindeki birimler küp sayısını bulma soruları (Battista ve Clements, 1996; Ben Chaim ve diğerleri, 1985; MGMP; Yolcu, 2008), belirli yapılara yönelik birim küp ekleme ve çıkarma testlerine (MGMP) de literatürde rastlanmaktadır. Literatürde karşılaşılan uzamsal yeteneği ölçmeye yönelik testlerden bazıları aşağıda listelenmiştir;

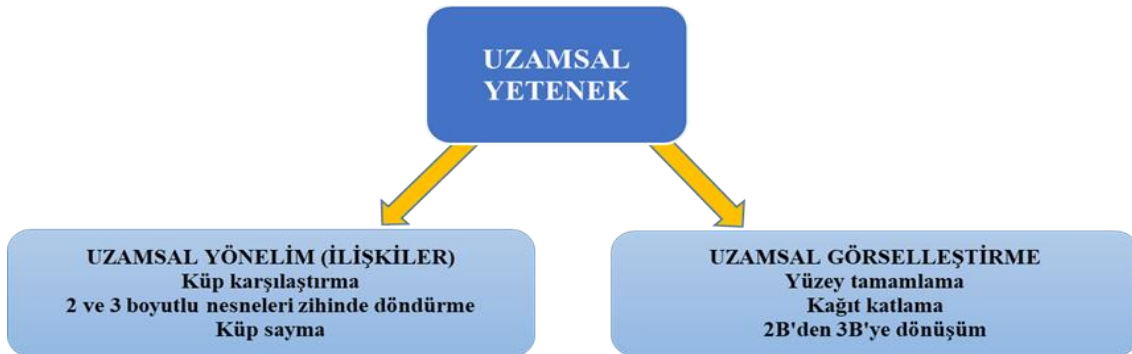
**Tablo 1.** Uzamsal Yetenek Ölçme Amaçlı Geliştirilen Testler

TEST ADI	TESTİ GELİŞTİREN ARAŞTIRMACI(LAR)
Zihinde Kesme Testi	CEEB (1939)
Kâğıt Katlama Testi	French ve diğerleri (1963)
Uzamsal Görselleştirme Testi	Lappan ve diğerleri (1983)
Minnesota Kâğıt Kaplama Testi	Likert (1970)
Purdue Uzamsal Görselleştirme Testi	Guay (1977)
Kesti Yüzey Testi	Cohen ve Hegarty (2012)
Uzamsal Görselleştirme Testi	Yüksel ve Bülbül (2014)
Kart Çevirme ve Küp Karşılaştırma Testi	Ekstrom ve diğerleri (1976)
Zihinde Döndürme Testi	Vandenberg ve Kuse (1978)
Wheatley Uzamsal Yetenek Testi	Wheatley (1978)
Guildford-Zimmerman Uzamsal Yönelim Testi	Guildford-Zimmerman (1948)
Santa Barbara Yön Algısı Ölçeği	Hegarty ve diğerleri (2002)
Revize Edilmiş Purdue Uzamsal Görselleştirme Testi	Yoon Yoon (2011)

Uzamsal yeteneği ölçmek için tasarlanan testlerde uzamsal yeteneğin alt boyutları konusunda tam bir fikir birliği yoktur. Uzamsal yeteneğin alt boyutlarını; McGee (1979), Lohman (1979), Clements (1998) ve Kayhan (2005) uzamsal görselleştirme ve uzamsal yönelim olarak; Linn ve Petersen (1985) ile Okagaki ve Frensch (1996) uzamsal görselleştirme, uzamsal algı ve zihinde döndürme olarak; Olkun ve Altun (2003) ve Turğut (2007) ise uzamsal yönelim (ilişkiler) ve uzamsal görselleştirme olarak; Contero (2005) uzamsal görselleştirme, uzamsal yönelim ve uzamsal ilişkiler olarak ayırmıştır. Literatürde uzamsal yeteneğin yaygın olarak kabul edilen boyutları

uzamsal ilişkiler ve uzamsal görselleştirme olarak kendini göstermektedir. Uzamsal yeteneklerin farklı isimlerle alt boyutları olmasına rağmen, uzamsal yeteneklerin kapsadığı işlemler genel olarak nesnelerin uzayda döndürülmesi ve bunların iki veya üç boyutlu olarak zihinde canlandırılmasını içermektedir. Bunun yanında söz konusu işlemlerin hızlı yapılması da önem arz etmektedir (Olkun, 2003). Bu nedenle uzamsal yeteneği ölçen testlerin zaman sınırlamaları bulunmaktadır.

Olkun ve Altun'a (2003) göre uzamsal ilişkiler; öğrencinin iki boyutlu ve üç boyutlu geometrik şekilleri bir bütün olarak zihinsel olarak döndürebilmesi ve bunları farklı konumlarda algılayabilmesidir. Ayrıca bu testler hızlı ve doğru karar verilmesini gerektiren testler olarak tanımlanmaktadır. Uzamsal görselleştirmeyi ise; bir veya daha fazla parça ve bu parçaların görüntülerinden oluşan, üç boyutlu uzayda nesnelerin iki ve üç boyutlu hareketi sonucu ortaya çıkan yeni durumları görselleştirme yeteneği olarak tanımlamışlardır. Bu beceriyi ölçen standart testlerdeki maddeler incelendiğinde hareketli parçalardan oluşan karmaşık şekiller ve/veya zihinde katlama ya da zihinsel bütünleme yoluyla 2-B'den 3-B'ye dönüştürme gibi zihinsel eylemleri gerektirdiği görülmektedir (Pellegrino ve diğerleri, 1984'den akt. Olkun ve Altun, 2003:2). Bu testler, uzamsal ilişkilerin aksine, giderek daha karmaşık hale gelen nesnelerin doğruluğuna hızdan daha çok odaklanmaktadır. Literatürde kabul edilen uzamsal yeteneğin alt boyutlarının içerikleri Şekil 1'de gösterilmektedir.



**Şekil 1.** Uzamsal Yetenek Alt Boyutları

Öğrenci performansının ölçülmesi, tüm öğrenci performansını kapsayan, iyi hazırlanmış bir ölçüm aracını gerektirmektedir. Öğrencilerin bilişsel düzeyleri, Benjamin Bloom'un (1956) geliştirdiği ve kendi adını verdiği Bloom Taksonomisi kullanılarak geliştirilen araçlar ile ölçülebilmektedir. Bloom taksonomisi ile öğrencilerin düşünme becerilerini ölçmeye yönelik sorular hazırlanabileceğini ve bunun sadece bilişsel alanda değil duyuşsal ve psiko-motor alanlarda da yapılabileceğini öngörmektedir (Çepni, 2003). Bilişsel alanların kategorileri, kuru ve ezberlenmiş bilginin hatırlanmasıyla başlayıp, kavrama, uygulama, analiz, sentez ve değerlendirmeye kadar uzanır. Bilgi aşamasından değerlendirme aşamasına geçtikçe davranışlar daha karmaşık hale gelir, öğrenilmesi ve geliştirilmesi zorlaşır (Tekin, 1991).

