



Development of Number Sense Test for Eighth-Grade Students: A Validity and Reliability Study*

Osman BİRĞİN^{a**} (ORCID ID - 0000-0003-3460-2731)

Elif Seval PEKER^b (ORCID ID - 0000-0002-5653-6352)

^aUşak Üniversitesi, Eğitim Fakültesi, Uşak/Türkiye

^bMilli Eğitim Bakanlığı, 31 Ağustos Ortaokulu, Uşak/Türkiye



Article Info

DOI: 10.14812/cufej.871367

Article history:

Received 30.01.2021

Revised 29.11.2021

Accepted 14.02.2022

Keywords:

8th grade,
Number sense,
Number sense components,
Test development

Research Article

Abstract

The aim of this study was to develop a test for determining the number sense performance of eighth-grade students. This study was conducted with 405 eighth-grade students. The SPSS 17.0 and LISREL 8.8 software were used for data analysis. Data were analyzed by using the item analysis, exploratory and confirmatory factor analysis, and Cronbach's alpha reliability tests. As a result of the analyzes, the number sense test was developed, which consists of 38 multiple-choice items and six number sense components and can explain 48.5% of the total variance. Cronbach's alpha reliability coefficients of the number sense test and its sub-components were found to be between 0.72 and 0.82, and the item-total correlations were between 0.30 and 0.47. The first and second-level confirmatory factor analysis results confirmed that the number sense test consisted of six components and that the established models fit well. As a result of this study, it was determined that the number sense test is a valid and reliable measurement tool in determining the number sense performance of eighth-grade students.

Sekizinci Sınıf Öğrencilerine Yönelik Sayı Duyusu Testinin Geliştirilmesi: Geçerlik ve Güvenirlik Çalışması*

Makale Bilgisi

DOI: 10.14812/cufej.871367

Makale Geçmişi:

Geliş 30.01.2021

Düzeltilme 29.11.2021

Kabul 14.02.2022

Anahtar Kelimeler:

Sekizinci sınıf,
Sayı duygusu,
Sayı duygusu bileşenleri,
Test geliştirme

Öz

Bu araştırmanın amacı, ortaokul sekizinci sınıf öğrencilerinin sayı duygusu performanslarını belirlemeye yönelik bir test geliştirmektir. Bu çalışma, 405 sekizinci sınıf öğrencisi ile yürütülmüştür. Verilerin analizi için SPSS 17.0 ve LISREL 8.8 paket programları kullanılmıştır. Veriler madde analizi, açıklayıcı ve doğrulayıcı faktör analizi ile Cronbach alpha güvenilirlik testleri kullanılarak analiz edilmiştir. Analizler sonucunda 38 çoktan seçmeli maddeden ve altı sayı duygusu bileşeninden oluşan ve toplam varyansın %48.5'ini açıklayabilen sayı duygusu testi geliştirilmiştir. Sayı duygusu testinin ve alt bileşenlerinin Cronbach alpha güvenilirlik katsayılarının 0.72 ile 0.82 arasında, madde-toplam korelasyonlarının 0.30 ile 0.47 arasında değer aldığı saptanmıştır. Yapılan birinci ve ikinci düzey doğrulayıcı faktör analizleri sayı duygusu testinin altı bileşenden oluştuğu ve kurulan modellerin iyi uyum gösterdiği doğrulanmıştır. Bu araştırma sonucunda sayı duygusu testinin sekizinci sınıf öğrencilerinin sayı duygusu performanslarını belirlemede geçerli ve güvenilir bir ölçme aracı olduğu belirlenmiştir.

*This study was derived from the second author's Master's thesis, supervised by the first author.

**Corresponding Author: osman.birgin@usak.edu.tr

Introduction

Numbers are among the most basic concepts of mathematics. Numbers play an important role in reasoning, flexible thinking, making relationships logically, and problem-solving related to the daily routine (Yang, 2019). The subject of numbers is also the basis of other learning areas of mathematics such as measurement, geometry, algebra, and data analysis. Number sense, which is an important skill for mathematics education, was first put forward in the National Council of Teachers of Mathematics [NCTM] (1989). In later mathematics education reform studies and curricula (NCTM, 2000; Turkish National of Minister Education [MEB], 2018), it is seen that numbers learning area and number sense skill are given importance.

NCTM (1989) emphasized that children with the number sense can develop the ability to comprehend the meanings of numbers, to establish multiple relationships between numbers, to recognize the relative sizes of numbers, to understand the effect of operations on numbers, and to determine the appropriate reference point for the measurement (p. 38). According to Hope (1989), number sense is the ability to make logical predictions about the various uses of numbers, recognize arithmetic errors, choose the most effective calculation method, and notice number patterns. Greeno (1991) defined number sense as flexible thinking, estimation skill in calculation, and the ability to reason with inference about numerical quantities. According to Reys et al. (1999), number sense means not only understanding the meaning of numbers and operations, but also meaning numerical situations, developing useful and effective strategies, and using numbers and operations with appropriate methods to form mathematical judgments. Kalchman, Moss, and Case (2001) defined number sense as estimating the given size, realizing the irrationality in the results, flexibility in mental calculation, and the ability to establish relationships between different representations. Yang (2019) expresses number sense as the ability to solve problems encountered in daily life by developing flexible and effective strategies.

Conceptual understanding of the numbers and arithmetic operations and the development of number sense skills play an important role in mathematics learning (NCTM, 2000). The development of students' number sense skills at an early age positively affects their mathematics learning in later years, whereas the lack of number sense makes it difficult to learn mathematics (Jordan, Glutting & Ramineni, 2010). In many studies conducted at different education levels (e.g. Akkaya, 2016; Cansız Aktaş et al., 2017; Çekirdekçi, Şengül & Doğan, 2016; Jordan et al., 2010; Harç, 2010; Kayhan Altay, 2010; Mohamed & Johnny, 2010; Tunali, 2018), it has been found that there is a positive and significant relationship between number sense and mathematics achievement. Moreover, it has been determined that the number sense is an important predictor of the student's mathematics achievement (e.g. Olkun, Mutlu & Sarı, 2017; Reys & Yang, 1998; Yang, Li, & Lin, 2008). Some studies in Turkey (e.g. Can, 2017; Çekirdekçi et al., 2016; Er & Dinç Artut, 2017; Harç, 2010; Tunali, 2018) have shown that students with high math achievement are better in terms of number sense skills than students with low mathematics achievement, however, they generally make rule-based solutions. This result indicates that students' number sense skills are not sufficiently developed.

Reys (1994) states that number sense is a difficult concept to define because it is a complex process involving numbers, operations, and the relationships between them. For this reason, there is no common consensus about the number sense components in the literature. Indeed, it is seen that there is various classification related to the number sense components in international studies (e.g. Greeno, 1991; Markovits & Sowder, 1994; McIntosh, Reys & Reys, 1992; Li & Yang, 2010; Reys et al., 1999; Yang, 1995; 2019; Yang, Li, & Li, 2008) and Turkey (e.g. Çekirdekçi Şengül & Doğan, 2017; İyemen İkizoğlu & Duatepe Paksu, 2016; Kayhan Altay & Umay, 2013; Soyuk, 2018; Şengül & Gülbağcı, 2012).

Number Sense Components

In the literature, it is seen that different number sense components are defined depending on the student's age, grade level, subject, and curriculum. The most comprehensive classification of the number sense components was made by McIntosh et al. (1992). McIntosh et al. (1992) defined three main components (*Knowledge of and facility with numbers, knowledge of facility with operations,*

applying knowledge of and the facility with numbers and operations to computational settings) and sub-components related to number sense skills. (a) The first main number sense component (*Knowledge of and facility with numbers*) includes the sense of orderliness of numbers, multiple representations for numbers, sense of the relative and absolute magnitude of numbers, and system of benchmarks. (b) The second main number sense component (*Knowledge of and facility with operations*) includes understanding the effects of operations, understanding mathematical properties, and understanding the relationships between operations. (c) The third main number sense component (*Applying knowledge of and the facility with numbers and operations to computational settings*) includes the understanding of the relationship between problem context and the necessary computation, awareness that multiple strategies exist, inclination to utilize an efficient representation and/or methods, and inclination to review data and results for sensibility.

Based on the components of the number sense proposed by McIntosh et al. (1992), five number sense components were created by Reys et al. (1999). These components are briefly as follows:

(a) *Understanding of the meaning and size of numbers*: This component requires recognizing the relative size of numbers, being able to compare two given fractions, being able to sort them, and finding the number between two numbers. For example; be able to recognize how many numbers between 1.52 and 1.53, to compare the fraction $\frac{2}{5}$ and the fraction $\frac{1}{2}$ in terms of magnitude.

(b) *Understanding and use of equivalent representations of numbers*: It refers to knowing the equivalents and equivalent representations of numbers and being able to use them when necessary. For example, being able to represent the fraction $\frac{2}{5}$ with different representations such as $\frac{4}{10}$, 40 %, 0.4, and realize which number multiplied by a number m will give the same result as dividing by 0.25

(c) *Understanding the meaning and effect of operations*: This component is related to the ability to realize how the result will change when the value of a number or operation changes in computational situations. That means realizing that multiplication does not always make numbers bigger and division does not always make numbers smaller. For example, being able to predict whether the outcome of the $70 \div 0.5$ operations could be equal to 70×2 , or whether the result of the $750 \div 0.98$ operations would be greater or less than 750.

(d) *Measurement benchmarks*: This component includes identifying and using reference/benchmarks that may be suitable for different situations. For example, the student can estimate the length of another object by reference to his height or the length of an object; the ordering of fractions and decimals by taking the fixed points 1, $\frac{1}{2}$, $\frac{1}{3}$, and $\frac{1}{4}$ as reference points.

(e) *Flexible computing and counting strategies for mental computation, written computation, and calculator use*: It requires being able to solve problems and predict the results without calculating using paper and pen, and calculating mentally. For example, being able to mentally calculate the approximate value of 6×98 ; "Ali is a 5th-grade student. Ali says he lives 30 000 days." Being able to decide whether Ali's proposition is logical or not.

The classification of number sense developed by Yang (1995) consists of six components. These components are briefly as follows: (a) *Understanding the basic meaning of numbers* (For example, this component requires an understanding of the basic meaning of numbers, the base-ten system, the concept of place-value, and number patterns. Just like a first-year primary school student can express the number 48 as "4 dozen eggs" or "4 tens plus 8 ones", (b) *Being able to compose and decompose numbers* (For example, being able to realize that the result of $3 + 3 + 3 + 3$ operation is the same as the 4×3 operation), (c) *Recognizing the magnitude of numbers* (For example, comparison two fraction, finding the third number closer to the given number, being able to sort the given numbers) (d) *Being able to use benchmarks appropriately* (For example, taking 1 as the closest value for 0.95 as reference; referencing a known point such as $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{1}{5}$ for another given fraction), (e) *Recognizing the relative effect of operations on numbers* (For example, knowing that multiplication does not always magnify the result, as the result of 199×0.9 is less than 199), (f) *Flexibility in applying numerical and operational knowledge to*

computational situations (For example, a school bus carries 45 students. How many buses would be needed if 915 students were to be taken to the museum?)

Yang, Li, and Li (2008) developed a two-tier number sense test for third-grade primary school students, consisting of multiple-choice questions and asking them to provide a reason for each question. This test was composed of five number sense components as follow: (a) *Understanding the meaning of numbers and operations*, (b) *Recognizing relative number size*, (c) *Being able to compose and decompose numbers*, (d) *Using multiple representations of numbers and operations*, (e) *Judging the reasonableness of computational results*. Yang (2019), in another study on fifth-grade primary school students in Taiwan, developed a three-tier number sense test that composed five components. In this test, unlike the previous study (Yang et al., 2008), there is a component of *"Being able to recognize the relative effect of operations on numbers"* instead of the component of *"Being able to compose and decompose numbers"*.

Although there have been many studies on number sense in the international literature for about thirty years, number sense skills have given importance in Turkey in recent years. Indeed, Birgin and Peker (2021) analyzed the studies on number sense in Turkey between 2000-2018 years by thematic content analysis. As a result of this study, it has been determined that studies on number sense in Turkey have accelerated since 2010, scale development studies for number sense are limited, and the number sense classification developed by Kayhan Altay and Umay (2013) is mostly used in the studies. Kayhan Altay and Umay (2013) developed a number sense test consisting of 17 questions and three components that require a written answer for 6th, 7th, and 8th-grade students. The number sense components of the test are briefly as follows: *Flexibility in the calculation* (How do you solve shortly 0.25×16 operation? Explain how you do it.), *conceptual thinking in fraction* (Thinking of points on a given number line, insert the fractions $\frac{1}{2}$, $2\frac{1}{2}$ and. Explain how you place it) and *using benchmark (reference points)* (How do you solve shortly 86424×500 operation? Explain how you think). İymen İkizoğlu and Duatepe Paksu (2016) developed a number sense test for 8th-grade students on exponential expressions, which requires 13 written answers and includes five number sense components. The number sense components in this test are *"equivalent representation"*, *"numerical estimation"*, *"number size"*, *"effect of operation"*, and *"use of reference point"*. On the other hand, Çekirdekçi, Şengül, and Doğan (2017) developed a number sense test consisting of 11 multiple-choice questions and three components (*knowing the equivalents of numbers and making quantitative inference*, *calculating the effects of operations using reference point*, and *knowing the meaning of numbers and flexible thinking*) for fourth-grade primary school students. It can be said that the grade level, subject, and number sense components on which number sense tests focus differ in Turkey.

Significance and Purpose of the Study

When the number sense classifications are examined in the literature, it is seen that a common terminology cannot be formed for the number sense components, and different names are used for the number sense components that express the same skill (e.g. Greeno, 1991; Markovits & Sowder, 1994; Yang, 1995). Although some number sense components (such as the use of reference points, the effect of operations, the meaning and size of numbers) are common in some studies (e.g. İymen İkizoğlu & Duatepe Paksu, 2016; McIntosh et al., 1992; Reys et al., 1999; Yang, 1995; 2019; Yang et al., 2008), some number sense components such as understanding the basic meaning of numbers (Li & Yang, 2010), being able to compose and decompose numbers (Reys et al., 1999; Yang, 2019), judging the reasonableness of computational results (İymen İkizoğlu & Duatepe Paksu, 2016), multiple representations of numbers (Kayhan Altay & Umay, 2013; Markovits & Sowder, 1994) are not included. In addition, it can be stated that the components in the number sense scales differ depending on the age group of the student and the subject content in the curriculum, and the priority regarding the number sense components changes. When the primary and middle school mathematics curriculum in Turkey (MEB, 2018) is examined, it is expected that the eighth-grade students will have number sense skills related to natural numbers, integers, rational numbers, irrational numbers, and different representations of numbers (decimal, exponent, percent, etc.). Some studies (e.g. Birgin & Peker, 2021;

Çekirdekçi et al., 2017; İymen İkizoğlu & Duatepe Paksu, 2016) show that tests and scales for number sense skills are limited in Turkey. Moreover, it can be said that these number sense tests and scales are not sufficient to determine the general number sense of eighth-grade students regarding integers, rational, irrational numbers, and different representations of numbers (decimal, exponent, etc.) in terms of their focus and grade level. Considering the changes in content, subjects, and gains in the mathematics curriculum (MEB, 2018) updated after 2013, instrument tools are needed to determine the general number sense skills of students after eight years of education. Therefore, this study aimed to develop a test to determine the number sense performance of eighth-grade students. Unlike previous number sense tests, this study includes students' number sense skills related to integers, rational and irrational numbers, and different representations of numbers at the end of eight years of education, and is based on the number sense components (Understanding the meaning of numbers, quantitative reasoning and inference, equivalent representation of numbers, the effect of operations on numbers, the use of reference points in a measurement, mental calculation, etc.) expressed in many number sense scales. In this respect, it is thought that this study will contribute to future research in the field of number sense.

Method

Study Group

Since the main purpose of this research was to develop a measurement tool, a convenience sampling method was preferred. The convenience sampling method is expressed as the sampling selected from individuals who are in the immediate vicinity, easy to reach, and willing to participate in the research voluntarily. In studies aiming to develop a measurement tool, it is emphasized that sampling the range of the measured feature is important and that participation should be voluntary (Erkuş, 2012). In this context, the necessary written institutional permissions were obtained from the Directorate of National Education regarding the schools where the instruments would be applied. In addition, this research was ethically approved by the Uşak University Social and Human Sciences Scientific Research and Publication Ethics Committee. Participants were informed about the research and informed consent was obtained from the participants. This research was conducted on the basis of voluntary participation. The study group of this research consisted of 405 eighth-grade students studying in different settlements of a province (province center, district center, and rural) in the western region of Turkey. 180 students participated in this research for item analysis and exploratory factor analysis, and 225 students participated in confirmatory factor analysis. 48.4% of the participants (n = 196) are female and 51.6% of them (n = 209) were male. 40.7% of the students (n = 165) were in the province center, 33.1% of them (n = 134) were in the district center and 26.2% of them (n = 106) were in the rural.

Development Process of Number Sense Test (NST)

The development process of the number sense test (NST) and the procedures performed in this process are summarized in Table 1. In the process of developing the NST, it was reviewed primarily on studies conducted in international (e.g. Facun & Nool, 2012; Greeno, 1991; McIntosh et al., 1992; Purnomo, Kowiyah, Alyani & Assiti, 2014; Li & Yang, 2010; Reys et al., 1999; Singh, 2009; Yang, 1995; 2019; Yang & Lin, 2015) and Turkey (e.g. Gülbağcı Dede, 2015; Harç, 2010; Kayhan Altay & Umay, 2013; Soyuk, 2018; Şengül & Gülbağcı Dede, 2013) to determine the characteristics of number sense and its components. The learning area of numbers and related acquisitions of the 2018 Mathematics Curriculum (MEB, 2018) were reviewed for NST to be suitable for eighth-grade students. Considering the achievements of the Turkish mathematics curriculum, it was decided to base the six number sense components, which are widely used in number sense definitions and number sense scales in the literature. These number sense components are (a) *Understanding the meaning of numbers*, (b) *Quantitative reasoning and inference*, (c) *Equivalent representation of numbers*, (d) *Effect of the operations on numbers* (e) *Use of reference points in the measurement*, and (f) *Mental calculation*.

