DEVELOPMENT AND VALIDATION OF SELECTED RESPONSE TEST TO MEASURE STUDENTS' DECLARATIVE KNOWLEDGE OF TRIANGLES

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

MELİKE GÜNEŞ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
MATHEMATICS EDUCATION IN MATHEMATICS AND SCIENCE EDUCATION

DECEMBER 2022

Approval of the thesis:

DEVELOPMENT AND VALIDATION OF SELECTED RESPONSE TEST TO MEASURE STUDENTS' DECLARATIVE KNOWLEDGE OF TRIANGLES

submitted by **MELİKE GÜNEŞ** in partial fulfillment of the requirements for the degree of **Master of Science** in **Mathematics Education in Mathematics and Science Education, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Mine Işıksal Bostan
Head of the Department, **Math. and Sci. Edu, METU**

Prof. Dr. Behiye Ubuz
Supervisor, **Math. and Sci. Edu, METU**

Examining Committee Members:

Prof. Dr. Didem Akyüz
Math. and Sci. Edu, METU

Prof. Dr. Behiye Ubuz
Math. and Sci. Edu, METU

Prof. Dr. Halil Giray Berberoğlu
Educational Sciences, Başkent University

Date: 02.12.2022

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name Last name : Melike Güneş

Signature :

ABSTRACT

DEVELOPMENT AND VALIDATION OF SELECTED RESPONSE TEST TO MEASURE STUDENTS' DECLARATIVE KNOWLEDGE OF TRIANGLES

Güneş, Melike
Master of Science, Mathematics Education in Mathematics and Science Education
Supervisor : Prof. Dr. Behiye Ubuz

December 2022, 68 pages

Currently, relatively limited selected-response tests are available to measure high school students' geometry knowledge. By adhering to Standards for Educational and Psychological Testing set by the American Educational Research Association (AERA et al., 2014), the current study seeks to develop an objective type of selected-response test for assessing 11-grade students' declarative knowledge of triangles, specifically definitions of types of triangles, congruent and similar triangles and identification of triangles and then determine the validity evidence for the test (DKTT) based on test content, internal structure, and reliability. Four factors emerged from using the Exploratory Factor Analysis (EFA) for the analysis of test dimensionality. These factors were named 'Identification of Triangles (Factor 1)', 'Definitions of Congruent and Similar Triangles (Factor 2)', Minimal Definitions of Types of Triangles (Factor 3)', and 'Non-minimal Definitions of Types of Triangles with Auxiliary Elements (Factor 4)'. The Cronbach's alpha and McDonald's omega coefficients for the entire test were 0.71 and 0.75, respectively. This test will contribute to thoroughly measuring the students' knowledge of definitions of types of triangles, congruent and similar triangles, and identification of triangles, measuring students' declarative knowledge of triangles.

Keywords: Test Development, Declarative Knowledge, Triangle

ÖZ

ÖĞRENCİLERİN ÜÇGENLER KONUSUNDA BİLDİRİME DAYALI BİLGİLERİNİ ÖLÇMEK İÇİN SEÇİLMİŞ YANIT TESTİNİN GELİŞTİRİLMESİ VE DOĞRULANMASI

Güneş, Melike
Yüksek Lisans, Matematik Eğitimi, Matematik ve Fen Bilimleri Eğitimi
Tez Yöneticisi: Prof. Dr. Behiye Ubuz

Aralık 2022, 68 sayfa

Şu anda, lise öğrencilerinin geometri bilgilerini ölçmek için nispeten sınırlı seçilmiş yanıtlı testler mevcuttur. Mevcut çalışma özellikle üçgen çeşitlerinin, eş ve benzer üçgenlerin tanımları ve üçgenlerin belirlenmesinde 11. sınıf öğrencilerinin bildirime dayalı üçgen bilgilerini ölçmek için nesnel bir seçilmiş yanıt testi türü geliştirmeyi Amerikan Eğitim Araştırmaları Derneği (AERA ve diğ., 2014) tarafından belirlenen Eğitimsel ve Psikolojik Test Standartlarına bağlı kalarak amaçlamaktadır ve ardından testin içeriği, iç yapısı ve güvenilirliğine dayalı olarak testin geçerlilik kanıtını (DKTT) belirlemektedir. Test boyutluluğunun analizi için Açımlayıcı Faktör Analizi'nin (AFA) kullanılmasından dört faktör ortaya çıkmıştır. Bu faktörler, 'Üçgenlerin Belirlenmesi (Faktör 1)', 'Eş ve Benzer Üçgenlerin Tanımları (Faktör 2)', 'Üçgen Çeşitlerinin Minimal Tanımları (Faktör 3)' ve ' Yardımcı Elemanlar ile Üçgen Çeşitlerinin Minimal Olmayan Tanımları (Faktör 4)' olarak adlandırılmıştır. Tüm test için Cronbach alfa ve McDonald's omega katsayıları sırasıyla 0.71 ve 0.75 olarak bulunmuştur. Bu test, öğrencilerin üçgen çeşitleri, eş ve benzer üçgenlerin tanımları ve üçgenleri belirleme bilgilerinin derinlemesine ölçerek öğrencilerin üçgenler hakkındaki bildirimsel bilgilerinin ölçülmesine katkıda bulunacaktır.

Anahtar Kelimeler: Tez Geliştirme, Bildirimsel Bilgi, Üçgen

To all my loved ones

ACKNOWLEDGMENTS

First and foremost, I would like to thank my supervisor, Prof. Dr. Behiye Ubuz, for all of her help, encouragement, patience, and support during my research. She supported me and helped me at every step throughout the research. She has always welcomed all of my questions and provided insightful commentary. Her enlightening remarks have extended my scholarly perspectives.

I would like to express my most profound appreciation to the thesis committee members, Prof. Dr. Didem Akyüz and Prof. Dr. Halil Giray Berberoglu, for their time, recommendations, and efforts in enhancing my study.

I would like to thank my darling, Hüseyin Avni Yaşar, who supported me and encouraged me, and gave technical assistance during my study.

Finally, I would like to send my deep gratitude and appreciation to all of the 11th graders who took part in this research for their enthusiasm, tolerance, and time. Without their contributions, this study would not have been able to be accomplished.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

Identifying triangles, defining triangle types, and congruent and similar triangles are important for success in higher-level mathematics courses. The ability to identify triangles by their characteristics and properties is essential for understanding more advanced concepts in geometry, such as congruence and similarity, transformations, spatial reasoning, and the properties of two- and three-dimensional shapes. In the teaching and learning of mathematics, definitions are crucially significant (Cansız Aktaş, 2016; Miller, 2018; Zazkis & Leikin, 2008; Zaslavsky & Shir, 2005). Researchers have indicated that teachers do not always possess sufficient knowledge of definitions, nor do they have the flexibility needed to consider alternate versions of definitions for a single concept (Fujita & Jones, 2007; Leikin & Winicki-Landman, 2001; Pickreign, 2007). Even students in senior high schools have difficulty learning plane geometry terms, defining and categorizing shapes, and properties of shapes (Atebe, 2008; Clements & Battista, 1992; Fuys et al., 1988; Usiskin, 1982). Triangles are one of the most frequently addressed topics in the mathematics curriculum (see Miyakawa, 2017; Otten et al., 2014).

Previous research has emphasized the definitions' importance for geometrical knowledge and the deductive structure of geometry since they play a crucial role in developing an understanding of the meanings of the concepts and are used as building blocks for the construction of geometrical theorems (Mariotti & Fischbein, 1997; Pimm, 1993; Wilson, 1990). Numerous studies have demonstrated that geometrical definitions are a challenge for both learners and many pre- and in-service teachers across all grade levels. They have trouble defining terms and using definitions to identify, classify, and generate examples and non-examples of the defined concept (Hershkowitz, 1987; Hershkowitz, 1989; Hershkowitz, 1990; Fujita & Jones, 2007; Marches, 2012; Pickreign, 2007; Vinner & Dreyfus, 1989). Other studies have investigated how students understand the definitions of some geometric figures, such as triangles and quadrilaterals (e.g., Fujita & Jones, 2007; Kaur, 2015; Usiskin, 2008). It is observed that studies put more emphasis on defining the concept of a triangle than types of triangles (e.g., Tsamir et al., 2015; Ulusoy, 2021; Zaslavsky & Shir, 2005). Thus, studies about definitions of types of triangles are scarce, and instruments used to collect data were comprised of open-ended questions such as "define right triangle" (Mayberry, 1983; Pielsticker, 2022; Ubuz & Aydın, 2018). Additionally, there have been a few studies that have clearly focused on definitions of congruent and similar triangles (e.g., Dündar & Gündüz, 2017; González & Herbst, 2009; Haj-Yahya, 2022). Beyond knowing the definitions, one of a concept's (triangle) roles is to enable someone to identify both examples and non-examples of the category. Therefore, non-examples result from concept acquisition (Tsamir et al., 2008). Students are encouraged to solely employ critical attributes when producing and identifying examples of geometric concepts in mathematics classes since acknowledging critical attributes is essential for correctly identifying figures (van Hiele, 1958). One of the primary goals of educators is to enable students to identify examples and develop geometrical conceptions using only critical attributes (Tsamir et al., 2008). Utilizing precise mathematical terminology can help to achieve this goal (Tsamir et al., 2015). Students are expected to select and identify triangles based on the following critical attributes: (a) being closed, (b) three sides, (c) three vertices, and (d) three angles. Declarative knowledge refers to information that is stored in an individual's memory and can be explicitly stated or declared, such as facts, definitions, and concepts. In this case, students' knowledge of definitions for different types of triangles, congruent and similar triangles, and the critical attributes that are used to identify triangles (examples and non-examples) are declarative knowledge about related geometry concepts. When the studies on these topics were reviewed, it was found that the majority of the studies used open-ended questions on their measurement tool following using interviews (e.g., Bernabeu et al., 2021; Burger & Shaugnessy, 1986; Clements et al., 1999; Ulusoy, 2022), whereas studies using selected-response (SR) opted for multiple-choice as the format for the SR (e.g., Usiskin, 1982; Senk, 1989; Ubuz, 2017). CR (constructed response) assessments frequently align with authentic scientific practices, such as constructing explanations or arguments based on evidence and provide teachers with greater insight into student thinking in their classrooms, allowing them to modify teaching strategies to meet learning objectives (Gerard & Linn, 2016). CR items can be straightforward or highly complex, and they can generate answers that are very brief or extremely long (Hogan, 2013). They make guessing almost impossible, are generally more complex than multiple-choice items, and are often rated by assessors using pre-determined criteria (Haladyna & Rodriguez, 2013). They are thought to allow for more in-depth and exhaustive measures, hence increasing content validity (Dennis & Newstead, 1994). However, inter-rater reliability has been recognized as an issue in scoring constructed-response items (Attali et al., 2013; Gierl et al., 2014; Kuo et al., 2016; Shermis & Burstein, 2013), as well as rater fatigue and drift (Almond, 2014). In measuring cognitive achievement and ability, selected-response formats provide a number of validity advantages. They promote content validity evidence by enabling a complete and representative sample of the cognitive domain; this representativeness of the content sampling procedure increases the validity evidence for domain inferences and minimizes one major threat to validity—construct underrepresentation (Haladyna & Downing, 2004; Messick, 1989). Objectivity is a crucial feature of every effective measurement, and it contributes to the test development process's validity evidence, defensibility, and efficiency. SR format aligns with the purpose of this study, which is to develop a selected-response test to measure 11th-grade students' declarative knowledge of definitions of types of triangles and congruent and similar triangles and to determine the validity evidence for the test based on its content, internal structure, and reliability because the SR format promotes content validity evidence by enabling a representative sample of the cognitive domain being measured (declarative knowledge) and is objective in nature, which will contribute to the overall validity of the test.

Compared to single-answer MC questions, MTF items have the ability to cover a wider breadth of topics (Dudley, 2006). Higher test reliabilities have been discovered when employing MTF compared to MC since this makes it possible for more information to be exposed per testing period (Frisbie & Sweeney, 1982; Javid, 2014; Kreiter & Frisbie, 1989; Mobalegh & Barati, 2012; Siddiqui et al., 2016). MTF questions encourage fast and precise scoring, require less reading than a corresponding multiple-choice question (Frisbie & Becker, 1991). Furthermore, they can assess higher cognitive abilities, such as comprehension, under Bloom's taxonomy (Downing & Yudkowsky, 2009; Richardson, 1992). Limitations and constraints for employing MTF questions include the fact that they have been demonstrated to have lower discrimination values than multiple-choice items; however, this disadvantage is mitigated by the fact that more items may be used concurrently (Frisbie & Becker, 1991). Another issue with MTF questions is that a high guess rate may add noise and damage an instrument's internal validity (Couch et al., 2018).

1.1 Purpose of the Study

The purposes of the study are as follows:

1. To develop a selected-response test to measure 11th-grade students' declarative knowledge of definitions of types of triangles, congruent and similar triangles, and identification of triangles by following the Standards for Educational and Psychological Testing set by the American Educational Research Association (AERA et al., 2014).
2. To determine the validity evidence for the test (DKTT) based on test content, internal structure, and reliability.

1.2 Research Questions of the Study

This study addressed the following research questions:

1. What are the dimensions of the Declarative Knowledge Test of Triangles (DKTT)?
2. What is the validity evidence of DKTT based on the test content?
3. Is DKTT reliable?