Bloom' un önerdiği taksonomiye göre öğrencilerin bilişsel düzeyleri şu şekilde ölçülebilmektedir:

Düşük bilişsel seviyeler (DBS) “bilgi”, “anlama” ve “uygulama” basamaklarını içerirken, daha yüksek bilişsel seviyeler (YBS) “analiz” ve “sentez” ve “değerlendirme” basamaklarını içermektedir (Colletta ve Chiappetta, 1989). Yenilenmiş Bloom Taksonomisinde ise bu basamaklar “Hatırlama, Anlama, Uygulama, Analiz Etme, Değerlendirme ve Yaratma” olarak değiştirilmiş ve güncellenmiştir.

2018 yılında yenilenen ve halen geçerli ve uygulanan yeni öğretim programı yapılandırmacı yaklaşımı temel almakta ve bilginin öğrenciler tarafından yapılandırıldığını kabul etmektedir. Öğretim programı değişikliği aynı zamanda ders içeriklerinde, öğretim yöntemlerinde, kullanılan araç ve gereçlerde, ölçme ve değerlendirme yöntemlerinde de değişiklikleri beraberinde getirmiştir (Gelbal ve Kelecioğlu, 2007). Uzamsal yetenek ile doğrudan ilişkili olan “Cisimlerin Farklı Yönlerden Görünümleri” ünitesi 7. sınıf düzeyinde matematik dersinin son konulardan biri olarak müfredatta yerini almaktadır. Konunun son ünitelerde yer alması, öğrencilerin yazılı sınavlarının bu konudan önce bitmesi, okullarda yapılan deneme sınavlarının da yine bu üniteden önce bitirilmesi, bu ünitenin ciddi şekilde değerlendirilmesinin ders öğretmenleri ile yapılan görüşmelerde çok zor hatta imkânsız olduğunu göstermiştir. Sürekli yenilenen ve gelişen eğitim programlarının da getirisi olarak ortaokul düzeyinde yeni ve güncel bir uzamsal yetenek ölçme aracına ihtiyaç olduğu yapılan alan yazın incelemesinde de açığa çıkmıştır. Bu yüzden geçerli-güvenilir özgün bir ölçme aracının literatürde bulunmaması, akademik çalışmalar yapan araştırmacıların ve öğretmenlerin ihtiyaçlarına cevap verebilecek şekilde ortaokul düzeyinde geçerliği ve güvenilirliği kanıtlanmış bir uzamsal yetenek başarı testinin geliştirilmesi çalışmanın amacıdır.

## **YÖNTEM**

### **Çalışma Deseni**

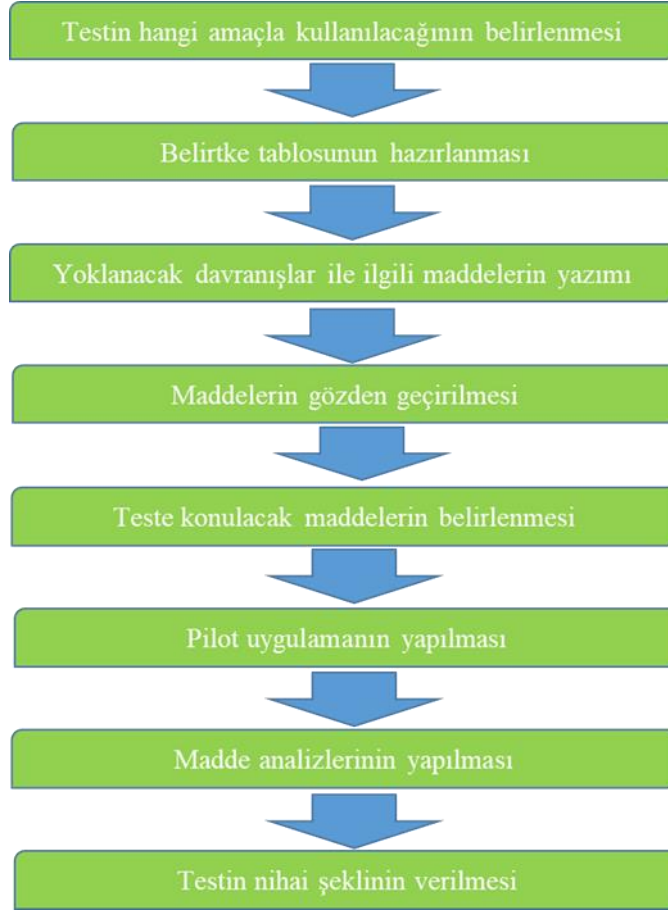
Bu çalışma ortaokul düzeyinde “Uzamsal Yetenek Testi” (UYT) geliştirilmeyi amaçlamış ve nicel araştırma yöntemlerinden biri olan tarama modeli kullanılarak gerçekleştirilmiştir. Bu model, geçmişte veya halen var olan bir durumun var olduğu şekli ile betimlenmesi, bir konuya ya da olaya ilişkin bireylerin görüş, yetenek ve beceri gibi niteliklerinin ortaya çıkarılması amaçlamaktadır (Fraenkel & Wallen, 1996; Karasar, 2011). Öğrencilerin çoktan seçmeli sorular aracılığıyla “uzamsal yetenek düzeyleri” belirlenmek istendiği için bu model tercih edilmiştir. Çalışmada, testin ölçtüğü özelliği ne kadar duyarlı ölçebileceği, sonuçların hatalardan ne kadar temizlendiğinin belirlenmesi için güvenlik analizi; testin ölçülmek istenen özelliği başka özellikler ile karıştırmadan ölçebilme seviyesini belirlemek için geçerlik analizi yapılmıştır (Küçükahmet, 2002).

### **Evren ve Örneklem**

Bu çalışmanın evrenini 2022-2023 eğitim-öğretim yılında Nevşehir ili Merkez ilçesinde öğrenim gören 7. sınıf öğrencileri oluşturmaktadır. Bu çalışmada örneklem belirlenirken kolay ulaşılabilir örneklem belirleme yöntemi ile Nevşehir ili Merkez ilçesine bağlı bir devlet ortaokulunda 7. sınıfta öğrenim gören 139 öğrenci ile çalışılmıştır. Örneklem oluşturulmasında hem hızlı hem de kolay olmasından dolayı kolay ulaşılabilir örneklem yöntemi tercih edilmiştir (Büyüköztürk ve diğerleri, 2012).

### Test Geliştirme Süreci

Bu çalışmada UYT' nin geliştirme sürecinde uygulanan adımlar Şekil 2'de verilmiştir (Turgut ve Baykul, 2014).



**Şekil 2.** UYT Geliştirme Adımları (Turgut ve Baykul, 2014).

Araştırmada öncelikle testin amacı belirlenmiştir. Bu çalışmadaki testin geliştirilmesiyle, ünite işlenmeden önce öğrencilerin uzamsal yetenek düzeylerini belirlemek, üniteyi işlendikten sonra öğrencilerin uzamsal yetenek düzeylerini ve öğrenme eksikliklerini belirlemek, öğretmenlerin ve araştırmacıların testten yararlanmalarını sağlamak, bu alanda çalışacak araştırmacılara yarar sağlamak ve alan yazına katkı sağlamak amaçlanmıştır.