In the process of creating the items of the NST, the questions used in the literature on number sense, national and international exam questions (High school entrance exam [TEOG, SBS, LGS], TIMMS, etc.),

mathematics textbooks, and resource books were used. Some of the items in the NST were taken from relevant sources exactly, and some items were changed or redesigned by the researchers. In the process of developing the items, it was based on the necessity of using number sense skills without using paper and pencil and the time constraint. In addition, in the process of writing multiple-choice items, students' errors and misconceptions expressed in the literature were taken into consideration. In this context, a draft multiple-choice item pool was created for NST.

Table 1.
Number Sense Test Development Stages and Procedures

Test Development Stages	Procedures
1. Determination of number sense components	<ul style="list-style-type: none"> Literature review, examination of number sense scales and tests, and mathematics curriculum.
2. Creating of the item pool	<ul style="list-style-type: none"> The literature, number sense scales, mathematics curriculum, textbook, and central exam questions (SBS, TEOG, LGS, TIMSS, etc.) were reviewed. An item pool was created for the test.
3. Ensuring content and appearance validity	<ul style="list-style-type: none"> Obtaining expert and teacher opinions regarding the suitability of draft items for language, student level, and number sense components. Making necessary corrections in the items in line with the comments and suggestions. Pilot implementation on 24 students and taking student opinions for the suitability of the draft items for the student
4. Application of the prepared number sense test	<ul style="list-style-type: none"> Application of the prepared draft test on a total of 405 eighth-grade students
5. Performing item analysis	<ul style="list-style-type: none"> Calculating item difficulty index and item distinctiveness index and removing unsuitable items.
6. Making structural validity analysis	<ul style="list-style-type: none"> KMO and Bartlett's test of sphericity Conducting exploratory factor analysis Conducting confirmatory factor analysis
7. Calculating reliability and giving the final version of the test	<ul style="list-style-type: none"> Calculation of the Cronbach's alpha internal consistency coefficient Giving the final version of the number sense test as a result of analysis

To ensure the content and appearance validity of the draft test, it was first reviewed by four experts in mathematics education in terms of content, language and expression, student's grade level, number sense component, and the curriculum, and various revisions were made in the items in line with expert opinions and suggestions. In addition, the opinions of four mathematics teachers who have a master's degree in mathematics education were taken and some revisions were made on the items in line with their suggestions. It was applied to 24 eighth-grade students to test the suitability of the draft NST for the student's level and the application time. Opinions of these students were taken to determine whether some expressions and visuals were not understood in the test items. Students stated that some questions were difficult, some were quite easy and the given time was sufficient. In line with the students' opinions, a visual in the test was made more prominent. The revised draft NST consists of a total of 47 multiple-choice questions for the six number sense components.

Number Sense Components and Properties of Number Sense Test

Some of the items in the NST are given as examples in Appendix-1. The distribution and properties of the items in the NST according to the number sense components are briefly as follows:

a) Understanding the meaning of numbers (UN): This component refers to understanding number size, comparing and ordering numbers, and understanding the relationship between numbers and their place values (McIntosh et al., 1997). The items M2, M4, M6, M8, M11, M22, M37, and M40 in NST were prepared for understanding numbers and knowledge of numbers. For example, item M6 (Appendix-1) in NST was written with inspiration from Singh (2009) and Purnomo et al. (2014). In this item, students are expected to realize that there can be infinite decimal fractions between 1.52 and 1.53, considering these numbers as 1.520 and 1.530 or 1.5200 and 1.5300.

b) Quantitative reasoning and inference (QR): This component includes making a judgment and inference about whether the numerical value obtained as a result of the calculation is reasonable. For this component, the student needs to know what the numerical value obtained means and whether it is logical or not (Kayhan Altay & Umay, 2013). The items M1, M7, M26, M28, M33, M39, and M44 in the NST were prepared in the context of the quantitative reasoning and inference component. For example, item M1 in this study was taken from the 2007 TIMMS. In this item, "One teacher for every 12 students will attend the trips organized by the school. If 110 students are going on the trip, how many teachers should attend at least?" When 110 students are divided into 12 for this question, 9 teachers are required. However, a teacher is also required for the remaining two students. The student is expected to notice this.

c) Equivalent representation of numbers (ER): The equivalent representation of numbers includes many different forms of representation of numbers. This component entails finding the equivalence of numbers, being able to reconstruct numbers by composing and decomposing, and transferring between multiple representations of numbers (percentage, exponential, decimal, shaded region, and number line representation, etc.) (McIntosh et al., 1997). The items M9, M12, M13, M15, M18, M21, M36, and M45 in NST were prepared in the context of the equivalent representation component of the numbers. For example, item M12 was taken from the 2015 High School Entrance Exam (TEOG) in Turkey (Appendix-1). In this item, the value of the 5^{-4} is assumed to be $\frac{1}{625}$ and is expected to be converted to a decimal fraction. Item M21 in the NST requires that the shaded region whose geometric shape is given be expressed as a percentage.

d) Effect of the operations on numbers (EO): This number sense component includes recognizing the effect and meaning of operations related to numbers. For example, dividing the number into equal sub-groups or multiplying the number by a number less than 1 to obtain a smaller number than itself. The items M10, M20, M23, M24, M34, M38, M43, M46 in the NST were prepared for recognizing the effect of operations on numbers. For example, item M24 (Find the closest value without calculating the result of $\frac{2}{3} \times \frac{5}{4} = ?$) and item M38 ($29 \div 0.8 = ?$, $29 \times 0.8 = ?$, $29 + 0.8 = ?$, and $29 - 0.8 = ?$ Which one gives the greatest result?) is a question that does not apply to the rule that multiplication increases the multiplied number, while division makes the number smaller. A student who had advanced number sense skills is expected to realize that when a number is multiplied by a decimal number smaller than 1, a smaller number is obtained than the number itself.

e) Use of reference points in the measurement (UR): This number sense component is the ability to use standard and non-standard units of measurement in a way that is compatible with daily life and logically. This number sense component also requires the ability to make predictions using properties such as mass, length, volume, time, and angle (McIntosh et al., 1997). The items M14, M17, M19, M25, M27, M30, M31, M32, and M41 in NST were prepared for the use of reference points in measurement. For example, the M27 item in NST was taken from the study conducted by Gülbağcı Dede (2015) and Harç (2010). This item requires realizing that the sum of two fractions greater than half is greater than 1 by using the half reference point without applying the denominator equalization rule. In item M19 (Appendix-1), a student who has developed number sense skills is expected to find that the circumference of the L shape is greater than the K shape by taking the AB line segment as a reference point.

f) Mental calculation (MC): The mental calculation component involves estimates without using a paper-pencil and a calculator and reaching the result by calculating mentally. The M3, M5, M16, M29, M35, M44, M47 items in NST were prepared for the mental calculation component. For example, item M3 was taken from the 2011 TIMMS (Which of the following is closest to the result of the $\frac{7.21 \times 3.86}{10.09}$?). In this item, the result is reached by rounding the decimal fraction 7.21 to 7, 3.86 decimal fraction to 4, and 10.09 decimal fraction to 10. M16 item was written inspired by the 2010 Placement Exam for Higher School (SBS). In this item, it is expected that the side length of a square garden with an area of $150 m^2$ is $\sqrt{150}m$ and this value is estimated to be between 12 m and 13 m.

Data Collection and Data Analysis

In this study, the participants were informed about the NST and its application form. Informed consent was obtained from all participants and the data were gathered based on a purely voluntary basis. Each item in the NST was presented to the students by using projection or smart board technology, and they were asked to determine the most appropriate answer within 35-40 seconds by using their number sense skills without calculating with paper and pencil. Each item has been completed independently. Total scores for the NST were obtained by giving 1 point for the correct answer to each item in the test and zero point for incorrect and blank answers. Before the data analysis, the suitability of the assumptions for analysis, the evaluation of missing data, normality test, determination of linearity, and extreme values (other than -3 and +3 values of standard deviation) were checked. Those with more missing data regarding the test and the extreme data were excluded from the analysis. It was determined that kurtosis (-0.596) and skewness (0.356) values of the test scores had values between +1 and -1 and showed a normal distribution.

The validity and reliability studies of the NST were carried out in two applications. The first application was carried out with 180 students selected from the schools that constitute the study group to determine the item difficulty index and item discrimination index for item analysis of the test. SPSS statistics program and Excel program were used for item analysis. The item-total correlation, item difficulty index, and item discrimination index value of each item were calculated based on the data obtained. In determining the item distinctiveness index, the method of 27% lower group and 27% upper group was used. As a result of item analysis, unsuitable items were eliminated. To determine the conformity of the samples for the construct validity, firstly the Kaiser-Meyer-Olkin (KMO) and Bartlett's test of sphericity were calculated (Büyükoztürk, 2007). Afterward, exploratory factor analysis was performed using Principal Components Analysis and Varimax orthogonal rotation technique. The sub-dimensions of the NST were determined based on the factor analysis results. In addition, the Cronbach Alpha internal consistency coefficient was calculated for the reliability of the overall test and sub-factors.

In the second stage of the NST, confirmatory factor analysis was performed on 225 eighth-grade students using the LISREL 8.8 program to determine the model fit of the number sense components, which emerged as a result of the exploratory factor analysis within the scope of the construct validity study, with the NST.

Findings

Findings Regarding Item Analysis

With item analysis, it is determined whether the items in the scale measure the feature they aim to measure without mixing other features (Tavşancıl, 2002). For item analysis, item-total correlation, item difficulty index (p_j), and item discrimination index (r_{jx}) values were calculated. It is stated in the literature that the item-total correlation value should be 0.20 and above (Kline, 1993).

The item difficulty index (p_j) is the ratio of the number of students who answered the question correctly to the number of all students. Item difficulty index takes values between 0 and 1. It is categorized that items with an item difficulty index of less than 0.20 were "very difficult", items between 0.21-0.40 were "hard", items between 0.41-0.60 were "moderately difficult", items between 0.61-0.80 were "easy", and items greater than 0.80 were "very easy". It is stated that the items at medium difficulty (0.50) increase the validity and reliability of the test (Turgut & Baykul, 2010).

Item discrimination index (r_{jx}) is the power to distinguish whether the item has the feature it aims to measure or not (Büyükoztürk, 2007). This index ranges from -1 and +1. Items with an item discrimination index lower than 0.20 and negative should be removed from the test and items between 0.20-0.29 should be corrected. Items with an item discrimination index of 0.30-0.39 have a good level of discrimination, while items with an item discrimination index of 0.40 and above have a very good level

of discrimination (Turgut & Baykul, 2010). In this study, the item difficulty index, item discrimination index, and item-total correlation values of the NST are presented in Table 2.

When Table 2 was examined in terms of item-total correlation values, it was determined that two items (M39, M43) had values between 0.132 and 0.174, three items (M40, M45, M46) had values between 0.203 and 0.233, and other items had values between 0.304 and 0.471. For this reason, it has been concluded that it would be appropriate to remove the NST for M39 and M43 items with item-total correlation values lower than 0.20. When Table 2 was examined in terms of item discrimination index criteria, it was decided to exclude five items in the NST (M39, M40, M43, M45, and M46) due to their discrimination index being lower than 0.20. Among the remaining items, 13 items (M1, M2, M4, M11, M12, M15, M17, M19, M21, M23, M26, M32, M38) were found to be “good” and the others were “very good”.

Table 2.
Item Analysis Results for the Number Sense Test

Item	<i>r</i>	<i>p_j</i>	<i>r_{jk}</i>	Item	<i>r</i>	<i>p_j</i>	<i>r_{jk}</i>	No	<i>r</i>	<i>p_j</i>	<i>r_{jk}</i>
M1	.342**	0.28	0.40	M17	.329**	0.43	0.36	M33	.423**	0.45	0.46
M2	.364**	0.44	0.35	M18	.420**	0.54	0.73	M34	.428**	0.40	0.51
M3	.381**	0.42	0.55	M19	.323**	0.25	0.35	M35	.357**	0.46	0.55
M4	.403**	0.31	0.34	M20	.388**	0.55	0.65	M36	.370**	0.40	0.45
M5	.318**	0.59	0.55	M21	.441**	0.30	0.33	M37	.358**	0.59	0.53
M6	.458**	0.33	0.48	M22	.423**	0.48	0.69	M38	.395**	0.18	0.36
M7	.381**	0.51	0.47	M23	.388**	0.28	0.30	M39	.132	0.72	0.10
M8	.367**	0.28	0.47	M24	.412**	0.57	0.43	M40	.224**	0.56	0.17
M9	.433**	0.43	0.61	M25	.325**	0.69	0.54	M41	.331**	0.45	0.40
M10	.471**	0.35	0.52	M26	.339**	0.32	0.33	M42	.356**	0.41	0.50
M11	.362**	0.21	0.31	M27	.394**	0.37	0.53	M43	.174*	0.81	0.13
M12	.435**	0.27	0.34	M28	.419**	0.47	0.51	M44	.324**	0.26	0.33
M13	.308**	0.60	0.58	M29	.375**	0.49	0.41	M45	.203**	0.84	0.21
M14	.375**	0.36	0.43	M30	.338**	0.62	0.55	M46	.233**	0.30	0.14
M15	.419**	0.63	0.37	M31	.361**	0.68	0.59	M47	.304**	0.37	0.33
M16	.324**	0.60	0.53	M32	.423**	0.40	0.40				

r: Item-total correlation, *p_j*: Item difficulty index, *r_{jk}*: Item discrimination index

p* < .05, *p* < .01

When Table 2 is examined in terms of item difficulty index, it is determined that four items are "easy" (M15, M25, M30, M31), 16 items are "difficult" (M1, M4, M6, M8, M10, M11, M12, M14, M19, M21, M23, M26, M27, M38, M44, M47), the remaining 22 items are "moderate" difficulty. In the light of all these findings, it can be said that the item discrimination index of the NST is higher than 0.30, and there are difficult, easy, and moderate items in the test in terms of item difficulty indices. As a result of item analysis, NST consisting of 38 items was obtained.

Findings Regarding Factor Analysis

Factor analysis is a statistical technique that aims to explain the measurement of the same structure or variables that measure the same quality together with a small number of factors (Büyüköztürk, 2007). For this reason, firstly the exploratory factor analysis was performed to obtain information about the nature of the factors measured by NST and to reveal the factor structure. The first step in factor analysis was to test the suitability of the samples for factor analysis and to determine whether the data are derived from the multivariate normal distribution. For this reason, Kaiser-Meyer-Olkin (KMO) and Bartlett's test of sphericity were conducted. If the Kaiser-Meyer-Olkin (KMO) value is higher than 0.60 and Bartlett's test of sphericity is significant, it means that the data are suitable for factor analysis (Büyüköztürk, 2007). In this study, the KMO value for NST was 0.675, and the chi-square value for Bartlett's test was 2109.922 (*p* < .001). These values are accepted as an indicator that factor analysis can be performed for the NST.

a) Exploratory Factor Analysis: The exploratory factor analysis of the scale was conducted using principal component analysis and varimax orthogonal rotation method. In the factor analysis process, the criteria were that the factor loading of the items was at least 0.40 and when the items were grouped under two separate factors, the difference between the factor loads was at least 0.10. The first results of the exploratory factor analysis showed that the NST was collected under 10 factors with an eigenvalue above 1.00 and could explain 58.66% of the total variance. However, it was observed that the items collected under sub-factors except for six sub-factors with an eigenvalue higher than 1.5 were very few in number or their loads under different components were close to each other. For this reason, in the process of developing the NST, factor analysis was limited to six sub-factors by considering six number sense components. As a result of the varimax orthogonal rotation method, it was determined that four items (M41, M42, M44, and M47) in the NST had high values in more than one factor and the difference between factor load values was less than 0.10. Therefore, it was decided to exclude these four items from the test, respectively, and the factor analysis was repeated for six sub-factors over the remaining 38 items. As a result of repeated factor analysis, it was determined that the factor load values collected under six sub-factors with an eigenvalue greater than 1.5 had values between 0.44 and 0.78 and could explain 48.5% of the total variance. The variance explained by the factor loading values of the sub-dimensions of NST resulting from varimax orthogonal rotation is presented in Table 3.