1.3 Significance of the Study

The majority of studies regarding the triangle definition or identification mentioned above collected validity evidence for the instrument based on test content (through judgmental evidence) (e.g., Atebe, 2008; Dündar & Gündüz, 2017; Fujita & Jones, 2007; Tsamir et al., 2015; Ulusoy, 2021). A few research (e.g., Ubuz & Aydın, 2018; Chinnappan et al., 2012; Gutiérrez & Jaime, 1998; Usiskin, 1982) that focused on developing a geometry test also included judgmental evidence and empirical evidence in addition to reliability estimates. When the literature was examined, the questions for defining triangles or types of a triangle were open-ended (i.e., asked students to write their definitions of triangles or types of a triangle) (see Altıparmak & Gürcan, 2021; Gutiérrez & Jaime, 1998; Tsamir et al., 2015; Ubuz & Aydın, 2018; Ulusoy, 2021), while studies on identification of triangles generally ask students to draw a triangle (e.g., Burger & Shaughnessy, 1986; Ulusoy, 2021) or participants choose a triangle (or cylinder or circle) from a series of figures (e.g., Tsamir et al., 2015; Senk, 1989; Ubuz, 2017). Researchers focused on the relationship between congruent triangle theorems and similar triangle theorems (e.g., Casanova, Cantoria & Lapinid, 2021; Haj-Yahya, 2022; Llinares & Clemente, 2019; López & Guzmán, 2011; Lutfi & Jupri, 2020; Parastuti et al., 2018; Patkin & Plaksin, 2011; Serow, 2006), proof in congruent and similar triangles (e.g., López & Guzmán, 2011; Leung et al., 2014; Sears & Chávez, 2015; Wang, Wang & An, 2018), students' difficulties in solving congruency and/or similarity questions that correspond to declarative and procedural knowledge level (Biber, 2020; Casanova et al., 2021; Poon & Wong, 2017; Wang et al., 2018; Wijaya et al., 2021), and the definitions of congruent and similar triangles using open-ended questions (e.g., González & Herbst, 2009; Haj-

Yahya, 2022; Dündar & Gündüz, 2017). This test will contribute to measuring students' declarative knowledge of triangles and, in particular, explore how these definitions and identifications are differentiated among students. Developing an accurate assessment of students' knowledge is essential for identifying areas of strength and weakness and for informing instructional decisions. By identifying the dimensions of the Declarative Knowledge Test of Triangles (DKTT), the study provides a detailed understanding of the content and structure of the test, which can be useful for educators and researchers. Using MTF questions in DKTT has several advantages. Firstly, they can be used to assess a large amount of content in a relatively short period of time. This can be particularly useful when testing a large number of students, as it allows for efficient assessment of their knowledge. Secondly, they are easy to score, as each option has only two possible responses. This can save time and resources compared to other types of questions that require more detailed scoring. MTF questions can be effective in measuring declarative knowledge as they require students to demonstrate their knowledge of concepts (triangle types, congruent and similar triangles, identification of triangles). They require students to think critically about the content and make judgments about its accuracy; thus, they measure students' ability to analyze and evaluate information.

CHAPTER 2

LITERATURE REVIEW

2.1 Declarative Knowledge

In cognitive psychology, there is a foundational difference between declarative (knowing about things) and procedural (knowing how) knowledge, as well as conditional knowledge (knowing when and where) knowledge (Kuhn, 2000). Nevertheless, these three types of knowledge have a hierarchical relation (Galeshi, 2014). According to Jüttner et al. (2013), mastery of content progresses from a low level of sophistication (declarative knowledge) to an intermediate level of sophistication (procedural knowledge) and eventually to the greatest degree of sophistication (conditional knowledge).

Anderson (1976, 1993, 2013) emphasizes that declarative acts, the conscious and control, are the foundations of knowledge and that this control opens the way for procedural processes. Furthermore, he contends that declarative knowledge is the base for knowledge transfers. Procedural knowledge, on the other hand, plays a crucial role in organizing concepts and gaining declarative knowledge (Lawson et al., 2000; Lawson, 1991). This is consistent with previous research, which has shown that procedural and declarative forms of knowledge are connected, and that one may be obtained from the other (Li et al., 1994; Ten Berge & Van Hezewijk, 1999; Dacin & Mitchell, 1986; Sahdra & Thagard, 2003; Willingham et al., 1989; Thagard, 2006; Hao et al., 2007; Lawson et al., 1991; Hanisch et al., 1991). Besides, declarative and/or conditional knowledge contribute to the procedural knowledge of students (Byrnes & Wasik, 1991; Engelbrecht et al., 2005; Hiebert & Wearne, 1996; Kuhn et al., 2016; Mack, 1990; Moss & Case, 1999). A reciprocal relation among declarative, conditional, and procedural knowledge has been reported in several research Aydın & Ubuz, 2010; Eryılmaz Çevirgen, 2012). According to the study of Aydın and Ubuz (2010), declarative knowledge improves conditional knowledge by enhancing the understanding of concepts and relationships between concepts. Likewise, declarative knowledge has a reciprocal relationship with procedural knowledge, implying that improving the knowledge of a concept contributes to improved algorithm selection and implementation or vice versa.

Declarative knowledge is defined by cognitive psychologists (e.g., Ragan & Smith, 1999) as long-term memory storage of facts and experiences (Yilmaz & Yalcin, 2012), and it is also referred to as "knowing-that" or "knowing what" (Aydın & Ubuz, 2010, p.443). It encapsulates both the knowledge of what something is and the knowledge of concepts and principles (Paris et al., 1983). Defined differently, declarative knowledge is information that can be recovered from memory without hesitation and is a key to achieving success in mathematics (Miller & Hudson, 2007). Recalling definitions of triangles (e.g., right, isosceles, and equilateral triangles) and congruence and similarity of triangles, their geometric properties, and notations for congruent and similar triangles are examples of declarative knowledge.

2.2 Mathematical Definitions

According to Tall and Vinner (1981), a definition is simply a set of words used to explain a concept. Unlike definitions generally, mathematical definitions "have the property that everything satisfying it belongs to the corresponding category and that everything belonging to the category satisfies the definition." (Alcock & Simpson, 2002, p. 28). Mathematical definitions have various distinctive characteristics. Definitions must be unambiguous (always understood in the same sense) (Zaslavsky & Shir, 2005) and only contain accurate mathematical terminology (Borasi, 1992; Levenson, 2012) since they are generated in a shared community (Zaslavsky & Shir, 2005). Furthermore, there may be different definitions of the same mathematical concept that are equivalent (de Villiers, 1998). Mathematicians and educators value this property of definitions because it allows for multiple equivalent ways to define a given object (Harel et al., 2006). According to van Hiele & van Hiele's third (1958) level: it is acceptable that one attribute can be derived from another or several attributes. For example, two different definitions of an equilateral triangle are given below.

- An equilateral triangle is a triangle in which all three sides are equal in length.
- An equilateral triangle is a triangle in which all three angles are equal and measure 60 degrees each.

Both definitions describe the equilateral triangle, and it is possible to derive one attribute (e.g., the equal side lengths) from the other attribute (e.g., the equal angles). This property is associated with the arbitrary nature of definitions, which are human creations (Linchevsky, 1992). Each definition is a component of a more elaborate, interconnected system of definitions (Van Dormolen & Zaslavsky, 2003). Alternative definitions range in minimality or form (e.g., textual vs. symbolic) (Linchevski, 1992; Van Dormolen & Zaslavsky, 2003; Zandieh & Rasmussen, 2010; Zaslavsky & Shir, 2005). . Only the descriptions necessary to ensure the object's identification are included in minimal definitions. Minimal definitions frequently follow a hierarchy, i.e., they incorporate definitions that have previously been formed by the community (Van Dormolen & Zaslavsky, 2003; Zaslavsky & Shir, 2005).

Understanding the development of meaning and the core of geometric concepts, theorems, and proofs depends heavily on definitions in geometry (Usiskin et al., 2008). A geometric concept's definition is seen to categorize appropriate examples and non-examples of that concept. The non-example of a concept stands for the geometric figure that does not meet the related attributes of the definition of the concept (Tsamir et al., 2015). The defining characteristics of a concept (called critical attributes) are what determine the necessary and sufficient conditions that define the concept's geometric properties (Tall & Vinner, 1981). To illustrate, the critical attributes of similar triangles are their angles and the lengths of their sides. In order for two triangles to be considered similar, they must have the same angles (this is a sufficient condition). Any two triangles that have the same angles will automatically meet the other essential condition for being similar as well, which is that the ratios of the lengths of their corresponding sides must be equal (these two conditions are both necessary for two triangles to be considered similar). In geometry, it is acceptable for definitions to include only the minimum set of necessary and sufficient attributes. The necessary attributes are those that are present in all examples of the concept, while sufficient attributes are a subset of necessary attributes that allow us to deduce the remaining necessary attributes. Vinner (1991) and Van Dormolen & Zaslavsky (2003) pointed out that it is ideal for mathematical definitions to include the minimum number of necessary attributes needed to define the concept. Leikin and Winicky-Landman also emphasized the importance of minimal definitions, stating that the essence of a mathematical definition is the identification of the necessary and sufficient conditions that define a concept, using the minimum number of such conditions possible (Leikin & Winicky-Landman, 2000, p. 64).

2.3 Identification of the Geometrical Shapes

Geometric concepts are derived from formal concept definitions and are abstract ideas (Tsamir et al., 2008). It is highlighted that analyzing the parts of concept images can help determine how well learners comprehend a concept (Vinner & Hershkowitz, 1980). This is why it has been so carefully explored how examples and non-examples may be used to comprehend learners' concept images, particularly in the domain of geometry (e.g., Cohen & Carpenter, 1980; Petty & Jansson, 1987; Vinner, 1991; Wilson, 1986). Beyond simply knowing its definition, the "concept image" is crucial to comprehending a concept: It is regarded to be "the total cognitive structure that is associated with the concept, which includes all the mental pictures and associated properties and processes" (Tall & Vinner, 1981, p. 152). Learners' concept images and definitions of two- and three-dimensional geometric shapes have received more focus in recent years. The majority of study has concentrated on the geometric solids knowledge of prospective math teachers (Bozkurt & Koç, 2012; Ertekin et al., 2014; Gökkurt & Soylu, 2016; Horzum & Ertekin, 2018; Kocak, et al., 2017; Ubuz & Gökbulut, 2015; Unlu & Horzum, 2018). Moreover, a few studies focused on how early-year teachers or students identify and exemplify cylinders and prisms (Tsamir et al., 2015), while others examined middle school students' concept images of geometric solids (Türnüklü & Ergin, 2016). According to studies, prospective math teachers at middle schools struggle to identify prisms (Horzum & Ertekin, 2018). According to Kocak et al. (2017), prospective middle school mathematics teachers only drew a right circular cylinder by identifying several critical attributes using a prototypical example (right circular cylinder). Accordingly, many studies have used particular identification tasks that include examples and non-examples to examine how well learners have learned geometric concepts, such as triangles (Burger & Shaughnessy, 1986; Tsamir et al., 2008), parallelograms (Fujita, 2012; Petty & Jansson, 1987; Ulusoy & Çakıroğlu, 2017), squares (Razel & Eylon, 1991), quadrilaterals (Vinner, 1991; de Villers, 1998; Currie & Pegg, 1998; Pratt & Davison, 2003; Zaslavsky & Shir, 2005). The concept image evolves and changes as time goes on due to experience. This process involves a concept's characteristics or properties, such as important geometric shape attributes, as well as an overview of all objects that fall under the concept in its entirety and the ability to identify connections between the concept and other concepts (Weigand et al., 2018). The prototypical examples of quadrilaterals are often identified correctly, but quadrilaterals in different orientations are not recognized (Fujita & Jones, 2007; Okazaki & Fujita, 2007; Fujita, 2012; Monaghan, 2000). Additionally, the equilateral and isosceles triangles are the most common prototypical triangles (Tsamir et al., 2008); children may not identify a triangle if it is not aligned with a horizontal side (e.g., Burger and Shaughnessy 1986). Non-prototypical examples are frequently viewed as non-examples (Hershkowitz, 1989; Schwarz & Hershkowitz, 1999; Wilson, 1990). Students are anticipated to solely employ critical attributes when constructing and identifying examples of geometric concepts in mathematics classes since understanding critical attributes are important for correctly identifying figures (van Hiele & van Hiele, 1958).

In their article on classifying examples and non-examples of triangles (Tirosh & Tsamir, 2008; Tsamir et al., 2008), they explained that intuitive examples of triangles and intuitive examples of non-triangles are figures that the tendency is to correctly identify them as triangles or as non-triangles, respectively, which is the basis for our question on measuring to identify triangles. Likewise, unintuitive examples of triangles and unintuitive examples of non-triangles are triangles and non-triangles, which are commonly misclassified as non-triangles and triangles, respectively. No formal definition establishes the differences between intuitive and non-intuitive examples and non-examples. They are based on replies from regular students for several tasks that demand them to identify examples and non-examples of the concept of a triangle. In this study, while intuitive non-examples include non-examples that were quickly identified as not being triangles (e.g., circles and trapezoids), non-intuitive non-examples included figures (i.e., missing a critical attribute of being a triangle, such as being an open or concave figure) that are not triangles and non-intuitive examples that are triangles. This study excludes prototypical examples (also known as intuitive examples) since children correctly identify them (Aslan & Aktaş Arnas, 2007; Tsamir et al., 2008).

Research studies on students understanding of identifying triangles mainly involved constructed-response questions (e.g., Bernabeu et al., 2021; Burger & Shaugnessy, 1986; Dağlı & Halat, 2016; Gutiérrez & Jaime, 1998; Lee, 2022; Mullis et al., 2000; Roldán-Zafra et al., 2022; Senk, 1989; Tsamir et al., 2008; Ubuz, 2017; Ulusoy, 2021, 2022; Usiskin, 1982; van Hiele, 1986) or used multiple-choice questions (e.g., Senk, 1989; Ubuz, 2017). We created multiple true-false questions where each option refers to a different figure that should be classified as a triangle or non-triangle which allowed us to more accurately identify students who had a complete, partial, or minimal understanding of the various answer statements (Brassil & Couch, 2019; Couch et al., 2015; Sands et al., 2018).