Testin amacı belirlendikten sonra belirtke tablosu hazırlanmıştır. Bu çalışmada 7. sınıf matematik dersi "Cisimlerin Farklı Yönlerden Görünümleri" ünitesi kazanımlarını da kapsayacak şekilde uzamsal yetenek alt boyutlarına uygun literatür araştırması sonucu elde edilen veriler ile sorular hazırlanmaya çalışılmıştır. 2009 yılında MEB müfredatında ünite 6., 7., ve 8. sınıf müfredatında parça parça var iken 2018' de yenilenen ve halen kullanılan müfredatta sadece 7. sınıf ünitesi olarak bulunmaktadır. Aşağıda Tablo 2'de 2009 yılı ünite kazanımlarına ve Tablo 3'de 2018 yılı ünitenin kazanımlarına yer verilmiştir.



**Tablo 2.** 2009 Yılı MEB Müfredatında Cisimlerin Farklı Yönlerden Görünümleri ile İlgili Kazanımları ve Sınıf Düzeyleri

SINIF	2009 YILI MÜFREDATI KAZANIMLARI
6. SINIF	Eş küplerle oluşturulmuş yapıların farklı yönlerden görünümünü çizer.
7. SINIF	Yüzlerinin farklı yönlerden görünümüne ait çizimleri verilen yapıları, birim küplerle oluşturur ve izometrik kâğıda çizer.
8. SINIF	Çizimleri verilen yapıları çok küplülerle oluşturur, çok küplülerle oluşturulan yapıların görünümünü çizer.

**Tablo 3.** 2018 Yılı MEB Müfredatında Cisimlerin Farklı Yönlerden Görünümleri ile İlgili Kazanımları ve Sınıf Düzeyi

SINIF	2018 YILI MÜFREDATI KAZANIMLARI
	<b>M.7.3.4. Cisimlerin Farklı Yönlerden Görünümleri</b>
	M.7.3.4.1. Üç boyutlu cisimlerin farklı yönlerden iki boyutlu görünümünü çizer.
	a) Eş küplerden oluşturulmuş yapılar ve bilinen geometrik cisimler kullanılır. Çizim için uygun kareli kâğıtlar kullanılır. Yapıların farklı yönlerden görünümünün ilişkilendirilmesi istenir (ön-arka ve sağ-sol görüntülerinin simetrik olması gibi).
7. SINIF	b) Uygun bilgi ve iletişim teknolojileriyle etkileşimli çalışmalara yer verilebilir.
	M.7.3.4.2. Farklı yönlerden görünümüne ilişkin çizimleri verilen yapıları oluşturur.
	a) Eş küplerden oluşturulmuş yapılar ve bilinen geometrik cisimler kullanılır. Eş küplerle oluşan yapıları çizmek için izometrik kâğıt kullanılabilir.
	b) Uygun bilgi ve iletişim teknolojileriyle etkileşimli çalışmalara yer verilebilir.

2018 yılı müfredatında cisimlerin farklı yönlerden görünümü konusu 2 kazanım ile 5 ders saati süre içinde öğretilmektedir. Bu kazanımları kapsayacak şekilde uzamsal yetenek alt boyutlarına yönelik hazırlanan maddeler “Yenilenmiş Bloom Taksonomisi” ne göre sınıflandırılmış ve belirtke tablosu oluşturulmuştur. UYT’ nin son haline ait belirtke tablosu Tablo 4’de sunulmuştur.

**Tablo 4.** Nihai Test İçin Belirtke Tablosu

KONULAR	BİLİŞSEL ALAN			
	Uygulama	Analiz Etme	Değerlendirme	Yaratma
Küp sayma	4, 14, 15, 16, 17, 21	1, 4, 5, 10, 11, 12, 13, 23	1, 2, 3, 5, 6, 7, 8, 9, 10, 12, 18, 19, 20	22, 24, 25
Küp Karşılaştırma				22, 24, 25
2B’ den 3B’ ye ve 3B’ den 2B’ ye Görselleştirme	14, 15, 16, 17	13	18, 19, 20	
2B ve 3B nesnelere Zihinde Döndürme	4, 14, 15, 16, 17	1, 4, 5, 10, 12, 23	1, 2, 5, 6, 7, 8, 9, 10, 12, 18, 19, 20	22, 24, 25
Yüzey Tamamlama (Zihinde Bütünleme)	4	4, 11	3	24, 25
Şekli Zihinde Ayırıştırma	14, 15, 16, 17, 21	1, 5, 10, 11, 12, 13	1, 2, 3, 5, 6, 7, 8, 9, 10, 12, 18, 19, 20	22, 24, 25

Tablo 4 incelendiğinde, soruların hatırlama ve anlama düzeyinde hiç yer almadığı, bir sorunun birden çok alt boyutta ve birçok düzeyde olduğu da görülmektedir.

Yoklanacak davranışlar ile ilgili maddelerin yazılması esnasında; sorular hazırlanmadan önce matematik dersi öğretim programı “Cisimlerin Farklı Yönlerden Görünümleri” ünitesine ait kazanımlar incelenip analiz edilmiştir.

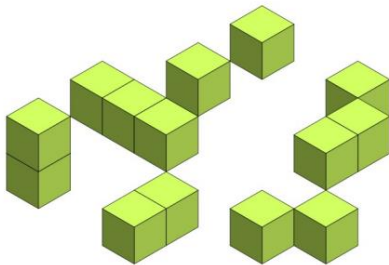
Daha sonra 7. sınıf matematik ders kitabı, yardımcı ders kaynakları ve alan yazında yer alan önceden yapılmış uzamsal yeteneğe yönelik testler incelenmiştir. Yapılan incelemeler sonucunda hiçbir soru aynı şekilde alınmayarak, soruların tamamı bizzat araştırmacı tarafından hazırlanmıştır. Her alt boyutun bulunduğu ve kazanımlara uygun 84 sorudan oluşan madde havuzu oluşturulmuştur.

Maddelerin gözden geçirilmesi yani testin geçerliği ve teste konulacak maddelerin belirlenmesi aşamasında kapsam geçerliğine bakmak için belirtke tablosu hazırlanmış ve uzman görüşü alınmıştır. Madde havuzunda yer alan sorularla ilgili Matematik Eğitimi alanında çalışan 3 öğretim üyesi, 1 araştırma görevlisi, 1 dil uzmanı ve 1 eğitim bilimleri uzmanı ile Milli Eğitim'e bağlı okullarda çalışan 2 matematik öğretmeninden görüş alınmıştır. Uzman görüşü sonrasında madde havuzundaki soru sayısının 64'e düşürülmesi kararı alınmıştır. Uzman görüşleri için aşağıdaki Tablo 5 hazırlanmış ve 84 soru için de uzmanlardan değerlendirmede bulunmaları istenmiştir.