Table 3.
Exploratory Factor Analysis and Orthogonal Rotation Results

Item	Common Factor Var.	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
M25	0.628	.781					
M31	0.598	.767					
M27	0.558	.739					
M30	0.524	.715					
M17	0.458	.609					
M14	0.428	.602					
M32	0.335	.498					
M19	0.324	.497					
M6	0.620		.760				
M4	0.521		.698				
M8	0.515		.663				
M11	0.417		.646				
M37	0.353		.556				
M2	0.348		.518				
M22	0.334		.440				
M18	0.533			.720			
M21	0.488			.647			
M36	0.468			.610			
M15	0.370			.593			
M9	0.384			.538			
M12	0.469			.529	.365		
M13	0.342			.525			
M38	0.550				.722		
M23	0.516				.710		
M34	0.452				.621		
M20	0.349				.546		
M10	0.459		.395		.536		
M24	0.341				.513		
M29	0.517					.676	
M3	0.483					.643	
M16	0.436					.640	
M5	0.412					.618	
M35	0.353					.499	
M26	0.503						.701

M1	0.393						.610
M28	0.439						.580
M33	0.374						.566
M7	0.399						.513
Eigenvalue:	5.703	4.305	2.921	2.222	1.720	1.560	
Explained Variance:	15.01%	11.33%	7.69%	5.85%	4.53%	4.11%	
Cronbach alpha:	.81	.79	.76	.73	.72	.75	

As a result of the principal component analysis, it was determined that the common factor variances of the items varied between 0.324-0.628, and the total variance explained by each factor after rotation was 15.01%, 11.33%, 7.69%, 5.85%, 4.53%, and 4.11%, respectively. When Table 3 is examined in terms of factor load values after factor rotation, it is determined that the first factor is between 0.497-0.781; the second factor is between 0.440-0.760; the third factor is between 0.525-0.720; the fourth factor is between 0.513-0.722; the fifth factor is between 0.499-0.676, and the sixth factor is between 0.513-0.701.

The sub-factors that emerged as a result of the exploratory factor analysis were named by considering the properties measured by the items and number sense components. Accordingly, the first sub-factor with eight items (M14, M17, M19, M25, M27, M30, M31, M32) was named as "Use of reference points in the measurement (UR)", the second sub-factor with seven items (M2, M4, M6, M8, M11, M22, M37) was named as "Understanding of the meaning of numbers (UN)", the third sub-factor with seven items (M9, M12, M13, M15, M18, M21, M36) was named as "Equivalent representation of numbers (ER)", the fourth sub-factor with six items (M10, M20, M23, M24, M34, M38) was named as "Effect of the operations on numbers (EO)", the fifth sub-factor with five items (M3, M5, M16, M29, M35) was named as "Mental calculation (MC)", and the sixth sub-factor with five items (M1, M7, M26, M28, M33) was named as "Quantitative reasoning and inference (QR)".

The correlation coefficients between the NST and the sub-factors were calculated (Table 4). As seen in Table 4, it was determined that there was a moderately positive significant relationship ($p < .01$). In addition, a positive and significant relationship was found between the sub-factors of the NST ($p < .05$).

Table 4.
Correlation Coefficients Between NST and Sub-factors

Variable	NST	QR	UN	MC	ER	UR	EO
Number sense test (NST)	1	.516**	.587**	.559**	.686**	.581**	.540**
Quantitative reasoning and inference (QR)		1	.204**	.194**	.274**	.155*	.103
Understanding the meaning of numbers (UN)			1	.296**	.264**	.144*	.097
Mental calculation (MC)				1	.292**	.134*	.196**
Equivalent representation of numbers (ER)					1	.212**	.286**
Use of reference points in the measurement (UR)						1	.212**
Effect of the operations on numbers (EO)							1

* $p < .05$, ** $p < .01$

b) Confirmatory Factor Analysis: Another way to test the construct validity of the scale is to perform the confirmatory factor analysis (Büyüköztürk, 2007). In the exploratory factor analysis, the process of finding factors is carried out based on the relationships between variables. In the confirmatory factor analysis, on the other hand, a hypothesis or theory determined before about the relationship between variables is tested (Çokluk et al., 2012; Sümer, 2000). In this study, to test whether the NST, which consists of 38 items and six sub-factors, developed with exploratory factor analysis, was applied on 225 eighth-grade students, and a first-order confirmatory factor analysis was performed using the LISREL 8.8 program (Figure 1).

Among the fit indices used to determine whether a model is accepting or not, there are chi-square (χ^2), RMSEA (Root Mean Square Error of Approximation), GFI (Goodness of Fit Index), AGFI (Adjusted Goodness of Fit Index), CFI (Comparative Fit Index) and NNFI (Non-Normed Fit Index), S-RMR

(Standardized Root Mean Square Residual) (Çokluk et al., 2012; Sümer, 2000). Chi-square/degrees of freedom (χ^2/df) ratio is less than 3 is a good value, but less than 5 is acceptable. If the RMSEA is less than 0.08, the GFI value greater than 0.90, the AGFI value greater than 0.80, the CFI value greater than 0.90, and the NNFI value greater than 0.90, the S-RMR less than 0.08, the model shows that it has a good fit (Çokluk et al., 2012; Kline, 2005). In this study, the fit index values of the first level confirmatory factor analysis of the NST are given in Table 5.

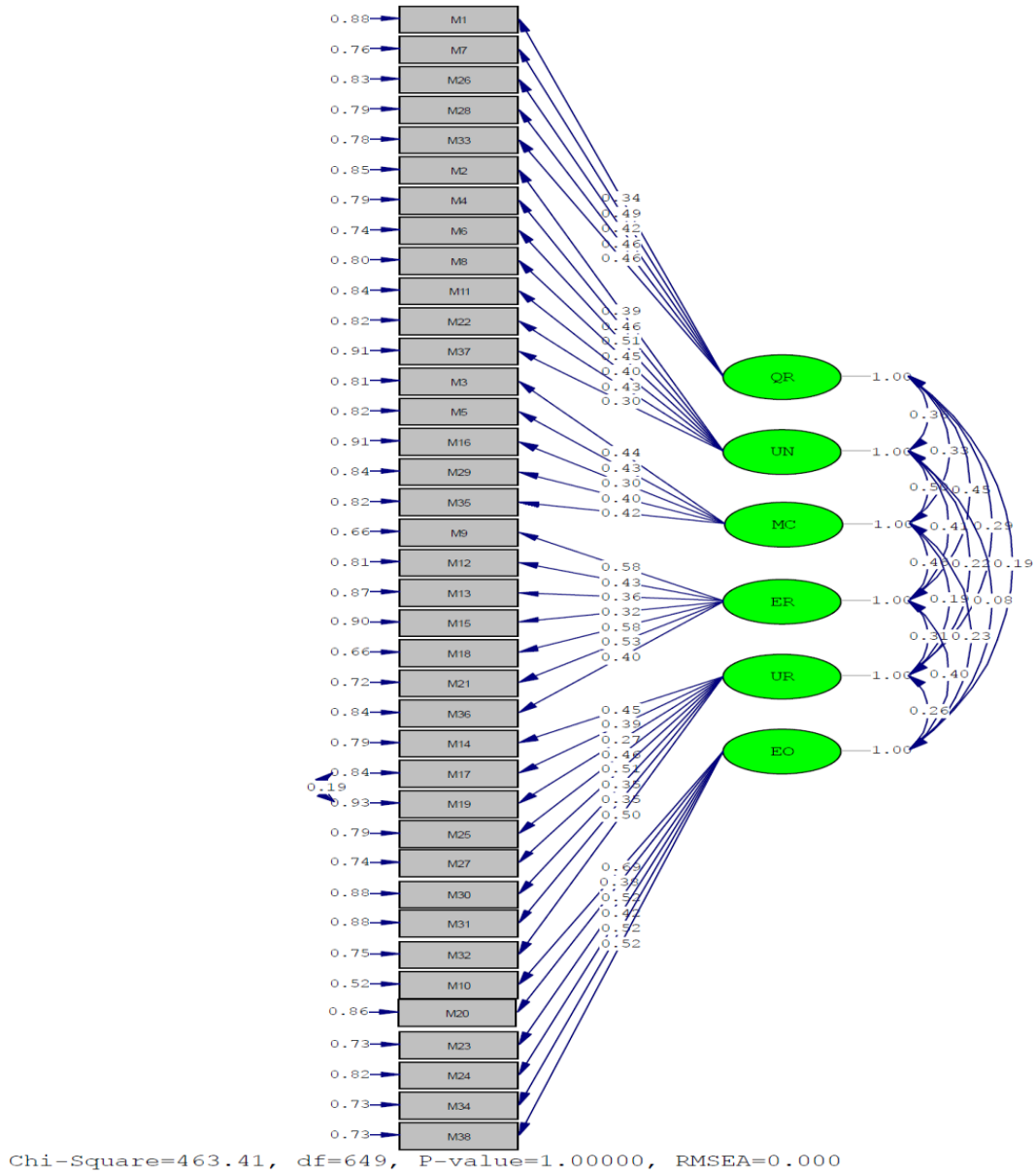


Figure 1. First-Order Confirmatory Factor Analysis for the NST (Standardized Values)

Table 5.
First-order Confirmatory Factor Analysis Results and Fit Indices for the NST

Fit Indices	Acceptable Value	Model Value
χ^2/sd	≤ 3.00	2.71
RMSEA	≤ 0.08	0.00
GFI	≥ 0.90	0.90
AGFI	≥ 0.80	0.89
CFI	≥ 0.90	1.00
NNFI	≥ 0.90	1.16
S-RMR	≤ 0.08	0.05

As seen in Table 5, the fit index values obtained from the first-order confirmatory factor analysis were calculated as $\chi^2 = 463.41$, $df = 649$, $\chi^2/df = 2.71$, $p < .001$, RMSEA = 0.00, GFI = 0.90, AGFI = 0.89, CFI = 1.00, NNFI = 1.16, S-RMR = 0.05 463.41 respectively, and these values were found to be at an acceptable level.

A second-order confirmatory factor analysis was conducted to reveal the structural relationship of the sub-dimensions that emerged as a result of the exploratory factor analysis with NST. As a result of the second-order confirmatory factor analysis (Table 6), it was determined that the fit index values were calculated as $\chi^2 = 472.08$, $df = 658$, $\chi^2/df = 2.72$, $p < .001$, RMSEA = 0.00, GFI = 0.90, AGFI = 0.89, CFI = 1.00, NNFI = 1.16, S-RMR = 0.051, respectively, and these values were at an acceptable level.

Table 6.
Results Regarding the Second-Order Confirmatory Factor Analysis of the NST

Second-order variable	First-order variables	Standardized value	t-value
NST	Quantitative reasoning and inference (QR)	0.57	3.29
	Understanding the meaning of numbers (UN)	0.56	3.62
	Mental calculation (MC)	0.63	3.74
	Equivalent representation of numbers (ER)	0.78	5.82
	Use of reference points in the measurement (UR)	0.41	3.34
	Effect of the operations on numbers (EO)	0.42	4.00

The standardized values for the sub-factors, which are the first-order variables, were determined to be 0.57 (QR), 0.56 (NK), 0.63 (MC), 0.78 (ER), 0.41 (RU) and 0.42 (EO), respectively. According to Kline (2005), the fact that these values are higher than 0.30 showed that the sub-factors are related to the test. The t-test was conducted to determine whether the standardized analysis values of each item related to the confirmatory factor analysis of the test were significant. Results of the t-test were 3.29 (QR), 3.62 (NK), 3.74 (MC), 5.82 (ER), 3.34 (RU) and 4.00 (EO), respectively, and it was found to be significant ($p < .01$). These results showed that there was a positive and significant relationship between the number sense test and its sub-factors.

Reliability: The reliability of a test is the level of the test free from accidental errors and the consistency of test items with the whole test. The reliability coefficient is expected to be above 0.70 in scales related to affective characteristics. In tests measuring cognitive skills, it is interpreted as reliability is not good for values below 0.60 (Büyüköztürk, 2007). For the reliability of the NST developed in this study, the Cronbach alpha internal consistency coefficient was found to be 0.82 for the whole test. Cronbach's alpha for sub-factors are 0.77 (NK), 0.72 (QR), 0.72 (ER), 0.76 (EO), 0.75 (MC), and 0.79 (RU), respectively. In this respect, the NST is considered to be reliable.

Discussion, Results, and Suggestions

This study aimed to develop an instrument to determine the number sense skills of eighth-grade students. To develop the number sense test (NST), firstly the literature and scales related to number sense were reviewed and six different number sense components were based in line with the gains of

the mathematics curriculum. To ensure the content validity and compatibility with the number sense components, opinions of four mathematics education experts and four mathematics teachers were taken. The draft NST consisting of a total of 47 multiple-choice items was developed by making corrections in line with their opinions. A pilot application was carried out on 24 students to test the suitability, appearance validity, and usage of NST for students. Some items were edited and revised in line with student views. For the validity and reliability of the final version of NST, item analysis and exploratory factor analysis were performed on 180 students, confirmatory factor analysis was performed on 225 students to test the construct validity.

Item-total correlation, item difficulty index, and item discrimination index were calculated for item analysis of the NST. It was decided to exclude 5 items with item-total correlation and item discrimination index below 0.20. It was determined that the remaining 42 items were between 0.30 and 0.47 in terms of item-total correlations, and were between 0.30 and 0.79 in terms of item discrimination index, and were between 0.18 and 0.69 in terms of item difficulty index, and the test consisted of items with difficult, medium and easy levels. These results showed that the item-total correlation, item difficulty index, and item discrimination index of the items in the NST were at acceptable levels (Büyüköztürk, 2007; Turgut & Baykul, 2010), and the items were interpreted as distinctive in measuring number sense skills.

For the construct validity of NST, exploratory factor analysis was performed using principal component analysis and varimax vertical rotation. As a result of the analysis, it was decided to exclude four items collected under more than one factor and lower than 0.10 among factor load values from the NST, and it was determined that the remaining 38 items could explain 48.5% of the total variance under 6 factors with eigenvalues greater than 1.5. The factor loads of the items took values between 0.44 and 0.78. It is stated that total variance ratios varying between 40% and 60% are considered sufficient in multi-factor scale structures in social sciences, and factor loads of 0.32 and above can generally be used (Çokluk et al., 2012; Tabachnick & Fidell, 2007). These results indicated that the sub-factors of NST and their items were acceptable in explaining the number sense components. As a result of the exploratory factor analysis, considering the characteristics of the items and the number sense components, the sub-factors were named as “Understanding the meaning of numbers”, “Equivalent representation of numbers”, “Effect of the operations in numbers”, “Use of reference points in the measurement”, “Mental calculation” and “Quantitative reasoning and inference”. In the literature, there is not a single theoretical framework for the number sense, the components of the number sense are defined and named in different ways. It is noteworthy that although some number sense components are similar in many studies (e.g. Facun & Nool, 2012; Greeno, 1991; Gülbağcı Dede, 2015; Kayhan Altay & Umay, 2013; McIntosh et al., 1992; Reys et al., 1999; Singh, 2009; Şengül & Gülbağcı Dede, 2013; Yang & Lin, 2015), there are different number sense components. It can be stated that this situation stems from the fact that the number sense scales differ according to age, grade level, curriculum acquisitions, and the subject of number sense (fractions, percentages, exponential numbers, natural numbers, rational numbers, etc.).

In the literature, it is stated that the sub-factors of the scale should not show a high level of relationship with each other, and the sub-factors that have a low relationship with the whole scale should be removed from the scale (Tavşancıl, 2002). In this study, it was determined that the correlation values of the sub-factor related to NST varied between 0.10 and 0.30, the correlation values between the sub-factors and NST were between 0.51 and 0.68, and there was a moderately significant positive relationship ($p < .01$). These results showed that the sub-factors were components of the NST and diverged from each other.

Another way to test the construct validity of scales is to perform confirmatory factor analysis (Çokluk et al., 2012; Tabachnick & Fidell, 2007). As a result of the confirmatory factor analysis performed using the LISREL program, it was determined that the standardized analysis values for each item were between 0.30 - 0.96, and the t-values calculated for these values varied between 3.35 and 7.95 and were significant. It was determined that the fit index values obtained as a result of confirmatory factor

analysis for NST ($\chi^2/df = 2.71$, $p < .001$, RMSEA = 0.00, GFI = 0.90, AGFI = 0.89, CFI = 1.00, NNFI = 1.16, S-RMR = 0.05) were at acceptable levels (Çokluk, et al., 2012; Kline, 2005; Sümer, 2000; Tabachnick & Fidell, 2007). In addition, as a result of the second-level confirmatory factor analysis performed for the sub-factor, the fit indexes of the model were found to be acceptable. These results showed that the factor structure of six number sense components related to NST was confirmed.

In this study, the Cronbach Alpha internal consistency coefficient for the NST was calculated as 0.82, and it was also determined that the internal consistency coefficients of the sub-factors ranged between 0.72 and 0.79. In the literature, a reliability coefficient calculated for achievement tests measuring cognitive skills of 0.60 and above is generally considered sufficient for the reliability of test scores (Büyükoztürk, 2007). These findings obtained from the reliability analysis showed that the NST had an acceptable level of reliability in terms of both the overall scale and the sub-factors.

The results of this study showed that NST, consisting of 36 multiple-choice questions and six number sense components, is a valid and reliable instrument tool for determining the number sense skills of eighth-grade students. In this respect, it is expected that this study will contribute to number sense tests, which are few in Turkey, and to future studies on number sense.

Research and Publication Ethics Statement

During the research process, all rules included in the “Directive for Scientific Research and Publication Ethics in Higher Education Institutions” have been adhered to, and none of the “Actions Contrary to Scientific Research and Publication Ethics” included in the second section of the Directive have been implemented.

Türkçe Sürümü

Giriş

Sayılar matematiğin en temel kavramları arasında yer almaktadır. Sayılar zihinsel akıl yürütme, esnek düşünme, mantıksal ilişki kurma ve günlük yaşam problemlerini çözmede önemli rol oynamaktadır (Yang, 2019). Bununla birlikte sayılar konusu matematiğin diğer öğrenme alanları olan ölçme, geometri, cebir ve veri analizinin temelini oluşturmaktadır. Matematik eğitimi için önemli bir beceri olan sayı duygusu ilk olarak Amerika'daki Ulusal Matematik Öğretmenleri Konseyinde (National Council of Teachers of Mathematics [NCTM], 1989) ortaya atılmıştır. Daha sonraki matematik eğitimi reform çalışmalarında ve öğretim programlarında da (NCTM, 2000; MEB, 2018) sayılar öğrenme alanına ve sayı duygusu becerisine önem verildiği görülmektedir.

