Research Article doi: 10.12973/ijem.5.4.547

International Journal of Educational Methodology
Volume 5, Issue 4, 547 - 565.
ISSN: 2469-9632
http://www.ijem.com/

Adaptation of the Test Developed to Measure Mathematical Knowledge of Teaching Geometry in Turkey*

Ozan Esendemir
Gaziantep University, TURKEY

Recep Bindak**
Gaziantep University, TURKEY

Received: June 9, 2019 ▪ Revised: September 7, 2019 ▪ September: April 20, 2019

Abstract: "Mathematical knowledge for teaching" is a concept indicating the requirement for a specific kind of knowledge required to teach mathematics. Mathematical knowledge for teaching necessitates a more complex structure than what is required to carry out mathematical tasks and the knowledge to do that. The purpose of this study is to realize the adaptation of "Mathematical Knowledge for Teaching -Geometry (MKT-G)" Test that was initially conceived in English to Turkish (or to Turkish culture). During the adaptation process; after the translations of the items, focus group interviews were held with a group consisting of mathematics teacher educators and experienced mathematics teachers, and then the data from 243 elementary mathematics teachers was analyzed via Item Response Theory (IRT). As a result of the analysis of the test items, psychometric values of the test items indicated that the items in the test performed well in Turkey. Besides, validity and reliability arguments were also tested. As a result, the Turkish version of the MKT-G test is highly reliable and valid to measure the teachers' knowledge of teaching geometry.

Keywords: *Teacher's knowledge in Geometry, Mathematical knowledge for teaching, pedagogical content knowledge.*

To cite this article: Esendemir, O., & Bindak, R. (2019). Adaptation of the test developed to measure Mathematical knowledge of teaching Geometry in Turkey. *International Journal of Educational Methodology*, *5*(4), 547-565. https://doi.org/10.12973/ijem.5.4.547

Introduction

The Trends in International Mathematics and Science Study (TIMSS) and Program for International Student Assessment (PISA) study conducted at the international level in the last 20 years are considered to be significant reference points in identifying the weak and strong aspects of students around the world. The results obtained from these studies impelled some researchers to examine the mathematical education policies of some countries successful in the subject (Stigler & Hiebert, 1999). The quality of the knowledge possessed by mathematics teachers is considered to be an important factor in the respective studies conducted (Ma, 2010). Although the TIMSS and PISA studies provided for culture and language adaptations, it has been considered that a similar process has not been followed in the interviews conducted with the teachers (Delaney, Ball, Hill, Schilling, & Zopf, 2008).

Teaching geometry is considered an important part of mathematics education, both on the national and international levels. In the new mathematics curriculum planned to be introduced in the 2018-2019 education year by the Ministry of National Education [MEB] (2018) at primary schools (K-8), geometry topics are planned to be presented at each grade level. Teaching geometry is an important part of mathematics teaching programs, not only on the national level but on the international level as well. In the Standards and Principles for Teaching Mathematics Book published in 2000 by the National Council of Teachers of Mathematics [NCTM], teaching of geometry is considered an important tool of explaining the axiomatic structure of mathematics and helping students gain reasoning skills. The studies conducted by NCTM are closely followed across the world, regardless of the fact that they focus on mathematics education in the U.S. A study conducted by Atweh, Clarkson and Nebres (2003) demonstrates that the reforms realized in a country start to show effects in other countries in several years. Kupari (2008) suggests that the studies conducted in Finland on teaching mathematics closely follow the developments in teaching mathematics in the U.S., and that the reforms in mathematics education made in this country closely follow the ones in the U.S. The essential importance of geometry in teaching mathematics in the Netherlands, where reforms are being realized on the basis of a "Realistic Mathematics Education [RME]" approach since 1980, is evident (van den Heuvel-Panhuizen, 2000). On the other hand, in Japan, where more traditional methods of teaching are adopted instead of an approach based on technology, geometry is

* This study is a part of a doctoral dissertation preparing by Ozan Esendemir under the advisory of Dr. Recep Bindak.

** **Corresponding author:**
Recep Bindak, Gaziantep University, Technical Sciences School. Gaziantep, Turkey. ✉ bindak@gantep.edu.tr

reported to be an important part of mathematics education starting from the first grade (Judson, 1999). Moreover, the idea of teaching geometry at all grade levels is supported by researchers and theoreticians as well (see NCTM, 2000). Within this context, geometry is considered to be an integral part of mathematics education. Highly qualified teachers are considered a must in providing education in geometry as part of the mathematics curriculum for all grade levels (Kupari, 2008).

The content knowledge of a mathematics teacher is expected to be strong and flexible. Teaching mathematics involves understanding the importance and place of reasoning in mathematics as well as the mathematical ideas and operations and the connections between these ideas and operations (Hill & Ball, 2004). However, in spite of the general acceptance of deep and strong knowledge of mathematics as an important part of a teacher's general knowledge, Hill, Blunk, Charalambous, Lewis, Phelps, Sleep and Ball (2008) point to the scarcity of studies on the measurement of teachers' knowledge and the effects of this knowledge on education and consider the scarcity of the studies in this field a deficiency. At this point, the need to consider the importance of studies on measuring the mathematical knowledge put into use in the teaching process and their effect on the teaching process arise.

"What should mathematics teachers know to teach geometry effectively?" has been seen a crucial and important question, but studies investigating this concept have been quite limited (Hill et al., 2008). Researchers as Guler (2014), Rosenshine and Furst (1986) and Turner and Meyer (2000) pointed out that the studies which focus on examining the teachers' knowledge were predominantly focusing on teachers' general pedagogical knowledge (e.g. teachers' general approaches to teaching mathematics, classroom management strategies, cognitive level of the language used, student responses, teacher-student communication) rather than examining teachers' mathematical subject matter knowledge and pedagogical content knowledge. Furthermore, there is an abundance of the researches that study teachers mathematical subject matter knowledge (Learning Mathematics for Teaching Project [LMTP], 2011). Since, mathematical knowledge for teaching is seen as a complex structure that neither only mathematical subject matter knowledge nor general pedagogical knowledge can separately explain (Hill et al., 2008). Hill and her colleagues defined a new knowledge type that both subject matter knowledge and pedagogical knowledge play significant roles, which mathematics teachers should have in order to teach mathematics effectively and call it "mathematical knowledge for teaching". Besides, they also developed a series of tests, including geometry, to measure the extent to which teachers had this type of knowledge. In this study, the main aim is the adaptation of a test developed in the U.S. to measure mathematics teachers' Mathematical Knowledge for Teaching in Geometry to Turkey.

Theoretical Background

"What kind of knowledge does a teacher need to be able to teach mathematics?" This important question has been the subject of many studies on mathematics throughout the years. Nevertheless, an exact answer has not yet been obtained. However, it is a commonly shared opinion that teaching mathematics is linked to knowledge in mathematics. According to a research conducted on preservice teachers assessed perspectives of teacher knowledge over one academic year can be improved by to create reflective exercises that encourage the development of mathematics content knowledge for teaching (Peace, Quebec Fuentes & Bloom, 2018). It is supported by many studies that the knowledge required to perform mathematical tasks is necessary but not sufficient in teaching the respective subject to the students (Ball, 1990; Ball, Hill & Bass, 2005; Ball, Lubienski & Mewborn, 2001). Moreover, teaching mathematics necessitates a more complex structure than what is required to carry out mathematical tasks. Based on studies conducted by Shulman and other researchers, Ball et al. (2005) attempted to demonstrate the types of mathematical knowledge required to teach mathematics. To serve this purpose, they subjected the teaching process to an analysis. Ball and Bass (2003) based their studies on previous research studies on teaching as well as the literature on mathematical knowledge and investigated the in-class teaching dimension of mathematics based on the observation of videotaped education processes. They constructed a new conceptual structure necessary for in-class teaching of mathematics based on their own findings and expressed this structure as "Mathematical Knowledge for Teaching" (MKT).