2.4 Triangle and Types of Triangles

The literature has focused on students' definition of a triangle (e.g., Burger & Shaughnessy, 1986; Ulusoy, 2021; Tsamir et al., 2015), identifying the properties of isosceles triangles(e.g., Chinnappan et al., 2012; Senk, 1989; Ubuz, 2017), and equilateral triangles (e.g., Gutiérrez & Jaime, 1998; Serow, 2006; Ubuz & Aydın, 2018; Usiskin, 1982; Altıparmak & Gürcan, 2021; Jin & Wong, 2021).In the study of Ubuz (2017), multiple-choice questions in the Van Hiele test asking to identify the triangle and analyze the properties of equilateral and isosceles triangles in declarative knowledge level were used. Additionally, students were asked to explain their rationale for each question based on their choice and establish a connection between equilateral triangle-isosceles triangle shape classes or concepts. In the question in which the properties of the isosceles triangle were asked, 57.5% of the students answered the question, "only two angles should be equal to each other." In the explanations, they stated that the two sides of the isosceles triangle are equal in length; therefore, the two angles corresponding to these sides must be equal in measure. Most students stated that they marked the option "must have at least two equal angles" because it was the closest to their thoughts. Altıparmak and Gürcan (2021) asked 4th-grade students to define a rectangle, square, equilateral triangle, isosceles triangle, and scalene triangle and the properties and relationship between these geometric shapes, identify them and explain the reason for their identification. The results showed that definitions of triangles are more accurate than those of rectangles and squares. The number of students who answered correctly to the definition of an equilateral triangle was the most among other geometric shapes. Equilateral triangles were correctly defined by 61 students, isosceles triangles by 56 students, and scalene triangles by 55 students. However, only five students out of 156 responded that the equilateral triangle is an isosceles triangle when asked which of the following triangles is isosceles (e.g., Ubuz, 2017). Despite creating a correct definition, the students were unable to depict this knowledge in their drawings for equilateral, isosceles, and scalene triangles. In a recent study by Ulusoy (2021), prospective early childhood teachers and prospective elementary school mathematics teachers wrote improper statements for triangle definitions by using necessary but insufficient conditions or neither necessary nor sufficient conditions with frequent incorrect terminology utilization. The necessary and sufficient conditions were provided in the formal statements for defining a triangle; however, they were mainly non-minimal and/or employed incorrect mathematical terminology. This finding is aligned with that of Tsamir's (2015) study, which found that only five out of the 31 definitions offered by early-year teachers were minimal and correct. When these studies were examined (Altıparmak & Gürcan, 2021; Gutiérrez & Jaime, 1998; Tsamir et al., 2015; Ubuz & Aydın, 2018; Ulusoy, 2021), the questions for defining triangles or types of a triangle were open-ended (i.e., asked students to write their own definitions of triangles or types of a triangle), and students were asked to draw the triangle or choose the triangle from the alternatives of the multiple-choice questions. These questions were useful for gathering detailed and qualitative information and exploring the respondent's knowledge and thoughts. However, one characteristic of these questions was that they could be time-consuming to analyze, as the responses may need to be coded and classified. In this case, it may be more effective to use closed-ended questions with a predetermined set of responses to measure students' declarative knowledge of triangles more objectively and reliably.

2.5 Congruent and Similar Triangles

An integral part of the foundational knowledge required to teach plane geometry is the concept of congruent triangles (Patkin & Plaksin, 2011), and it has a significant place due to its connection with similarity since if two triangles are congruent, they are also similar (Haj-Yahya, 2022).

Numerous research has examined how students perceive the theorems pertaining to congruent and/or similar triangles (Hadas et al., 2000; Hoyles, 1998; Jones et al., 2013). These research investigated students' understanding of how necessary and sufficient the criteria in the congruent and similar triangle theorems are for constructing and generating congruent or similar triangles. The theorems for congruent and similar triangles may be used as formal definitions for congruent and similar triangles since they have all the necessary and sufficient characteristics of mathematical definitions, such as minimalism and elegance (Van Dormolen & Zaslavsky, 2003; Zaslavsky & Shir, 2005). According to a recent research by Haj-Yahya (2022), participants did not consider the two theorems of congruent and similar triangles as the concepts' formal definitions. Others rejected the similar triangle theorem, which had only angles in its formal definition, but they agreed with the congruent triangle theorem, which had three equal sides. In order to attain van Hiele's (1958) third level, students were considered to have reached it if they agreed with both minimum and non-minimal definitions of congruent and similar triangles. At this level, students are aware that definitions are arbitrary, and they are able to acknowledge equivalence and accept equivalent definitions as in Gutiérrez and Jaime's (1998) study. Additionally, outcomes showed a strong inclination towards non-minimal definitions of congruent and similar triangle concepts. The majority of the students provided minimum definitions that only contained sides, according to an analysis of their replies and a comparison with those of the students who provided minimal definitions that also included angles.

Researchers used various instruments mainly consisting of constructed questions to gather data on congruent and similar triangles and the relationship between congruent triangle theorems and similar triangle theorems (e.g., Casanova, Cantoria & Lapinid, 2021; Haj-Yahya, 2022; Llinares & Clemente, 2019; López & Guzmán, 2011; Lutfi & Jupri, 2020; Parastuti, Usodo et al., 2018; Patkin & Plaksin, 2011; Serow, 2006), proof in congruent and similar triangles (e.g., López & Guzmán, 2011; Leung et al., 2014; Sears & Chávez, 2015; Wang, Wang & An, 2018), the definitions of congruent and similar triangles (e.g., González & Herbst, 2009; Haj-Yahya, 2022; Dündar & Gündüz, 2017), students' difficulties in solving congruency and/or similarity questions (Biber, 2020; Casanova et al., 2021; Poon and Wong, 2017; Wang et al., 2018; Wijaya et al., 2021). Students' difficulties were found to be stemmed from the fact that they did not master the basic concept of congruence and similarity (Ngirishi & Bansilal, 2019; Wijaya et al., 2021) and finding the similarity in overlapped triangles and Angle-Angle type questions when questions are examined in terms of similarity types (e.g., Biber, 2020; Poon & Wong, 2017), correspondence conceptions of congruency(Jones & Fujita, 2013), also in math language, natural language (Wang et al., 2018). The purposes of Haj-Yahya's (2022) study were how participants defined similar triangles and congruent triangles concepts; and what the characteristics of the definitions of congruent and similar triangles were by using open-ended questionnaires and interviews. In the first stage, the students were asked to define the congruent and similar triangles concepts. Categories of responses to the questions in Task 1 were generated by the results of the first stage of the questionnaire (e.g., non-minimal definition, minimal definition including only sides). The second stage of the questionnaire examined the participants' perceptions of the mathematical definitions of congruent and similar triangles. In the first question of the second stage, there were two definitions for similar triangles. One definition is a non-minimal definition, while the other is a minimal definition and similar triangles theorem based including only angles(AA theorem). In the second question of the second stage, there were two definitions for congruent triangles. One is a non-minimal definition, while the other is a minimal and congruent triangle theorem based including only sides (SSS theorem). The results of the study revealed that only a third of the participants agreed that the Angle-Angle Similarity Theorem was a formal definition of similar triangles, while half of the participants believed that the Side-Side-Side Congruence Theorem was a formal definition of congruent triangles. The students in the study who did not consider the similarity and congruence theorems to be formal definitions provided reasoning, such as the distinction between definitions and theorems due to the essence of the concept (see Okazaki, 2013) or general acceptance of definitions. On the other hand, those who did consider the theorems to be formal definitions gave explanations such as the theorems being equivalent to definitions or containing all necessary and sufficient attributes to define the concept, which aligns with the formal deductive level in van Hiele and van Hiele's hierarchy. The participants in the study believed that the equality or proportionality of a triangle's side lengths was more fundamental to the concepts of congruence and similarity than the angles of the triangles, which explains why more participants accepted the minimal definition of congruent triangles, which is based on the Congruence Theorem and only includes side lengths, rather than the minimal definition of similar triangles, which is based on the Similarity Theorem and only includes angles.

This research adopted the categorization for definitions of triangle types, congruent and similar triangles, by Zazkis & Leikin (2008). The definition is regarded as correct (appropriate) if it has necessary and sufficient conditions, incorrect (inappropriate) if necessary but not necessary, or not necessary and not sufficient. Only those descriptions are included in minimal definitions that are required to ensure object identification. Minimal definitions frequently follow a hierarchy, i.e., they comprise definitions that have already been created by the community (Van Dormolen & Zaslavsky, 2003; Zaslavsky & Shir, 2005). A hierarchical relationship between several concepts has been proposed by van Dormolen and Zaslavsky (2003). It can be concluded that definitions of an equilateral triangle based on triangles and the definition of a right triangle based on a polygon are hierarchical in nature when it is distinguished between levels of hierarchy (Shir & Zaslavsky, 2001, 2002). As Shir and Zaslavsky (2001) stated, the degree of the hierarchy of a concept can vary, and the degree decreases as one goes back. When defining the types of triangles following Shir & Zaslavsky (2001), it can be thought of as level 1 to define based on triangles, level 2 to define them based on polygons, and level 3 to define based on geometric shapes.

Based on the framework of van Dormolen and Zaslavsky (2003), Zazkis & Leikin (2008), we developed a Declarative Knowledge Test of Triangles (DKTT) and categorized the options for the definitions for the types of triangles and congruence and similar to those depicted in Table 2.1. Since triangle types may also be defined by auxiliary elements of a triangle, employing such elements were also included in the framework. Likewise, the familiarity or unfamiliarity of the definitions was introduced to the framework since it can affect students' responses in terms of picking the correct and incorrect definitions.

Table 2.1 Framework for the Options of DKTT

Content	The Hierarchy of the Concept	Conditions	Minimality	Intuitive /Non-Intuitive- Examples and Non-examples of Triangles	Using Auxiliary element	Familiarity of Definitions
Definitions of Types of Triangles	Level 1 (based on triangle)	Necessary and Sufficient	Minimal		(e.g., median, altitude)	
	Level 2(based on polygon)	Necessary and not sufficient	Non-minimal			
	Level 3 (based on geometric shape)	Not necessary and not sufficient				
Identification of Triangles				Intuitive non-example		
				Non-intuitive non-example		
				Non-intuitive example		

Table 2.1 Framework for the Options of DKTT (continued)

Definitions of Congruent and Similar Triangles	Necessary and Sufficient	Minimal		
	Necessary and not sufficient	Non-minimal		
	Not necessary and not sufficient			

CHAPTER 3

METHOD AND RESULTS

There are two studies in this chapter. Test development is the focus of Study 1, whereas instrument testing is the focus of Study 2.

3.1 Study 1: Test Development

The purposes of the study are to develop a selected-response test to measure 11th-grade students' declarative knowledge of definitions of types of triangles, congruent and similar triangles, and identification of triangles by following the Standards for Educational and Psychological Testing set by the American Educational Research Association (AERA et al., 2014) and then determine the validity evidence for the test (DKTT) based on test content, internal structure, and its reliability. The test development phases are detailed below under validity evidence based on test content.

3.1.1 Validity Evidence Based on Test Content

According to the Standards for Educational and Psychological Testing (AERA et al., 2014), content-based evidence constitutes the topic, wording, administration, and format of the items or questions on a test. It may also emerge from professional judgments of the connection between test components and the construct. Sources of validity based on test content are explained in Study 1: Test Development.

3.1.1.1 Generating Multiple True-False Questions on Triangles

The test development process began with the generation of multiple true-false questions based on essay-type questions from the Geometry Knowledge Test about

Triangles (GKT-T), which was developed by Ubuz and Aydın (2018) in light of learning objectives on triangles. Only questions in the declarative knowledge domain asking the definition of triangle types (equilateral, isosceles, right triangle), congruence and similarity of triangles, and identification of triangles were built on from the GKT-T to generate multiple true-false questions.It was determined that the Multiple True-False (MTF) item format would be appropriate for this developed test after considering the study's purpose, the level of knowledge it measures, its content, and the possible correct and incorrect responses that students can provide to open-ended questions. In other words, the following supports our decision to use MTF as the item format:

- The concepts under consideration can be defined differently.
- More figures can be identified as a triangle.
- Each response must be marked as either true or false by the student. MTF questions can therefore be used to identify students who have mixed understanding (Parker et al., 2012).
- MTF enables us to provide more options compared to the multiple-choice (MC) format.
- MTF question format covers a wider range of subjects than the MC format, which only collects a single student response, whereas MTF questions give numerous data on student understanding (Couch et al., 2018).
- MTF items are easy to administer and score, making them a convenient and efficient testing method.
- MTF items can be less subjective than other items, such as essay questions, which can be evaluated differently by different graders. This can make multiple true-false items a more fair and reliable way to test knowledge.

Three steps were followed in the question generation process:

3.1.1.1.1 Writing Different Responses to the Questions

In this step, we tried to write different correct and incorrect responses to the concepts by considering possible errors and misconceptions that students might have. Table 3.1 depicts the options written in this step.

Table 3.1 Options for the Concepts Defined by Us

The concepts defined	Options for the corresponding concepts
An equilateral triangle	1, 2, 3, 5, 6, 7
A right triangle	12, 13, 15, 16
An isosceles triangle	22, 23, 24, 25, 26
Congruence of triangles	32, 33, 34, 35, 37, 38
Similarity of triangles	42, 43, 44

3.1.1.1.2 Taking Possible Responses from Literature

The literature was also helpful in providing students' responses to the questions which asked about the identification of triangles and definitions for congruent and similar triangles. This helped to ensure that the test options were representative of the full range of the constructs being measured and that they were appropriate for the intended test-taking population (11-grade students). We conducted a thorough review of the literature in the areas in which the identification of triangles, definitions for congruent and similar triangles, and types of triangles related to the declarative knowledge that the test is intended to measure. The options taken from the literature and their question content was given in Table 3.2.

Identification of triangles: In Ubuz and Aydın (2018)'s article, students were asked to identify which figures represent a triangle. Figure 3.1 shows the figures that are triangles or non-triangles. Figure 3.1 was used as students' responses to the identifying triangles.

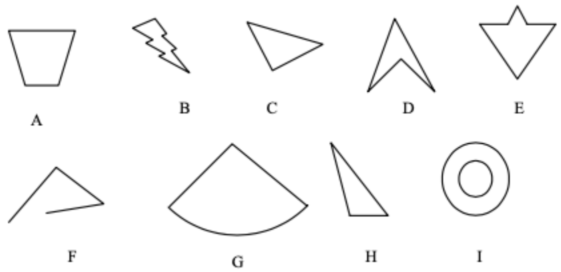

Figure 3.1 Figures Used in Identifying Triangles

Intuitive non-examples were non-examples that were quickly identified as not being triangles (i.e., A and I in Figure 3.1), non-intuitive non-examples were missing critical attributes of being triangles (i.e., B, D, E, F, and G in Figure 3.1) that were not triangles and non-intuitive examples were triangles (i.e., C and H in Figure 3.1). We excluded prototypical examples (also known as intuitive examples) since children correctly identify them (Aslan et al., 2007; Tsamir et al., 2008).