NCTM'in (1989) raporunda sayı duygusuna sahip çocukların sayıların anlamlarını kavrama, sayılar arasında çoklu ilişkiler kurabilme, sayıların göreceli büyüklüklerini fark edebilme, işlemlerin sayılar üzerindeki etkisini anlayabilme, nesnelerin ölçümleri için uygun kıyaslama (referans) noktası belirleme becerisini geliştirebilecekleri ifade edilmektedir (s. 38). Hope'a (1989) göre sayı duygusu, sayıların çeşitli kullanım alanları hakkında mantıklı tahminler yapabilme, aritmetik hataları fark edebilme, en etkili hesaplama yolunu seçebilme ve sayı örüntülerini fark edebilme becerisidir. Greeno (1991), sayı duygusunu esnek düşünme, hesaplamada tahmin becerisi ve sayısal miktarlar hakkındaki çıkarım ile muhakeme yeteneği olarak tanımlamıştır. Reys ve diğerlerine (1999) göre sayı duygusu, sayı ve işlemlerin genel anlamının yanı sıra, sayısal durumları yönetmek, yararlı ve etkin stratejileri geliştirmek ve matematiksel yargıları oluşturmak için sayı ve işlemlerin uygun yöntemlerle kullanılması anlamına gelmektedir. Kalchman, Moss ve Case (2001), sayı duygusunu verilen büyüklüğü tahmin etme, sonuçlardaki mantıksızlığı fark etme, zihinsel olarak hesaplamada esneklik ve farklı temsiller arasındaki ilişkileri kurabilme olarak tanımlamıştır. Yang (2019) ise sayı duygusunu günlük hayatta karşılaşılan problemlerde esnek ve etkili strateji geliştirerek başa çıkabilme becerisi olarak ifade etmektedir.

Matematik öğrenmede sayılar ve aritmetik işlemleri kavramsal düzeyde anlama ile sayı duygusunun gelişimi önemli rol oynamaktadır (NCTM, 2000). Öğrencilerin erken yaşlarda sayı duygusu becerisinin gelişmesi daha sonraki yıllarda matematik öğrenmesini olumlu yönde etkilemekte, buna karşın sayı duygusundaki eksikliği matematik öğrenmeyi zorlaştırmaktadır (Jordan, Glutting ve Ramineni, 2010). Farklı öğretim kademelerinde yapılan birçok araştırmada (Akkaya, 2016; Cansız Aktaş vd., 2017; Çekirdekçi, Şengül ve Doğan, 2016; Jordan vd., 2010; Harç, 2010; Kayhan Altay, 2010; Mohamed ve Johnny, 2010; Tunalı, 2018) sayı duygusu ile matematik başarısı arasında pozitif yönlü anlamlı bir ilişki olduğu saptanmıştır. Üstelik sayı duygusunun öğrencinin matematik başarısının önemli bir yordayıcısı olduğu (Olkun, Mutlu ve Sarı, 2017; Reys ve Yang, 1998; Yang, Li ve Lin, 2008) belirlenmiştir. Türkiye'deki bazı araştırmalar (Can, 2017; Çekirdekçi vd., 2016; Er & Dinç Artut, 2017; Harç, 2010; Tunalı, 2018) matematik başarısı yüksek olan öğrencilerin düşük olan öğrencilere göre sayı duygusu becerileri bakımından daha iyi olduklarını, buna karşın genellikle kural temelli çözüm yaptıklarını göstermiştir. Bu sonuç, öğrencilerin sayı duygusu becerilerinin yeterince gelişmediğini işaret etmektedir.

Reys (1994) sayı duygusunun, sayılar, işlemler ve bunlar arasındaki ilişkileri içeren karmaşık bir süreç olması nedeniyle tanımlanmasının zor bir kavram olduğunu ifade etmektedir. Alanyazında sayı duygusu bileşenleriyle ilgili ortak bir fikir birliği yoktur. Üstelik uluslararası çalışmalarda (Greeno, 1991; Markovits ve Sowder, 1994; McIntosh, Reys ve Reys, 1992; Li ve Yang, 2010; Reys vd., 1999; Yang, 1995; 2019; Yang, Li ve Li, 2008) ve Türkiye'de yapılan çalışmalarda (Çekirdekçi, Şengül ve Doğan, 2017; İyemen İkizoğlu ve Duatepe Paksu, 2016; Kayhan Altay ve Umay, 2013; Soyuk, 2018; Şengül ve Gülbağcı, 2012) sayı duygusu bileşenleri ile ilgili çeşitli sınıflandırmaların yapıldığı görülmektedir.

Sayı Duyusu Bileşenleri

Alanyazında öğrencinin yaş grubu, sınıf düzeyi, konu ve müfredata bağlı olarak farklı sayı duyusu bileşenlerinin tanımlandığı görülmektedir. Sayı duyusu bileşenleri konusunda en kapsamlı sınıflandırma McIntosch ve diğerleri (1992) tarafından yapılmıştır. McIntosch ve diğerleri (1992) sayı duyusu becerisi için “sayıları etkili kullanabilme”, “işlemleri etkili kullanabilme” ve “sayı ile işlem bilgisiyle hesaplamalar yapabilme” şeklinde üç ana bileşen ve bu bileşenlere ilişkin alt bileşenleri oluşturmuştur. (a) Sayıları etkili kullanabilme ana bileşeni kapsamında sayıların düzenli olması, göreceli ve mutlak büyüklüğünü algılama ve ölçüm referansları alt bileşenleri yer almaktadır. (b) İşlemleri etkili kullanabilme ana bileşeni kapsamında işlemlerin etkilerini anlama, matematiksel özellikleri anlama, işlemler arasındaki ilişkileri anlama becerileri yer almaktadır. (c) Sayı ve işlem becerisiyle mantıklı hesaplamalar yapabilme ana bileşeni kapsamında problem içeriği ve içerisindeki ilişkiyi anlama, birden fazla stratejinin var olduğunu fark edebilme, etkili bir temsil ya da metot kullanma, veri ve sonuçları mantıklı biçimde gözden geçirme bileşenleri yer almaktadır.

McIntosch ve diğerleri (1992) tarafından ortaya atılan bileşenler temel alınarak Reys ve diğerleri (1999) tarafından beş sayı duyusu bileşeni oluşturulmuştur. Bu bileşenler kısaca şöyledir:

(a) *Sayıların anlam ve büyüklüklerini anlama*: Sayıların göreceli büyüklüğünü fark etmeyi, verilen iki kesri karşılaştırabilmeyi, sıralayabilmeyi ve iki sayı arasındaki sayıyı bulmayı gerektirir. Örneğin; 1,52 ile 1,53 arasında ne kadar sayı olduğunu fark edebilme, $\frac{2}{5}$ kesri ile $\frac{1}{2}$ kesrini büyüklük bakımından kıyaslayabilme.

(b) *Sayıların denk gösterimlerini anlama ve kullanma*: Sayıların eşdeğerlerini ve denk gösterimlerini bilmeyi ve gerektiğinde bunu kullanabilmeyi ifade etmektedir. Örneğin, $\frac{2}{5}$ kesrinin $\frac{4}{10}$, %40 ya da 0,4 gibi farklı gösterim biçimleri ile ifade edebilme, bir m sayısının hangi sayı ile çarpımının 0,25 ile bölümüyle aynı sonucu vereceğini fark edebilme.

(c) *İşlemlerin sayılar üzerindeki etkisini anlama*: Bu bileşen hesaplama durumlarında bir sayının veya işlemin değeri değiştiği zaman sonucun nasıl değişeceğini fark etme becerisi ile ilgilidir. Yani, çarpma işleminin daima sayıları büyütmeceği ve bölme işleminin sayıları daima küçültmeceğini hissedebilmeyi ifade eder. Örneğin, $70 \div 0,5$ işleminin sonucunun 70×2 işlemine eşit olup olmayacağını, ya da $750 \div 0,98$ işleminin sonucunun 750’den büyük mü yoksa küçük mü olabileceğini tahmin edebilme.

(d) *Ölçmede kıyaslama (referans noktası kullanma)*: Bu beceri farklı durumlara uygun olabilecek referans/kıyaslama noktalarını belirlemeyi ve kullanmayı içermektedir. Örneğin, öğrencinin kendi boyunu ya da bir nesnenin uzunluğunu referans olarak bir başka nesnenin uzunluğunun tahmin edebilmesi; 1, $\frac{1}{2}$ ve $\frac{1}{4}$ gibi sabit noktaları referans noktası alınarak kesirlerin ve sayıların sıralanması.

(e) *Zihinden hesaplama ve hesaplamada esneklik*: Kağıt kaleme dayalı hesaplama yapmaksızın problem çözebilme ve sonucun uygunluğunu sorgulamak için tahmin etme, zihinden işlem yapabilmeyi gerektirir. Örneğin, 6×98 işleminin yaklaşık değerini zihinden hesaplayabilme; “Ali, beşinci sınıf öğrencisidir. Ali, 30.000 gün yaşadığını söylüyor.” Ali’nin bu önermesinin mantıklı olup olmadığını karar verebilme.

Yang (1995) tarafından geliştirilen sayı duyusu sınıflaması altı bileşenden oluşmaktadır. Bu bileşenler kısaca şöyledir: (a) *Sayıların anlamını kavrama* (Örneğin, bu bileşensayıların temel anlamını, onluk sayma sistemini, basamak değeri kavramını ve sayı örüntülerini anlamayı gerektirir. İlkokul birinci sınıf bir öğrencinin 48 sayısını “4 düzine yumurta” ya da “4 onluk ve 8 birlik” şeklinde ifade edebilmesi gibi), (b) *Sayıları ayırıştırma ve yeniden birleştirme* (Örneğin, $3+3+3+3$ işleminin sonucunun 4×3 işlemi ile aynı olduğunu fark edebilme), (c) *Sayı büyüklükleri* (Örneğin, iki kesrin karşılaştırılması, verilen sayıları uygun bir şekilde sıralayabilme), (d) *Kıyaslama (referans) noktası kullanımı* (Örneğin, 0,95 sayısı için en yakın değeri olan 1’in; verilen bir başka kesir için $\frac{1}{4}$, $\frac{1}{2}$ ve $\frac{1}{5}$ gibi bilinen bir noktanın referans alınması), (e) *İşlemlerin sayılar üzerindeki etkisi* (Örneğin, çarpmanın her zaman sonucu büyütmediğini bilmek, $199 \times 0,9$ ’un sonucunun 199’da küçük olması gibi), (f) *Sayı ve işlem bilgisini hesaplama durumlarına*

uygulamadaki esneklik (Örneğin, bir okul otobüsü 45 öğrenciyi taşımaktadır. Eğer 915 öğrenci müzeye götürülmek istenirse, kaç tane otobüse ihtiyaç olur?).

Yang, Li ve Li (2008) ilköğretim üçüncü sınıf öğrencileri için çoktan seçmeli sorulardan oluşan ve cevapların gerekçesini isteyen iki aşamalı sayı duyusu testi geliştirmişlerdir. Beş bileşenden oluşan bu testin sayı duyusu bileşenleri sırasıyla *“Sayıların ve işlemlerin anlamını anlama”*, *“Göreceli sayı büyüklüğünü kavrama”*, *“Sayıları birleştirme ve ayırıştırma”*, *“Sayıların ve işlemlerin çoklu gösterimlerini kullanabilme”* ve *“İşlemsel sonuçların akla uygunluğunu yargılama”* şeklindedir. Yang (2019) Tayvan’daki beşinci sınıf öğrencileri üzerinde yaptığı başka bir çalışmada beş sayı duyusu bileşeni içeren üç aşamalı test geliştirmiştir. Bu testte daha önceki çalışmasından (Yang vd., 2008) farklı olarak *“Sayıları birleştirme ve ayırıştırma”* bileşeni yerine, *“Sayılar üzerinde işlemlerin etkisini fark etme”* bileşeni yer almaktadır.

Uluslararası alanyazında sayı duyusu konusunda yaklaşık otuz yıldır birçok çalışma yapılmasına rağmen, Türkiye’de sayı duyusu becerileri son yıllarda önem kazanmıştır. Nitekim, Birgin ve Peker (2021) yaptıkları çalışmada 2000-2018 yılları arasında Türkiye’de sayı duyusu alanında yapılan çalışmaları tematik içerik analizi ile incelemişlerdir. Bu çalışma sonucunda Türkiye’de sayı duyusu alanındaki araştırmaların 2010 yılından itibaren hız kazandığı, sayı duyusuna yönelik ölçek geliştirme çalışmalarının sınırlı olduğu ve yapılan çalışmalarda Kayhan Altay ve Umay (2013) tarafından geliştirilen sayı duyusu sınıflandırmasının daha çok kullanıldığı belirlenmiştir. Kayhan Altay ve Umay (2013) ortaokul altıncı, yedinci ve sekizinci sınıf öğrencilerine yönelik yazılı cevap gerektiren 17 soru ve 3 bileşenden oluşan sayı duyusu testi geliştirmişlerdir. Bu testin sayı duyusu bileşenleri kısaca şöyledir: *“Hesaplama esneklik”* (0,25 x 16 işlemi kısa yoldan nasıl çözersiniz? Nasıl yaptığınızı gösterin.) *“Kesirlerde kavramsal düşünme”* (Verilen bir sayı doğrusundaki noktaları düşünerek, $\frac{1}{2}$, $2\frac{1}{2}$ ve $\frac{1}{4}$ kesirlerini yerleştiriniz. Nasıl yerleştirdiğinizi açıklayınız.) ve *“Referans noktası kullanımı”* (86424 x 500 işlemi kısa yoldan nasıl çözersiniz? Nasıl düşündüğünüzü gösteriniz.). İymen İkizoğlu ve Duatepe Paksu (2016), üslü ifadeler konusuyla ilgili ortaokul sekizinci sınıf öğrencilerine yönelik 13 yazılı cevap gerektiren ve beş sayı duyusu bileşenini içeren sayı duyusu testi geliştirmişlerdir. Bu testte yer alan sayı duyusu bileşenleri *“denk gösterim”*, *“sayısal tahmin”*, *“sayı büyüklüğü”*, *“işlem etkisi”* ve *“referans noktası”* şeklindedir. Diğer taraftan Çekirdekçi, Şengül ve Doğan (2017), ilkokul dördüncü sınıf öğrencilerine yönelik 11 çoktan seçmeli soru ve üç bileşenden (*“Sayıların eşdeğerlerini bilme ve niceliksel muhakeme-çıkarımda bulunma”*, *“Referans noktası kullanarak işlemlerin etkilerini hesaplama”* ve *“Sayıların anlamını bilme ve esnek düşünme”*) oluşan sayı duyusu testi geliştirmişlerdir. Ulusal alanyazındaki sayı duyusu testlerinin odaklandıkları sınıf düzeyleri ile odaklandıkları konu ve sayı duyusu bileşenlerinin farklılık gösterdiği söylenebilir.

Çalışmanın Önemi ve Amacı

Alanyazında sayı duyusuna ilişkin sınıflamalar incelendiğinde sayı duyusu bileşenleri için ortak bir terminoloji oluşturulamadığı, aynı beceriyi ifade eden sayı duyusu bileşenleri için farklı isimlendirmeler (Greeno, 1991; Markovits ve Sowder, 1994; Yang, 1995) kullanıldığı görülmektedir. Bazı çalışmalarda bazı sayı duyusu bileşenleri (Referans noktası kullanımı, işlemlerin etkisi, sayıların anlamı ve büyüklüğü gibi) ortak olmasına karşın (İymen İkizoğlu ve Duatepe Paksu, 2016; McIntosh vd., 1992; Reys vd., 1999; Yang, 1995; 2019; Yang vd., 2008), sayıların temel anlamını kavrama (Li ve Yang, 2010), sayıları birleştirme ve ayırıştırma (Reys vd., 1999; Yang, 2019), niceliksel muhakeme ve çıkarım (İymen İkizoğlu ve Duatepe Paksu, 2016), sayıların çoklu gösterimi (Kayhan Altay ve Umay, 2013; Markovits ve Sowder, 1994) gibi sayı duyusu bileşenleri yer almamaktadır. Ayrıca sayı duyusu ölçeklerinde yer alan bileşenlerin öğrencinin yaş grubu ve öğretim programlarındaki konu içeriklerine bağlı olarak farklılık gösterdiği ve sayı duyusu bileşenlerine ilişkin önceliğin değiştiği ifade edilebilir. Türkiye’de ilkokul ve ortaokul matematik dersi öğretim programları (MEB, 2018) incelendiğinde sekizinci sınıf düzeyindeki öğrencilerin doğal sayılar, tamsayılar, rasyonel sayılar, irrasyonel sayılar ile sayıların farklı gösterimine (ondalık, üslü, yüzde, vb.) ilişkin sayı duyusu becerilerine sahip olmaları beklenmektedir. Yapılan çalışmalar (Birgin ve Peker, 2021; Çekirdekçi vd., 2017; İymen İkizoğlu ve Duatepe Paksu, 2016) Türkiye’de sayı duyusu becerisine yönelik testlerin ve ölçeklerin sınırlı sayıda olduğunu göstermektedir. Üstelik bu sayı duyusu testleri ve ölçeklerinin odaklandıkları konu ve sınıf düzeyi bakımından sekizinci sınıf öğrencilerinin

tamsayılar, rasyonel, irrasyonel sayılar ile sayıların farklı gösterimlerine (ondalık, üslü, vb.) ilişkin genel sayı duygusunu belirlemede yeterli olmadığı söylenebilir. 2013 yılı sonrasında güncellenen matematik dersi öğretim programlarındaki (MEB, 2018) içerik, konu ve kazanımlardaki değişimler düşünüldüğünde sekiz yıllık eğitim sonucunda öğrencilerdeki genel sayı duygusu becerilerini belirlemeye yönelik ölçme araçlarına ihtiyaç vardır. Bu nedenle bu çalışma, sekizinci sınıf öğrencilerinin sayı duygusu performanslarını belirlemeye yönelik test geliştirmeyi amaçlamıştır. Bu çalışma, önceki sayı duygusu testlerinden farklı olarak sekiz yıllık eğitim sonunda öğrencilerin tamsayılar, rasyonel ve irrasyonel sayılar ile sayıların farklı gösterimlerine ilişkin sayı duygusu becerilerini içermekte ve birçok sayı duygusu ölçeğinde ifade edilen sayı duygusu bileşenlerini (Sayı bilgisi, niceliksel muhakeme ve çıkarım, sayıların denk gösterimi, sayılarda işlemlerin etkisi, ölçmede referans noktası kullanma, zihinsel hesaplama gibi) temel almaktadır. Bu yönüyle bu çalışmanın sayı duygusu alanında gelecekte yapılacak araştırmalara katkı sağlaması düşünülmektedir.