Ball and her colleagues investigated the knowledge which should be possessed by mathematics teachers under "The Learning Mathematics for Teaching Project" implemented in 2000. Within the scope of this project, they defined four different types of teachers' knowledge that they built on two different categories of teachers' knowledge given by Shulman (1986) in his study, which are Subject Matter Knowledge and Pedagogical Content Knowledge (Delaney, 2008). While Common Content Knowledge (CCK) and Specialized Content Knowledge (SCK) are addressed within the scope of Subject Matter Knowledge, Knowledge of Content and Student (KCS) and Knowledge of Content and Teaching (KCT) are addressed within the scope of Pedagogical Content Knowledge.

Common Content Knowledge (CCK), a type of knowledge addressed under Subject Matter Knowledge, is defined as "the knowledge containing mathematical knowledge and skills used in cases other than the purpose of teaching act" (Ball, Thames & Hill, 2008). This type of knowledge lets individuals perform a mathematical operation correctly or solve a mathematical problem. Teachers should be able to provide the right mathematical knowledge and definitions for their students, select the right operations and perform such operations correctly. For example, common content knowledge is needed so that area of a triangle can be found correctly. An example to a right triangle is shown in Figure 1:

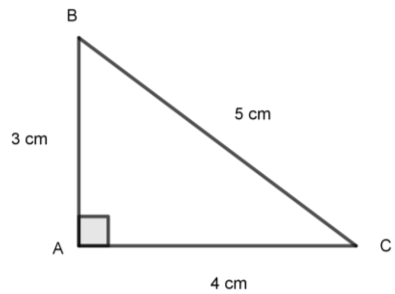


Figure 1. A triangle with a right angle

Many adults can recall the way of finding the area given in Figure 1. From here, eventually, the following common solution is found:

$$\frac{4cm \times 3cm}{2} = 6cm^2$$

Figure 2. The common way of finding the area of a right triangle

It is evident that the ability to correctly find the area of the triangle like the one in Figure 1 is a critical and necessary type of knowledge for the mathematics teaching process. However, this is not enough for teaching mathematics. Teachers should be able to easily recognize typical student mistakes and know the cause of such mistakes:

$$3cm + 4cm + 5cm = 12cm$$

Figure 3. A typical student mistake in finding the area of a triangle

First of all, it is important to notice the incorrectness of this solution. However, it is also meaningful to be familiar with the source of the mistake so that one can provide an efficient teaching process. The incorrectness of the solution in Figure 3 results from the student failing to "calculate the area of the triangle." In addition to analyzing student mistakes, teachers should be able to answer "why" questions from students, use proper models to allow them notice mistakes in their solutions and analyze the correctness or incorrectness of non-standardized solutions proposed by students. In this context, teaching is not only about telling that "to find the area two perpendicular sides of a right triangle should be multiplied and then divided by 2." Teaching entails explaining that the primary reason for finding the perimeter instead of area is that the student confused the area with perimeter.

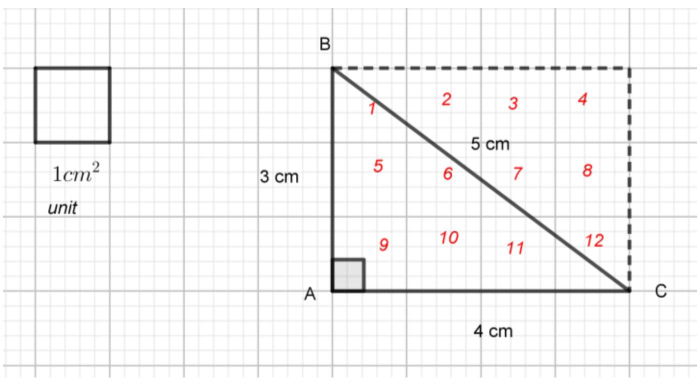


Figure 4. Finding the area of a right triangle with using unit of area

In this sense, the knowledge of how to eliminate student mistakes also has an important place in teaching process. For this purpose, explaining the area of the triangle on a rectangle having a side length of 3 and 4 centimeters is a potential proper method for clarifying area concept.

Teachers' encouragement of their students to develop and use various methods of solutions is considered a critical component of mathematics teaching. In this context, a teacher should be capable of deciding whether non-standardized methods of solutions, which are likely to be proposed by students, are acceptable and have an understanding of mathematical idea underlying this method of solution (Ball, Hill & Bass, 2005). Various methods of solutions for area of the triangle in Figure 1can be seen in Figure 5.

Student A	Student B	Student C
$3cm + 4cm + 5cm = 12cm$	$\frac{4cm \times 3cm}{2} = 6cm^2$	$\frac{12 \times 1cm^2}{2} = 6cm^2$

Figure 5. Various methods of solutions proposed by students

The opinion is that teachers should have a different knowledge from that needed for calculating the area of a triangle so that they can judge whether student solutions presented in Figure 5 are mathematically correct. Moreover, the judgment concerning whether various ways of solutions are mathematically significant entails a substantial amount of teachers' knowledge for teaching mathematics. This is referred to as Specialized Content Knowledge (SCK).

SCK is a type of teachers' knowledge solely needed during teaching process (Ball, Thames & Hill, 2008). SCK is a critical type of teachers' knowledge having influence on a teacher's presenting mathematical ideas, answering "why" questions from students, selecting a notation for the representation of a mathematical idea, establishing relationships between notations, associating subjects with those which are already taught or to be taught subsequently, adapting the knowledge contained in course books to the level of the class, providing descriptions of mathematical operations and rules or assessing students' descriptions and judging whether students' descriptions are rational (Ball, Thames & Hill, 2008).

Knowledge of Content and Student

The third component of the framework of mathematical knowledge for teaching is Knowledge of Content and Student (KCS) consisting of the combination of mathematical knowledge and knowledge of student (Ball, Thames & Bass, 2008). It is considered important that teachers have the foresight concerning what kind of ideas their students may have and what they may find difficult. The method or example of notation selected within in-class teaching process is expected to attract the attention of students. In this context, teachers should decide what is interesting and motivating for their students. When, in other words, a teacher assigns a task to their students, they are expected to have an opinion about what their students can do to complete that task, and predict whether the students will find the said task difficult or easy (Ball, Thames & Bass, 2008). Sometimes, students are likely to use incomplete or unclear descriptions when building their ideas. In this context, teachers should understand their students' ideas and have the ability to interpret mathematical knowledge. At this point, teachers are expected to have an understanding of students and their mathematical ideas. A teacher who knows what kind of mistakes students are likely to make is considered to be able to prepare the course content to minimize or eliminate such mistakes.

Knowledge of Content and Teaching

The last component of the said framework is "Knowledge of Content and Teaching (KCT)." This type of knowledge contains a domain with which mathematical knowledge is intertwined. Many tasks used for mathematics teaching result from the planning of teaching to enable realization of the learning process. In this sense, a teacher should, when planning the subject of a course, consider which notations or examples should be used to start a course and which notations or examples can be used to achieve a deeper understanding. At this very point, it is considered important for a teacher to be familiar with advantages and disadvantages of any notation or example selected for teaching a concept in the context of mathematical teaching process and also be aware that different methods will promote teaching. The entire process entails the awareness of pedagogical elements having an effect on mathematical knowledge and students' learning. For example, knowledge of content and teaching comes into play for a teacher, who has the intention of preparing visual material about geometric objects to improve students' ideas of such objects, to also include notations of geometric objects other than those which are known and used the most and do not fall under the category of associated geometric objects (Clements & Sarama, 2000).