Definitions of congruent triangles: For definitions of congruent triangles, we used one incorrect response from Haj-Yahya (2022)'s study, which is "Two congruent triangles are similar if each of the triangles covers the other" (p. 11).

Table 3.2 Options Taken From the Literature and Their Question Content

Options	The question content of options is…
50, 51, 52, 53, 54, 55, 56, 57, 58	the identification of triangles
36	the definitions of congruent triangles

3.1.1.1.3 The First Field Testing of the 11th Graders by Essay-type Questions

383 11th-grade students from five schools in Ankara—three Anatolian and two Science High Schools—were given essay-type questions whose learning objectives

are shown in Table 3.3. The frequency of the answers given by the students to the questions was determined. Most of the responses the students provided were asked to define triangle types, and congruent and similar triangles were those we had previously written. Students who gave different responses to these questions are used to create new options. Students' responses were not considered for identifying triangles since figures were used as options.

Table 3.3 Learning Objectives of Essay-type Questions

Learning Objectives
Define equilateral triangle
Define right triangle
Define isosceles triangle
Identify triangles
Define congruency of triangles
Define similarity of triangles

Specifically, options 4, 8, 9, 10, and 11 for the definition of an equilateral triangle; options 14, 17, 18, 19, 20, and 21 for the definition of a right triangle; options 27, 28, 29, 30, 31 for the definition of congruent triangles; options 39, 40, and 41 for the definition of similar triangles are extracted from students' responses. The learning objectives of developing the "Declarative Knowledge Test of Triangles (DKTT)" are shown in Table 3.4.

Table 3.4 Learning Objectives of DKTT

Learning Objectives	Options
Select correct and incorrect definitions for an equilateral triangle	1-11
Select correct and incorrect definitions for a right triangle	12-21
Select correct and incorrect definitions for an isosceles triangle	22-31

Table 3.4 Learning Objectives of DKTT (continued)

Select correct and incorrect definitions for congruent triangles	32-41
Select correct and incorrect definitions for similar triangles	42-49
Select figures that represent and do not represent triangles	50-58

3.1.1.2 Item Refinement

Expert evaluation and developmental field tests were employed to refine the questions and options. Before testing the instrument, the developed questions were sent to four university staff members and two high school mathematics teachers for the subject matter expert evaluation in order to evaluate the instrument's content, context, and construct, along with checking each item's readability, clarity, and congruency with the related geometrical knowledge construct. The three university staff members were who conducted research on geometry and had extensive university-level mathematics teaching expertise; one member had considerable experience and research in mathematics education. The mathematics teachers had notable amount of experience in teaching mathematics and geometry at the high school level. Revisions made by subject matter experts are given below:

- The university staff members who work in geometry and teachers accentuated the usage of side lengths and angle measures rather than only side and angle when defining triangle types, congruent and similar triangles.
- An option created by the student's definition, "A right triangle is a triangle formed by joining the non-adjacent sides of two perpendicular line segments," was removed from the test because subject matter experts thought it was problematic because joining the non-adjacent sides of two perpendicular line segments can create more than one triangle, how these lines join were not clear (e.g., linear, curve).

- The option, which is similar to Euclid's method of superposition, "Two triangles are congruent if one can cover the other and vice versa," was decided to be kept in the test by most experts, although "cover" has no mathematical definition.
- The answer key for option 21, "A right triangle is a triangle whose product of the slopes of two perpendicular sides is -1.", is determined to be incorrect by experts since it was not comprehensive when the right triangle was placed in the Cartesian coordinate system. Assume that we place one of the right sides of a right triangle on the x-axis in the Cartesian coordinate system; the x-axis has a slope of 0. When we place the other perpendicular side on the y-axis, the slope of the y-axis is infinity. The product of the slopes of these two perpendicular sides, $0.\infty$ is indeterminate and does not equal -1; therefore, this definition is not comprehensive.

Also, a second developmental field test was carried out to refine and assure the clarity and comprehensibility of the options. Four students were interviewed, two of them were middle achievers, and the remaining two were high achievers based on their previous semester's mathematics grades, and their thought processes were analyzed as they answered each option. The reason for choosing these four students was to get a diverse sample of test-takers with different levels of achievement in mathematics which could help to prevent Type 1 and Type 2 errors. Including both middle and high achievers in the second field test allowed us to see how the options being tested were understood by a range of students and to identify any potential issues with clarity or comprehensibility that might not have been apparent with a more homogenous group of test-takers and then to make appropriate modifications to improve the clarity and comprehensibility of the options being tested. The following modifications are implemented to the options after the interview:

- Option 7," Equilateral triangle is a triangle with equal side lengths and congruent interior angle measures.", was rewritten because it was misunderstood by some interviewed students who understood that the length of the sides is equal to the measure of the interior angles.

- Options 10 and 11 are understood by the students as interior angle measures and exterior angle measures of an equilateral triangle are equal. We rewrote these options as “an equilateral triangle is a triangle in which all interior angles are congruent, all exterior angles are congruent, and side lengths are equal in length.”.
- Option 21, “a right triangle is a triangle whose slopes of two perpendicular sides are opposite multiplicative inverses of each other.”, was altered as “A right triangle is a triangle whose product of the slopes of two perpendicular sides is -1.” because students were not familiar with the former definition. Hence, the definition was revised to be what learners see in theorems found in mathematics textbooks stating that if two nonvertical lines are perpendicular, the product of their slopes is -1.

After making changes to the questions and their options based on the findings from the second field testing, third field testing was conducted to decide the average amount of time students required to complete the test. The test was administered to 6 students, including middle and high achievers, without time limitations. When the students were given the test, the time it took each student to complete it was recorded to predict how long the test would take. Each student's time spent on the test was recorded, along with an estimation of how long it would take them to finish the test. The test was completed by the students at the following times: 15, 16, 17, 17, 20, and 20 minutes. The time limitation for this test was established at 20 minutes since the student who took the longest finished it in that amount of time.

3.2 Study 2: Instrument Testing

Instrument testing is an essential part of test development, as it ensures that the measurement tool being used to conduct the test is accurate and reliable. The following sections describe the instrument testing, as well as identify the underlying dimensions of the DKTT.

3.2.1 Declarative Knowledge Test of Triangles (DKTT)

In DKTT, Question 1 has 11 options; Questions 2, 3, and 4 have ten options each; Questions 5 and 6 have eight and nine options, respectively. There are 6 question stems and a total of 58 options based on these questions. Questions 1, 2, and 3 ask students to select correct and incorrect definitions for the types of triangles (equilateral triangle, right triangle, and isosceles triangle). Questions 4 and 5 ask to select correct and incorrect definitions for congruent triangles and similar triangles, respectively. Question 6 asks students to select figures that represent a triangle.

3.2.2 Test Administration and Scoring

Declarative Knowledge Test of Triangles (DKTT), a test of six question stems and 58-option, was administered to 379 grade-11 students, and 20 minutes were given to them to complete the test. Table 3.5 describes the information on the gender and high school types of the students participating in the study. Half of the students are male, and half are female. While the rate of Anatolian High School students was 39.3%, the rate of Science High School students was 60.7%. Anatolian High Schools are educational institutions that aim to prepare students for higher education programs according to their interests, abilities, and achievements. Middle Schools or İmam-hatip Middle Schools are institutions that provide boarding and/or daytime education with a four-year education period, and students are placed on the basis of the central exam score or address information. Course schedules and curricula approved by the Ministry are applied in these schools. Education is done in Turkish. Science High Schools are educational institutions that aim to be a source for the education of students in the fields of science and mathematics as scientists. They are institutions that provide boarding and/or daytime education in a Middle School or İmam-hatip Middle school with a four-year education period based on the central exam score. While students with high scores are placed in Science High Schools, those with lower scores are placed in Anatolian High Schools. In Science High Schools, course schedules and curricula approved by the Ministry are applied. Laboratory and application studies are emphasized in science programs. Education is done in Turkish. Moreover, science high schools aim to be a resource for students to be trained as scientists in the fields of science and mathematics. The curricula of Science and Anatolian High Schools are the same, however, the student at Science High School studies Physics, Chemistry, and Biology for 15 hours per week, while the student at Anatolian High School studies 6 hours per week.

Table 3.5 Gender and School Type Distribution of Students

			Frequency (f)	Percent (F %)	Valid Percent
Gender		Male	180	47.5	50
		Female	180	47.5	50
		Total	360	95	100
		Missing	19	5	
	Total		379	100	
School		Anatolian	149	39.3	
		Science	230	60.7	
	Total		379	100	

The response to each option is coded as "1" if the response is correct and "0" if incorrect. There were six questions in DKTT, with 8 and 11 options under each. The students received two distinct scores for Questions 1, 2, 3, 4, and 5, and three distinct scores for Question 6. Each option under Questions 1 to 5 was categorized as either a minimal or non-minimal definition of the concept corresponding to the question. For each question, 1 to 5, the total score of students' responses to the options with minimal definitions were added and named as sub-score items with the minimal definition for the question. Likewise, the total score of students' responses to the options with non-minimal definitions was added and named as sub-score items that are non-minimal definitions for the question. To clarify, students' answers for each option under Question 1 were categorized as either minimal or non-minimal definitions. The total score from options with minimal definitions was added and named as the minimal definitions for Question 1. The new score is named sub-score item S1.1. Similarly, the total score from options with non-minimal definitions was added, which is the non-minimal definitions for Question 1, and the new score is named sub-score item S1.2. Question 6 asks students to identify triangles and non-triangles. There are two types of non-valid triangle examples (intuitive non-examples and non-intuitive non-examples) and one atypical triangle example: non-intuitive example (Tsamir et al., 2008). Three separate scores were created by adding the responses to the options from the provided figures corresponding to these classes (intuitive non-examples, non-intuitive non-examples, and non-intuitive examples). Sub-scored item S6.1 is the total score for options that fall in intuitive non-examples(non-triangle), sub-scored item S6.2 is the total score for options that fall in non-intuitive non-examples (non-triangle), and sub-scored item S6.3 is the total score for options that fall in non-intuitive examples (triangle).This scoring method is utilized because many options had the same theoretical justification, and the use of factor analysis in validating instruments containing dichotomous data is contentious. While some authors (Polit, 1996; Streiner, 1994) advise against using factor analysis with dichotomous data, other statisticians advise using it with discretion (Tabachnick & Fidell, 2001). Information about the sub-scores used in factor analysis was given in Table 3.6.

Table 3.6 Sub-scored Items of DKTT

Sub-scored Items	The total score of options involving …
S1.1	minimal definitions for an equilateral triangle
S1.2	non-minimal definitions for an equilateral triangle
S2.1	minimal definitions for a right triangle
S2.2	non-minimal definitions for a right triangle

Table 3.6 Sub-scored Items of DKTT (continued)

S3.1	minimal definitions for an isosceles triangle
S3.2	non-minimal definitions for an isosceles triangle
S4.1	minimal definitions for congruent triangles
S4.2	non-minimal definitions for congruent triangles
S5.1	minimal definitions for similar triangles
S5.2	non-minimal definitions for similar triangles
S6.1	intuitive non-example
S6.2	non-intuitive non-example
S6.3	non-intuitive examples

3.2.3 Validity Evidence Based on Test Content of the DKTT

Item analysis of DKTT (i.e., item difficulty and discriminations index) was calculated to provide evidence based on test content. Findings regarding DKTT's item difficulty and item discrimination indices for sub-scored items were given in detail.

3.2.3.1 Item Discrimination and Difficulty Index for Sub-scored Items

We used the recognized range of 0.24-0.91 for the item difficulty indices for DKTT’s sub-scored items, as proposed by Downing and Yudkowsky (2009). Sub-scored items under 0.24 were classified as extremely difficult, which DKTT did not have, whereas options over 0.91 were classified as extremely simple sub-scored items (S6.1, S6.2, and S6.3). The remaining sub-scored items were determined to be of moderate difficulty (S1.1, S1.2, S2.1, S2.2, S3.1, S3.2, S4.1, S4.2, S5.1, S5.2).

Ebel and Frisbie (1991) classified sub-scored items according to the following rules: sub-scored items with an index of 0.2 were poor sub-scored items (S1.1, S1.2, S2.1, S2.2, S3.1, S5.1, S6.1, and S6.3), but indices of greater than 0.4 show excellent discrimination (S3.2), between 0.30 and 0.39 show good discrimination (S4.1, S5.2, and S6.2), and 0.20 to 0.29 indicate acceptable discrimination (S4.2).

Based on the item discrimination and difficulty indices provided in Table 3.7, the following interpretations can be made:

The item discrimination index for sub-scored items S1.1, S1.2, S3.1, S5.1, S6.1, and S6.3 indicated that these items were not able to differentiate between high- and low-ability test takers effectively. These non-discriminant sub-scored items were definitions of an equilateral triangle (S1.1 and S1.2), minimal definitions for an isosceles triangle (S3.1) and similar triangles (S5.1), and intuitive non-example (S6.1) and non-intuitive examples (S6.3) of triangles. Among these items, minimal definitions of an equilateral triangle (S1.1), isosceles triangle (S3.1), and similar triangles (S5.1) were of moderate difficulty, while intuitive non-example (S6.1) and non-intuitive examples of triangles (S6.3), and non-minimal definitions of an equilateral triangle (S1.2) were found to be relatively easy. The item discrimination index for sub-scored items S2.1 (minimal definitions for a right triangle), S2.2 (non-minimal definitions for a right triangle), S5.2 (non-minimal definitions for similar triangles), S6.2 (non-intuitive non-example), S4.1 (minimal definitions for congruent triangles), S4,2 (non-minimal definitions for congruent triangles) indicated that these items were able to differentiate between high- and low-ability test takers to some extent. Among these sub-scored items, S2.1, S2.2, S5.2, and S4.1 were of moderate difficulty; S6.2 and S4.2 were relatively easy. S3.2 (non-minimal definitions for an isosceles triangle) was able to effectively differentiate between high- and low-ability test takers, and it was relatively easy. Lastly, sub-scored items

(S6.1, S6.2, and S6.3) about the identification of triangles were found to be relatively easy; thus, students were able to choose the correct and incorrect figures representing triangles.