Yöntem

Çalışma Grubu

Bu araştırmanın temel amacı ölçme aracının geliştirmek olduğundan uygun örnekleme (convenience sampling) yöntemi tercih edilmiştir. Uygun örnekleme yöntemi, yakın çevrede bulunan, ulaşılması kolay ve araştırmaya gönüllü olarak katılmak isteyen bireylerden seçilen örnekleme olarak ifade edilmektedir. Ölçme aracı geliştirmeyi amaçlayan çalışmalarda ölçülen özelliğin ranjının örneklenmesinin önemli olduğu ve katılımın gönüllük esasına dayanması gerektiği vurgulanmaktadır (Erkuş, 2012). Bu bağlamda ölçme araçlarının uygulanacağı okullara ilişkin Milli Eğitim Müdürlüğü'nden gerekli yazılı kurumsal izinler alınmıştır. Ayrıca bu araştırma etik açıdan Uşak Üniversitesi Sosyal ve Beşeri Bilimler Bilimsel Araştırma ve Yayın Etiği Kurulu tarafından onaylanmıştır. Katılımcılara araştırma hakkında bilgilendirme yapılmış ve katılımcılardan bilgilendirilmiş onam alınmıştır. Bu araştırma, gönüllü katılım esasına göre yürütülmüştür. Bu araştırmanın çalışma grubunu Türkiye'nin batı bölgesindeki bir ilin farklı yerleşim yerlerinde (il merkezi, ilçe merkezi ve kırsal) öğrenim gören 405 sekizinci sınıf öğrencisi oluşturmaktadır. Bu araştırmaya madde analizi ve açımlayıcı faktör analizi için 180 öğrenci, doğrulayıcı faktör analizi için 225 öğrenci katılmıştır. Katılımcıların %48.4'ü (n=196) kız öğrenci, %51.6'ı (n=209) erkek öğrencidir. Öğrencilerin %40.7'i (n=165) il merkezi, %33.1'i (n=134) ilçe merkezi ve %26.2'i (n=106) kırsalda öğrenim görmektedir.

Sayı Duyusu Testinin (SDT) Geliştirilme Süreci

Sayı duygusu testinin (SDT) geliştirilme süreçleri ve bu süreçte yapılan işlemler Tablo 1'de özetlenmiştir. SDT'nin tasarlanması sürecinde öncelikle sayı duygusunun özelliklerinin ve sayı duygusu bileşenlerini belirlemek amacıyla yurt dışı (Facun ve Nool, 2012; Greeno, 1991; Li ve Yang, 2010; McIntosh vd., 1992; Purnomo, Kowiyah, Alyani ve Assiti, 2014; Li ve Yang, 2010; Reys vd., 1999; Singh, 2009; Yang, 1995; 2019; Yang ve Lin, 2015) ve yurt içi (Gülbağcı Dede, 2015; Harç, 2010; Kayhan Altay ve Umay, 2013; Soyuk, 2018; Şengül ve Gülbağcı Dede, 2013) alanyazın incelemesi yapılmıştır. SDT'nin sekizinci sınıf öğrencilerine uygun olması için 2018 Matematik Dersi Öğretim Programının (MEB, 2018) sayılar öğrenme alanı ve ilgili kazanımları gözden geçirilmiştir. Matematik dersi öğretim programının kazanımları dikkate alınarak alanyazındaki sayı duygusu tanımlamalarında ve sayı duygusu ölçeklerinde yaygın olarak kullanılan altı sayı duygusu bileşeninin temel alınmasına karar verilmiştir. Bu sayı duygusu bileşenleri sırasıyla (a) *Sayı bilgisi*, (b) *Niceliksel muhakeme ve çıkarım*, (c) *Sayıların denk gösterimi*, (d) *Sayılarında işlemlerin etkisi* (e) *Ölçmede referans noktası kullanma* ve (f) *Zihinsel hesaplama* şeklindedir.

SDT'deki maddelerin oluşturulmasında alanyazın çalışmalarda kullanılan sorulardan, ulusal ve uluslararası sınav sorularından (TEOG, SBS, LGS, TIMMS, vb.), matematik ders kitabı ve kaynak kitaplarındaki sorulardan yararlanılmıştır. SDT'de yer alan maddelerin bazıları ilgili kaynaklardan aynen alınmış, bazı maddeler araştırmacılar tarafından değiştirilmiş ya da yeniden tasarlanmıştır. Maddelerin geliştirilmesi sürecinde kâğıt ve kalem kullanmadan sayı duygusu becerilerini kullanmayı gerekli kılmaması ve zaman kısıtlamasının olması temel alınmıştır. Çoktan seçmeli maddelerin yazılması sürecinde ilgili

alanyazın çalışmalarında ifade edilen öğrenci hata ve kavram yanlışları da dikkate alınmıştır. Bu bağlamda SDT için taslak çoktan seçmeli madde havuzu oluşturulmuştur.

Tablo 1.

Sayı Duyusu Testini Geliştirme Aşamaları ve Yapılan İşlemler

Test Geliştirme Aşamaları	Yapılan İşlemler
1. Sayı Duyusu Bileşenlerinin Tespiti ve Belirlenmesi	<ul style="list-style-type: none"> Alanyazın taraması, sayı duyusu ölçek ve testler ile matematik dersi öğretim programının incelenmesi
2. Madde Havuzunun Oluşturulması	<ul style="list-style-type: none"> Alanyazın, sayı duyusu ölçekleri, matematik dersi öğretim programı, ders kitabı ve merkezi sınav soruları (SBS, TEOG, LGS, TIMSS, vb.) referans alınmıştır.
3. Kapsam ve Görünüş Geçerliğinin Sağlanması	<ul style="list-style-type: none"> Taslak maddelerin dil, öğrenci düzeyi ve sayı duyusu bileşenine uygunluğuna ilişkin alan uzmanların ve öğretmenlerin görüşleri alınması Görüş ve öneriler doğrultusunda maddelerde gerekli ekleme ve düzeltmeler yapılması Taslak maddelerin öğrenciye uygunluğu için 24 öğrenci üzerinde pilot uygulama yapılması ve görüşlerinin alınması
4. Hazırlanan Sayı Duyusu Testinin Uygulanması	<ul style="list-style-type: none"> Hazırlanan taslak testin toplam 405 sekizinci sınıf öğrencisi üzerinde uygulanması
5. Madde Analizinin Yapılması	<ul style="list-style-type: none"> Madde güçlüğü ve madde ayırt ediciliğinin hesaplanması, uygun olmayan maddelerin çıkartılması
6. Yapı Geçerlik Analizlerinin Yapılması	<ul style="list-style-type: none"> KMO ve Bartlett küresellik testinin yapılması Açımlayıcı faktör analizi yapılması Doğrulamalı faktör analizinin yapılması
7. Güvenirlik Hesabı ve Testin Son Şeklinin Verilmesi	<ul style="list-style-type: none"> Cronbach alpha iç tutarlılık güvenirlilik katsayısı hesaplanması Analizler sonucunda sayı duyusu testinin son şeklinin verilmesi

Hazırlanan taslak SDT'nin kapsam ve görünüş geçerliğini sağlamak amacıyla öncelikle dört matematik alan eğitimi uzmanı tarafından içerik (kapsam), sayı duyusu bileşeni ve öğretim programına uygunluk, dil ve ifade, öğrenci düzeyine uygunluk bakımından incelenmiş ve uzman görüşleri doğrultusunda maddelerde çeşitli düzenlemeler yapılmıştır. Ayrıca matematik eğitimi alanında yüksek lisans yapan dört matematik öğretmenin de görüşleri alınmış ve öneriler doğrultusunda maddeler üzerinde bazı düzeltmeler yapılmıştır. Taslak SDT'nin öğrenci düzeyine uygunluğunu ve uygulama süresini test etmek için 24 sekizinci sınıf öğrencisine uygulanmıştır. Test maddelerinde anlaşılmayan ifade ve görsel olup olmadığını belirlemek amacıyla bu öğrencilerin görüşleri alınmıştır. Öğrenciler bazı soruların zor, bazılarının ise oldukça kolay olduğunu ve verilen sürenin yeterli olduğunu belirtmişlerdir. Öğrenci görüşleri doğrultusunda testteki bir görsel daha belirgin hale getirilmiştir. Revize edilen taslak SDT, altı sayı duyusu bileşenine yönelik toplam 47 çoktan seçmeli sorudan oluşmaktadır.

Sayı Duyu Testinin Sayı Duyusu Bileşenleri ve Özellikleri

SDT'de yer alan maddelerden bazıları Ek 1'de örnek olarak sunulmuştur. SDT'deki maddeler ve sayı duyusu bileşenlerine göre dağılımı ve özellikleri kısaca şöyledir:

a) Sayı bilgisi (Sayıların anlamını kavrama): Sayı bilgisi bileşeni sayı büyüklüğünü anlamayı, sayıların karşılaştırmayı ve sıralamayı, sayılar ve sayıların basamak değerleri arasındaki ilişkiyi anlamayı ifade etmektedir (McIntosh vd., 1997). SDT'deki M2, M4, M6, M8, M11, M22, M37 ve M40 maddeleri sayıları anlama ve sayı bilgisi kapsamında hazırlanmıştır. Örneğin, SDT'deki M6 maddesi Singh (2009) ile Purnomo ve diğerlerinin (2014) çalışmalarından esinlenilerek yazılmıştır. Bu soruda öğrencilerin 1,52 ile 1,53 arasındaki sayılar için bu sayıları 1,520 ile 1,530 ya da 1,5200 ile 1,5300 gibi düşünülerek bu sayılar arasında sonsuz ondalık kesir olabileceğini fark etmeleri beklenmektedir.

b) Niceliksel muhakeme ve çıkarım: Bu bileşen, hesaplama sonucunda elde edilen sayısal değer mantıklı olup olmadığına ilişkin bir yargıda ve çıkarımda bulunmayı içerir. Bu bileşen için öğrencinin elde ettiği sayısal değer ne ifade ettiğini, mantıklı olup olmadığını bilmesi gerekir (Kayhan Altay ve Umay, 2013). SDT'deki M1, M7, M26, M28, M33, M39 ve M44 maddeleri niceliksel muhakeme ve çıkarım

bileşeni bağlamında hazırlanmıştır. Örneğin, bu araştırma kapsamındaki M1 maddesi 2007 TIMMS matematik testinden alınmıştır. Bu soruda “Okulun düzenlediği gezilere her 12 öğrenci için 1 öğretmen katılacaktır. Geziye 110 öğrenci gidiyorsa en az kaç öğretmen katılmalıdır?” sorusu için 110 öğrenci 12’ye bölündüğünde 9 öğretmen gerekmektedir. Fakat kalan 2 öğrenci için de bir öğretmen gerekmektedir. Öğrencinin bunu fark etmesi beklenmektedir.

c) Sayıların denk gösterimi: Sayıların denk gösterimi sayıların birçok farklı temsil biçimlerini içerir. Bu bileşen, sayıların eş değerini bulmayı, ayrıştırma ve birleştirme ile sayıları yeniden oluşturabilmeyi ve sayıların çeşitli temsil biçimleri (yüzde, üstel, ondalık, taralı bölge ve sayı doğrusundaki gösterimi gibi) arasında geçiş yapmayı gerektirir (McIntosh vd., 1997). SDT’deki M9, M12, M13, M15, M18, M21, M36 ve M45 maddeleri sayıların denk gösterimi bileşenine bağlamında hazırlanmıştır. Örneğin, M12 maddesi 2015 TEOG sınavı matematik testinden alınmıştır (Ek-1). Bu soruda 5^{-4} ifadesinin değerinin $\frac{1}{625}$ olduğu düşünülüp ondalık kesre çevrilmesi beklenmektedir. Testteki M21 maddesi geometrik şekli verilen taralı bir bölgenin yüzde olarak ifade edilmesini gerektirmektedir.

d) İşlemlerin etkisi: İşlemlerin etkisi bileşeni belirli sayılarla ilgili işlemlerin etki ve anlamını içerir. Örneğin bölmenin sayıyı eşit alt gruplara ayırması veya sayıyı 1’den daha küçük bir sayı ile çarparak kendinden daha küçük sayı elde edilmesi gibi. SDT’deki M10, M20, M23, M24, M34, M38, M43, M46 maddeleri sayılarda işlemlerin etkisi bileşenine ilişkin hazırlanmıştır. Örneğin, M24 maddesi ($\frac{2}{3} \times \frac{5}{4} = ?$ işleminin sonucunu hesaplama yapmadan en yakın cevabı işaretleyiniz.) ve M38 maddesi ($29 \div 0,8 = ?$, $29 \times 0,8 = ?$, $29 + 0,8 = ?$ ve $29 - 0,8 = ?$ işlemlerinden hangisi en büyük sonucu verir?) çarpma işleminin çarpılan sayıyı büyüttüğü, bölme işleminin ise sayıyı küçülttüğü kuralına uymayan bir sorudur. Sayı duyusu gelişmiş bir öğrencinin bir sayı ile 1’den daha küçük ondalık sayıyı çarpıldığında sayının kendisinden daha küçük sayı elde edildiği cevabını vermesi beklenmektedir.

e) Ölçmede referans noktası kullanımı: Bu sayı duyusu bileşeni standart ve standart olmayan ölçüm birimlerini günlük yaşamla uyumlu ve mantıklı olarak kullanabilme yeteneğidir. Bu sayı duyusu bileşeni kütle, uzunluk, hacim, zaman ve açı gibi özellikleri kullanarak tahmini olarak ölçme yapabilmeyi de gerektirmektedir (McIntosh vd., 1997). SDT’deki M14, M17, M19, M25, M27, M30, M31, M32 ve M41 maddeleri ölçmede referans noktası kullanımı bileşeni kapsamında hazırlanmıştır. Örneğin SDT’deki M27 maddesi Gülbağcı Dede (2015) ve Harç (2010) tarafından yapılan çalışmadan alınmıştır. Payda eşitleme kuralı uygulanmadan yarım referans noktası kullanılarak yarımından büyük iki kesrin toplamının 1’den büyük olduğu fark edilmelidir. Testin M19 maddesinde (Ek-1) sayı duyusu gelişmiş öğrencinin AB doğru parçasını referans noktası olarak L şeklinin çevresinin K şeklinden büyük olduğunu bulması beklenmektedir.

f) Zihinden Hesaplama: Zihinden hesaplama bileşeni hesaplamada kâğıt-kalem ve hesap makinesi kullanmadan çözümü uygularken tahmin yapıp zihinden hesaplama yaparak sonuca ulaşmayı içerir. SDT’deki M3, M5, M16, M29, M35, M44, M47 maddeleri zihinden hesaplama bileşeni kapsamında hazırlanmıştır. Örneğin, M3 maddesi 2011 TIMMS matematik testinden alınmıştır. Bu soruda 7,21 ondalık kesri 7’ye 3,86 ondalık kesri 4’e ve 10,09 ondalık kesri 10’a yuvarlanarak sonuca ulaşılır. M16 maddesi 2010 SBS matematik testinden esinlenilerek yazılmıştır. Bu soruda alanı $150 m^2$ olan karenin bir kenarının uzunluğunun $\sqrt{150}m$ ve bu değer 12 m ile 13 m arasında olduğunun tahmin edilmesi beklenmektedir.

Verilerin Toplanması ve Veri Analizi

Bu çalışmada SDT ve uygulama konusunda katılımcılara bilgilendirme yapılmıştır. Tüm katılımcılardan bilgilendirilmiş onam alınmış ve veriler gönüllük esasına göre toplanmıştır. SDT’deki her bir soru projeksiyon ya da akıllı tahta teknolojisi kullanılarak çalışma grubundaki öğrencilere sunulmuş, kâğıt ve kaleme dayalı aritmetik bir hesaplama yapmadan sayı duyusu becerilerini kullanarak 35-40 sn içinde en uygun cevabı belirlemeleri istenmiştir. Böylece her bir soru birbirinden bağımsız şekilde tamamlamaları sağlanmıştır. SDT’deki her bir doğru cevap için 1 puan, yanlış ve boş cevaplar için sıfır puan verilerek toplam puanlar elde edilmiştir. Veri analizi öncesinde analizlerin uygunluğu ve

varsayımların kontrolü için boş verilerin değerlendirilmesi, normallik testi, doğrusalılık ve uç değerlerin belirlenmesi (standart sapmanın -3 ile +3 değerleri dışında kalan) işlemleri yapılmıştır. SDT'ye ilişkin boş verisi fazla olanlar ile uç veriler analizden çıkartılmıştır. SDT puanlarının basıklık (-0.596) ve çarpıklık (0.356) değerlerinin +1 ile -1 arasında değer aldığı ve normal dağılım gösterdiği belirlenmiştir.