The framework of mathematical knowledge for teaching may shortly be described as follows: the recognition of a wrong answer is addressed under common content knowledge, the understanding of the cause of a mistake under specialized content knowledge, the knowledge of common mistakes and students' knowledge of which mistakes they are likely to make under knowledge of content and student, and the knowledge of which teaching material (methods and examples of notation, etc.) will help students avoid or correct their mistakes under knowledge of content and teaching (Ball, Thames & Phelps, 2008). Ball, Thames and Phelps (2008) illustrated the domains of MKT as in Figure 6.

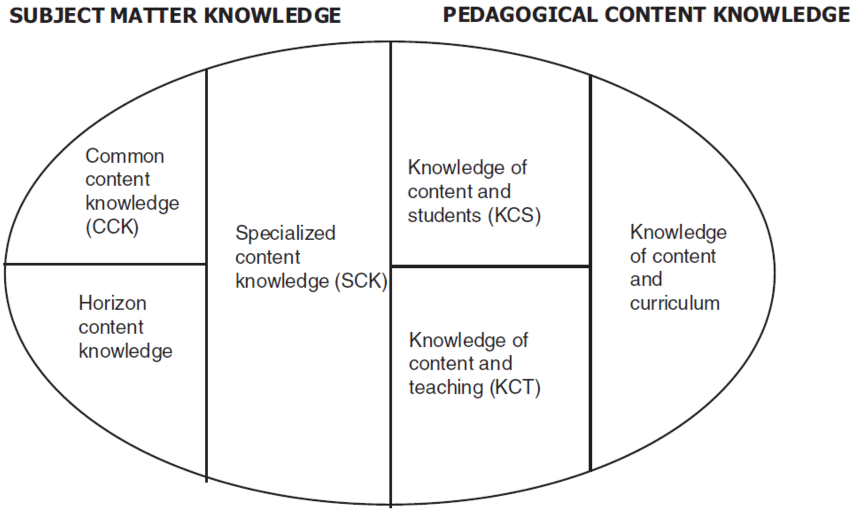


Figure 6. Domains of Mathematical Knowledge for Teaching

As seen in Figure 6, Mathematical Knowledge for Teaching is built upon Subject Matter and Pedagogical Content Knowledge. Subject Matter is composed of common content knowledge, specialized content knowledge and horizon content knowledge. This third type of knowledge (horizon content knowledge) is defined by Ball (1993) as a type of knowledge containing information about the relationship between mathematical subjects across curricula.

Methodology

During the research process, the adaptation of test items was analyzed by both quantitative and qualitative methods. In qualitative aspects, an analysis was carried out on focus group interviews for the adaptation of test items to Turkish, and four individual mathematics teachers were interviewed after they completed the test to obtain their responses to those items included in the test. In quantitative aspects, the results obtained following the implementation of the test were analyzed in psychometric aspects via Item Response Theory.

Mathematical Knowledge for Teaching Geometry Measures: The Instrument

Mathematical Knowledge for Teaching Geometry test (MKT-G) was developed by the researchers of the Learning Mathematics for Teaching (LMT) project conducted by the University of Michigan (see http://www.umich.edu/~lmtweb/history.html for more details). During the development of such measure researchers of the project gathered the information about; mathematics learning-teaching theories, studies in this field, as well as research on teaching materials and student work, and in the context of their own experience, the mathematical knowledge that teachers use in teaching mathematics primary education and in the interpretation of student responses or procedures and elimination of student misconceptions with materials.

Table 1. Distribution of MKT-G test items according to MKT domains

Domains of MKT	Number of the Items	Percentage
Common Content Knowledge	23	76%
Specialized Content Knowledge	4	13%
Knowledge of content and teaching	2	7%
Knowledge of content and students	1	3%

The test consists of 30 items. Table 1 shows how items included within the framework of Mathematics Knowledge for Teaching.

Adaptation process of the MKT-G Test

The items of MKT-G test were originally developed at the University of Michigan in such conditions that the items should have captured the teachers' mathematical knowledge for teaching geometry in the United States. In this study, the very first phase of adaptation of the test items includes translation of the items from English to Turkish. Translation of test items from English to Turkish has been made by making one-to-one translation without changing the mathematical content of the test items.

After the translation, in order to implement the test in Turkey, a meticulous adaptation process for applicability in Turkey were undergoing. Before the implementation of the test to teachers, focus group interviews were carried out with a group that consists of two mathematics teacher educators whom were lecturing in the area of geometry teaching, four PhD candidate studying in mathematics education and one 16-year experienced elementary mathematics teacher. During the interview, every item was discussed according to the five important topics: (i) Translation of the test items into Turkish, (ii) Cultural adaptation of the items, (iii) Implementation of the test to middle school mathematics teachers and interviews after implementation of the test, (iv) validity and reliability studies, and (v) psychometric analysis of the test items. The amendments made to ensure that the items were compatible with the Turkish culture were discussed by the group and a final decision was reached for each change.

(*i*) *Translation of the test items into Turkish*: The test was translated into Turkish by the researcher. The opinion of the instructors at the Department of English Language Teaching was sought during the translation process. The cultural suitability of the concepts was not taken into account when translating the texts. Therefore, the test items are assumed not to have undergone any change during this process.

(ii) Cultural adaptation of the items: In addition to translation into another language, the existing test should also be adapted culturally. Materials lacking cultural items are reported to potentially cause participants to focus their attention to something else (Hambleton, 1994). Such distraction may adversely affect the achievement of the participants in relation to items (Yen, 1993). Focus group interviews were held within this process. Focus group interviews were carried out to make items suited to the Turkish culture following the translation of the test from English into Turkish. To that end, focus group interviews were conducted with a group of two specialized mathematics teacher trainers, two experienced middle school mathematics teachers, four experienced teachers doing a doctorate in mathematics teaching and working at state schools affiliated to the Ministry of National Education and one researcher conducting doctorate studies in mathematics teaching. The four following criteria, put forth by Delaney, Ball, Hill, Schilling and Zopf (2008), were primarily taken into consideration during focus group interviews conducted during the cultural adaptation of items included in the test: (a) changes in the overall cultural context, (b) changes in school culture context, (c) changes in mathematical structure and (d) other. However, this structure was subsequently revised by Ng, Mosvold and Fauskanger (2012) who claimed that these items could necessarily be changed in terms of the language culture of the country of adaptation. In this case, changes in Turkish language culture were also considered another important criterion.

Sample and Data Collection

The application was carried out in the fall and spring semesters of the 2015-2016 academic year. The researcher has visited 38 middle schools by himself in order to reach mathematics teachers. Those who were willing to participate in the study from the mathematics teachers working in the schools were the sample of the MKT-G test. The teachers who were willing to participate consist of 243 mathematics teachers (131 women and 112 men) whose seniority changes between 1 and 21 years (see Table 2).

Table 2. Participants of the Study

Seniority	Gender		Total
	Women N (%)	Men N (%)	N (%)
1 year	6 (2,5)	7 (2,9)	13 (5,3)
2-5 years	30 (12,3)	25 (10,3)	55 (22,6)
6-10 years	49 (20,2)	37 (15,2)	86 (35,4)
11+ years	46 (18,9)	43 (17,7)	89 (36,6)
total	131 (53,9)	112 (46,1)	243 (100,0)

In addition, the teachers' ability scores were examined according to gender and years of experience in the teaching profession within the study. We thought that providing information about the participants would help to explain the data. Therefore, we would share the information about the teacher according to the ability levels examined in the study. Average ability score of teachers was found 0,096 which is very close to 0,000 value which is accepted as normal value. How teachers' ability scores were distributed according to the ability levels is seen in Table 3.