Table 3.7 Item Difficulty and Discrimination Indices for Sub-scored Items

Sub-scored Items	Item Difficulty Index	Item Discrimination Index
S1.1	0.54	0.13
S1.2	0.82	0.15
S2.1	0.75	0.17
S2.2	0.56	0.17
S3.1	0.64	0.13
S3.2	0.86	0.4
S4.1	0.74	0.3
S4.2	0.85	0.29
S5.1	0.61	0.13
S5.2	0.53	0.3
S6.1	0.95	0.16
S6.2	0.93	0.32
S6.3	0.95	0.13

3.2.4 Validity Evidence Based on Internal Structure: Assessing Dimensionality of DKTT

The degree to which the relationships between test items and test components fit the construct upon which the suggested test scores interpretations are based can be shown through an analysis of a test's internal structure (AERA et al., 2014). A part of validating a test's internal structure is assessing the test's dimensionality. A prevalent statistical technique to assess the dimensionality of a data set is factor analysis (Bollen, 1989; Brown, 2006; Kline, 2015; Thompson, 2004). This is a statistical method that is used to identify underlying patterns or relationships among a set of variables. In the context of a test, factor analysis can be used to identify the underlying dimensions or factors that are measured by the test items and to assess the extent to which the test items are related to these factors. This can provide evidence of the internal structure of the test and can help to demonstrate the validity of the test. In this study, Exploratory Factor Analysis (EFA) was used to provide evidence to support the validity of an internal structure of a measurement instrument (DKTT) by verifying the number of underlying dimensions and the pattern of item-to-factor relationships (i.e., factor loadings). The reason for using EFA is that compared to other techniques, such as item response theory and latent variable modeling, EFA is relatively easy to perform and interpret, allows us to investigate a wide range of structures, including simple structures with one or two factors, as well as more complex structures with multiple factors. It is a widely accepted and commonly used technique in the field of psychology and education research and can help us identify patterns in a large dataset, especially useful when working with a large number of variables and looking for ways to simplify the data. The Exploratory Factor Analysis (EFA) included the following:

1. Measures of fit, such as the KMO measure and Bartlett's test of sphericity, which can be used to assess the suitability of the data for factor analysis.

2. The total variance explained is included in the results of exploratory factor analysis. It indicates the proportion of the total variance in the data that is accounted for by the factors. This can be useful for determining the adequacy of the factor solution and for comparing different factor solutions.
3. The number of factors extracted and the eigenvalues of the factors. The eigenvalues can be used to determine the relative importance of each factor.
4. A scree plot, which is a graph that shows the eigenvalues of the factors on the y-axis and the number of factors on the x-axis. The point at which the plot levels off indicates the number of factors that should be retained.
5. The factor loadings, which are the correlations between the factors and the individual variables in the data. High factor loadings indicate that a particular variable is strongly associated with a particular factor.
6. The parallel analysis method is used to determine the number of factors to retain. It involves generating random data with the same characteristics as the original data, and then performing factor analysis on the random data. The number of factors with eigenvalues greater than those in the random data is the number of factors that should be retained in the original data. The parallel analysis method is useful because it provides a more objective way to determine the number of factors to retain, compared to methods such as the scree plot, which can be subject to interpretation.
7. An unrotated factor matrix was included in the results of EFA. This matrix shows the factor loadings before the factors have been rotated, and can be useful for identifying the underlying structure of the data, and provide valuable information.
8. A rotated factor matrix, which shows the factor loadings after the factors have been rotated to a more interpretable orientation.
9. An interpretation of the factors and their meaning. This involved examining the variables with high loadings on each factor and determining what they have in common, in order to come up with a label or name for the factor.

3.2.4.1 KMO and Bartlett’s Test

Before utilizing factor analysis, sampling adequacy and sphericity were tested by Kaiser-Meyer- Olkin and Bartlett’s test, respectively, by using SPSS 25, shown in Table 3.8.

The Kaiser-Meyer-Olkin (KMO) measure is a statistic that assesses the suitability of data for factor analysis. It ranges from 0 to 1, with values closer to 1 indicating that the data is more suitable for factor analysis. In this case, the KMO measure is 0.77, which suggests that the data is suitable for factor analysis.

Bartlett's test of sphericity is a statistical test used to assess the null hypothesis that the correlation matrix of the data is an identity matrix, which would indicate that the variables are not related. In this case, the approximate chi-square value is 1487.889, the degrees of freedom is 78, and the significance value is 0.000. These results indicate that the null hypothesis can be rejected, and that the variables are related.

Table 3.8 KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.77
Bartlett's Test of Sphericity	Approx. Chi-Square	1487.89
	df	78
	Sig.	0

3.2.4.2 Total Variance Explained

The following table contains the number of extracted dimensions, the initial eigenvalues connected to the specified dimensions, the percentage of the total variance, and the cumulative percentage of every dimension. Principal Axis Factoring (PAF) is employed as an extraction method because it does not assume multivariate normality and is less likely to encounter estimation issues than, say, maximum likelihood extraction (Fabrigar & Wegener, 2012). Additionally, PAF allows for the factors to be rotated in order to achieve a more interpretable solution. Only factors having eigenvalues are greater than one retained for DKTT, as shown in Table 3.9 and Table 3.10. The tables presented the results of exploratory factor analysis, including the total variance explained by each factor. The initial eigenvalues and the extraction sums of squared loadings were two different ways of quantifying the amount of variance explained by each factor. The total and cumulative percentages of variance explained by each factor were also shown. The numbers under the "factor" column indicate the specific factor or dimension that each row of data.

The results indicate that the first factor explains the largest amount of variance, with an initial eigenvalue of 3.584 and a total percentage of variance explained of 27.569. The second factor explains the next largest amount of variance, with an initial eigenvalue of 1.809 and a total percentage of variance explained of 13.915. The remaining factors each explain progressively less variance.

Overall, the results suggest that the first two factors explain the majority of the variance in the data, while the remaining factors explain progressively less variance.

Table 3.9 Total Variance Explained with Unrotated Matrix

	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Factor	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.584	27.569	27.569	3.222	24.788	24.788
2	1.809	13.915	41.483	1.431	11.004	35.792
3	1.355	10.420	51.903	.706	5.431	41.223
4	1.163	8.948	60.851	.627	4.827	46.049

Table 3.9 Total Variance Explained with Unrotated Matrix (continued)

5	.956	7.358	68.209
6	.877	6.743	74.951
7	.706	5.427	80.379
8	.648	4.982	85.361
9	.563	4.333	89.694
10	.505	3.884	93.578
11	.428	3.293	96.871
12	.254	1.955	98.826
13	.153	1.174	100.000

Extraction Method: Principal Axis Factoring.

It is typically more informative to interpret the total variance explained from the rotated factor matrix, rather than from the unrotated matrix. This is because the rotated matrix provides a more interpretable representation of the factors, which can make it easier to understand the results and make inferences about the data. Therefore, the total variance explained from the rotated factor matrix is given below. When the matrix is rotated, factor 4 may not be retained.

Table 3.10 Total Variance Explained with Rotated Matrix

	Initial Eigenvalues			Rotation Sums of Squared Loadings		
Factor	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.584	27.569	27.569	2.312	17.786	17.786
2	1.809	13.915	41.483	1.712	13.166	30.952

Table 3.10 Total Variance Explained with Rotated Matrix (continued)

3	1.355	10.420	51.903	1.290	9.924	40.877
4	1.163	8.948	60.851	.672	5.173	46.049
5	.956	7.358	68.209			
6	.877	6.743	74.951			
7	.706	5.427	80.379			
8	.648	4.982	85.361			
9	.563	4.333	89.694			
10	.505	3.884	93.578			
11	.428	3.293	96.871			
12	.254	1.955	98.826			
13	.153	1.174	100.000			

Extraction Method: Principal Axis Factoring.

3.2.4.3 Scree Test Method

The scree test is a method used in exploratory factor analysis to determine the number of factors to retain in the analysis. This approach plots the eigenvalues' magnitudes (vertical axis) against their ordinal numbers (whether the first eigenvalue, the second, etc.). The magnitude of subsequent eigenvalues often decreases abruptly before seeming to level off. It is suggested that all eigenvalues (and hence factors) be kept in the abrupt decline that comes before the line when they begin to level out. According to the scree plot in Figure 3.1, the magnitude of successive eigenvalues is steeped and then tends to level off after the fourth-factor number.

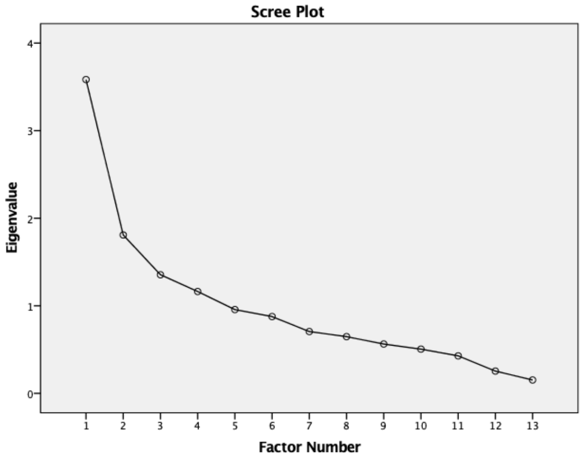

Figure 3.2 Scree Plot

3.2.4.4 Parallel Analysis Method

The parallel analysis method is a statistical technique used in exploratory factor analysis (EFA) to determine the number of factors that should be extracted from a dataset. This method involves generating random data with the same characteristics as the original dataset and then conducting an EFA on the random data. The number of factors extracted from the random data was then used as a benchmark to determine the appropriate number of factors to extract from the original dataset. Thus, only factors whose eigenvalues are higher than their random counterparts will be kept from the original data set. Alternately, the 95th percentile of these duplicated values can be used as the comparison value instead of the average of the eigenvalues for a particular component, offering a slightly more rigorous evaluation of factor significance. (Pituch & Stevens, 2015). After applying the parallel analysis, the number of factors to be retained was specified as four factors. The naming of the factors will be explained in the next section.

3.2.4.5 Unrotated and Rotated Component Matrix

The matrix represents the results of a principal axis factoring analysis on a set of variables on declarative knowledge of triangles. The analysis was able to extract four factors, and a total of 42 iterations were required.

The sub-scored items in the matrix shown in Table 3.11 represent total scores for different options involving various definitions of types of triangles, congruent and similar triangles, and identification of triangles. The options were divided into two categories: those involving minimal definitions (S1.1, S2.1, S3.1, S4.1, S5.1 etc.) and those involving non-minimal definitions (S1.2, S2.2, S3.2, S4.2, S5.3 etc.). The definitions were equilateral triangles, right triangles, isosceles triangles, congruent triangles, and similar triangles. The final three items in the matrix (S6.1, S6.2, and S6.3) involve non-examples, with S6.1 being intuitive non-examples and S6.2 and S6.3 being non-intuitive non-examples and examples, respectively.

Factor loadings represent the strength and direction of the association between each item (variable or measure) and a particular factor in a factor analysis. Factor loadings can range from -1 to 1, with negative values indicating a negative association and positive values indicating a positive association. The absolute value of the factor loading indicates the strength of the association, with larger values indicating a stronger association.

In table 3.11, the item S6.1 (intuitive non-examples) has a loading of .79 on Factor 1, -.45 on Factor 2, and .14 on Factor 3. This indicates that S6.1 has a strong positive association with Factor 1, a moderate negative association with Factor 2, and a weak positive association with Factor 3.

S6.2 (non-intuitive non-examples) has a loading of .79 on Factor 1, -.40 on Factor 2, and .17 on Factor 3. This indicates that S6.2 has a strong positive association with Factor 1, a moderate negative association with Factor 2, and a weak positive association with Factor 3.

S6.3 (non-intuitive example) has a loading of .73 on Factor 1 and -.39 on Factor 2. This indicates that S6.3 has a strong positive association with Factor 1 and a moderate negative association with Factor 2. Overall, question asking identification of triangles was loaded onto factor 1.

S4.1(minimal definitions for congruent triangles) has a loading of .60 on Factor 1, .281 on Factor 2, -.352 on Factor 3, and .129 on Factor 4. This indicates that S4.1 has a strong positive association with Factor 1, a moderate positive association with Factor 2, a moderate negative association with Factor 3, and a weak positive association with Factor 4.

S5.1 (minimal definitions for similar triangles) has a loading of .55 on Factor 1, .25 on Factor 2, -.22 on Factor 3, and .22 on Factor 4. This indicates that S5.1 has a moderate positive association with Factor 1, a moderate positive association with Factor 2, a moderate negative association with Factor 3, and a moderate positive association with Factor 4.

S4.2 (non-minimal definitions for congruent triangles) has a loading of .48 on Factor 1, .16 on Factor 2, and -.33 on Factor 3. This indicates that S4.2 has a moderate positive association with Factor 1, a weak positive association with Factor 2, and a moderate negative association with Factor 3.

S3.2 (non-minimal definitions for an isosceles triangle) has a loading of .339 on Factor 1, .31 on Factor 2, and -.14 on Factor 4. This indicates that S3.2 has a moderate positive association with Factor 1, a moderate positive association with Factor 2, and a weak negative association with Factor 4.

S5.2 (non-minimal definitions for similar triangles) has a loading of .26 on Factor 1 and -.15 on Factor 3. This indicates that S5.2 has a weak positive association with Factor 1 and a moderate negative association with Factor 3.

S3.1 (minimal definitions for an isosceles triangle) has a loading of .43 on Factor 1, .63 on Factor 2, .24 on Factor 3, and -.41 on Factor 4. This indicates that S3.1 has a moderate positive association with Factor 1, a strong positive association with Factor 2, a moderate positive association with Factor 3, and a moderate negative association with Factor 4.