**Tablo 5.** Uzman Soru Değerlendirme Tablosu

SORU NUMARASI	Uygun	Uygun Değil	Düzeltilme Gerekli	Düzeltilme Önerisi
---------------	-------	-------------	--------------------	--------------------

Uzman görüşleri sonucunda aynı özellikte olan veya aynı kazanımı ölçmeye yönelik sorular madde havuzundan çıkarılarak 64 soruluk test elde edilmiştir. 20 soru havuzdan çıkartılmış bazı sorular da ise düzeltmeler yapılmıştır. Uzmanlar, hazırlanan soruların ilgili kazanımlara uygun olup olmadığı, belirtilen bilişsel alan seviyesine uygunluğu, soruların öğrenci seviyesine uygunluğu ve anlaşılabilirliği, dilbilgisi kurallarına uygunluğu vb. konularda görüş bildirmişlerdir. Örnek olarak Şekil 3'de verilen soru maddesindeki ifade uzman dönütü olarak parantez içindeki "Şekilleri birim küplere ayırarak işlem yapmanız istenmektedir" ifadesi eklenerek revize edilmiştir.



Yalnızca yandaki şekilde verilen birim küpler kullanılarak en büyük hacimli bir küp elde etmek istenirse kaç adet birim küp artar? (Şekilleri birim küplere ayırarak işlem yapmanız istenmektedir.)

- A) 12
- B) 10
- C) 8
- D) 6

**Şekil 3.** Pilot uygulama örnek revize edilmiş test maddesi.

Verilerin toplanmasında öğrenciler not kaygısı gütmeden soruları gönüllü olarak yanıtlamaları sağlanmıştır. Böylece denek ön yargısının önüne geçildiği varsayılmıştır. Öğrencilere test cevaplama süresi olarak 40+40 dakika yani 2 ders saati verilmiştir. Test öğrencilere 2 bölüm şeklinde uygulanmıştır. Uzamsal yeteneğin ölçülmesi için uygulanan teste bilişsel stratejiler kullanarak çözüme ulaşılamaması için uzman görüşlerine göre 80 dakika zaman sınırı getirilmiştir (Linn ve Petersen, 1985).

Pilot uygulamanın yapılması adımında öncelikle soruların anlaşılır olup olmadığını ve soruların cevaplanması için verilen sürenin yeterli olup olmadığını test etmek için 9 yedinci sınıf öğrencisine uzamsal yetenek testi ön pilot uygulama olarak uygulanmıştır. Gelen dönütler doğrultusunda soruların öğrenciler tarafından anlaşıldığı ve verilen sürenin testi cevaplamak için yeterli olduğu görülmüştür. Ardından pilot uygulama, çoktan seçmeli 64 sorudan oluşan UYT, 2022-2023 eğitim-öğretim yılında Nevşehir ili, Merkez ilçesinde öğrenim gören 139 yedinci sınıf öğrencisine uygulanmıştır. Uygulama için Nevşehir Hacı Bektaş Veli Üniversitesi Rektörlüğü Bilimsel Araştırmalar ve Yayın Etik Kurulu'ndan 27.03.2023 tarih ve 2023.03.84. karar sayılı izin alınmıştır. Gereken izinler aynı zamanda İl Milli Eğitim Müdürlüğü'nden de başvuru yapılarak alınmıştır.

Madde analizlerinin yapılması ve teste son şeklinin verilmesi aşamasında; testin güvenilirlik analizi, KR-20 iç tutarlılık katsayısına bakılarak hesaplanmıştır. 64 sorudan oluşan UYT puanlanırken her doğru cevaba "1", her yanlış cevaba "0" puan verilerek değerlendirmeler toplam 64 puan üzerinden yapılmıştır. Toplam örneklem büyüklüğü 139 olduğundan alt grup üst grup şeklinde gruplandırma yapılmamış tüm öğrencilerin sınav sonuçları üzerinden değerlendirme yapılmıştır.

## BULGULAR

### UYT Madde Analizine İlişkin Bulgular

Testin 64 soruluk pilot halinin madde güçlük ve madde ayırt edicilik analizleri TAP Madde ve Test Analizi Programında yapılmıştır. UTY' nin pilot uygulamasına ait madde analizi sonuçları Tablo 6'da verilmiştir.

**Tablo 6.** Uzamsal Yetenek Testi Pilot Uygulama Neticesinde Elde Edilen Madde Analizi Değerleri

Soru	Madde Güçlük İndeksi (Pj)	Madde Ayırtedicilik İndeksi (rjx)	Soru	Madde Güçlük İndeksi (Pj)	Madde Ayırtedicilik İndeksi (rjx)
Madde 1*	0.87	0.28	Madde 33*	0.87	0.41
Madde 2*	0.98	0.05	Madde 34*	0.91	0.28
Madde 3*	0.93	0.13	Madde 35	0.77	0.42
Madde 4*	0.86	0.21	Madde 36*	0.78	0.42
Madde 5*	0.88	0.13	Madde 37*	0.81	0.29
Madde 6*	0.93	0.03	Madde 38*	0.90	0.33
Madde 7*	0.81	0.41	Madde 39*	0.86	0.34
Madde 8*	0.10	0.12	Madde 40*	0.91	0.26
Madde 9	0.53	0.41	Madde 41	0.52	0.55
Madde 10*	0.74	0.44	Madde 42*	0.76	0.64
Madde 11	0.38	0.49	Madde 43	0.61	0.60
Madde 12*	0.47	0.46	Madde 44*	0.60	0.60
Madde 13*	0.41	0.41	Madde 45	0.62	0.50
Madde 14	0.42	0.53	Madde 46*	0.76	0.37

Madde 15*	0.81	0.23	Madde 47	0.58	0.53
Madde 16*	0.88	0.11	Madde 48*	0.63	0.52
Madde 17	0.63	0.57	Madde 49	0.45	0.61
Madde 18*	0.60	0.30	Madde 50	0.53	0.68
Madde 19*	0.42	0.34	Madde 51*	0.50	0.34
Madde 20	0.49	0.60	Madde 52	0.37	0.44
Madde 21*	0.61	0.57	Madde 53*	0.80	0.51
Madde 22	0.54	0.60	Madde 54	0.67	0.42
Madde 23*	0.63	0.80	Madde 55*	0.83	0.46
Madde 24	0.55	0.68	Madde 56*	0.81	0.49
Madde 25	0.42	0.52	Madde 57	0.61	0.82
Madde 26*	0.68	0.45	Madde 58*	0.71	0.32
Madde 27*	0.71	0.49	Madde 59*	0.34	0.18
Madde 28	0.65	0.50	Madde 60	0.64	0.40
Madde 29	0.65	0.62	Madde 61	0.62	0.55
Madde 30*	0.65	0.57	Madde 62*	0.59	0.18
Madde 31	0.38	0.47	Madde 63	0.53	0.58
Madde 32	0.34	0.37	Madde 64*	0.58	0.55

\*İşaretleli maddeler testten çıkarılmıştır.