Table 3. Distribution of teachers according to ability levels

Ability level	Number of teachers	Percentage
Very high	21	8,6%
High	51	21,0%
Middle	104	42,8%
Low	55	22,6%
Very low	12	%4,9

When the ability levels of the participating teachers are examined, it is seen that 42.8% of the teachers take an ability score in the range of -0.600 to 0.600 and take part in the 'middle' ability level (see Table 3). On the other hand, it is seen that teachers with 'very high' and 'high' ability levels constitute 29.6% of the group while 'very low' and 'low' ability teachers constitute 27.6% of the group.

While the implementation of the MKT-G test was going on, the interviews were conducted with four teachers who completed the test and wanted to share their opinions on the items included in the test. The interviews have two main objectives. The first of these objectives is to test the validity assumptions of the test. The other objective of the interviews is to obtain Turkish mathematics teachers' opinions in how well the items in the MKT-G test reflect mathematics education in Turkey. In this context, seven questions were prepared with reference to the questions used by Delaney et al. (2008) in their study.

1. Why did you select the option you marked? What did you think?

2. To what extent are the questions linked to the geometry area in the secondary school mathematics program? Is the geometry as measured by this question associated with junior high school geometry we teach in elementary schools?

3. Were the narratives of the questions suitable for situations encountered in the real classroom environment?

4. Was there an unclear language?

5. What questions do you think mathematics teachers can find difficult?

6. Were there any situations that cannot be included in the options of the questions?

7. What is your general view of the questions?

This phase of the study includes the implementation of the MKT-G test to the mathematics teacher of 243 secondary schools who work in secondary schools attached to the Ministry of National Education in Gaziantep.

Data Analysis

The first data came from the focus group interviews. After the interviews, the suggested changes were presented in tables for each of the five topics.

When it comes to the analysis of the data from interviews with four teachers we followed two different approaches. Primarily, with the first question, it was aimed to understand the mathematics knowledge that teachers use when responding to items. In that way, we can see whether the items may provoke the mathematics knowledge intended by test designers. We followed the steps identified by the work of Schilling and Hill (2007). In that study, three different validity assumptions were discussed for the sake of the researches: elemental validity, structural validity and ecological validity arguments. In this study, we studied only two validity arguments because both we saw the validity arguments gathered from elemental and structural assumptions point out that the test has strong validity and also the ecological validity assumption would give an immense body of findings that exceeds the limitations of an article.

Elemental validity assumption relies on the idea that the items reflect teachers' mathematical knowledge for teaching and not extraneous factors such as test-taking strategies or idiosyncratic aspects of the items (e.g., flaws in items) (Hill, Dean & Goffney, 2007). The said validity assumption proposes that the items in the test should reflect teachers' knowledge of teaching rather than their knowledge of exam strategy or indication of errors in items (Schilling & Hill, 2007).

A. Inference: Teachers' reasoning for a particular item will be consistent with the multiple-choice answer they selected.

Structural validity assumption is another method used to investigate the validity of MKT-G test employed during the study is through the investigation of structural validity. Structural assumption: The domain of mathematical knowledge for teaching can be distinguished by both subject matter area (e.g., number and operations, algebra, geometry) and the types of knowledge deployed by teachers. The latter types include the following: content knowledge (CK), which contains both common content knowledge (CCK), or knowledge that is common to many disciplines and the public at large and specialized content knowledge (SCK) or knowledge specific to the work of teaching; and knowledge of content and students (KCS), or knowledge concerning students' thinking around particular mathematical topics. Implications of this include:

A. Teachers' reasoning for a particular item will reflect the type of reasoning (either CK or KCS) that the item was designed to reference (see Hill, Dean & Goffney, 2007).

The elemental and structural assumptions and its inferences asserts that scores reflect specific types of mathematical knowledge for teaching. Analyzing the thinking processes underlying the answers may provide insight into such assumptions. Based on this idea in order to understand patterns in respondents' thought processes Hill, Dean and Goffney (2007) constructed a set of codes that reflect teachers' thinking process (see Table 4).

Table 4. The codes that reflect teachers' thinking process

Code Category	Code abbreviation	Code name	Code description
Mathematical methods	MJ	Mathematical justification	This code reflects *correct* mathematical reasoning about an item, usually with recourse to definitions, definitive examples or counter-examples, or the consideration of unusual cases (e.g., zero or negative numbers).
	MR	Memorized rules or algorithm	Respondent refers to memorized rule or algorithm for primary justification for answer
	DF	Definitions	Correct uses of definitions are included in mathematical justification, above. This code captures inaccurate or incomplete uses of definitions (e.g., "I can't remember the definition exactly").
	EX	Examples/counterexamples or pictures	This code reflects a respondent choosing numbers, cases, figures or shapes (including drawing figures and shapes) to assist in reasoning through a problem. If an example is used as proof (e.g., by counterexample, by unusual case) it is coded as mathematical justification above. This code mainly comprises examples that do not rise to level of justification, or examples used to support incorrect answers.
	OR	Other mathematical reasoning	Mathematical thinking that does not fall into the category of justification, definitions, examples or pictures. The respondent may use mathematical deduction, inference, or other type of thinking to support her answer (e.g., "looking at these numbers listed in the problem, (c) must be true.") The response can either reflect correct thinking or incorrect thinking; the main feature is that there is something mathematical about the thought process, rather than a non-mathematical process such as test-taking, guessing, etc.
	KS	Knowledge of student and content	Respondent invokes knowledge of students as partial or complete explanation for selecting their answer (e.g., "my students do this all the time.")
Non-mathematical methods	GU	Guessing	Respondent reports guessing.
	TT	Test-taking skills	Respondent uses information in the stem, matches response choices to information in the stem, or works to eliminate answers as method for solving problem (e.g., "I knew it wasn't (a), thus (d) 'all of the above' couldn't be true.").
	OT	Other	Other non-mathematical thinking.
	NP	Not present	It represents the cases where no reasoning process is encountered.

This coding scheme is supposed to provide some evidence for both the elemental and structural validity assumptions. In the former assumption, it is expected to have information about the amount of test-taking, guessing, and other non-mathematical thinking that occurred in response to the questions. In the latter one, it would give an important understanding of whether respondents appeared to present the consistent knowledge with the construct demanded by the item.

Analyzing the answers to the remaining interview questions were gone straightforward. We transcribed the entire interview but presented only parts with opposite opinions.

Reliability Study

Test reliability analyses were conducted in accordance with the IRT. The reliability calculation for the IRT is made from a perspective distinct from that of the classical test theory. According to Hambleton, Swaminathan and Rogers (1991), one of the most important advantages of IRT is that no parallel test is needed to measure the reliability of the test (p. 5).

The reliability of the data obtained from the study was calculated with BILOG-MG 3.0 program (Zimowski, Muraki, Mislevy, & Bock, 2003). While it gets close to value 0, the reliability diminishes and the closer the value to 1, the higher the reliability level is.

Psychometric Analysis of the Test Items

The data collected from 243 mathematics teachers was analyzed quantitatively using psychometric methods. Through the quantitative analysis process, we calculated the point-biserial correlation estimates, item difficulties and item discrimination parameters by two-parameter IRT model in each phase. The two parameters IRT model were chosen according to the model-data fit evidence. For this, the likelihood ratio statistics, also Akaike Information Criterion [AIC] (Akaike, 1974) and Bayesian Information Criterion [BIC] (Schwarz, 1978) criteria have been taken into consideration.