S2.1, S1.1, and S1.2 represent the total scores for options involving minimal definitions for a right triangle, minimal definitions for an equilateral triangle, and non-minimal definitions for an equilateral triangle, respectively. The factor loadings for these items indicate how strongly they are associated with each of the four factors that were extracted in the principal axis factoring analysis. Specifically, S2.1 has a factor loading of 0.322 on Factor 1, 0.374 on Factor 2, and 0.295 on Factor 3, but low loading on Factor 4. S1.1 has a factor loading of 0.212 on Factor 2 and 0.20 on Factor 3, but low loading on Factors 1 and 4. S1.2 has low loading on Factors 1, 2, and 3, but a factor loading of 0.245 on Factor 4.

Table 3.11 Unrotated Component Matrix of DKTT

Sub-scored Items	Factor			
	1	2	3	4
S6.1	.799	-.452	.138	
S6.2	.787	-.407	.164	
S6.3	.740	-.390		
S4.1	.602	.286	-.344	.141
S5.1	.552	.258	-.210	.226
S4.2	.485	.170	-.330	
S3.2	.339	.319		-.148
S5.2		.264		-.149
S3.1		.426	.619	.225
S2.1		.322	.377	.305
S1.1			.212	.209

Table 3.11 Unrotated Component Matrix (continued)

S2.2		.186	.210	.325
S1.2				.262

Extraction Method: Principal Axis Factoring.

Suppressed small coefficients below absolute value of .10.

Table 3.12 Rotated Factor Matrix of DKTT

Sub-scores	Factor			
Items	1	2	3	4
S6.1	.906	.206		
S6.2	.878	.195		
S6.3	.794	.269		
S4.1	.153	.732	.153	
S5.1	.156	.631	.147	.151
S4.2	.151	.592		
S5.2		.305		
S3.1		.220	.866	-.127
S2.1		.157	.483	.270
S3.2		.305	.367	
S1.1			.291	.106
S2.2		.115	.168	.590
S1.2				.437

Extraction Method: Principal Axis Factoring.

Rotation Method: Varimax with Kaiser Normalization.

Rotation converged in 5 iterations.

Suppressed small coefficients below absolute value of .10.

The rotated factor matrix showed the results of a principal axis factoring analysis on a set of variables or data related to declarative knowledge of triangles. The analysis was able to extract four factors, and a varimax rotation with Kaiser normalization was applied. The rotated factor matrix shows the factor loadings for each item on each of the four factors. Based on Table 3.12, factor 1 appears to be primarily associated with items S6.1, S6.2, and S6.3, which are related to the identification of triangles. Factor 2 is associated with items S4.1, S5.1, S4.2, and S5.2, which are related to minimal definitions for congruent triangles and similar triangles and non-minimal definition of congruent triangles. Factor 3 is associated with items S3.1, S2.1, and S3.2, which are related to minimal definitions for isosceles triangles and right triangles. Factor 4 is associated with items S2.2 and S1.2, which are related to minimal definitions for equilateral triangles. S1.1 (minimal definitions for an equilateral triangle) does not have any significant factor loadings, indicating that it is not strongly associated with any of the four factors. S5.2 is included in Factor 2 (Definitions of Congruent and Similar Triangles) due to its theoretical construct and item characteristics, the option under S5.2 (i.e., Option 45) being an ambiguous statement that might have caused low factor loads of S5.2.

However, S3.2 is not included any of the factors identified since options under S3.2 were involved definitions of an obtuse-angled triangle and a scalene triangle instead of an isosceles triangle and consequently having low factor load.

Factor 1: Identification of Triangles

S6.1 (sub-score of intuitive non-example), S6.2 (sub-score of non-intuitive non-example), and S6.3 (sub-score of non-intuitive examples) fell under factor 1 and were named as identification of triangles because it measures to identify triangles among diverse objects using triangle attributes and/or visual figures.

Factor 2: Definitions of Congruent and Similar Triangles

S4.1(minimal definitions for congruent triangles), S5.1 (minimal definitions for similar triangles), S4.2(non-minimal definitions for congruent triangles), and S5.2 (non-minimal definitions for similar triangles) fell under factor 2 and were named as definitions of congruent and similar triangles. S4.1, S4.2, and S5.1 are differentiated because each option falling in these sub-scores constitutes AA (Angle-Angle), SSS (Side-Side-Side), and ASA (Angle-Side-Angle) theorems, along with using the Turkish meaning of congruent and similar (i.e., Mason, 1989).

Factor loadings greater than .40 represent a strong association; however, the factor load for 5.2 is .30. Options in S5.2 are 42, "Two triangles are similar if and only if their corresponding angles are the same size and the lengths of their corresponding sides are in the same proportion." and option 45; "Two triangles are similar if their corresponding angles are congruent, and the lengths of the corresponding sides are different." They are non-minimal definitions for similar triangles. The only difference in these options is that while 42 says the lengths of two triangles corresponding sides have the same proportion, option 45 says they are different. When we look at the theoretical construct for this sub-scored item (S5.2), we can include it as the definition of congruent and similar triangles. When item characteristics were examined, it has found that S5.2 was able to differentiate between high- and low-test takers and had a moderate difficulty. Hence these may not contribute to low factor loading. The complex structure of the Option 42 and 45 involving 2 sentences that are not similarity theorem could lead to low factor loading in S5.2.

Factor 3: Minimal Definitions of Types of Triangles

S3.1 (minimal definitions for an isosceles triangle) and S2.1 (minimal definitions for a right triangle) are called minimal definitions of types of triangles.

Factor loading for S3.2 (sub-score for the non-minimal definitions of an isosceles triangle) is below .40. S3.2 is sub-score of Option 27, "An isosceles triangle is a triangle with two sides of equal length and two angles of equal measure which are less than 90 degrees." and Option 29, "An isosceles triangle is a triangle that has no measures of congruent angles or lengths of sides.". The reason for low factor loading might be that Option 27 is the definition of an obtuse-angled triangle while option 29 is the definition of a scalene triangle.

S1.1, the sub-score of minimal definitions of an equilateral triangle, has a low factor loading (.29). Options in S1.1 involves minimal definitions for an equilateral triangle including sides, except option 6. S1.1 were found to be not able to differentiate between high- and low-ability test takers effectively and has moderately difficult. Besides that, the explanation for the low loading may be linguistic in nature because the name of the equilateral triangle may hint at the participants' replies. To conclude, having low discrimination index and linguistic issue could lead to low factor loading for S1.1.

Factor 4: Non-minimal Definitions of Types of Triangles with Auxiliary Elements

S2.2 is the sub-score of the non-minimal definition of a right triangle, and S 1.2 is the sub-score of the non-minimal definition of an equilateral triangle. Thus, this factor 4 is named for non-minimal definitions of types of triangles with auxiliary elements. The common characteristics of the sub-scored items in DKTT is that they include basic geometry concepts (e.g., point) and auxiliary elements of a triangle (e.g., median, altitude of a triangle). Furthermore, options in S1.2 involves points, equal distance, having three equal interior angle measure of 60 degrees, median, the length of altitude, exterior angles along with angle measure and side lengths of a triangle.

3.3 Reliability

To evaluate the internal consistency for the entire DKTT and each factor independently, we computed Cronbach's alpha and McDonald's omega coefficients (see Table 3.13). In Table 3.13, the McDonald's omega values range from 0.746 to 0.931. This indicates that the DKTT has good reliability, as the omega values are all above 0.7. Cronbach's alpha is another measure of reliability that is commonly used. The Cronbach's alpha values in Table 3.13 range from 0.67 to 0.727. This also indicates good reliability, as alpha values above 0.7 are generally considered acceptable (Bagozzi, & Yi, 2012). The Cronbach's alpha value is acceptable for DKTT because of the nature of the test, that is, the declarative knowledge construct being multifaceted. This might make it more challenging to achieve a high level of internal consistency on the test as the sub-scored items measure slightly different dimensions of the declarative knowledge. Overall, the results in Table 3.13 suggest that the DKTT (Declarative Knowledge Test of Triangles) has good reliability, as indicated by the high values for both McDonald's omega and Cronbach's alpha.

Table 3.13 Reliability of DKTT

Frequentist Individual Item Reliability Statistics

Sub-scored Item	If item dropped	
	McDonald's ω	Cronbach's α
S1.2	0.931	0.727
S2.1	0.919	0.699
S2.2	0.929	0.721
S3.1	0.909	0.69
S4.1	0.896	0.67
S4.2	0.924	0.704
S5.1	0.904	0.677

Table 3.13 Reliability of DKTT (continued)

S5.2	0.928	0.716
S6.1	0.926	0.714
S6.2	0.916	0.695
S6.3	0.927	0.715
DKTT	0.746	0.71

CHAPTER 4

CONCLUSION

The purpose of this study is to develop a selected-response test to measure 11th-grade students' declarative knowledge of definitions of types of triangles, congruent and similar triangles, and the identification of triangles by following the Standards for Educational and Psychological Testing set by the American Educational Research Association (AERA et al., 2014) and then determine the validity evidence for the test (DKTT) based on test content, internal structure, and reliability. In the current study, in addition to the reliability of the DKTT, we further supplied three sources of validity evidence based on test content, internal structure, and reliability in the method and results chapter of this thesis.

4.1 Identification of Triangles

Many studies have examined learners' acquirement of the triangle concept using particular identification tasks that include examples and non-examples (Burger & Shaughnessy,1986; Tsamir et al.,2008). In the DKTT, students were asked to identify triangles among given figures between options 50 and 58 and S6.1, S6.2, and S6.3. From rotated factor matrix of DKTT, it has found that Factor 1 is primarily related to the identification of triangles, as the highest loadings for this factor are found in S6.1, S6.2, and S6.3. S6.1. Sub-scored item S6.1 is the total score for options that fall in intuitive non-examples (non-triangle), sub-scored item S6.2 is the total score for options that fall in non-intuitive non-examples (non-triangle), and sub-scored item S6.3 is the total score for options that fall in non-intuitive examples (triangle). However, it has been found that students can erroneously identify triangular non-examples with convex or concave curved sides as triangles (Clements et al.,1999; Tsamir et al., 2008). They pointed out that intuitive examples of triangles and intuitive examples of non-triangles are figures and that students are more likely to identify them correctly as triangles or non-triangles, respectively. Similarly, unintuitive examples of triangles and unintuitive examples of non-triangles are triangles and non-triangles, respectively, which are commonly misidentified as non-triangles and triangles (Tirosh & Tsamir, 2008; Tsamir et al., 2008). The results of this study revealed that students were able to select the figures that are triangle and non-triangle by considering their critical attributes (e.g., closed, two-dimensional). The DKTT provided a measurement tool to identify students' partial or incorrect knowledge of identifying triangles. As a result, we stressed that intuitive and non-intuitive examples and non-examples should be included in teaching triangles.

4.2 Types of Triangles

Three sides, three vertices, three angles, the sum of every two sides being greater than the third, being closed, and the total of the interior angles being 180° are all crucial attributes of a triangle. Apart from these, we found that types of triangles can be defined by using auxiliary elements of triangles during the test development process. For instance, "an isosceles triangle is a triangle in which the orthocenter is a vertex of the triangle.". Additionally, it was found that students have difficulty acknowledging the equilateral triangle as an isosceles triangle (Altıparmak et al., 2021; Ubuz, 2017). This is in line with the result of this study showed that almost 48% of students selected that an isosceles triangle is a triangle whose at least two sides are equal in length. Furthermore, the present study revealed that students tended to choose non-minimal definitions over minimal ones correctly. This result is consistent with Tsamir's (2015) research, which revealed that only five of the 31 definitions provided by early-year teachers were minimal and correct. In DKTT, two of the four factors were loaded based on the minimality and non-minimality of triangle types as in Minimal Definitions for Types of Triangles (Factor 3)' and Non-Minimal Definitions of Triangle Types Including Auxiliary Elements (Factor 4)'. Similar to what previous research has shown (e.g., Hannibal, 1999; Tsamir et al., 2015), this study also demonstrated that certain inappropriate terminology is employed in the definitions of types of triangles that are written (e.g., shape, joining non-adjacent sides). Option 14 states, for instance, "An isosceles triangle is a geometric object formed by the joining of three non-linear points, any two of them at right angles." Using the terminology shape instead of 3-sided polygon or triangle generates equivocacy in the definition since a shape may be open or closed and two-dimensional or three-dimensional (e.g., (Hannibal, 1999; Tsamir et al., 2015), and how lines join was not clearly explained (e.g., linear, curve).

4.3 Congruent and Similar Triangles

If the options in DKTT meet all the necessary and sufficient characteristics of mathematical definitions, such as minimality/non-minimality, theorems for congruent and similar triangles are utilized as formal definitions for congruent and similar triangles(Van Dormolen & Zaslavsky, 2003; Zaslavsky & Shir, 2005). Furthermore, students' responses to DKTT questions were unaffected by the fact that some of the options were theorems for congruent and similar triangles during the interview. For the definition of congruent triangles, 20% and 69% of students correctly selected minimal definitions and non-minimal definitions, respectively. In addition, for similar triangles, 6% and 25% of them correctly chose minimal and non-minimal definitions, respectively. As also revealed in Haj-Yahya's (2022) article, students tended to select non-minimal definitions for congruent and similar triangles more correctly than minimal ones. The DKTT results revealed that 15% of students were considered to have attained van Hiele's (1958) third level for defining congruent triangles and 2% for similar triangles since they selected both minimal and non-minimal definitions of congruent and similar triangles. The result of EFA depicted that sub-scored items under the congruent and similar triangles were loaded in Factor 2.

To conclude, four factors emerged from using the EFA to analyze test dimensionality. These factors were named 'Identification of Triangles (Factor 1)', 'Definitions of Congruent and Similar Triangles Factor 2', 'Minimal Definitions for a Right and an Isosceles Triangle (Factor 3)', and 'Non-minimal Definitions for Equilateral and Right Triangles (Factor 4)'. During the EFA process, some sub-scored items with poor loadings were not included in the factors.

The DKTT could be administered to 11th-grade students to measure their knowledge of these concepts (types of triangles, congruent and similar triangles, and identifying triangles), with the results being used to identify areas of strength and weakness in students' knowledge of these topics. The test could also be used as part of a more extensive research study on the effectiveness of different teaching methods or educational interventions on students' knowledge of triangles.