Test maddeleri üzerinde yapılan madde güçlük analizi, maddenin uygulanan grup için ne kadar kolay ya da zor olduğu ile ilgili bilgi vermektedir. Bu değer 1.00'a yaklaştıkça doğru cevaplama sayısının arttığını ve maddenin kolaylaştığını, 0.00'a yaklaştıkça yanlış cevaplama sayısının arttığını ve maddenin zorlaştığını ifade etmektedir (Kan, 2011). Tablo 6'da verilenlere göre sonuç olarak elde edilen UYT' de yer alan maddelerin güçlük indekslerinin 0.34 ile 0.77 arasında değer alarak testin ortalama güçlük indeksinin 0.647 çıktığı görülmektedir. Bu bilgilere göre testte zor ve kolay maddelerin her düzeyinden soru yer aldığı ve testin ortalama güçlüğü ise orta düzeyde olduğu söylenebilir.

Test maddeleri üzerinde yapılan madde ayırt edicilik değeri, maddelerinin ölçülmek istenen özelliği ne derecede ölçtüğünün belirlenmesini sağlamaktadır. Bu değer +1'e yaklaştıkça maddenin testin ölçmeyi amaçladığı özelliği daha iyi ölçtüğü, 0'a yaklaştıkça maddenin testin ölçmeyi amaçladığı özelliği ölçmede yetersiz kaldığı ifade edilmektedir. Bu değerinin eksi çıkması halinde maddenin amaçlanandan başka bir özelliği ölçtüğü düşünülebilir (Kan, 2011). Tablo 6'da verilenlere göre sonuç olarak elde edilen UYT' de yer alan maddelerin ayırt edicilik indekslerinin 0.33 ile 0.82 arasında değer alarak testin ortalama ayırt edicilik indeksinin 0.429 çıktığı görülmektedir. Bu sonuca göre testin ayırt ediciliğinin yeterli düzeyde olduğu söylenebilir. UYT' nin pilot ve nihai haline ilişkin TAP analizi sonuçları Tablo 7'de verilmiştir.

**Tablo 7.** Uzamsal Yetenek Testinin Pilot ve Nihai Uygulamalarına İlişkin TAP Analizi Sonuçları

İstatistikler	Pilot Test	Nihai Test
Madde sayısı	64	25
Uygulanan Kişi Sayısı	139	139
Ortalama	41.432	13.489
Standart Sapma	11.396	5.638
Varyans	129.871	31.789
Minimum	11	2
Maksimum	60	24
Çarpıklık	-0.447	-0.043
Basıklık	-0.434	-0.906
Ortalama madde güçlüğü	0.647	0.540
Ortalama madde ayırt ediciliği	0.429	0.527

Tablo 7 incelendiğinde; testin pilot halinin ortalaması 41.432, standart sapması 11.396, varyansı 129.871, 64 sorudan oluşan testten alınan minimum puan 11, maksimum puan 60, çarpıklık değeri -0.447, basıklık değeri -0.434, testin ortalama güçlük indeksi 0.647, testin ayıt edicilik indeksinin ise 0.429 olduğu görülmektedir. Testin nihai halinin ise ortalaması 13.489, standart sapması 5.638, varyansı 31.789, 25 sorudan oluşan nihai testten alınan minimum puan 2, alınan maksimum puan 24, çarpıklık değeri -0.043, basıklık değeri -0.906, testin ortalama güçlük indeksi 0.540, testin ayırt edicilik indeksinin ise 0.527 olduğu görülmektedir.

Yapılan analiz ve değerlendirmeler sonrasında madde ayırt edicilik indeksi. 30'un altında kalan sorular ile ortalama güçlüğü. 20'den küçük ve 0.80'den büyük sorular testten çıkarılmıştır. Testlerin kapsam geçerliğinin sağlanması için hem testlerle ölçülecek kritik kazanımların belirlenmesi, hem de test maddelerinin yazılması aşamalarında uzman görüşleri alınmıştır. Testlerin kapsam geçerliği analizinde her bir test ile ölçülen kazanımları ve bu kazanımların Bloom taksonomisindeki bilişsel düzeylerini gösteren belirtke tabloları hazırlanmıştır. Benzer boyutları ölçen sorular da uzman görüşleri doğrultusunda çıkarılarak 25 soruluk nihai UYT elde edilmiştir.

#### UYT Güvenirliğine İlişkin Bulgular

Güvenirlik, bireylerin test maddelerine verdikleri cevaplar arasındaki ilişki güvenilirlik olarak tanımlanmaktadır (Büyüköztürk, 2018). Turgut ve Baykul'a (2014) göre bir ölçme aracının güvenilirliğini tek uygulama ile belirlemenin bir yöntemi olarak ölçme aracındaki maddelerin birbiri ile tutarlılığını ortaya çıkaran Kuder-Richardson-20 (KR-20) testi uygulanabilir. Başarı testlerinin geliştirildiği pek çok çalışmada güvenirlığın KR-20 güvenilirlik katsayısı hesaplanarak rapor edildiği görülmektedir. Test maddelerine verilecek cevapların doğru/yanlış, evet/hayır gibi iki seçeneği olması durumunda KR-20 katsayısının kullanılmasının uygun olacağı ve bu katsayının .70 ve üzerinde olmasının test puanlarının güvenirligi için yeterli olacağı Büyüköztürk (2017) tarafından belirtilmiştir. Kehoe (1995) KR-20 değerinin 10-15 civarında maddeden oluşan çoktan seçmeli testler için 0.50; 50 ve üzeri maddeden oluşan testler için ise en az 0.80 olması gerektiğini belirtmektedir (akt. Tan, Kayabaşı ve Erdoğan, 2003). Başarı testinin pilot ve nihai ölçümlerine ilişkin KR-20 değerleri Tablo 8'de verilmiştir.

**Tablo 8.** Uzamsal Yetenek Başarı Testi Güvenirlik Analizi Sonuçları

TEST TİPİ	SORU SAYISI	KR-20 KATSAYISI
PİLOT TEST	64	.848
NİHAİ TEST	25	.918

139 kişi ile yapılan pilot uygulama sonucunda UYT' nin 64 soruluk ön uygulamasının KR-20 güvenilirlik katsayısı .848 olarak bulunmuştur. 39 sorunun çıkarılması ile elde edilen 25 soruluk testin son halinin KR 20 güvenilirlik katsayısı ise .918 olarak hesaplanmıştır. Bu sonuçlara göre geliştirilen UYT' nin güvenirliginin yeterli düzeyde olduğu söylenebilir.