SDT'nin geçerlik ve güvenirlik çalışmaları iki aşamada yürütülmüştür. İlk aşamada SDT'nin madde analizleri kapsamında madde güçlük indeksi ve madde ayırt edicilik indeksini belirlemek amacıyla çalışma grubunu oluşturan okullardan seçilen 180 öğrenci ile gerçekleştirilmiştir. Madde analizinde SPSS istatistik programı ve Excel programı kullanılmıştır. Elde edilen veriler üzerinden her bir maddenin madde-toplam korelasyon, madde güçlük indeks ve madde ayırt edicilik indeks değeri hesaplanmıştır. Madde ayırt edicilik indeksinin belirlenmesinde %27'lik alt ve %27'lik üst grup yönteminden yararlanılmıştır. Madde analizi sonucunda uygun olmayan maddelerin elemesi yapılmıştır. SDT'nin yapı geçerliğine uygunluğunu belirlemek amacıyla öncelikle Kaiser-Meyer-Olkin KMO örneklem uygunluğu testi ve Bartlett küresellik testi hesaplanmıştır (Büyükoztürk, 2007). Sonrasında Temel Bileşenler Analizi ve Varimax Dik Döndürme Tekniği kullanılarak açımlayıcı faktör analizi gerçekleştirilmiştir. Faktör analizi sonuçlarına dayalı olarak madde elemesi yapılarak testin alt boyutları belirlenmiştir. Ayrıca SDT'nin tümü ve her bir alt boyutun güvenirliği için ilişkin Cronbah alpha iç tutarlılık katsayısı hesaplanmıştır.

SDT'nin ikinci aşamasında yapı geçerlik çalışması kapsamında açımlayıcı faktör analizi sonucunda ortaya çıkan sayı duygusu bileşenlerinin (yapının) sayı duygusu testi ile model uyumunu belirlemek amacıyla 225 sekizinci sınıf öğrenci üzerinde LISREL 8.8 programı yardımıyla doğrulayıcı faktör analizi gerçekleştirilmiştir.

Bulgular

Madde Analizlerine İlişkin Bulgular

Madde analizi ile ölçekteki maddelerin ölçmeyi amaçladığı özelliği başka özellikleri karıştırmadan ölçüp ölçmediği belirlenmektedir (Tavşancıl, 2002). Madde analizi kapsamında madde-toplam korelasyonu, madde güçlük indeksi (p_i) ve madde ayırt edicilik indeks (r_{jk}) değerleri hesaplanmıştır. Alanyazında madde-toplam korelasyon değerinin 0.20 ve üzerinde olması gerektiği belirtilmektedir (Kline, 1993).

Madde güçlük indeksi (p_i), soruya doğru cevap veren öğrencilerin sayısının, tüm öğrencilerin sayısına oranıdır. Madde güçlük indeksi 0 ile 1 arasında değerler almaktadır. Madde güçlük indeksi 0,20'den küçük olan maddelerin "çok zor", 0,21 - 0,40 arasında olan maddelerin "zor", 0,41 - 0,60 arasında olan maddelerin "orta düzey zorlukta", 0,61 - 0,80 arasında olan maddelerin "kolay" ve 0,80'den büyük olan maddeler ise "çok kolay" olduğu belirtmektedir. Testlerde maddelerin orta güçlükte (0,50) olması istenilen bir düzeydir ve testin geçerlilik ve güvenirliğini artırmaktadır (Turgut ve Baykul, 2010).

Madde ayırt edicilik indeksi (r_{jk}), madde ile ölçülmesi amaçlanan özelliğe sahip olan ve olmayanları birbirinden ayırabilme gücüdür (Büyükoztürk, 2007). Bu indeks -1 ile +1 arasında değişmektedir. Madde ayırt edicilik indeksi 0.20'den düşük ve negatif olan maddelerin testten çıkartılması ve 0,20-0,29 arasında olan maddelerin düzeltilmesi gerektiği, madde ayırt edicilik indeksi 0,30-0,39 olan maddelerin "iyi", 0,40 ve üzerinde olan maddelerin ise "çok iyi" düzeyde ayırt edici madde olduğu ifade edilmektedir (Turgut ve Baykul, 2010). Bu araştırma kapsamında SDT'deki maddelere ilişkin madde güçlük indeksi, madde ayırt edicilik indeksi ve madde-toplam korelasyon değerleri Tablo 2'de sunulmuştur.

Tablo 2 madde-toplam korelasyon değerleri bakımından incelendiğinde iki maddenin (M39, M43) 0.132 ile 0.174 değerini aldığı, üç maddenin (M40, M45, M46) 0.203 ile 0.233 arasında, diğer maddelerin 0.304 ile 0.471 arasında değer aldığı belirlenmiştir. Bu nedenle madde-toplam korelasyon değeri 0.20'den düşük olan M39 ve M43 maddelerinin SDT'nin çıkartılmasının uygun olacağı kanaatine varılmıştır. Ayrıca Tablo 2 madde ayırt edicilik indeks kriterleri bakımından incelenmiş SDT'deki beş maddenin (M39, M40, M43, M45 ve M46) ayırt edicilik indeksi 0.20'den düşük olması nedeniyle testten çıkartılmasına karar verilmiştir. Geri kalan maddeler arasından 13 maddenin (M1, M2, M4, M11, M12,

M15, M17, M19, M21, M23, M26, M32, M38) “iyi”, diğerlerinin ise “çok iyi” nitelikte ayırt edici olduğu belirlenmiştir.

Tablo 2 madde güçlük indeks kriterleri bakımından incelendiğinde dört maddenin (M15, M25, M30, M31) “kolay”, 16 maddenin (M1, M4, M6, M8, M10, M11, M12, M14, M19, M21, M23, M26, M27, M38, M44, M47) “zor”, geriye kalan 22 maddenin “orta” güçlüğüne sahip olduğu belirlenmiştir. Tüm bu bulgular ışığında SDT’nin maddelerin ayırt ediciliğinin 0,30’dan yüksek olduğu, madde güçlük indeksleri bakımından test içerisinde zor, kolay ve orta düzeyde sorular olduğu söylenebilir. Madde analizi sonucunda toplam 38 maddeden oluşan SDT elde edilmiştir.

Tablo 2.
Sayı Duyusu Testi İçin Yapılan Madde Analizi Sonuçları

No	r	p _j	r _{jk}	No	r	p _j	r _{jk}	No	r	p _j	r _{jk}
M1	.342**	0.28	0.40	M17	.329**	0.43	0.36	M33	.423**	0.45	0.46
M2	.364**	0.44	0.35	M18	.420**	0.54	0.73	M34	.428**	0.40	0.51
M3	.381**	0.42	0.55	M19	.323**	0.25	0.35	M35	.357**	0.46	0.55
M4	.403**	0.31	0.34	M20	.388**	0.55	0.65	M36	.370**	0.40	0.45
M5	.318**	0.59	0.55	M21	.441**	0.30	0.33	M37	.358**	0.59	0.53
M6	.458**	0.33	0.48	M22	.423**	0.48	0.69	M38	.395**	0.18	0.36
M7	.381**	0.51	0.47	M23	.388**	0.28	0.30	M39	.132	0.72	0.10
M8	.367**	0.28	0.47	M24	.412**	0.57	0.43	M40	.224**	0.56	0.17
M9	.433**	0.43	0.61	M25	.325**	0.69	0.54	M41	.331**	0.45	0.40
M10	.471**	0.35	0.52	M26	.339**	0.32	0.33	M42	.356**	0.41	0.50
M11	.362**	0.21	0.31	M27	.394**	0.37	0.53	M43	.174*	0.81	0.13
M12	.435**	0.27	0.34	M28	.419**	0.47	0.51	M44	.324**	0.26	0.33
M13	.308**	0.60	0.58	M29	.375**	0.49	0.41	M45	.203**	0.84	0.21
M14	.375**	0.36	0.43	M30	.338**	0.62	0.55	M46	.233**	0.30	0.14
M15	.419**	0.63	0.37	M31	.361**	0.68	0.59	M47	.304**	0.37	0.33
M16	.324**	0.60	0.53	M32	.423**	0.40	0.40				

r: Madde-toplam korelasyonu, p_j: madde güçlük indeksi, r_{jk}: madde ayırt edicilik indeksi
*p < .05, **p < .01

Faktör Analizine İlişkin Bulgular

Faktör analizi, aynı yapıyı veya aynı niteliği ölçen değişkenleri bir araya toplayarak ölçmeyi az sayıda faktör ile açıklamayı amaçlayan bir istatistiksel tekniktir (Büyüköztürk, 2007). Bu nedenle SDT’nin ölçtüğü faktörlerin doğası hakkında bilgi edinmek ve faktör yapısını ortaya koymak için öncelikle açımlayıcı faktör analizi (exploratory factor analysis) yapılmıştır. Faktör analizindeki ilk adım ölçeğin uygulandığı grubun faktör analizi için uygunluğunu test etmek ve verilerin çok değişkenli normal dağılımdan gelip gelmediği belirlemektir. Bu nedenle Kaiser-Meyer-Olkin (KMO) ve Bartlett’s küresellelik testi yapılmıştır. Kaiser-Meyer-Olkin (KMO) değerinin 0.60’dan yüksek olması ve Bartlett küresellik testinin anlamlı çıkması verilerin faktör analizi için uygun olduğunu ifade etmektedir (Büyüköztürk, 2007). Bu araştırma kapsamında SDT için KMO değeri 0.675, Bartlett testi için ki-kare değeri 2109.922 (p < .001) bulunmuştur. Bu değerler SDT ilişkin faktör analizinin yapılabileceğinin bir göstergesi olarak kabul edilmiştir.

a) Açımlayıcı Faktör Analizi: Ölçeğin açımlayıcı faktör analizi temel bileşenler analizi ve varimax dik döndürme yöntemi kullanılarak yürütülmüştür. Faktör analizi işleminde maddelerin faktör yükünün en az 0.40 olması ve maddeler iki ayrı faktör altında toplandığında faktör yükleri arasındaki farkın en az 0.10 olması ölçüt alınmıştır. Açımlayıcı faktör analizinin ilk sonuçları, SDT’nin özdeğeri 1.00’in üzerinde olan 10 faktör altında toplandığını ve toplam varyansın %58.66’sını açıklayabildiğini göstermiştir. Ancak özdeğeri 1.5’ten yüksek olan altı faktör dışındaki faktörler altında toplanan maddelerin sayıca çok az olduğu ya da farklı bileşenler altındaki yüklerinin birbirine yakın olduğu görülmüştür. Bu nedenle SDT’nin geliştirilmesi sürecinde altı sayı duyusu bileşeni dikkate alınarak faktör analizi altı faktör ile sınırlandırılmıştır.

Varimax dik döndürme metodu sonucunda, SDT’de yer alan M41, M42, M44 ve M47 maddelerinin birden fazla faktörde yüksek değere sahip olduğu ve faktör yük değerleri arasında farkın 0.10’dan düşük

olduğu tespit edilmiştir. Bu nedenle bu dört maddenin sırasıyla testten çıkartılmasına karar verilmiş ve geri kalan 38 madde üzerinden faktör analizi altı faktör için tekrarlanmıştır. Tekrarlanan faktör analizi sonucunda öz değeri 1.5'den büyük altı faktör altında toplanan faktör yük değerlerinin 0.44 ile 0.78 arasında değer aldığı ve toplam varyansın %48.5'ini açıklayabildiği belirlenmiştir. SDT'nin varimax dik döndürme sonucunda oluşan alt boyutlarına ilişkin faktör yük değerleri açıkladıkları varyans Tablo 3'te sunulmuştur.

Tablo 3.
Açımlayıcı Faktör Analizi ve Dik Döndürme Sonuçları

Madde	Faktör Ortak Varyansı	Faktör 1	Faktör 2	Faktör 3	Faktör	Faktör 5	Faktör 6
M25	0.628	.781					
M31	0.598	.767					
M27	0.558	.739					
M30	0.524	.715					
M17	0.458	.609					
M14	0.428	.602					
M32	0.335	.498					
M19	0.324	.497					
M6	0.620		.760				
M4	0.521		.698				
M8	0.515		.663				
M11	0.417		.646				
M37	0.353		.556				
M2	0.348		.518				
M22	0.334		.440				
M18	0.533			.720			
M21	0.488			.647			
M36	0.468			.610			
M15	0.370			.593			
M9	0.384			.538			
M12	0.469			.529	.365		
M13	0.342			.525			
M38	0.550				.722		
M23	0.516				.710		
M34	0.452				.621		
M20	0.349				.546		
M10	0.459		.395		.536		
M24	0.341				.513		
M29	0.517					.676	
M3	0.483					.643	
M16	0.436					.640	
M5	0.412					.618	
M35	0.353					.499	
M26	0.503						.701
M1	0.393						.610
M28	0.439						.580
M33	0.374						.566
M7	0.399						.513
Öz-değer:		5.703	4.305	2.921	2.222	1.720	1.560
Varyans:		%15.01	%11.33	%7.69	%5.85	%4.53	%4.11
Cronbach alpha:		.81	.79	.76	.73	.72	.75

Temel bileşen analizi sonucunda maddelerin ortak faktör varyanslarının 0.324-0.628 arasında değiştiği ve döndürme sonrası her bir faktörün açıkladıkları toplam varyans sırasıyla %15.01, %11.33, %7.69, % 5.85, %4.53 ve %4.11 olduğu belirlenmiştir. Tablo 3'deki faktör döndürme sonrasında birinci faktörün 0.497-0.781 arasında; ikinci faktörün 0.440-0.760 arasında; üçüncü faktörün 0.525-0.720

arasında; dördüncü faktörün 0.513-0.722 arasında; beşinci faktörün 0.499-0.676 arasında; altıncı faktörün 0.513-0.701 arasında yük değerleri aldığı görülmektedir.

Yapılan açımlayıcı faktör analizi sonucunda ortaya çıkan alt faktörlere içinde barındırdıkları maddelerin ölçtükleri özellikler ve sayı duygusu bileşenleri göz önünde bulundurularak isimlendirme yapılmıştır. Buna göre faktör 1'deki sekiz madde "Ölçmede Referans Noktası Kullanımı (RK)" (M14, M17, M19, M25, M27, M30, M31, M32), faktör 2'deki yedi madde "Sayı Bilgisi (SB)" (M2, M4, M6, M8, M11, M22, M37), faktör 3'deki yedi madde "Sayıların Denk Gösterimi (DG)" (M9, M12, M13, M15, M18, M21, M36), faktör 4'teki altı madde "Sayılarda İşlemlerin Etkisi (İE)" (M10, M20, M23, M24, M34, M38), faktör 5'deki beş madde "Zihinden Hesaplama (ZH)" (M3, M5, M16, M29, M35), faktör 6'daki beş madde "Niceliksel Muhakeme ve Çıkarım (NM)" (M1, M7, M26, M28, M33) olarak isimlendirilmiştir.

SDT ile alt faktör arasındaki korelasyon katsayıları hesaplanmıştır (Tablo 4). Tablo 4'te görüldüğü gibi SDT ile alt faktörler arasında orta düzeyde pozitif yönlü anlamlı ilişki olduğu belirlenmiştir ($p < .01$). Ayrıca SDT alt faktörlerinin de kendi arasında pozitif yönlü ve anlamlı ilişki olduğu saptanmıştır ($p < .05$).

Tablo 4.
Faktörler Arasındaki Korelasyon Katsayıları

Değişkenler	SDT	NM	SB	ZH	DG	RK	İE
Sayı duygusu testi (SDT)	1	.516**	.587**	.559**	.686**	.581**	.540**
Niceliksel muhakeme ve çıkarım (NM)		1	.204**	.194**	.274**	.155*	.103
Sayı bilgisi (SB)			1	.296**	.264**	.144*	.097
Zihinden hesaplama (ZH)				1	.292**	.134*	.196**
Sayıların denk gösterimi (DG)					1	.212**	.286**
Ölçmede referans noktası kullanımı (RK)						1	.212**
Sayılarında işlemlerin etkisi (İE)							1

* $p < .05$, ** $p < .01$

b) Doğrulayıcı Faktör Analizi: Ölçek geliştirme çalışmalarında ölçeğin yapı geçerliğini sınamanın bir diğer yolu da doğrulayıcı faktör analizinin yapılmasıdır (Büyüköztürk, 2007). Açımlayıcı faktör analizinde, değişkenler arası ilişkilerden hareketle faktör bulmaya yönelik bir işlem; doğrulayıcı faktör analizinde ise değişkenler arası ilişki konusunda daha önce saptanan bir hipotezin ya da kuramın test edilmesi söz konusudur (Çokluk vd., 2012; Sümer, 2000). Bu araştırma kapsamında açımlayıcı faktör analizi ile geliştirilen 38 maddeli ve altı faktörlü SDT'nin model olarak doğrulanıp doğrulanmadığını test etmek amacıyla 225 sekizinci sınıf öğrencisi üzerinde uygulanmış ve LISREL 8.8 programı kullanılarak birinci düzey doğrulayıcı faktör analizi yapılmıştır (Şekil 1).