In terms of future studies on the DKTT, it may be helpful to conduct additional research further to establish the validity and reliability of the test. This could include collecting data from a larger sample of 11th-grade students and examining the test's performance in different educational settings or with different groups of students. It may also be helpful to compare the DKTT to other declarative knowledge measures or assess the test's validity evidence based on relations to another variable by examining its relationship to other measures of achievement or performance. Future studies could also delve deeper into how students with different linguistic abilities perform on DKTT.

REFERENCES

Almond, R. G. (2014). Using automated essay scores as an anchor when equating constructed response writing tests. *International Journal of Testing*, *14*(1), 73-91.

Altıparmak, K., & Gürcan, G. (2021). Examination of 4th Grade Students' Definitions for Square, Rectangle and Triangle Geometric Shapes. *Education Quarterly Reviews*, *4*(3).

Anderson, J. R. (1976). Language, memory, and thought. Hillsdale, New Jersey: L.

Anderson, J. R. (1993). Problem solving and learning. *American psychologist*, *48*(1), 35.

Anderson, J. R. (2013). The architecture of cognition. Psychology Press.

Aslan, D., & Arnas, Y. A. (2007). Three-to six-year-old children's recognition of geometric shapes. *International Journal of Early Years Education*, *15*(1), 83-104.

Atebe, H. U. (2008). *Student's van Hiele levels of geometric thought and conception in plane geometry: a collective case study of Nigeria and South Africa* (Doctoral dissertation, Rhodes University).

Attali, Y., Lewis, W., & Steier, M. (2013). Scoring with the computer: Alternative procedures for improving the reliability of holistic essay scoring. *Language Testing*, *30*(1), 125-141.

Aydın, U., & Ubuz, B. (2010). Structural model of metacognition and knowledge of geometry. *Learning and Individual Differences*, *20*(5), 436-445.

Bagozzi, R. P., & Yi, Y. (2012). Specification, evaluation, and interpretation of structural equation models. *Journal of the academy of marketing science*, *40*(1), 8-34.

Bernabeu, M., Moreno, M., & Llinares, S. (2021). Primary school students 'understanding of polygons and the relationships between polygons. *Educational Studies in Mathematics*, *106*(2), 251-270.

Biber, A. Ç. (2020). Students' difficulties in similar triangle questions. *Kıbrıslı Eğitim Bilimleri Dergisi, 15*(5), 1146-1159.

Bollen, K. A. (1989). *Structural equations with latent variables*(Vol. 210). John Wiley & Sons.

Borasi, R. (1992). *Learning mathematics through inquiry*. Heinemann.

Bozkurt, A., & Koc, Y. (2012). Investigating First Year Elementary Mathematics Teacher Education Students' Knowledge of Prism. *Educational Sciences: Theory and Practice, 12*(4), 2949-2952.

Brassil, C. E., & Couch, B. A. (2019). Multiple-true-false questions reveal more thoroughly the complexity of student thinking than multiple-choice questions: a Bayesian item response model comparison. *International Journal of STEM Education, 6*(1), 1-17.

Brown, T.A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.

Burger, W. F., & Shaughnessy, J. M. (1986). Characterizing the van Hiele levels of development in geometry. *Journal for research in mathematics education, 17*(1), 31-48.

Byrnes, J. P., & Wasik, B. A. (1991). Role of conceptual knowledge in mathematical procedural learning. *Developmental psychology*, *27*(5), 777.

Casanova, J. R., Cantoria, C. C. C., & Lapinid, M. R. C. (2021). Students' geometric thinking on triangles: much improvement is needed. *Infinity Journal, 10*(2), 217-234.

Chinnappan, M., Ekanayake, M. B., & Brown, C. (2012). Knowledge use in the construction of geometry proof by Sri Lankan students. *International Journal of Science and Mathematics Education*, *10*(4), 865-887.

Clements, D. H., & Battista, M. T. (1992). Geometry and spatial reasoning. *Handbook of research on mathematics teaching and learning*, *420*, 464.

Clements, D. H., Swaminathan, S., & Hannibal, M. A. Z., & Sarama, J.(1999). Young children's concepts of shape. *Journal for Research in Mathematics Education, 30*(2), 192-212.

Cohen, M. P., & Carpenter, J. (1980). The effects of non-examples in geometrical concept acquisition. *International Journal of Mathematical Educational in Science and Technology, 11*(2), 259-263.

Couch, B. A., Wood, W. B., & Knight, J. K. (2015). The molecular biology capstone assessment: a concept assessment for upper-division molecular biology students. CBE—*Life Sciences Education*, *14*(1), ar10.

Couch, B. A., Hubbard, J. K., & Brassil, C. E. (2018). Multiple–true–false questions reveal the limits of the multiple–choice format for detecting students with incomplete understandings. *BioScience*, *68*(6), 455-463.

Currie, P., & Pegg, J. (1998). Investigating students understanding of the relationships among quadrilaterals. In *Teaching Mathematics in New Times: Conference Proceedings. Melbourne: Mathematics Education Research Group of Australasia Incorporated*.

Dacin, P. A., & Mitchell, A. A. (1986). The measurement of declarative knowledge. *ACR North American Advances*.

Dağlı, Ü. Y., & Halat, E. (2016). Young Children's Conceptual Understanding of Triangle. *Eurasia Journal of Mathematics, Science and Technology Education, 12*(2), 189-202.

Dennis, I., & Newstead, S. E. (1994). The strange case of the disappearing sex bias. *Assessment & Evaluation in Higher Education*, *19*(1), 49-56.

De Villiers, M. (1998, July). To teach definitions in geometry or teach to define?. In *PME conference* (Vol. 2, No. 8).

Downing, S. M., & Yudkowsky, R. (2009). Introduction to assessment in the health professions. In *Assessment in health professions education* (pp. 21-40). Routledge.

Dudley, A. (2006). Multiple dichotomous-scored items in second language testing: Investigating the multiple true-false item type under norm-referenced conditions. *Language Testing*, *23*(2), 198-228.

Dündar, S., & Gündüz, N. (2017). Justification for the Subject of Congruence and Similarity in the Context of Daily Life and Conceptual Knowledge. *Journal on Mathematics Education, 8*(1), 35-54.

Ebel, R., & Frisbie, D. A. (1991). *Essentials of educational measurement Englewood Cliffs*, NJ: Prentice-Hall.

Engelbrecht, J., Harding, A., & Potgieter, M. (2005). Undergraduate students' performance and confidence in procedural and conceptual mathematics. *International journal of mathematical education in science and technology*, *36*(7), 701-712.

Ertekin, E., Yazici, E., & Delice, A. (2014). Investigation of primary mathematics student teachers' concept images: cylinder and cone. *International Journal of Mathematical Education in Science and Technology*, *45*(4), 566-588.

Eryılmaz Çevirgen, A. (2012). Casual relations among 12th grade students' geometry knowledge, spatial ability, gender and school type.

Fabrigar, l.r., & Wegener, D.t. (2012*). Factor analysis*. newyork, ny: oxford University press.

Frisbie, D. A., & Sweeney, D. C. (1982). The relative merits of multiple true-false achievement tests. *Journal of Educational Measurement*, 29-35.

Frisbie, D. A., & Becker, D. F. (1991). An analysis of textbook advice about true-false tests. *Applied Measurement in Education*, *4*(1), 67-83.

Fujita, T., & Jones, K. (2007). Learners' understanding of the definitions and hierarchical classification of quadrilaterals: towards a theoretical framing. *Research in Mathematics Education*, *9*(1), 3-20.

Fujita, T. (2012). Learners' level of understanding of the inclusion relations of quadrilaterals and prototype phenomenon. *The Journal of Mathematical Behavior*, *31*(1), 60-72.

Fuys, D., Geddes, D., & Tischler, R. (1988). The van Hiele model of thinking in geometry among adolescents. *Journal for Research in Mathematics Education. Monograph*, *3*, i-196.

Galeshi, R. (2014). Are We There Yet? A Comparative Study of Eighth Grade Mathematics Performance. *Global Education Journal*, 2014(3).

Gerard, L. F., & Linn, M. C. (2016). Using automated scores of student essays to support teacher guidance in classroom inquiry. *Journal of Science Teacher Education*, *27*(1), 111-129.

Gierl, M. J., Latifi, S., Lai, H., Boulais, A. P., & De Champlain, A. (2014). Automated essay scoring and the future of educational assessment in medical education. *Medical education*, *48*(10), 950-962.

González, G., & Herbst, P. G. (2009). Students' conceptions of congruency through the use of dynamic geometry software. *International Journal of Computers for Mathematical Learning*, *14*(2), 153-182.

Gökkurt, B., & Soylu, Y. (2016). Examination of middle school mathematics teachers' mathematical content knowledge: The sample of prism. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, *16*(2), 451-481.

Gutiérrez, A., & Jaime, A. (1998). On the assessment of the Van Hiele levels of reasoning. *Focus on learning problems in mathematics, 20*, 27-46.

Hadas, N., Hershkowitz, R., & Schwarz, B. B. (2000). The role of contradiction and uncertainty in promoting the need to prove in dynamic geometry environments. *Educational Studies in Mathematics*, *44*(1), 127–150. https://doi.org/10.1023/A:1012781005718

Haj-Yahya, A. (2019). Can Classification Criteria Constitute a Correct Mathematical Definition? Preservice and In-Service Teachers' Perspectives. *International Journal of Research in Education and Science*, *5*(1), 88-101.

Haj-Yahya, A. (2022). Students' conceptions of the definitions of congruent and similar triangles. *International Journal of Mathematical Education in Science and Technology, 53*(10), 2703-2727.

Haladyna, T. M., & Downing, S. M. (2004). Construct‐irrelevant variance in high‐stakes testing. *Educational Measurement: Issues and Practice*, *23*(1), 17-27.

Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items*. Routledge.

Hanisch, K. A., Kramer, A. F., & Hulin, C. L. (1991). Cognitive representations, control, and understanding of complex systems: a field study focusing on components of users' mental models and expert/novice differences. *Ergonomics*, *34*(8), 1129-1145.

Hao, T., Li, H., & Wenyin, L. (2007, October). Acquiring Procedural Knowledge Historical Text. In *Third International Conference on Semantics, Knowledge and Grid (SKG 2007)* (pp. 491-494). IEEE.

Harel, G., Selden, A., & Selden, J. (2006). Advanced mathematical thinking: Some PME perspectives. In *Handbook of research on the psychology of mathematics education* (pp. 147-172). Brill.

Hershkowitz, R., & Vinner, S. (1980). Concept images and common cognitive paths in the development of some simple geometrical concepts. In *Proceedings of the 4th PME conference* (pp. 177-184). The Weizmann Institute of Science.

Hiebert, J., & Wearne, D. (1996). Instruction, understanding, and skill in multidigit addition and subtraction. *Cognition and instruction, 14*(3), 251-283.

Hogan, T. (2013). Constructed-response approaches for classroom assessment. *Sage handbook of research on classroom assessment*, 275-293.

Horzum, T., & Ertekin, E. (2018). Prospective mathematics teachers' understanding of the base concept. *International Journal of Mathematical Education in Science and Technology*, *49*(2), 176-199.

Hoyles, C. (1998). A culture of proving in school mathematics?. In *Information and communications technologies in school mathematics* (pp. 169-182). Springer, Boston, MA.

Javid, L. (2014). The comparison between multiple-choice (MC) and multiple true-false (MTF) test formats in Iranian intermediate EFL learners' vocabulary learning. *Procedia-Social and Behavioral Sciences*, *98*, 784-788.

Jin, H., & Wong, K. Y. (2021). Complementary measures of conceptual understanding: a case about triangle concepts. *Mathematics Education Research Journal*, 1-22.

Jones, K., & Fujita, T. (2013). Characterising triangle congruency in lower secondary school: the case of Japan.

Jones, K., Fujita, T., & Miyazaki, M. (2013). Learning congruency-based proofs in geometry via a web-based learning system. *Proceedings of the British Society for Research into Learning Mathematics*, *33*(1), 31-36.

Jüttner, M., Boone, W., Park, S., & Neuhaus, B. J. (2013). Development and use of a test instrument to measure biology teachers' content knowledge (CK) and pedagogical content knowledge (PCK). *Educational assessment, evaluation and accountability*, *25*(1), 45-67.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.

Kocak, M., Özdemir, B. G., & Soylu, Y. (2017). An Investigation the Pedagogical Content Knowledge of Primary Mathematics Prospective Teachers about the Concept of Cylinder. *Cukurova University Faculty of Education Journal*, *46*(2), 711-765.

Kreiter, C. D., & Frisbie, D. A. (1989). Effectiveness of multiple true-false items. *Applied Measurement in Education*, *2*(3), 207-216.

Kuhn, D. (2000). Metacognitive development. *Current directions in psychological science*, 9(5), 178-181.

Kuhn, C., Zlatkin-Troitschanskaia, O., Pant, H. A., & Hannover, B. (2016). Valide Erfassung der Kompetenzen von Studierenden in der Hochschulbildung. *Zeitschrift für Erziehungswissenschaft*, *19*(2), 275-298.

Kuo, B. C., Chen, C. H., Yang, C. W., & Mok, M. M. C. (2016). Cognitive diagnostic models for tests with multiple-choice and constructed-response items. *Educational Psychology*, *36*(6), 1115-1133.

Lawson, A. E. (1991). Constructivism and domains of scientific knowledge: A reply to Lythcott and Duschl. *Science Education*, 75(4), 481-488.

Lawson, A. E., McElrath, C. B., Burton, M. S., James, B. D., Doyle, R. P., Woodward, S. L., ... & Snyder, J. D. (1991). Hypothetico-deductive reasoning skill and concept acquisition: Testing a constructivist hypothesis. *Journal of Research in Science teaching*, *28*(10), 953-970.

Lawson, A. E., Alkhoury, S., Benford, R., Clark, B. R., & Falconer, K. A. (2000). What kinds of scientific concepts exist? Concept construction and intellectual development in college biology. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *37*(9), 996-1018.