Item Response Theory with two-parameter model tries to explain the data with two different parameters; item discrimination parameter and item difficulty parameter. Besides these two parameters, also it is important to study the point-biserial correlation parameter too. The point-biserial correlation indicates how an item correlates with all other items (de Ayala, 2009, p.365). A high point-biserial correlation refers to a strong relationship between the item and the underlying construct of the test (Delaney et al, 2008). Furthermore, higher point-biserial correlation also a parameter to discriminate the individuals who performs closely in the test. On the other hand, a low point-biserial correlation indicates a poorly discriminating item that is measuring nothing. Item discrimination parameters reflect the item's capacity to differentiate among the individuals across the continuum (Baker, 2001). The labels used to describe an item's discrimination could be related to ranges of values of the parameters as such:

Table 5. Labels for item discrimination parameter values

Label	Range of Values
none	0
Very low	0.01-0.34
Low	0.35-0.64
Moderate	0.65-1.34
High	1.35-1.69
Very high	>1.70

The item difficulty is a term for defining the location of an item on the continuum (de Ayala, 2009, p. 98). The BILOG-MG 3.0 program reports item difficulties in standard deviations, on a scale where 0 represents the average teacher ability. Difficulties lower than 0 describe easier items and higher difficulties reflect more difficult items. The value obtained in this parameter is equal to the ability score of the individual or individuals who have a 50% probability of responding correctly to the item to be measured. In theory, the item difficulty parameter value is calculated in the range with -4.000 and 4.000.

Table 6. Labels for item difficulty parameter values

Range of Values	Label
<-1.801	Very easy
-1.800 to -0.601	Easy
-0.600 to 0.600	Moderate
0.601 to 1.800	Hard
>1.801	Very hard

Here, when the item difficulty value approaches to -4.000, the level of difficulty should get lower, on the other hand approaching to 4.000 indicates that the level of difficulty becomes higher.

Results

The first phase of the adaptation process was the translation of the test from English into Turkish. However, while translating the items the main idea was 'do not change the essence of the content'. In Table 7, it can be seen however an item has been translated to Turkish (Hill, Schilling & Ball, 2004).

Table 7. An example for the translation process of a test item

An original released item from LMT	**Translation of the item in Turkish**
At the close of a lesson on reflection symmetry in polygons, Mr. White gave her students several problems to do. He collected their answers and read through them after class. For the problem below, several of her students answered that the figure has two lines of symmetry and several answered that it has four.	Cokgenlerde yansima simetrisi uzerine olan bir dersin kapanisinda, Bay White ogrencilerine cozmeleri icin bir kac problem verir. Asagidaki problem icin, ogrencilerinden bir kaci seklin iki simetri dogrusu oldugunu ve bazilarinin ise dort tane oldugunu soylemistir.
How many lines of symmetry does this figure have?	Bu sekil kac tane simetri dogrusuna sahiptir?
	
Which of the following is the most likely reason for these incorrect answers? (Circle ONE answer.)	Asagidakilerden hangisi yanlis yanitlarin en muhtemel nedenidir? (BIR yanit isaretleyin.)
a) Students were not taught the definition of reflection symmetry.	a) Ogrencilere yansima simetrisinin tanimi ogretilmemistir.
b) Students were not taught the definition of a parallelogram.	b) Ogrencilere bir paralelkenarin tanimi ogretilmemistir.
c) Students confused lines of symmetry with edges of the polygon.	c) Ogrenciler simetri dogrusunu cokgenin kosesi ile karistirmistir.
d) Students confused lines of symmetry with rotating half the figure onto the other half.	d) Ogrenciler simetri dogrusunu seklin yarinin diger yarisi uzerine dondurulmesi ile karistirmistir.

During focus group interviews, items were adapted in terms of the five following key criteria: (i) changes in overall cultural context, (ii) changes in the context of school culture, (iii) changes in mathematical structure, (iv) changes in the translation from the original language into Turkish and (v) other.

As indicated by Delaney et al. (2008), the topic of overall cultural context includes the alteration of the names and words in the test which are not mathematical but used in daily language so that they are suited to the said culture. Some examples of the changes made accordingly are presented in Table 8.

Table 8. Actual changes in overall cultural context

Type of change	Example of use in the original test	Example of use in the test adapted to Turkish	Retranslation of Turkish phrases into English	Number of actual changes
Persons' names	**Ms**. Madison	Melek **ogretmen**	Melek **teacher**	23
	Mr. Grant	Gokhan **ogretmen**	Gokhan **teacher**	
	Carrie	Canan	(a popular female name)	
	Kip	Koray	a popular male name	
	Lily	Leyla	a popular female name	
Non-mathematical language	to stump	Sasirtmak	to confuse	2
	video game	Oyun	game	

The second change category for adaptation process includes changes in school culture or generally in educational system. This category considered the use of language at schools, or in a more general sense the use of language in the educational system.

Table 9. Changes in the context of school culture or generally in educational system

Type of change	Example of use in the original test	Example of use in the test adapted to Turkish	Retranslation of Turkish phrases into English	Number of actual changes
School language	Ms. Marcos' **class**	Merve ogretmenin **ogrencileri**	Teacher Merve's **students**	12
	textbook	Ders kitabi	Course book	
	resource book	**Ders** kitabi	Course book	
	class discussion	Derste uzerinde konusmak	to discuss during the course	
	unit	Konu	subject	
	raise hand	parmak kaldirmak	Raise finger	
Structure of teaching system	**math methods** book	**Matematik ogretim yontemleri** kitabi	**Mathematics teaching methods** book	1

The changes in the context of overall culture and school culture of the country to which items are adapted are important and necessary for preventing the distraction of teachers by foreign terms and context (Delaney et al., 2008). Items have a very slight chance of being affected in terms of their measuring mathematical knowledge since the changes contained herein do not disrupt the mathematical structure of items.

Table 10. Changes in the mathematical structure

Type of change	Example of use in the original test	Example of use in the test adapted to Turkish	Retranslation of Turkish phrases into English	Number of actual changes
Symbolic Notations	Bh	$a \times h$		2
	bh	$c \times h$		
	1	$1br^2$	$1 squarredunit$	
Mathematical Language	Area of a **circle**	**Dairenin** alani	**Circle's** area	8
	All of the **angles** of a hexagon are equal.	Bir altigenin tum **ic acilarinin** olculeri esittir.	All of the **interior angles** of a hexagon are equal.	
	A trapezoid's two bases (**B and b**)	Bir yamugun iki tabani (a **ve** c)	A trapezoid's two bases (***a* and *c***)	
	Base (e.g. polygon)	Taban	Base	
	Irregularly shaped regions	**Kapali egrisel** bolgeler	Closed curvilinear areas	
	Rotation	donme hareketi	Rotation movement	
	Draw a square with an area of **twenty** centimeters **squared**	Alani **20** santimetrenin **karesi olan** bir kare ciziniz	Draw a square whose area is square of 20 centimeters.	
Units of measurement	inch	cm	centimeter	1

Unlike two earlier categories, the third category relates to the mathematical structures of items (see Table 10). The changes to this category were altered to suit the mathematics culture of Turkey. Units of measurement were converted into metric units used in Turkey. Table 3 includes the examples of the changes. The changes introduced in this category do not only alter the mathematical structure of items but also creates the risk of the test becoming mathematically easier or more difficult (Delaney et al., 2008; Ng, Mosvold & Fauskanger, 2012). In this sense, any changes under this category have been planned with great care.

Adoption of a structure unique to the Turkish language during the adaptation of the test to Turkish is critical for eliminating ambiguity during translation (Kwon, Thames & Pang, 2012; Ng, Mosvold & Fauskanger, 2012). Table 4 lists the changes in the language structure during the adaption of the test to Turkish.