**UYT Yapı Geçerliliği**

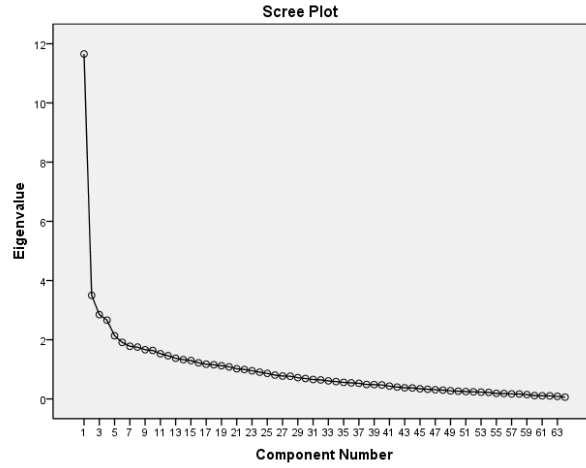
Ölçme araçlarının açıklayıcı faktör analizlerini yapabilmek için Kaiser-Meyer-Olkin (KMO) Değeri ve Barlett's testi sonuçları önem arz etmektedir. KMO değerinin 0.7 ve üzerinde olması iyi, 0.5-0.7 arasında olması ise yeterli ilişkiyi sağlayacak örneklem anlamına gelmektedir. Barlett's testi sonuçlarının anlamlı olması ise örneklemin normal dağılımı hakkında bilgi vermektedir (Seçer, 2013). Bu değerler ışığında çalışma yapılan 64 soruluk test için 139 kişilik örneklem büyüklüğünün yeterli olduğu söylenebilmektedir. Analiz sonucunda KMO değeri .706 olarak, Barlett's testi ise .000 ( $p < .05$ ) olarak anlamlı bulunmuştur. KMO ve Barlett's testi sonuçları dağılımın normal olduğunu göstermekte ve faktör analizine devam edilebilirliğe izin vermektedir.

Seçer' e (2013) göre yapılan test sonucunda faktör sayısını belirlemek için her bir faktörün öz değerinin "1" in üzerinde olması ve toplam varyansının en az %5' ini açıklaması gerekmektedir. Tablo 9'da testin toplam varyans değerleri verilmiştir.

**Tablo 9.** Uzamsal Yetenek Testi Toplam Varyans Değerleri

Faktör	Öz Değer	Varyansın % si	Toplam %
1	11.655	18.211	18.211
2	3.497	5.464	23.675
3	2.850	4.452	28.128
4	2.662	4.160	32.287
5	2.134	3.335	35.622
6	1.910	2.984	38.607
7	1.779	2.780	41.387
...	...	...	...
...	...	...	...
63	.087	.136	99.904
64	.061	.096	100.000

Tablo 9 incelendiğinde öz değerleri "1" in üzerinde olan ve toplam varyansın %5' ini açıklayan iki faktörün olduğu görülse de ölçme araçları geliştirme çalışmalarında sadece toplam varyans tablosuna bakmak aracın kaç faktörlü olduğuna karar vermek için yeteli değildir. "ScreePlot" grafiği de aracın faktör yapısı hakkında bilgi vermektedir (Seçer, 2013). Şekil 4'de testin ScreePlot grafiği verilmiştir.



Şekil 4. Uzamsal Yetenek Testi "ScreePlot" Grafiği

"ScreePlot" grafiğinde iki nokta arası bir faktöre işaret etmektedir. Grafikte görüldüğü üzere test tek faktörlüdür. Buna göre testin tek faktörlü olduğu söylenebilir. Tetrakorik korelasyon analizi sonucunda, tek faktörlü 25 maddeden oluşan bir test çalışmada kullanılmak üzere geliştirilmiştir. Faktör olmadığı için testin faktör yüklerine bakmaya gerek kalmamıştır.

#### UYT Puanlarının Yorumlanması

Geliştirilen UYT, uygulama, analiz etme, değerlendirme, yaratma düzeylerinden olan 25 sorudan oluşmaktadır. Testte; küp sayma, küp karşılaştırma, 2-B'den 3-B'ye ve 3B' den 2B' ye görselleştirme, 2 ve 3 boyutlu nesnelere zihinde döndürme, yüzey tamamlama ve şekli zihinde ayırıştırma olmak üzere 6 alt boyuttan sorular bulunmaktadır. Test çoktan seçmeli ve 7. sınıf seviyesine uygun bir şekilde 4 seçenektir. UYT' nin değerlendirilmesinde yanlış ve boş cevaplara "0", doğru cevaplara ise "1" puan verilmiştir. Testten alınabilecek en düşük puan 0, en yüksek puan 25'tir. Testten alınan puanın artması uzamsal yeteneğin güçlülüğü, azalması ise uzamsal yeteneğin zayıflığı anlamına gelmektedir.

Araştırmada elde edilen bulgular; çalışmanın amacını ve problemini destekler nitelikte olmalıdır. Bulgular kısmında sadece bulgular sunulup açıklanmalıdır. Asla yorum yapılmamalıdır. Yorum, tartışma ve sonuç kısmında yapılmalıdır. Bulgular bölümünde gerek görüldüğünde tablo, şekil, grafik veya resimlerle kullanılıp açıklama yapılabilir.

#### TARTIŞMA ve SONUÇ

UYT; küp sayma, küp karşılaştırma, 2B' den 3B' ye ve 3B' den 2B' ye görselleştirme, 2B ve 3B nesnelere zihinde döndürme, yüzey tamamlama (zihinde bütünleme) ve şekli zihinde ayırıştırma alt boyutlarını içeren uzamsal yeteneği belirlemeye yönelik nihai hali dört seçenekli 25 çoktan seçmeli maddeden oluşan bir testtir. Test geliştirilirken güncel matematik dersi 7. sınıf "Cisimlerin Farklı Yönlerden Görünümleri" ünitesi kazanımları dikkate alınmıştır. Alan yazında ortaokul düzeyi uzamsal yetenek testleri incelendiğinde; Michigan State Üniversitesi matematik bölümü öğretim elemanları tarafından II. kademe öğrencileri için geliştirilmiş MGMP

Uzamsal Görselleştirme Testi karşımıza çıkmaktadır. Bu test Turğut (2007) tarafından Türkçe diline çevrilmiş ve uyarlama çalışması yapılmıştır. Orijinal test 5 şıklı, 10 farklı tip 32 sorudan oluşurken, Türkçeye uyarlanan yeni test 6 tip 31 sorudan oluşmaktadır. Testin pilot uygulaması 4 ilköğretim okulunda, 128 altıncı sınıf, 150 yedinci sınıf ve 104 sekizinci sınıf olmak üzere toplam 382 öğrenci ile yapılmıştır. Toplanan verilerin analizi sonucunda güvenilirlik katsayısı 0,830 olarak bulunmuştur. Test, uzamsal yeteneğin iki alt bileşenini ölçmektedir ve Türkçeye uyarlanmış halinin ismi MGMP Uzamsal Yetenek Testi olarak değiştirilmiştir. Yine alan yazında yer alan ve çalışmalarda sıklıkla kullanılan Yıldız ve Tüzün' ün (2011) çalışmalarında kullandıkları, ilköğretim II. Kademe için A.B.D.' de gerçekleştirilen "Middle Grades Mathematics Project" adlı proje için hazırlanmış olan Uzamsal Görselleştirme Testidir. Test 15 sorudan oluşmakta ve her sorunun 5 cevap şıklı bulunmaktadır. Test soruları birim küplerden oluşturulmuş yapıların izometrik görünümüne ve sağdan, soldan, önden ve arkadan görünümüne yönelik sorular içermektedir. Testin güvenilirliği son- test puanlarına göre hesaplandığında 0,971 (N=108) olarak bulunmuştur. Yıldız ve Tüzün' ün (2011) çalışmalarında kullandıkları Peters ve arkadaşlarının (1995) geliştirdiği "Mental Rotation Test" ise her sorunun niteliği aynı olan toplam 24 sorudan oluşmaktadır. Birim küplerden oluşturulmuş bir şeklin farklı yönlerde ve farklı açılarla döndürüldüğünde oluşacak yeni halini bulmaya yönelik sorular içermektedir. Her sorunun 4 şıklı vardır.