Lee, T. N. (2022). Justifying triangle shapes through their properties in argumentation. *European Journal of Psychology of Education*, 1-17.

Leikin, R., & Winicki-Landman, G. (2001). Defining as a vehicle for professional development of secondary school mathematics teachers. *Mathematics Teacher Education and Development*, *3*(2001), 62-73.

Leung, K. C., Ding, L., Leung, A. Y. L., & Wong, N. Y. (2014). Prospective teachers' competency in teaching how to compare geometric figures: the concept of congruent triangles as an example. *Research in Mathematical Education*, *18*(3), 171-185.

Li, J., Ang, J. S. K., Tong, X., & Tueni, M. (1994). AMS: A declarative formalism for hierarchical representation of procedural knowledge. *IEEE transactions on knowledge and data engineering*, 6(4), 639-643.

Llinares, S., & Clemente, F. (2019). Characteristics of the shifts from configural reasoning to deductive reasoning in geometry. *Mathematics Education Research Journal, 31*(3), 259-277.

López, J. L., & Guzmán, J. (2011, October). Proof and argumentation in the solution of triangle congruence problems: A study with high school. *In Psychology of Mathematics Education* (p. 1895).

Lutfi, M. K., & Jupri, A. (2020, April). Analysis of junior high school students' spatial ability based on Van Hiele's level of geometrical thinking for the topic of triangle similarity. *In Journal of Physics: Conference Series* (Vol. 1521, No. 3, p. 032026). IOP Publishing.

Mack, N. K. (1990). Learning fractions with understanding: Building on informal knowledge. *Journal for research in mathematics education, 21*(1), 16-32.

Mason, M. M. (1989). Geometric Understanding and Misconceptions among Gifted Fourth-Eighth Graders.

Messick, S. (1989). Validity. em r. linn (org.), educational measurement.(13-103). *New York, NY: American Council on Education and Macmillan Publishing Company*.

Miller, S. P., & Hudson, P. J. (2007). Using evidence-based practices to build mathematics competence related to conceptual, procedural, and declarative knowledge. *Learning Disabilities Research & Practice, 22*(1), 47-57.

Miller, S. M. (2018). An analysis of the form and content of quadrilateral definitions composed by novice pre-service teachers. *The Journal of Mathematical Behavior*, *50*, 142-154.

Miyakawa, T. (2017). Comparative analysis on the nature of proof to be taught in geometry: The cases of French and Japanese lower secondary schools. *Educational Studies in Mathematics*, *94*(1), 37-54.

Mobalegh, A., & Barati, H. (2012). Multiple true-false (MTF) and multiple-choice (MC) test formats: A comparison between two versions of the same test paper of Iranian NUEE. *Journal of Language Teaching and Research*, *3*(5), 1027.

Monaghan, J. J. (2000). SPH without a tensile instability. *Journal of computational physics*, *159*(2), 290-311.

Moss, J., & Case, R. (1999). Developing children's understanding of the rational numbers: A new model and an experimental curriculum. *Journal for research in mathematics education*, *30*(2), 122-147.

Mullis, I. V., Martin, M. O., Gonzalez, E. J., Gregory, K. D., Garden, R. A., O'Connor, K. M., ... & Smith, T. A. (2000). TIMSS 1999 international mathematics report.

Ngirishi, H., & Bansilal, S. (2019). An Exploration of High School Learners' Understanding of Geometric Concepts. *Problems of Education in the 21st Century*, *77*(1), 82-96.

Okazaki, M., & Fujita, T. (2007, July). Prototype phenomena and common cognitive paths in the understanding of the inclusion relations between quadrilaterals in Japan and Scotland. In *Proceedings of the 31st Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 41-48).

Otten, S., Gilbertson, N. J., Males, L. M., & Clark, D. L. (2014). The mathematical nature of reasoning-and-proving opportunities in geometry textbooks. *Mathematical Thinking and Learning*, *16*(1), 51-79.

Parastuti, R. H., Usodo, B., & Subanti, S. (2018). Student's error in writing mathematical problem solving associated with corresponding angles of the similar triangles. *Pancaran Pendidikan*, *7*(1).

Paris, S. G., Lipson, M. Y., & Wixson, K. K. (1983). Becoming a strategic reader. *Contemporary educational psychology*, *8*(3), 293-316.

Parker, J. M., Anderson, C. W., Heidemann, M., Merrill, J., Merritt, B., Richmond, G., & Urban-Lurain, M. (2012). Exploring undergraduates' understanding of photosynthesis using diagnostic question clusters. CBE—*Life Sciences Education*, *11*(1), 47-57.

Patkin, D., & Plaksin, O. (2011). Congruent triangles sufficient and insufficient conditions suggested milestones for inquiry and discussion. *Research in Mathematical Education*, *15*(4), 327-340.

Petty, O. S., & Jansson, L. C. (1987). Sequencing examples and nonexamples to facilitate concept attainment. *Journal for Research in Mathematics Education*, *18*(2), 112-125.

Pickreign, J. (2007). Rectangles and Rhombi: How Well Do Preservice Teachers Know Them?. *Issues in the undergraduate mathematics preparation of school teachers, 1*.

Pituch, K. A., & Stevens, J. P. (2015). *Applied multivariate statistics for the social sciences: Analyses with SAS and IBM's SPSS*. Routledge.

Polit, D. F. (1996). *Data analysis & statistics for nursing research*. Appleton & Lange.

Poon, K. K., & Wong, K. L. (2017). Pre-constructed dynamic geometry materials in the classroom–how do they facilitate the learning of 'Similar Triangles'?. *International Journal of Mathematical Education in Science and Technology*, *48*(5), 735-755.

Pratt, D., & Davison, I. (2003). Interactive Whiteboards and the Construction of Definitions for the Kite. *International Group for the Psychology of Mathematics Education*, *4*, 31-38.

Razel, M., & Eylon, B. S. (1991, July). Developing mathematics readiness in young children with the Agam Program. In *fifteenth conference of the International Group for the Psychology of Mathematics Education, Genova, Italy*.

Richardson, R. (1992). The multiple choice true/false question: what does it measure and what could it measure?. *Medical teacher, 14*(2-3), 201-204.

Rios, J., & Wells, C. (2014). Validity evidence based on internal structure. *Psicothema, 26*(1), 108-116.

Roldán-Zafra, J., Perea, C., Polo-Blanco, I., & Campillo, P. (2022). Design of an Interactive Module Based on the van Hiele Model: Case Study of the Pythagorean Theorem. *International Electronic Journal of Mathematics Education, 17*(1), em0672.

Sahdra, B., & Thagard, P. (2003). Procedural knowledge in molecular biology. *Philosophical Psychology, 16*(4), 477-498.

Sands, D., Parker, M., Hedgeland, H., Jordan, S., & Galloway, R. (2018). Using concept inventories to measure understanding. *Higher Education Pedagogies, 3*(1), 173-182.

Schwarz, B. B., & Hershkowitz, R. (1999). Prototypes: Brakes or levers in learning the function concept? The role of computer tools. *Journal for research in mathematics education*, 362-389.

Sears, R., & Chávez, Ó. (2015, February). Students of two-curriculum types Performance on a proof for congruent triangles. In *CERME 9-Ninth Congress of the European Society for Research in Mathematics Education* (pp. 192-197).

Senk, S. L. (1989). Van Hiele levels and achievement in writing geometry proofs. *Journal for research in mathematics education, 20*(3), 309-321.

Serow, P. (2006, July). Triangle-property relationships: Making the connections. In *Proceedings of the 30th conference of the International Group for the Psychology of Mathematics Education* (Vol. 5, pp. 89-969).

Shermis, M. D., & Burstein, J. (Eds.). (2013). *Handbook of automated essay evaluation: Current applications and new directions*. Routledge.

Shir, K., & Zaslavsky, O. (2001). What constitutes a (good) definition? The case of a square. In *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education, v. 4* (pp. 161-168).

Shir, K., & Zaslavsky, O. (2002). Students' conceptions of an acceptable geometric definition. In *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education, v4* (pp. 201-208).

Siddiqui, N. I., Bhavsar, V. H., Bhavsar, A. V., & Bose, S. (2016). Contemplation on marking scheme for Type X multiple choice questions, and an illustration of a practically applicable scheme. *Indian journal of pharmacology, 48*(2), 114.

Streiner, D. L. (1994). Figuring out factors: the use and misuse of factor analysis. *The Canadian Journal of Psychiatry, 39*(3), 135-140.

Ragan, T. J., & Smith, P. L. (1999). Instructional design. *New York: Macmillan Publishing Company*.

Tabachnick, B.G. and Fidell, L.S. (2001) *Using Multivariate Statistics*. 4th Edition, Allyn and Bacon, Boston.

Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational studies in mathematics, 12*(2), 151-169.

Ten Berge, T., & Van Hezewijk, R. (1999). Procedural and declarative knowledge: An evolutionary perspective. *Theory & Psychology*, 9(5), 605-624.

Thagard, P. (2006). How to collaborate: Procedural knowledge in the cooperative development of science. *The Southern journal of philosophy, 44*(S1), 177-196.

Thompson, B. (2004). Exploratory and confirmatory factor analysis: Understanding concepts and applications. *Washington, DC*, *10694*(000).

Tirosh, D., & Tsamir, P. (2008). Triangles and non-triangles in preschool. *Mispar Hazak*, 15, 47–53.

Tsamir, P., Tirosh, D., & Levenson, E. (2008). Intuitive nonexamples: The case of triangles. *Educational Studies in Mathematics, 69*(2), 81-95.

Tsamir, P., Tirosh, D., Levenson, E., Barkai, R., & Tabach, M. (2015). Early-years teachers' concept images and concept definitions: triangles, circles, and cylinders. *ZDM, 47*(3), 497-509.

Türnüklü, E., & Ergin, A. S. (2016). 8. sınıf öğrencilerinin cisimleri görsel tanıma ve tanımlamaları: cisim imgeleri. *İlköğretim Online*, *15*(1).

Ubuz, B., & Gökbulut, Y. (2015). Primary prospective teachers' knowledge on pyramid: generating definitions and examples. *Kırşehir Eğitim Fakültesi Dergisi*, *16*(2), 335-351.

Ubuz, B. (2017). Dörtgenler arasındaki ilişkiler: 7. sınıf öğrencilerinin kavram imajları [Relations among quadrilaterals: 7th grade stu- dents 'concept images]. *Yaşadıkça Eğitim Dergisi*, *31*(1), 55–68.

Ubuz, B., & Aydın, U. (2018). Geometry knowledge test about triangles: Evidence on validity and reliability. *ZDM*, *50*(4), 659-673.

Ulusoy, F., & Çakıroğlu, E. (2017). Middle school students" types of identification for parallelogram: Underspecification and overgeneralization. *Abant İzzet Baysal University Journal of Faculty of Education*, *17*(1), 457-475.

Ulusoy, F. (2021). Prospective early childhood and elementary school mathematics teachers ' concept images and concept definitions of triangles. *International Journal of Science and Mathematics Education*, *19*(5), 1057-1078.

Ulusoy, F. (2022). Middle school students ’reasoning with regards to parallelism and perpendicularity of line segments. *International Journal of Mathematical Education in Science and Technology*, 1-20.

Unlu, M., & Horzum, T. (2018). Mathematics Teacher Candidates' Definitions of Prism and Pyramid. *International Journal of Research in Education and Science*, *4*(2), 670-685.

Usiskin, Z. (1982). Van Hiele levels and achievement in secondary school geometry.

Van Dormolen, J., & Zaslavsky, O. (2003). The many facets of a definition: The case of periodicity. *The Journal of Mathematical Behavior, 22*(1), 91-106.

Van Hiele, P. M. (1958). A method of initiation into geometry at secondary schools. *Report on methods of initiation into geometry*.

Van Hiele, P. M. (1986). *Structure and insight: A theory of mathematics education*. Academic press.

Vinner, S., & Tall, D. (1991). Advanced mathematical thinking. *Dordrecht Kluwer*.

Wang, Z., Wang, Z., & An, S. (2018). Error analysis of 8th graders’ reasoning and proof of congruent triangles in China. *Journal of Mathematics Education*, *11*(2), 85-120.

Weigand, H. G., Filler, A., Hölzl, R., Kuntze, S., Ludwig, M., Roth, J., ... & Wittmann, G. (2018). *Didaktik der Geometrie für die Sekundarstufe I*. Springer Spektrum.

Willingham, D. B., Nissen, M. J., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of experimental psychology: learning, memory, and cognition, 15*(6), 1047.

Wilson, P. S. (1986). Feature frequency and the use of negative instances in a geometric task. *Journal for Research in Mathematics Education*, *17*(2), 130-139.

Wilson, E. O. (1990). *Success and dominance in ecosystems: the case of the social insects* (Vol. 2, pp. I-XXI). Oldendorf/Luhe: Ecology Institute.

Winicki-Landman, G., & Leikin, R. (2000). On equivalent and non-equivalent definitions: Part 1. *For the learning of Mathematics, 20*(1), 17-21.

Wijaya, T. T., Mutmainah, I. I., Suryani, N., Azizah, D., Fitri, A., Hermita, N., & Tohir, M. (2021, October). Nineth grade students mistakes when solving congruence and similarity problem. *In Journal of Physics: Conference Series* (Vol. 2049, No. 1, p. 012066). IOP Publishing.

Willingham, D. B., Nissen, M. J., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of experimental psychology: learning, memory, and cognition*, *15*(6), 1047.

Yilmaz, I., & Yalcin, N. (2012). The relationship of procedural and declarative knowledge of science teacher candidates in Newton's laws of motion to understanding. *American International Journal of Contemporary Research*, *2*(3), 50-56.

Zandieh, M., & Rasmussen, C. (2010). Defining as a mathematical activity: A framework for characterizing progress from informal to more formal ways of reasoning. The Journal of Mathematical Behavior, 29(2), 57-75.

Zazkis, R., & Leikin, R. (2008). Exemplifying definitions: a case of a square. *Educational Studies in Mathematics*, *69*(2), 131-148.

Zaslavsky, O., & Shir, K. (2005). Students' conceptions of a mathematical definition. *Journal for Research in Mathematics Education*, *36*(4), 317-346.