Table 11. Changes in Turkish language culture

Type of change	Example of use in the original test	Example of use in the test adapted to Turkish	Retranslation of Turkish phrases into English	Number of actual changes
Language structure	**I'm** not sure	Emin degilim	Not sure	8
	Choose the shape below that, no matter how you rotate it, will not match any of the shapes above.	Asagida verilen sekillerden hangisinin donme hareketi altindaki goruntusu yukarida verilen sekillerden birisi olamaz**?**	Which image of the following shapes under rotation movement cannot be one of the shapes given above**?**	
	Only Carrie's.	Sadece Cananin **yontemi kabul edilebilirdir**	Only Canan's **method is acceptable**	
	What do you call all quadrilaterals whose two diagonals are both lines of symmetry?	Her iki kosegeni de simetri dogrusu olan tum dortgenlere **ne ad verilir?**	**What are** all quadrilaterals whose two diagonals are both lines of symmetry **called?**	
	Count the number of whole squares completely inside the region and the number of squares partly inside the region. **Divide** the number of squares partly inside the region by two and **add** it to the number of whole squares.	Tamami bolgenin icinde kalan kareler ile bir kismi bolgenin icinde kalan kareleri **saydim**. Bir kismi bolgenin icinde kalan karelerin sayisini ikiye **boldum** ve buldugum sayiyi tamami bolgenin icinde kalan karelerin sayisina **ekledim**.	**I counted** the number of squares completely inside the region and the number of squares partly inside the region. **I divided** the number of squares partly inside the region by two and **added** it to the number of squares completely inside the region	

The elimination of deficiencies which are likely to occur in language structure when translating the test items into a different culture is considered of importance to ensure intelligibility of the items. The wording "I'm not sure" used in the original test would be translated into Turkish as "Ben emin degilim". However, since Turkish is an agglutinative language, the wording will appear to give the same meaning when the affix "-m" follows as a personal pronoun. Using the personal pronoun "Ben" and the suffix "-m" at the same time would cause incoherence. Therefore, the said wording was used as "Emin degilim". Moreover, the wording "Choose the shape below that, no matter how you rotate it, will not match any of the shapes above" does not appear to be a question in general when translated into Turkish. During focus group interviews, it was stated that it would make more sense to replace this wording with a question sentence and the sentences was paraphrased as "Asagida verilen sekillerden hangisinin donme hareketi altindaki goruntusu yukarida verilen sekillerden birisi olamaz?" (Which image of the following shapes under rotation movement cannot be one of the shapes given above?) A similar process was used for other changes and given under the table 'Changes related to Turkish language culture.'

The changes under the fifth category contain arrangements which are not included in the item adaptation process, but which have been made to improve readability, such as arrangement of the layout of questions on the page, re-numbering of the items, enlarging font sizes, or highlighting using bold font to emphasize sentences. In this sense, the changes under this category are assumed not to affect the overall culture, school culture or mathematical culture in the test, but to improve them ergonomically (in terms of form/convenience).

(iii) Interviews after implementation of the test

In the interview, the first question was to capture the evidence of the validity of the test. Below, evidence for elemental and structural validity and then reliability of the test shared.

Evidence for the Elemental Assumption

Tables 12 shows the types of thinking that occurred in answers to the items in MKT-G test. Results show that the most commonly executed processes include mathematical justification, examples or counterexamples or other mathematical reasoning which are mathematical in nature. In other words, respondents mostly used mathematical knowledge or thinking rather than non-mathematical skills or strategies. Throughout the interviews, respondents claimed no use of guessing while minimal use of test-taking skills or non-mathematical reasoning. In addition, 5% of the responses had no reasoning et al. The participants only responded to the answer without and reasoning.

Table 12. Explanations for the items in the test

teachers		MJ	MR	DF	EX	OR	KS	GU	TT	OT	NP
M1	n	13	0	5	2	5	2	0	0	0	1
	%	43.3	0	16.7	6.7	23.3	6.7	0	0	0	3.3
M2	n	21	2	0	3	2	1	0	0	0	1
	%	70.0	6.7	0	10.0	6.7	3.3	0	0	0	3.3
M3	n	13	2	3	5	5	2	0	0	0	0
	%	43.3	6.7	10.0	16.7	16.7	6.7	0	0	0	0
M4	n	16	1	1	4	1	1	0	1	1	4
	%	53.3	3.3	3.3	13.3	3.3	3.3	0	3.3	3.3	13.3
Average	n	15.75	1.25	2.25	3.50	3.75	1.5	0	0.25	0.25	1.5
	%	52.5	4.2	7.5	11.7	12.5	5.0	0	0.8	0.8	5.0

It is very important for respondents to use predominantly mathematical justifications along with examples or counterexamples and other mathematical reasoning. In that way, we can understand the items probed the mathematical knowledge and thinking successfully.

Evidence for the Structural Assumption

Coherence between the mathematical thinking that the teachers participated in answering the items and the information planned to be revealed with the items were examined. More precisely, it is examined whether the items allow the teachers to reveal the type of mathematical knowledge that is planned to be measured. In this sense, the relation between the mathematical ways of thinking revealed by the teachers who are interviewed in Table 13 and the type of knowledge that the items aim to reveal is seen.

Table 13. Consistency of answers

Teachers	Frequency of Inconsistent answers	Percentage of inconsistent answers
M1	3	%10
M2	0	%0
M3	9	%30
M4	1	%3.3
average	3.25	%10.8

Table 13 shows that the average inconsistency responses of the MKT-G test have a 10% range in all responses. In other words, it is seen that the ratio of teachers' mathematical knowledge in responding to the items is close to 90%.

Furthermore, Table 12 shows that teachers used their knowledge of students while answering the items in the test through interviews. It is also seen that only 5% of the strategies used by the teachers in answering items are based on KCS. However, considering that only one of the items in the test was designed to reference KCS, it is important that teachers use the right method in the right place. For example, in an item, a concave quadrilateral and a false idea based on the wrong measurement of the angles were given. In the item, the respondents were wanted to find the most reasonable appraisal for student' s thinking. In discussing this item, a teacher commented:

In fact, we [teachers] expect them [the students] to use a formula to find the answer. The formula doesn't say anything, it just *does* something. However, I'm giving you the basic logic under it. In my class, *I want them to draw a triangle and make sense with their angles*.

In this particular item, the teacher gave an explanation right from his class. He replied how he would act to solve that false reasoning. Likewise, in the items included in the test, it is expected that both the mathematical knowledge that common other subject areas and teacher-specific knowledge too. In this sense, success in revealing information specific to individuals teaching in mathematics provides important information about the structural validity of the test.

Reliability of the Test

The reliability value of the test was calculated from the data of 243 teachers who participated in the study. BILOG MG-3.0 software was used for the calculation.

Table 14. Reliability of MKT-G test

Test	Number of the items	Reliability
MKT-G	N=30	0.8102

The reliability value of the data obtained as a result of the application of MKT-G test was calculated as 0.8102. This value is found significantly high.

Finding about the Relevance of the Items to Turkish Mathematics Classrooms

In interviews with teachers, the mathematical content contained in test items teachers in Turkey had asked that it is the degree to relations with the mathematics they teach in school question, in general, they stated their opinion that it is compatible. However, one of the teachers expressed his opinion about the content of ITEM 28 *"...I have not encountered anything like this*." All of the items apart from that item were described as the items were related to the mathematics that are taught in Turkish middle schools. During the interviews, one teacher expressed his opinion about the relevance of an item that includes properties of polygons in Turkish mathematics classes:

"Yes [it is related]. I'm doing this in class. I tell my students that what properties the polygons must have or should not have. Like; 'the inner angles of a regular polygon cannot be different'; or the sum of angles in triangles cannot be greater than 180 degrees...'

Another teacher expresses the relationship between the items in the test and the content of mathematics teaching in Turkey as follows:

[items] actually correspond to the misconceptions that students often encounter. For example, if the circumference [of polygons] increases, the area will also increase and vice versa. [Items and mathematics contents] match each other."

It was observed that the teachers did not give an opinion indicating the contrary. For example, an opinion on this situation was recorded as follows:

"...let me say this: I had drawn a rectangle [in the classroom], I asked whether the figure was the proper manager or not. We talked a lot about it in the classroom…"

During the interviews with the teachers, they were asked whether the items included in the test had a narrative disorder or not. The opinions submitted in the context of this question showed that ITEM 28 contained a language problem. One of the teachers expressed this situation as follows: *"the term 'how far around' is a question mark is in my head."* However, the term "how far around" was changed to "circumference" and the teacher's approval about the correction was taken. Therefore, it was thought that the other items in the MKT-G test were not problematic in terms of language.