Testin güvenilirliği son-test puanlarına göre hesaplandığında 0,661 (N=108) olarak bulunmuştur. Dokumacı Sütçü ve Oral (2019) yaptığı araştırmada 29 maddeden oluşan; yüksek ayırt ediciliğe sahip maddelerden oluşan orta güçlükte bir test geliştirmişleridir. İki faktörden oluşan testin 14 maddeden oluşan birinci faktör için KR-20 iç tutarlılık katsayısı .77 ve 15 maddeden oluşan ikinci faktör için .78 olarak bulunmuştur. Testin geneli için ise bu değer .78 olarak bulunmuştur. Testin pilot uygulaması, 2016–2017 eğitim-öğretim yılında Diyarbakır İli Merkez İlçelerinde bulunan beş farklı ortaokulda yapılmıştır. Uygulama 301 yedinci sınıf öğrencisiyle gerçekleştirilmiştir. Test "Uzamsal Görselleştirme Testi" olarak literatüre kazandırılmıştır. Ekstrom, French, Harmon ve Derman (1976) tarafından geliştirilen ve Delialioğlu (1996) tarafından Türkçe'ye çevrilen Kağıt Katlama ve Yüzeysel Geliştirme Testlerinin güvenilirlik katsayıları sırasıyla .84 ve .82 iken Olkun (2003) tarafından geliştirilen İki Boyutlu Geometride Uzamsal Görselleştirme Testinin ise .77'dir. Diğer testlerden farklı olarak Yüksel ve Bülbül (2014) tarafından üniversite öğrencileri için geliştirilen uzamsal görselleştirme testinde madde analizi, geçerlik-güvenirlik analizi yapılmış, tek faktörlü yeterli uyum indekslerine sahip olduğu görülmüştür. Güvenirlik katsayısı .84 olarak hesaplanmıştır.

Uzamsal yetenek ve bileşenleri ile ilgili farklı tanımlamalar yapılsa ve farklı ölçme araçlarının geliştirildiği araştırmalar yapılmış olsa da bu çalışmaların ortak noktası uzamsal yeteneğin önemidir. Araştırmalarda çalışılan öğrenci gruplarının farklı özelliklere sahip olmasına rağmen aynı test ile değerlendirilmeleri bu önemi ortaya koymaktadır. Bu çalışmada 7. sınıf ünitesi kazanımları dikkate alınarak hazırlansa da, ortaokul düzeyinde kullanılabilecek son zamanlarda yapılmış güncel ve kapsamlı bir uzamsal yetenek testinin alan yazında bulunmaması bu çalışmanın temelini oluşturmuştur. UYT' nin güvenilirlik katsayısı .918 olarak hesaplanmıştır.



## ÖNERİLER

Bu çalışmada geliştirilen test, yedinci sınıf öğrencilerinin uzamsal yeteneklerini matematiksel bağlam kullanarak değerlendirmek için tasarlanmıştır. Araştırmacılar farklı sınıf düzeylerine yönelik yeni ve güncellenmiş testler geliştirebilirler. Araştırmacıların ayrıca diğer derslerde farklı konularda farklı sınıf düzeylerinde uzamsal yetenekleri ele alan testler geliştirmeleri önerilmektedir. Öğretmenlerin güvenirliliği ve geçerliliği çalışılan başarı testlerinden yararlanabilmeleri için bu testlerin ders kitaplarında yer alması ve Milli Eğitim Bakanlığı Ölçme, Değerlendirme ve Ölçme Hizmetleri Genel Müdürlüğü'nün internet sitesinde yayınlanması sağlanabilir. Performans testlerinde çoktan seçmeli soruların yanı sıra performansa ve sınıf düzeyine bağlı olarak açık uçlu sorular da yer alabilmektedir. Açık uçlu sorulardan oluşmuş başarı testlerinin de geliştirilebileceği araştırmalara önerilmektedir. Başarı testi geliştirme sürecine ilişkin öğretmen ve öğrenci bilgilerini belirlemek amacıyla da araştırmalar yapılabilir.

## Etik Metni

Uygulama için Nevşehir Hacı Bektaş Veli Üniversitesi Rektörlüğü Bilimsel Araştırmalar ve Yayın Etik Kurulu'ndan 27.03.2023 tarih ve 2023.03.84. karar sayılı izin alınmıştır. Gereken izinler aynı zamanda İl Milli Eğitim Müdürlüğü'nden de başvuru yapılarak alınmıştır. Bu makalede dergi yazım kurallarına, yayın ilkelerine, araştırma ve yayın etiği kurallarına, dergi etik kurallarına uyulmuştur. Makale ile ilgili doğabilecek her türlü ihlallerde sorumluluk yazara aittir.

**Yazarın Katkı Oranı Beyanı:** Yazarın katkı oranı %100'dür.

## KAYNAKÇA

- Aksu, B. (2018). *Ölçme ve Değerlendirme*. Cinius Yayınları, 1. Baskı.
- Alias, M., Black, T. R., & Gray, D. E. (2002). Effect of instructions on spatial visualization ability in civil engineering students. *International Education Journal*, 3(1), 1-12.
- Başol, G. (2019). *Eğitimde Ölçme ve Değerlendirme*. Pegem Akademi, 6. Baskı.
- Burnett, S., Lane, D. (1980). Effects of Academic Instruction on Spatial Visualization, *Intelligence*, Sayı 4, 233-342.
- Büyüköztürk, Ş., (2018). *Sosyal bilimler için veri analizi el kitabı istatistik, araştırma deseni SPSS uygulamaları ve yorum*. Pegem A. yayıncılık.
- Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş. & Demirel, F. (2012). *Bilimsel araştırma yöntemleri*. Pegem Akademi, 15. Baskı.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York: Cambridge University Press.
- Cizek, G. J. (1997). *Learning, Achievement and Assessment: Constructs at a Crossroads*. In G. D. Phye (Ed.), *Handbook of Classroom Assessment. Learning, Achievement and Adjustment*. California: Academic Press.
- Clements, D.H. (1998). *Geometric and spatial thinking in young children*. (ERIC Servis No. ED436232)