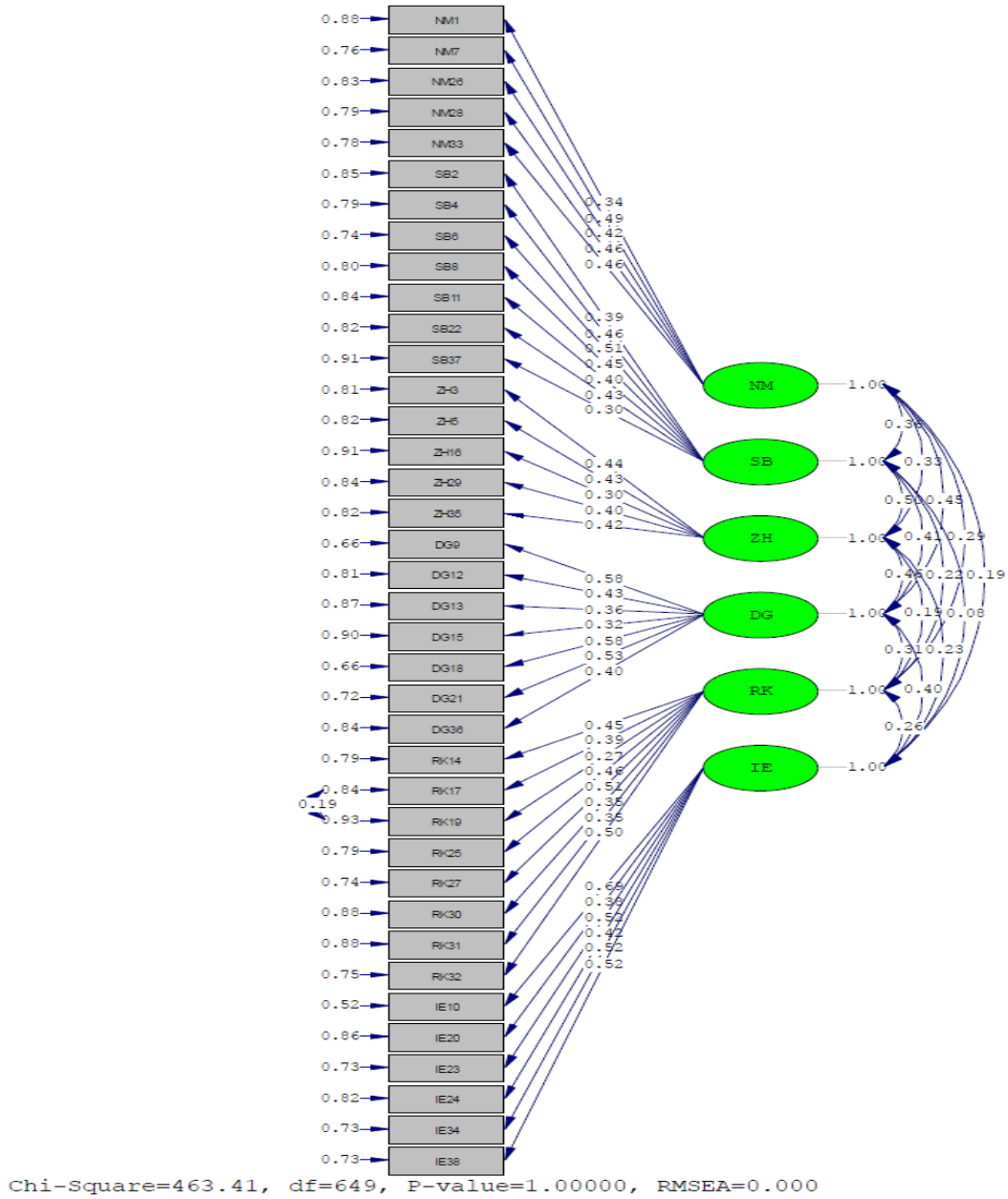
Bir modelin kabul edilip edilmemesinde kullanılan uyum indeksleri arasında ki-kare (χ^2), RMSEA (Root Mean Square Error of Approximation), GFI (Goodness of Fit Index), AGFI (Adjusted Goodness of Fit Index), CFI (Comparative Fit Index) ve NNFI (Non-Normed Fit Index), S-RMR (Standardized Root Mean Square Residual) yer almaktadır (Çokluk vd., 2012; Sümer, 2000). Ki-kare/serbestlik derecesi (χ^2/sd) oranının 3'ten küçük olması iyi bir değerdir, ancak 5'ten küçük olması kabul edilebilir. RMSEA'nın 0.08'den küçük, GFI değerinin 0.90'dan büyük, AGFI değerinin 0.80'den büyük, CFI değerinin 0.90'dan büyük ve NNFI değerinin 0.90'dan büyük olması, S-RMR'nin 0.08'den küçük olması modelin veriyeye uyumunun iyi olduğunu göstermektedir (Çokluk vd., 2012; Kline, 2005). Bu çalışma kapsamında SDT'nin birinci düzey doğrulayıcı faktör analizi sonucun ortaya çıkan uyum indeks değerleri Tablo 5'te verilmiştir.

Tablo 5.
SDT'ye İlişkin Birinci Düzey Doğrulayıcı Faktör Analizi Sonuçları ve Uyum İndeksleri

Uyum İndeksi	Kabul Edilebilir Değer	Model Değeri
χ^2/sd	≤ 3.00	2.71
RMSEA	≤ 0.08	0.00
GFI	≥ 0.90	0.90
AGFI	≥ 0.80	0.89
CFI	≥ 0.90	1.00

NNFI	≥ 0.90	1.16
S-RMR	≤ 0.08	0.05

Tablo 5'te görüldüğü gibi birinci düzey doğrulayıcı faktör analizinden elde edilen uyum indeks değerleri sırasıyla $\chi^2 = 463.41$, $sd = 649$, $\chi^2/sd = 2.71$, $p < .001$, RMSEA = 0.00, GFI = 0.90, AGFI = 0.89, CFI = 1.00, NNFI = 1.16, S-RMR = 0.05 olarak hesaplanmış ve bulunan bu değerlerin kabul edilebilir düzeyde olduğu belirlenmiştir.



Şekil 1. SDT'ye İlişkin 1.DüzyeDoğrulayıcı Faktör Analizi (Standartlaştırılmış Değerler)

Açımlayıcı faktör analizi sonucunda ortaya çıkan alt boyutlar ile SDT ile olan yapısal ilişkisini ortaya koymak için ikinci düzey doğrulayıcı faktör analizi yapılmıştır. Yapılan ikinci düzey doğrulayıcı faktör analizi sonucunda (Tablo 6) uyum indeks değerleri sırasıyla $\chi^2 = 472.08$, $sd = 658$, $\chi^2/sd = 2.72$, $p < .001$, RMSEA = 0.00, GFI = 0.90, AGFI = 0.89, CFI = 1.00, NNFI = 1.16, S-RMR = 0.051 olduğu ve bulunan bu

değerlerin de kabul edilebilir düzeyde uyum gösterdiği belirlenmiştir. Birinci düzey değişken olan alt faktörlere ilişkin standardize çözümlenme değerlerinin sırasıyla 0.57 (NM), 0.56 (SB), 0.63 (ZH), 0.78 (DG), 0.41 (RK) ve 0.42 (İE) olduğu belirlenmiştir. Kline'e (2005) göre bu değerlerin 0.30'dan yüksek olması alt faktörlerin SDT'nin kendisi ile ilgili olduğunu göstermiştir. SDT'de doğrulayıcı faktör analizine ilişkin her bir maddeye ait standardize çözümlenme değerlerinin anlamlı olup olmadığını belirlemek için hesaplanan t-testi değerlerinin sırasıyla 3.29 (NM), 3.62 (SB), 3.74 (ZH), 5.82 (DG), 3.34 (RK) ve 4.00 (İE) olduğu ve anlamlı ($p < .01$) olduğu bulunmuştur. Bu durum sayı duyusu testi ile alt boyutları arasında pozitif yönlü anlamlı bir ilişki olduğunu göstermiştir.

Tablo 6.
SDT'nin İkinci Düzey Doğrulayıcı Faktör Analizine İlişkin Sonuçlar

İkinci Düzey Değişken	Birinci Düzey Değişkenler	Standart Çözümleme Değeri	t-değeri
Sayı Duyusu Testi	Niceliksel muhakeme ve çıkarım (NM)	0.57	3.29
	Sayı bilgisi (SB)	0.56	3.62
	Zihinsel hesaplama (ZH)	0.63	3.74
	Sayıların denk gösterimi (DG)	0.78	5.82
	Ölçmede referans noktası kullanma (RK)	0.41	3.34
	Sayılarla işlemlerin etkisi (İE)	0.42	4.00

Güvenirlilik: Bir testin güvenirliliği, o testin tesadüfi hatalardan arınıklık derecesi ve test maddelerinin testin tümüyle olan tutarlılığıdır. Duyuşsal özellikler ile ilgili ölçeklerde güvenirlilik katsayısının 0.70 değerinin üzerinde olması beklenir. Bilişsel becerileri ölçen testlerde 0.60'ın altındaki değerler için, güvenirliliğin iyi olmadığı şeklinde yorumlanmaktadır (Büyüköztürk, 2007). Bu araştırmaya kapsamında geliştirilen SDT'nin güvenirliliği için Cronbach alpha iç tutarlılık katsayısı hesaplanmış testin tümü için 0.82'dir. Alt faktörler için Cronbach alpha iç tutarlılık katsayıları sırasıyla 0.77 (SB), 0.72 (NM), 0.72 (DG), 0.76 (İE), 0.75 (ZH) ve 0.79 (RK) olarak hesaplanmıştır. Bu yönüyle teşhis testinin güvenilir olduğu kabul edilmiştir.

Tartışma, Sonuç ve Öneriler

Bu araştırma, ortaokul sekizinci sınıf öğrencilerinin sayı duyusu becerilerini belirlemeye yönelik ölçme aracının geliştirilmesini amaçlamıştır. Sayı duyusu testinin (SDT) geliştirilmesi amacıyla öncelikle sayı duyusu ile ilgili alanyazın ve ölçekler incelenmiş, matematik dersi öğretim programının kazanımları doğrultusunda altı farklı sayı duyusu bileşeni temel alınmıştır. Kapsam geçerliği ve sayı duyusu bileşenlerine uygunluğunu sağlamak amacıyla dört matematik eğitimi uzmanı ve dört matematik öğretmenin görüşleri alınarak ve gerekli düzeltmeler yapılarak toplam 47 çoktan seçmeli maddeden oluşan taslak SDT oluşturulmuştur. SDT'nin öğrenciye uygunluğunu, görünüş geçerliğini ve kullanımını test etmek için 24 öğrenci üzerinde pilot uygulaması yapılmıştır. Son halini alan SDT'nin geçerlik ve güvenirliliği için 180 öğrenci üzerinde madde analizi ve açımlayıcı faktör analizi gerçekleştirilmiş, yapı geçerliğini test etmek için 225 öğrenci üzerinde doğrulayıcı faktör analizi gerçekleştirilmiştir.

SDT'nin madde analizi kapsamında madde-toplam korelasyonu, madde güçlük indeksi ve madde ayırt edicilik indeksleri hesaplanmıştır. Madde-toplam korelasyonu ile madde ayırt edicilik indeksi 0.20'nin altında değer alan beş maddenin testten çıkarılmasına karar verilmiştir. Geri kalan 42 maddenin madde-toplam korelasyonları 0.30 ile 0.47 arasında, madde ayırt edicilik indekslerinin 0.30 ile 0.79 arasında, madde güçlük indeksleri bakımından 0.18 ile 0.69 arasında değer aldığı, testin zor, orta ve kolay düzeyde olan maddelerden oluştuğu belirlenmiştir. Bu sonuçlar SDT'de yer alan madde-toplam korelasyon, madde güçlük ve maddelerin ayırt edicilik değerlerinin kabul edilebilir düzeyde olduğunu (Büyüköztürk, 2007; Turgut ve Baykul, 2010) göstermiş ve maddelerin sayı duyusu becerilerini ölçmede ayırt edici nitelikte olduğu şeklinde yorumlanmıştır.

SDT'nin yapı geçerliği için temel bileşenler analizi ve varimax dik döndürme işlemleri kullanılarak yapılan açımlayıcı faktör analizi yapılmıştır. Analiz sonucunda birden fazla faktör altında toplanan ve faktör yük değerleri arasında 0.10'dan düşük olan dört maddenin testten çıkartılmasına karar verilmiş,

geriye kalan 38 maddenin öz değeri 1.5’den büyük 6 faktör altında toplam varyansın %48.5’ni açıklayabildiği ve maddelerin faktör yüklerinin 0.44 ile 0.78 arasında değer aldığı saptanmıştır. Sosyal bilimlerde yapılan çok faktörlü ölçek yapılarında %40 ile %60 arasında değişen toplam varyans oranlarının yeterli kabul edildiği ve genel bir kural olarak 0.32 ve üzeri faktör yüklerinin kullanılabilmesini belirtmektedir (Çokluk vd., 2012; Tabachnick ve Fidell, 2007). Bu sonuçlar SDT’nin alt faktörlerinin ve ilgili maddelerin sayı duygusu bileşenlerini açıklamada kabul edilebilir nitelikte olduğunu göstermiştir. Açıklayıcı faktör analizi sonucunda ortaya çıkan alt faktörler içinde barındırdıkları maddelerin ölçtükleri özellikler ve sayı duygusu bileşenleri göz önünde bulundurularak “Sayı Bilgisi”, “Sayıların Denk Gösterimi”, “Sayılarda İşlemlerin Etkisi”, “Ölçmede Referans Noktası Kullanımı”, “Zihinden Hesaplama” ve “Niceliksel Muhakeme ve Çıkarım” olarak isimlendirilmiştir. Alanyazında sayı duygusuna yönelik tek bir teorik çerçeve bulunmamakta, sayı duygusu bileşenleri farklı şekillerde tanımlanmakta ve isimlendirilmektedir. Sayı duygusu ölçek geliştirme çalışmalarında (Facun ve Nool, 2012; Greeno, 1991; Gülbagcı Dede, 2015; Kayhan Altay ve Umay, 2013; McIntosh vd., 1992; Reys vd., 1999; Singh, 2009; Şengül ve Gülbagcı Dede, 2013; Yang ve Lin, 2015) bazı sayı duygusu bileşenlerinin benzer olmasına karşın, farklı sayı duygusu bileşenlerinin olduğu dikkat çekmektedir. Bu durum sayı duygusu ölçeğinin odaklandıkları yaş, sınıf düzeyi, öğretim programı kazanımları ve sayı duygusu konusuna (kesirler, yüzdeler, üslü sayılar, doğal sayılar, rasyonel sayılar, vb.) göre farklılık göstermesinden kaynaklanmaktadır.

Alanyazında ölçek alt faktörlerin birbirleriyle yüksek düzeyde ilişki göstermemesi, ölçeğin tümü ile düşük ilişki gösteren alt faktörlerin ise ölçekten çıkartılması gerektiği belirtilmektedir (Tavşancıl, 2002). Bu çalışmada SDT’ye ilişkin alt faktörün kendi arasındaki korelasyon değerlerinin 0.10 ile 0.30 arasında değiştiği, alt faktörler ile SDT arasındaki korelasyon değerlerinin ise 0.51 ile 0.68 arasında değer aldığı ve pozitif yönlü orta düzeyde anlamlı ilişki olduğu ($p < .01$) belirlenmiştir. Bu sonuçlar, alt faktörlerin SDT’nin bileşenleri olduğunu ve birbirinden ayrıştığını göstermiştir.

Ölçeklerin yapı geçerliğini test etmenin bir diğer yolu da doğrulayıcı faktör analizinin yapılmasıdır (Çokluk vd., 2012; Tabachnick ve Fidell, 2007). LISREL programı kullanılarak yapılan doğrulayıcı faktör analizi sonucunda her bir maddeye ilişkin standardize çözümlenme değerlerinin 0.30 - 0.96 arasında değer aldığı ve bu değerler için hesaplanan t -değerlerinin de 3.35 - 7.95 arasında değiştiği ve anlamlı olduğu saptanmıştır. SDT için doğrulayıcı faktör analizi kapsamında elde edilen uyum indeks değerlerinin ($\chi^2/sd = 2.71$, $p < .001$, RMSEA = 0.00, GFI = 0.90, AGFI = 0.89, CFI = 1.00, NNFI = 1.16, S-RMR = 0.05) kabul edilebilir düzeyde olduğu (Çokluk, vd., 2012; Kline, 2005; Sümer, 2000; Tabachnick ve Fidell, 2007) belirlenmiştir. Ayrıca alt faktör için yapılan ikinci düzey doğrulayıcı faktör analizi sonucunda modelin uyum indekslerinin kabul edilebilir nitelikte olduğu saptanmıştır. Bu sonuçlar, SDT’ye ilişkin altı sayı duygusu bileşenlerine ilişkin faktör yapısının doğrulandığını göstermiştir.

Bu çalışmada SDT’nin güvenilirlik çalışmaları kapsamında hesaplanan Cronbach alpha iç tutarlık katsayıları ölçeğin geneli için 0.82, alt faktörler açısından iç tutarlık katsayılarının 0.72 ile 0.79 arasında değiştiği belirlenmiştir. İlgili alanyazında bilişsel becerileri ölçen testler için hesaplanan güvenilirlik katsayısının 0.60 ve üzerinde olması test puanlarının güvenilirliği için genel olarak yeterli görülmektedir (Büyükköztürk, 2007). Güvenirlik analizinden elde edilen bu bulgular, hem ölçeğin bütününde hem de faktörler bazında kabul edilebilir düzeyde güvenilirliğe sahip olduğunu göstermiştir.