Psychometric Characteristics of MKT-G Test Items

The psychometric values of the MKT-G test items, point-biserial correlation, item discrimination, item difficulty parameters are presented below (see Table 15). In addition, the correlation between point-biserial correlation values obtained in the U.S. study and the values reached in Turkish study was also calculated.

Table 15. Psychometric characteristics of MKT-G items

Items	Point-biserial correlation-Turkey	Item discrimination parameter	Item difficulty parameter	Point-biserial correlation in U.S.
ITEM 1	0.213	**0.525**	-2.666	0.318
ITEM 2	0.174	**0.552**	-2.601	0.473
ITEM 3	0.174	**0.525**	-3.951	0.489
ITEM 4	**0.074**	**0.377**	-1.652	0.056
ITEM 5	0.285	0.606	-0.694	0.321
ITEM 6	0.289	0.695	-1.065	0.615
ITEM 7	0.311	0.656	-0.261	0.536
ITEM 8	0.296	**0.569**	0.354	0.457
ITEM 9	0.488	0.934	-1.971	0.685
ITEM 10	0.402	0.880	-1.852	0.700
ITEM 11	0.697	1.711	-1.476	0.679

Table 15. Continued

Items	Point-biserial correlation-Turkey	Item discrimination parameter	Item difficulty parameter	Point-biserial correlation in U.S.
ITEM 12	0.434	0.956	-0.858	0.593
ITEM 13	0.579	1.274	-1.806	0.611
ITEM 14	0.506	1.125	-0.912	0.505
ITEM 15	0.506	1.082	-2.200	0.596
ITEM 16	0.345	0.779	-1.359	0.646
ITEM 17	0.518	1.122	-1.531	0.422
ITEM 18	0.567	1.207	-2.045	0.426
ITEM 19	0.648	1.503	-1.667	0.616
ITEM 20	0.511	1.069	-3.020	0.778
ITEM 21	0.613	1.771	-1.219	0.864
ITEM 22	0.578	1.646	-1.242	0.853
ITEM 23	0.296	0.701	-2.057	0.776
ITEM 24	0.331	0.776	-2.191	0.604
ITEM 25	0.492	1.092	-1.424	0.657
ITEM 26	0.434	1.000	-2.283	0.422
ITEM 27	0.222	0.607	2.314	0.480
ITEM 28	0.167	**0.528**	2.605	0.610
ITEM 29	0.160	**0.481**	0.625	0.577
ITEM 30	0.391	0.861	0.010	0.514
Average	0.390	0.920	-1.270	0.562

Point-biserial correlation values of items were found to be 0.074 (ITEM 4) the lowest and 0.697 (ITEM 11) the highest. The correlation value of a reliable item is expected to be close to 1 since point-biserial correlation is a measure that shows the correlation between the responses given by individuals to each item included in the test and their responses to the whole test (Hambleton, Swaminathan & Rogers, 1991, p. 94). To put it more clearly, an individual with a high score in the point-biserial correlation value is expected to correctly respond to the item, for which the said value is calculated, while an individual with a low score is expected to respond incorrectly. Additionally, point-biserial correlation values for the other items were investigated. The intention here is to demonstrate that items having negative point-biserial correlation values fail to allow measurement in consistence with other items included in the test. The lack of items having a negative point-biserial correlation value indicates that the test items measures consistently.

The significant relationship ($r = 0.525$; $n = 30$; $p < 0.01$) between point-biserial correlations scores of the results in U.S. application and obtained in this study is another proof of the validity of the MKT-G test.

When Table 15 is examined in terms of item discrimination parameter values, it was seen that while 60% of the items have moderate and 13% of them have “high” or “very high” discrimination powers. On the other hand, it was noteworthy that the average distinguishing value of the test items was found as 0.920 and that means a greater value than the critical value of 0.800.

Thereafter, it was observed that 80% of the items were found very easy and easy, while 10% is found moderate by the teachers.

Discussion

The adaptation of a test developed in a country to any other country is considered to involve a highly challenging process. However, in this study, it was observed that a test developed with the intent to measure mathematical knowledge for teaching of middle school mathematics teachers serving at schools in the United States could successfully be adapted to Turkey. A critical question arises in relation to the capability of adapting such a test developed in another country to another country: To what extent does the test match with the geometry knowledge of mathematics teachers in the country to which the test is going to be adapted? It is considered that the answer to this question could be best given by mathematics teacher trainers and experienced mathematics teachers. Therefore, focus group interviews included both mathematics teacher trainers and mathematics teachers. They were found to make key contributions to understanding the extent to which each item included in the test was consistent with the knowledge of mathematics teachers in Turkey. Basically, it was concluded that the area calculation of geometric objects, the relationship between area and circumference concepts, basic characteristics of linear geometric objects, and angle concepts in polygons were important components of the content of geometry teaching in Turkey.

Although it was seen that there are similarities between mathematics teaching and learning cultures in the United States and Turkey, it was found that there are variances in terms of overall and school cultures and mathematical language. Another significant difference in the adaptation process during translation from English into Turkish is that the necessary avoidance of disrupting coherence as well as the necessary preservation of the content as is turned out to be an essential criterion. In this context, the steps developed by Delaney et al. (2008) and finalized by Ng, Mosvold and Fauskanger (2012) have been shown to provide significant support for cultural adaptation of the items without changing the mathematical content that is intended to be measured. The interviews with four teachers during the implementation of the test also provides valuable information in terms of suitability of the test items to middle school mathematics in Turkey. Though, the teachers gave brief but clear messages that includes confirmation of usability of the test items.

Validity and reliability results of the adapted MKT-G test seems also promising. Teachers' answers to the questions were mostly seen as a reference to the mathematics teaching knowledge that intended to be measured with the use of the test items. In addition, it was seen that teachers' solution methods were observed very little in test-taking and other non-mathematical strategies categories and also where no strategies were observed was significant. In addition, it was observed that teachers' thinking and the answers they give to items were highly consistent. Since the reliability analysis gave a quite high value, the test is found also a reliable measurement. Therefore, the test can be accepted as a valid and reliable instrument.

The findings and psychometric results from the semi-structured interviews conducted with four teachers following the pilot application of the test indicate that the adaptation process was successful in general terms.

Psychometric values contain important information about the characteristics of the test items. Point-biserial correlation values indicate how well the dichotomous response on a specific item correlates with the total score in a test. In that sense, negative or very close to zero values indicate, respectively, an inverse or weak relationship with the score from the whole test. There are very few items that have a point-biserial correlation close to zero, and none have a negative value. Therefore, along with the idea that as stated in de Ayala (2009) the items with negative correlation values should be discarded, none of the items were considered to be discarded. In addition, the positive significant relationship between point-biserial correlations of the items obtained in Turkish adaptation with values from the U.S. version of the research gives promising results in the favor of adaptation. The relation between two data sets show a specific relationship with Turkish adaptation and the U.S. version of the MKT-G test. Besides, item discrimination parameter values of the test items indicate that mostly items discriminate the individuals "moderately" and very few instances the discrimination occurs as "high" or "very high". That means; the test discriminates the individuals not quite strong nor very weak but in moderate manners. It was found that the teachers that participated in the study found the items mostly "very easy" or "easy". In a very few instances, the items were considered as "very hard" or "moderate".