- Colletta, A.T. ve Chiappetta, E.L. (1989). *Science Introduction in The Middle and Secondary Schools* (2nd edn.). Ohio, USA: Merrill Publishing Company
- Contero, M., Naya, F., Company, P., Saorin, J.K., & Conesa, J.(2005). Improving visualization skills in engineering education. *IEEE Computer Graphics and Applications* (5), 24-31.
- Çepni, S. (2003). An Analysis of University Science Instructors' Examination Questions According to The Cognitive Levels. *Educational Sciences: Theory&Practice*, 3(1), 65-84.
- Delialioğlu, Ö. (1996). *Contribution of students' logical thinking ability, mathematical skills and spatial ability on achievement in secondary school physics*. Yayınlanmamış yüksek lisans tezi. Orta Doğu Teknik Üniversitesi.
- Doğan, N. (2019). *Eğitimde Ölçme ve Değerlendirme*. Pegem Akademi, 1. Baskı.
- Dokumacı Sütçü, N. & Oral, B. (2019). Uzamsal Görselleştirme Testinin Geliştirilmesi: Geçerlik ve Güvenirlik Çalışmaları. *Kastamonu Eğitim Dergisi*, 27 (3), 1179-1195. DOI: 10.24106/kefdergi.2826.
- Ekstrom, R. B., Dermen, D., & Harman, H. H. (1976). *Manual for kit of factor-referenced cognitive tests (Vol. 102)*. Princeton, NJ: Educational Testing Service.
- Fraenkel, J. K., & Wallen, N. E. (1996). *How to design and evaluate research in education (3. edition)*. New York: McGraw-Hill, Inc.
- Gelbal, S. ve Kelecioğlu, H. (2007). *H. Ü. Eğitim Fakültesi Dergisi (H. U. Journal of Education)*, 33, 135-145.
- Guay, R. B. (1976). *Purdue Spatial Visualization Test*. West Lafayette, Purdue Research Foundation.
- Gür, H. & Kobak Demir, M. (2017). Pergel Cetvel Kullanarak Temel Geometrik Çizimlerin Öğretmen Adaylarının Geometrik Düşünme Düzeylerine Ve Tutumlarına Etkisi. *Eğitimde Kuram ve Uygulama*, 13(1), 88-110.
- Kan, A. (2011). *Ölçme aracı geliştirme*. S. Tekindal (Ed.), Eğitimde ölçme ve değerlendirme (ss. 249-257). Pegem Akademi.
- Karasar, N. (2011). *Bilimsel araştırma yöntemleri*. Nobel Yayınları, 3. Baskı.
- Karip, E. (Ed.) (2012). *Ölçme ve değerlendirme*. Pegem Akademi, 5. Baskı.
- Kayhan, E.B. (2005). *Investigation of high school students' spatial ability*. Yayınlanmamış doktora tezi, ODTÜ, Ankara.
- Kehoe, J. (1995). Basic Item Analysis for Multiple-Choice Tests. (ERIC Servis No. ED398237)
- Kyllonen, P. C., Lohman, D. F., & Snow, R. E. (1984). Effects of aptitudes, strategy training, and task facets on spatial task performance. *Journal of Educational Psychology*, 76(1), 130.
- Küçükahmet, L. (2002). *Öğretimde planlama ve değerlendirme*. Nobel Yayın Dağıtım, 13. Baskı.
- Linn, M. C., & Petersen, A. C. (1985). Emergence and characterization of gender differences in spatial abilities: A meta- analysis. *Child Development*, 56, 1479-1498.
- Lohman, D. F. (1979). *Spatial ability: Individual differences in speed and level*. Aptitude Research Project, School of Education, Stanford University.
- McGee, M.G. (1979). *Human spatial abilities: Sources of sex differences*. Praeger.
- Odell, R.L. (1993). *Relationship Among Three Dimensional Laboratory Models, Spatial Visualization Ability, Gender and Earth Science Achievement*. Yayınlanmamış Doktora Tezi School of Education, Indiana University, Indiana.

- Okagaki, L.R., & Frensch, P.A. (1996). *Effects of video game playing on measures of spatial performance: Gender effects in late adolescents*. In P. Greenfield & R. Cocking (Eds.), *Interacting with video* (pp. 115-140) Norwood, NJ: Ablex Corporation.
- Olkun S. (2003). Making connections: Improving spatial abilities with engineering drawing activities. *International Journal of Mathematics Teaching and Learning*, 3(1), 1-10. <https://pdfs.semanticscholar.org/0328/1231a0411b83f44268d5c1145940067e8556.pdf> adresinden 13/09/2023 tarihinde erişilmiştir.
- Olkun, S., Altun, A. (2003). İlköğretim Öğrencilerinin Bilgisayar Deneyimleri ile Uzamsal Düşünme ve Geometri Başarıları Arasındaki ilişki, *Turkish Journal of Educational Technology*, 2 (4), Article 13.
- Özçelik, D. A. (2016). *Ölçme ve Değerlendirme*. Pegem Akademi, 5. Baskı.
- Özdemir, S. M., (2009). Eğitimde program değerlendirme ve Türkiye’de eğitim programlarını değerlendirme çalışmalarının incelenmesi. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 6(2), 126-149.
- Pellegrino, J.W., Alderton, D.L., Shute, V.J. (1984). Understanding Spatial Ability, *Educational Psychologist*, 19, 239-253.
- Seçer, İ. (2013). *SPSS ve LISREL ile pratik veri analizi: Analiz ve raporlaştırma*. Anı Yayıncılık.
- Tan, Ş., & Kayabaşı, Y. (2003). *Öğretimi Planlama ve Değerlendirme*. Anı Yayıncılık, 4. Baskı.
- Tekin, H. (1991). *Eğitimde Ölçme ve Değerlendirme*. Yargı Yayınları.
- Turğut, M. (2007). *İlköğretim II. kademedeki öğrencilerin uzamsal yeteneklerinin incelenmesi*. Yayımlanmamış yüksek lisans tezi, Dokuz Eylül Üniversitesi, İzmir.
- Turğut, M. & Yılmaz, S. (2012). İlköğretim 7. Ve 8. Sınıf Öğrencilerinin Uzamsal Yeteneklerinin İncelenmesi. *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi*, (19), 69-79. <https://dergipark.org.tr/en/pub/zgefd/issue/47945/606597> adresinden 18/09/2023 tarihinde erişilmiştir.
- Turgut, M. F., & Baykul, Y. (2014). *Eğitimde ölçme ve değerlendirme*. Pegem Akademi, 6. Baskı.
- Winter, J. W., Lappan, G., Fitzgerald, W. & Shroyer, J. (1989). *Middle Grades Mathematics Project: Spatial Visualization*. NY: Addison-Wesley.
- Yıldız, B. ve Tüzün, H. (2011). Üç Boyutlu Sanal Ortam Ve Somut Materyal Kullanımının Uzamsal Yeteneğe Etkileri. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi (H. U. Journal of Education)* 41: 498-508.
- Yılmaz, H. (1998). *Eğitimde Ölçme ve Değerlendirme*. Mikro Yayınları.
- Yolcu, B. & Kurtuluş, A. (2010). 6. Sınıf Öğrencilerinin Uzamsal Görselleştirme Yeteneklerini Geliştirme Üzerine bir Çalışma. *İlköğretim Online*, 9 (1), 256-274. <https://dergipark.org.tr/en/pub/ilkonline/issue/8596/106967> adresinden 18/09/2023 tarihinde erişilmiştir.
- Yüksel, N. S., & Bülbül, A. (2014). Uzamsal görselleştirme üzerine test geliştirme çalışması. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 8(2), 124-142.