Bu çalışmanın sonuçları, 36 çoktan seçmeli soru ve altı sayı duygusu bileşeninden oluşan SDT’nin sekizinci sınıf öğrencilerinin sayı duyu becerilerini belirlemede geçerli ve güvenilir bir araç olduğunu göstermiştir. Bu bağlamda bu çalışmanın Türkiye’deki sınırlı sayıda olan sayı duygusu testlerine ve sayı duygusu konusunda yapılacak olan çalışmalara katkı sunması beklenmektedir.

Araştırma ve Yayın Etiği Beyanı

Araştırma sürecinde Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesinde yer alan tüm kurallara uyulmuş ve yönergenin ikinci bölümünde yer alan Bilimsel Araştırma ve Yayın Etiğine Aykırı Eylemlerden hiçbiri gerçekleştirilmemiştir.

References


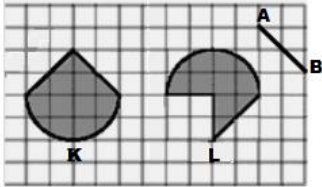
- Akkaya, R. (2016). An investigation into the number sense performance of secondary school students in Turkey. *Journal of Education and Training Studies*, 4(2), 113-123.
- Berch, D. B. (2005). Making sense of number sense implications for children with mathematical disabilities. *Journal of Learning Disabilities*, 38(4), 333-339.
- Birgin, O., & Peker, E. S. (2021). Thematic content analysis of studies on number sense in Turkey. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi-Hacettepe University Journal of Education*, 36(3), 593-609.
- Büyüköztürk, Ş. (2007). *Sosyal bilimler için veri analizi el kitabı* (7. Baskı). Ankara: Pegem Akademi Yayıncılık.
- Can, D. (2017). *İlkokul dördüncü sınıf öğrencilerinin sayı duyularının bağlam temelli ve bağlam temelli olmayan problem durumlarında incelenmesi* (Yayınlanmamış doktora tezi). Hacettepe Üniversitesi, Ankara.
- Cansız Aktaş, M., Tuğrul Özdemir, E., & Yavuz Mumcu, H. (2017). Examination of the number sense skills of secondary school students (6th-8th grades). *Journal of Education and Practice*, 8(25), 199-207.
- Çekirdekci, S., Şengül, S., & Doğan, M. C. (2016). 4. sınıf öğrencilerinin sayı hissi ile matematik başarıları arasındaki ilişkinin incelenmesi. *Qualitative Studies (NWSAQs)*, 11(4), 48-66.
- Çekirdekçi, S., Şengül, S., & Doğan, M. C. (2017). 4. sınıf sayı hissi testi'nin geliştirilmesi. *Kalem Eğitim ve İnsan Bilimleri Dergisi*, 7(2), 441-473.
- Çokluk, O., Şekercioğlu, G., & Büyüköztürk, Ş. (2012). *Sosyal bilimler için çok değişkenli istatistik SPSS ve LISREL uygulamaları* (2. Baskı). Ankara: Pegem Akademi Yayıncılık.
- Er, Z., & Dinç Artut, P. (2017). Investigation of number sense strategies used by eight grade on the subject of natural numbers, decimal numbers, fractions, percentages of eight grade students. *International Journal of Social Sciences and Education Research*, 3(1), 218-229.
- Erkuş, A. (2012). *Psikolojide ölçme ve ölçek geliştirme-1: Temel kavramlar ve işlemler*. Ankara: Pegem Akademi.
- Facun, R. D., & Nool, N. R. (2012). Assessing the number sense of grade 6 pupils. *International Conference on Education and Management Innovation IPEDR 30* (pp. 297-301). Singapore: IACSIT Press.
- Greeno, J. G. (1991). Number sense as situated knowing in a conceptual domain. *Journal for Research in Mathematics Education*, 22(3), 170-218.
- Gülbağcı Dede, H. (2015). *İlköğretim ve ortaöğretim matematik öğretmen adaylarının sayı hissini incelenmesi* (Yayınlanmamış doktora tezi). Marmara Üniversitesi, Eğitim Bilimleri Enstitüsü, İstanbul.
- Harç, S. (2010). *6. sınıf öğrencilerinin sayı duygusu kavramı açısından mevcut durumlarının analizi* (Yayınlanmamış yüksek lisans tezi). Marmara Üniversitesi, Eğitim Bilimleri Enstitüsü, İstanbul.
- Hope, J. (1989). Promoting number sense in school. *Arithmetic Teacher*, 36(6), 12-16.
- İymen İkizoğlu, E., & Duatepe Paksu, A. (2016). Üslü ifadeler ile ilgili sayı duygusu ölçeğinin geliştirilmesi. *Buca Eğitim Fakültesi Dergisi*, 41, 42-65.
- Jordan, N. C., Glutting, J., & Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. *Learning and Individual Differences*, 20(2), 82-88.
- Kalchman, M., Moss, J., & Case, R. (2001). Psychological models for the development of mathematical understanding: Rational numbers and functions. In S. Carver & D. Klahr (Eds.), *Cognition and instruction* (pp. 1-38). Mahwah, NJ: Erlbaum

- Kayhan Altay, M. (2010). *İlköğretim ikinci kademe öğrencilerinin sayı duyularının; sınıf düzeyine, cinsiyete ve sayı duyusu bileşenlerine göre incelenmesi* (Yayınlanmamış doktora tezi). Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Kayhan Altay, M., & Umay, A. (2013). The development of number sense scale towards middle grade students. *Education and Science*, 38(167), 241-245.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd Edition ed.). New York: The Guilford Press.
- Li, M. N. F., & Yang, D. C. (2010). Development and validation of a computer-administered number sense scale for fifth-grade children in Taiwan. *School Science and Mathematics*, 110(4), 220-230.
- Markovits, Z., & Sowder, J. (1994). Developing number sense: An intervention study in grade 7. *Journal for Research in Mathematics Education*, 25, 4-29.
- McIntosh, A. B., Reys, B. J., & Reys, E. R. (1992). A proposed framework for examining basic number sense. *For the Learning of Mathematics*, 12(3), 2-8.
- Milli Eğitim Bakanlığı (MEB) (2018). *Ortaokul matematik dersi (1-8.sınıflar) öğretim programı*. Ankara: Milli Eğitim Basımevi.
- Mohamed, M., & Johnny, J. (2010). Investigating number sense among students. *Procedia-Social and Behavioral Sciences*, 8, 317-324.
- National Council of Teachers of Mathematics (NCTM). (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author
- National Council of Teachers of Mathematics (NCTM). (2000). *The Principles and Standards for School Mathematics*. Reston, VA: NCTM.
- Olkun, S., Mutlu, Y., & Sarı, M. H. (2017). The relationships between number sense and mathematics achievement. *International Conference on Education and New Developments*, Lisbon, Portugal.
- Purnomo, Y. W., Kowiyah, K., Alyani, F., & Assiti, S. (2014). Assessing number sense performance of Indonesian elementary school students. *International Education Studies*, 7(8), 74-84.
- Reys, B. J. (1994). Promoting number sense in middle grades. *Teaching Mathematics in the Middle School*, 1(2), 114-120.
- Reys, R. E., & Yang, D. C. (1998). Relationship between computational performance and number sense among sixth and eighth-grade students in Taiwan. *Journal for Research in Mathematics Education*, 29(2), 225-237.
- Reys, R., Reys, B., Emanuelsson, G., Johansson, B., McIntosh, A., & Yang, D. C. (1999). Assessing number sense of students in Australia, Sweden, Taiwan, and the United States. *School Science and Mathematics*, 99(2), 61-70.
- Şengül, S., & Gülbağcı Dede, H. (2013). Sayı hissi bileşenlerine ait sınıflandırmaların incelenmesi. *The Journal of Academic Social Science Studies*, 8(6), 645-664.
- Şengül, S., & Gülbağcı, H. (2012). An investigation of 5th grade Turkish students' performance in number sense on the topic of decimal numbers. *Procedia - Social and Behavioral Sciences*, 46, 2289-2293.
- Singh, P. (2009). An assessment of number sense among secondary school students. *International Journal for Mathematics Teaching and Learning*, 155, 1-29.
- Soyuk, R. (2018). *Sekizinci sınıf öğrencilerinin kareköklü ifadeler konusunda sayı duyularının incelenmesi* (Yayınlanmamış yüksek lisans tezi). Osmangazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Eskişehir.
- Sümer, N. (2000). Yapısal eşitlik modelleri: Temel kavramlar ve örnek uygulamalar. *Türk Psikoloji Yazıları*, 3(6), 49-73.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Pearson/Allyn & Bacon.
- Tavşancıl, E. (2002). *Tutumların ölçülmesi ve SPSS ile veri analizi*. Ankara: Nobel Yayın Dağıtım.

- Tunalı, C. (2018). *Özel yetenekli öğrencilerin sayı duygusu düzeylerinin belirlenmesi* (Yayınlanmamış yüksek lisans tezi). Dokuz Eylül Üniversitesi, İzmir.
- Turgut, M. F., & Baykul, Y. (2010). *Eğitimde ölçme ve değerlendirme* (2.Baskı). Ankara: Pegem Akademi.
- Verschaffel, L., Greer, B., & De Corte, E. (2007). Whole number concepts and operations. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 2, pp. 557-628). Charlotte, NC: Information Age Publishing.
- Yang, D. C. (1995). *Number sense performance and strategies possessed by sixth and eighth grade students in Taiwan* (Unpublished doctoral dissertation). University of Missouri, Columbia.
- Yang, D. C. (2019). Development of a three-tier number sense test for fifth-grade students. *Educational Studies in Mathematics*, 101, 405-424.
- Yang, D. C., & Lin, Y. C. (2015). Assessing 10- to 11-year-old children's performance and misconceptions in number sense using a four-tier diagnostic test. *Educational Research*, 57(4), 368-388.
- Yang, D. C., Li, M. N., & Li, W. J. (2008). Development of a computerized number sense scale for 3rd graders: Reliability and validity analysis. *International Electronic Journal of Mathematics Education*, 3(2), 110-124.
- Yang, D. C., Li, M. N., & Lin, C. I. (2008). A study of the performance of 5th graders in number sense and its relationship to achievement in mathematics. *International Journal of Science and Mathematics Education*, 6(5), 789-807.


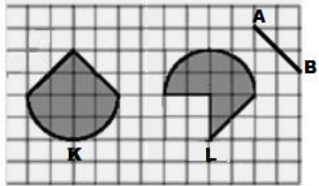
Appendix 1.

Examples Items of the Number Sense Test

Number Sense Component	Sample Item
C1. Understanding the meaning of numbers (7 items)	<p>M2. Which of the numbers below is the biggest? (Facun & Nool, 2012)</p> <p>A) $\frac{499}{500}$ B) $\frac{500}{501}$ C) $\frac{4568}{4569}$ D) $\frac{4569}{4570}$</p> <p>M11. Which of the following is between $0.\overline{45}$ and $0.\overline{452}$ repeating decimal fractions? (SBS 2012)</p> <p>A) 0.451 B) 0.455 C) 0.453 D) 0.457</p>
C2. Equivalent representation of numbers (7 items)	<p>M12. Which of the following is the number equal to 5^{-4}?</p> <p>A) 0.0005 B) 0.0002 C) 0.0016 D) 0.0025</p> <p>M21.  What is the percentage of the shaded area on the right?</p> <p>A) 30 % B) 60% C) 80 % D) 160 %</p>
C3. Effect of the operations on numbers (6 items)	<p>M24. Please mark the closest answer without calculating the result of the $\frac{2}{3} \times \frac{5}{4} = ?$ operation?</p> <p>A) Less than $\frac{2}{3}$ B) $\frac{2}{3}$ C) More than $\frac{2}{3}$ D) I cannot do without calculation</p> <p>M38. Which of the following operations gives the greatest result?</p> <p>A) $29 \div 0.8$ B) 29×0.8 C) $29 + 0.8$ D) $29 - 0.8$</p>
C4. Use of reference points in the measurement (8 items)	<p>M19.  Considering the length of AB, which of the following would be true when the perimeters of the K and L shapes are compared?</p> <p>A) The circumference of K is smaller than the circumference of L. B) The circumference of L will be smaller than the circumference of K. C) The circumference of K is equal to the circumference of L. D) The circumference of K is 2 times the circumference of L.</p> <p>M27. Which of the following fraction sums is greater than 1? (Gülbağcı Dede, 2015)</p> <p>A) $\frac{5}{11} + \frac{3}{7}$ B) $\frac{7}{15} + \frac{5}{12}$ C) $\frac{1}{2} + \frac{4}{9}$ D) $\frac{5}{9} + \frac{8}{15}$</p>
C5. Mental calculation (5 items)	<p>M3. Which of the following is closest to the result of the $\frac{7,21 \times 3,86}{10,09} = ?$</p> <p>A) $\frac{7 \times 3}{10}$ B) $\frac{7 \times 4}{10}$ C) $\frac{7 \times 3}{11}$ D) $\frac{7 \times 4}{11}$</p> <p>M16. What is the approximate value of the side length of a square garden with an area of 150 m^2? (SBS 2010)</p> <p>A) $7m - 8m$ B) $9m - 10m$ C) $11m - 12m$ D) $12m - 13$</p>
C6. Quantitative reasoning and inference (5 items)	<p>M1. One teacher for every 12 students will attend the school trip. If 110 students are going on the trip, how many teachers should attend at least? (TIMSS 2007)</p> <p>A) 7 B) 9 C) 10 D) 11</p> <p>M26. With a ticket for a toy in a playground, the game can be played for a maximum of 4 minutes. How many tickets are required to play the game for 9 minutes?</p> <p>A) 2 B) $2\frac{1}{4}$ C) 3 D) Not enough information is given.</p>

Ek 1.

Sayı Duyusu Testinde Yer Alan Bileşenler ve Soru Örnekleri

Sayı Duyusu Bileşeni	Örnek Madde
F1. Sayı Bilgisi (7 madde)	<p>M2. Aşağıdaki sayılardan hangisi en büyüktür? (Facun ve Nool, 2012)</p> <p>A) $\frac{499}{500}$ B) $\frac{500}{501}$ C) $\frac{4568}{4569}$ D) $\frac{4569}{4570}$</p> <p>M11. Aşağıdakilerden hangisi $0,\overline{45}$ ile $0,\overline{452}$ devirli ondalık kesirleri arasında yer alır? (SBS 2012)</p> <p>A) 0,451 B) 0,455 C) 0,453 D) 0,457</p>
F2. Sayıların Denk Gösterimi (7 madde)	<p>M12. 5^{-4} sayısı aşağıdakilerden hangisine eşittir?</p> <p>A) 0,0005 B) 0,0002 C) 0,0016 D) 0,0025</p> <p>M21.  Yandaki taralı bölgenin tamamının ifade ettiği yüzdelik değer (%) bakımından karşılığı nedir?</p> <p>A) % 30 B) % 60 C) % 80 D) % 160</p>
F3. Sayılarda İşlemlerin Etkisi (6 madde)	<p>M24. $\frac{2}{3}x\frac{5}{4}=?$ işleminin sonucunu hesaplama yapmadan en yakın cevabı işaretleyiniz?</p> <p>A) $\frac{2}{3}$ den daha az B) $\frac{2}{3}$ e eşittir C) $\frac{2}{3}$ den daha fazla D) Hesaplamadan yapamam</p> <p>M38. Aşağıdaki işlemlerden hangisi en büyük sonucu verir?</p> <p>A) $29 \div 0.8$ B) 29×0.8 C) $29 + 0.8$ D) $29 - 0.8$</p>
F4. Ölçmede Referans Noktası Kullanımı (8 madde)	<p>M19.  AB uzunluğunu göz önüne alarak K ve L şekillerinin çevrelerini karşılaştırsak aşağıdakilerden hangisi doğru olur?</p> <p>A) K'nin çevresi, L'nin çevresinden küçük olur. B) L'nin çevresi, K'nin çevresinden küçük olur. C) K'nin çevresi, L'nin çevresine eşit olur. D) K'nin çevresi, L'nin çevresinin 2 katıdır.</p> <p>M27. Aşağıdaki hangi kesirlerin toplamı 1'den büyüktür? (Gülbağcı Dede, 2015)</p> <p>A) $\frac{5}{11} + \frac{3}{7}$ B) $\frac{7}{15} + \frac{5}{12}$ C) $\frac{1}{2} + \frac{4}{9}$ D) $\frac{5}{9} + \frac{8}{15}$</p>
F5. Zihinsel Hesaplama (5 madde)	<p>M3. Aşağıdakilerden hangisi $\frac{7,21 \times 3,86}{10,09}$ işlemin sonucuna en yakındır?</p> <p>A) $\frac{7 \times 3}{10}$ B) $\frac{7 \times 4}{10}$ C) $\frac{7 \times 3}{11}$ D) $\frac{7 \times 4}{11}$</p> <p>M16. Alanı 150 m^2 olan kare şeklindeki bir bahçenin bir kenar uzunluğu, hangi metreler arasındadır? (SBS, 2010)</p> <p>A) $7m - 8m$ B) $9m - 10m$ C) $11m - 12m$ D) $12m - 13m$</p>
F6. Niceliksel Muhakeme ve Çıkarım (5 madde)	<p>M1. Okulun düzenlediği gezilere her 12 öğrenci için 1 öğretmen katılacaktır. Geziye 110 öğrenci gidiyorsa en az kaç öğretmen katılmalıdır? (TIMSS, 2007)</p> <p>A) 7 B) 9 C) 10 D) 11</p> <p>M26. Bir oyun parkındaki oyuncak için alınan bir bilet ile en fazla 4 dakika oyun oynanabilmektedir. 9 dakika oyun oynayabilmek için kaç bilet almak gerekir?</p> <p>A) 2 B) $2\frac{1}{4}$ C) 3 D) Yeterli bilgi verilmemiştir.</p>