Knowledge for teaching geometry test is intended to measure both the existing knowledge of teachers and the knowledge which teachers are expected to possess in the light of advancements. The options included that the variances between the content of geometry teaching of both countries would make test adaptation difficult. The most fundamental cause of this was considered to be the likelihood of dealing with the geometry knowledge in the test in line with mathematical content at various school levels. However, given the fact that the content of mathematics teaching in Turkey closely follows the advancements in the U.S., it was considered that the geometry knowledge which mathematics teachers should possess would advance in parallel with the said advancements. In fact, the interviews conducted revealed that four mathematics teachers were found to be familiar with the content of the items on the test although the mathematics teachers in Turkey found some items in the test easier or more difficult.

References

Akaike, H. (1974). A new look at the statistical identification model. *IEEE Transaction Automatic Control*, *19*(6), 716–723.

Atweh, B., Clarkson, P., & Nebres, B. (2003). Mathematics education in international and global contexts. In A. J. Bishop, M. A. Clements, C. Keitel, J. Kilpatrick, & F. K. S. Leung (Eds.), *Second international handbook of mathematics education* (pp. 185-229). Dordrecht, The Netherlands: Springer.

Baker, F. B. (2001). *The basics of item response theory* (2nd ed). Washington, DC: ERIC clearinghouse on assessment and evaluation.

Ball, D. L. (1990). The mathematical understandings that prospective teachers bring to teacher education. *The Elementary School Journal*, *90*(4), 449-466.

Ball, D. L. (1993). With an eye on the mathematical horizon: dilemmas of teaching elementary school mathematics. *The Elementary School Journal*, *93*(4), 373-397.

Ball, D. L., & Bass, H. (2003). Toward a practice-based theory of mathematical knowledge for teaching. In B. Davis & E. Simmt (Eds.), *Proceedings of the 2002 annual meeting of the Canadian Mathematics Education Study Group* (pp. 3–14). Edmonton, Canada: CMESG/GCEDM.

Ball, D. L., Hill, H. C., & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator, 29*(1), 14-17, 20-22, 43-46.

Ball, D. L., Lubienski, S. T., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching (4th Edition)* (pp. 433-456). New York, NY: Macmillan.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of teacher education*, *59*(5), 389-407.

Clements, D., & Sarama, J. (2000). Young children's ideas about geometric shapes. *Teaching children mathematics, 6*(8), 482 – 488.

de Ayala, R. J. (2013). *The theory and practice of item response theory*. New York, NY: Guilford Press.

Delaney, S. F. (2008). *Adapting and using US measures to study Irish teachers' mathematical knowledge for teaching* (Unpublished doctoral dissertation). University of Michigan, Michigan, USA.

Delaney, S., Ball, D. L., Hill, H. C., Schilling, S. G., & Zopf, D. (2008). "Mathematical knowledge for teaching": Adapting US measures for use in Ireland. *Journal of mathematics teacher education*, *11*(3), 171-197.

Guler, M. (2014). Ogretmen adaylarinin matematik ogretme bilgilerinin incelenmesi: Cebir ornegi [Investigating pre-service teachers' knowledge for teaching mathematics: The sample of algebra] (Unpublised master's thesis). Karadeniz Techniqal University, Trabzon, Turkey.

Hambleton, R. K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment*, *10*(3), 229-244.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory* (Vol. 2). Newbury Park, CA: Sage.

Hill, H. C., & Ball, D. L. (2004). Learning mathematics for teaching: Results from California's mathematics professional development institutes. *Journal for Research in Mathematics Education*, *35*(5), 330-351.

Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L., & Ball, D. L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, *26*(4), 430-511.

Hill, H. C., Dean, C., & Goffney, I. M. (2007). Assessing Elemental and Structural Validity: Data from Teachers, Non-teachers, and Mathematicians Measurement. *Interdisciplinary Research and Perspectives*, 5(2-3), 81-92, doi: 10.1080/15366360701486999.

Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *Elementary School Journal, 105*(1), 11-30.

Judson, T. W. (1999). Japan: A different model of mathematics education. In E. A. Gavasto, S. G. Krantz & W. McCallum (Eds.), *Contemporary issues in mathematics education* (pp. 75-81). Cambridge, UK: Cambridge University Press.

Kupari, P. (2008). Mathematics education in Finnish comprehensive school: characteristics contributing to student success. *Proceedings of the XI International Congress in Mathematics Education*, ICME, Mexico.

Learning Mathematics for Teaching [LMT]. (2008). *Mathematical Knowledge for Teaching (MKT) Measures: Mathematics Released Items*. Retrieved from http://www.sitemaker.umich.edu/lmt/files/LMT_sample_items.pdf

Learning Mathematics for Teaching Project [LMTP]. (2011). Measuring the mathematical quality of instruction. *Journal of Mathematics Teacher Education*, *14*(1), 25-47.

Ma, L. (2010). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States* (2nd ed.). New York, NY: Routledge.

Milli Egitim Bakanligi [MEB]. (2018). *Matematik dersi ogretim programi (Ilkokul ve Ortaokul 1-8. Siniflar)* [Primary schools' math education program for grades 1-8]. Ankara, Turkey: Milli Egitim Bakanligi.

National Council of Teachers of Mathematics [NCTM]. (2000). *Principles and standards for school mathematics* (Vol. 1). Raston, VA: National Council of Teachers of Mathematics.

Ng, D., Mosvold, R., & Fauskanger, J. (2012). Translating and Adapting the Mathematical Knowledge for Teaching (MKT) Measures: The Cases of Indonesia and Norway. *The Mathematics Enthusiast, 9*(1), 149-178.

Peace, H., Quebec Fuentes, S., & Bloom, M. (2018). Preservice teachers' transforming perceptions of science and mathematics teacher knowledge. *International Journal of Educational Methodology, 4*(4), 227-241.

Rosenshine, B., & Furst, N. (1986). The use of direct observation to study teaching. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 122–183). New York, NY: Macmillan.

Schilling, S. G., & Hill, H. C. (2007). Assessing measures of mathematical knowledge for teaching: A validity argument approach. *Measurement*, *5*(2-3), 70-80.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461-464.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4-14.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York, NY: Free Press.

Turner, J. C., & Meyer, D. K. (2000). Studying and understanding the instructional contexts of classrooms: Using our past to forge our future. *Educational Psychologist*, *35*(2), 69-85.

van den Heuvel-Panhuizen, M. (2000). *Mathematics education in the Netherlands: A guided tour. FI-ICME-9 cd-rom*. Utrecht, The Netherlands: Freudenthal Instituut.

Yen, W. M. (1993). Scaling performance assessments: Strategies for managing local item dependence. *Journal of Educational Measurement*, *30*(3), 187-213.

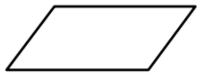
Zimowski, M. F., Muraki, E., Mislevy, R. J., & Bock, R. D. (2003). *Bilog-MG 3.0; Item analysis and test scoring with binary logistic models for multiple groups*. Mooresville, IN: Scientific Software International.

Appendix

Some released items from MKT- test

18. At the close of a lesson on reflection symmetry in polygons, Ms. White gave her students several problems to solve. She collected their answers and read through them after class. For the problem below, several of her students answered that the figure has two lines of symmetry and several answered that it has four.

How many lines of symmetry does this figure have?



Which of the following is the most likely reason for these incorrect answers? (Circle ONE answer.)

a) Students were not taught the definition of reflection symmetry.

b) Students were not taught the definition of a parallelogram.

c) Students confused lines of symmetry with edges of the polygon.

d) Students confused lines of symmetry with rotating half the figure onto the other half.

27. After investigating where π comes from, a student says, “Why is it just circles that have this special number, π?” Mrs. Bell wants to give her students other shapes that have a constant ratio between “how-far-around” and “how-far-across” (analogous to circumference and diameter). Which of the following shapes would <u>best</u> fit her purpose? (Circle ONE answer.)

a) Squares

b) Rectangles

c) Hexagons

d) All of these would work equally well.

e) None of these would work