DOKUZ EYLÜL ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

TIMSS Matematik Alt Test Performansının Orta Ölçekli Bir Şehirdeki Yansıması[1]

The Reflection of TIMSS Mathematics Subtest Performance In a Middle-Sized City[1]

Feride ÖZYILDIRIM GÜMÜŞ, Doç. Dr., Aksaray Üniversitesi, ferideozyildirim@aksaray.edu.tr

Oğuzhan ATILGAN, Mezun Öğrenci, Aksaray Üniversitesi, oguzhan.atilgan@asu.edu.tr

Geliş tarihi - *Received:* 30 Mart 2023
Kabul tarihi - *Accepted*: 21 Eylül 2023
Yayın tarihi - *Published*: 28 Aralık 2023

[1] This study was obtained from the project of TUBITAK 2209-A University Students Research Projects Support Program with 1919B012105919 project number which was carried out under the supervision of the first author and direction of second author.

Özyıldırım Gümüş, F. ve Atılgan, O. (2023). The reflection of TIMSS mathematics subtest performance in a middle-sized city. *Batı Anadolu Eğitim Bilimleri Dergisi, 14*(2), 949-968.
DOI. 10.51460/baebd.1273731

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

Öz.
Bu çalışma kapsamında İç Anadolu bölgesinde yer alan orta ölçekli bir ilin merkezinde öğrenim gören 5.sınıf öğrencilerinin TIMSS matematik alt testlerinin maddelerinden oluşan bir başarı testine dair performanslarının öğrenme alanı, bilişsel seviye ve cinsiyet değişkenleri çerçevesinde incelenmesi amaçlanmıştır. Bu amaç doğrultusunda da nicel araştırma yöntemleri kullanılarak ilişkisel tarama yoluna gidilmiştir. TIMSS matematik alt testlerinin dördüncü sınıf düzeyi için resmi kurumlarca açıklanan maddelerinden öğrenme alanı ve bilişsel seviye dağılımı dikkate alınarak oluşturulan bir başarı testi için öncelikle 154 tane beşinci sınıf öğrencisi ile geçerlik güvenirlik çalışması yapılmış ve ardından farklı 333 öğrenci ile asıl çalışma gerçekleştirilmiştir. Elde edilen verilere göre öğrenciler her üç öğrenme alanında (sayılar ve işlemler, geometri ve ölçme, veri) en yüksek performansı, en üst bilişsel seviye olan akıl yürütme seviyesindeki maddelerde göstermişleridir. Öğrenme alanı bazında elde edilen sonuçlara göre ise öğrenciler en yüksek performansı veri, en düşük başarı ise sayılar ve işlemler öğrenme alanında sergilemişlerdir. Sonuçlar cinsiyet temelinde incelendiğinde ise sayılar ve işlemler öğrenme alanının uygulama düzeyinde, geometri öğrenme alanının bilme ve uygulama düzeyine ait puanlarda cinsiyete göre anlamlı farklar olduğu belirlenmiştir.
Anahtar Kelimeler: Bilişsel seviye, matematik öğrenme alanı, TIMSS matematik alt testi.

Abstract.
The aim of this study was to examine the performance of fifth grade students, who are studying in a middle-sized city in the Central Anatolia region, on an achievement test consisting of TIMSS mathematics subtest items, in terms of learning area, cognitive level, and gender variables. To achieve this aim, a quantitative research method was used, and a relational survey was conducted. Firstly, validity and reliability study were conducted with 154 fifth grade students for an achievement test created based on the distribution of learning area and cognitive level of the items announced by official institutions for the fourth grade level of TIMSS mathematics subtests. Then, the actual study was carried out with 333 fifth grade students. According to the results, the students showed the highest performance in items at the highest cognitive level, which is the reasoning level, in all three learning areas (numbers and operations, geometry and measurement, data). In terms of learning area, the students showed the highest performance in data, while the lowest performance was exhibited in numbers and operations. When the results were examined according to gender, it was found that there were significant gender differences in the scores belonging to the applying level of the numbers and operations learning area and the knowing and applying levels of the geometry and measurement learning area.
Keywords: Cognitive level, mathematics learning area, TIMSS mathematics subtest.

Özyıldırım Gümüş, F. ve Atılgan, O. (2023). The reflection of TIMSS mathematics subtest performance in a middle-sized city. *Batı Anadolu Eğitim Bilimleri Dergisi, 14*(2), 949-968.
DOI. 10.51460/baebd.1273731

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

Genişletilmiş Özet

Giriş. TIMSS (Trends in International Mathematics and Science Study) uluslararası düzeyde gerçekleştirilen, öğrencilerin fen bilimleri ve matematik alanındaki bilgilerini ve becerilerini inceleyen bir araştırmadır. 4. ve 8. sınıf öğrencilerinin katıldığı TIMSS, dört yılda bir uygulanmaktadır. Türkiye TIMSS'e 8. sınıf düzeyine 1999, 2007, 2011, 2015 ve 2019 yıllarında katılmıştır. 4. sınıf düzeyinde ise 2011 ve 2015 yıllarında katılmıştır. 2019 yılında ise Türkiye TIMSS'e 8. sınıfların yanında 5. sınıf düzeyi ile katılmıştır. Çünkü MEB (2020) aldığı son kararla gerek yaş ortalaması nedeniyle, gerekse sınavın çerçevesine uygunluk açısından TIMSS'in 4.ve 8. sınıflar yerine 5. ve 8. sınıflarda uygulanmasının daha uygun olduğuna karar vermiştir.

TIMSS, konu bilgisinin yanında bilişsel becerileri de ölçmeye yönelik bir sınavdır. Bu bağlamda TIMSS matematik alt testinde yer alan sorular bilme, uygulama ve akıl yürütmeye dair bilişsel seviyelere yönelik hazırlanmaktadır (MEB, 2020). Ancak ders kitaplarındaki bazı (örneğin veri işleme öğrenme alanında) görevlerin bilişsel alanları ile öğretim programındaki kazanımların ait oldukları bilişsel alanlarının farklı olduğu belirlenmiştir (Yilmaz, N., Ay, Z., & Aydin, Ş., 2021). Bununla birlikte TIMSS her okulda uygulanmadığından, daha küçük çaptaki illerde öğrenim gören öğrencilerin bu teste ilişkin durumları görmek mümkün olmayabilir. Bir başka ifade ile ülke çapından coğrafi bölge bazında (MEB, 2020) örneklem seçilerek gerçekleştirilen TIMSS, tüm şehirlerde, her okulda uygulanamadığından küçük lokasyonlar hakkında detaylı bilgi sağlayamayabilir. Bu nedenle de gerçekleştirilen bu çalışmada, araştırmanın yürütüldüğü ildeki okullardan örneklemler seçilmiş ve elde edilen verilerle uluslararası düzeyde gerçekleştirilen bir sınavın, orta ölçekli bir ildeki yansıması incelenmiştir.

Yöntem. Nicel araştırma yöntemlerinden ilişkisel tarama deseninin benimsendiği çalışmada İç Anadolu bölgesinde yer alan orta ölçekli bir ilin merkezindeki devlet okullarının beşinci sınıfında öğrenim gören öğrenciler arasından küme örnekleme yöntemi ile pilot çalışma için 154 öğrenci, ana uygulama için 333 öğrenci olmak üzere toplam 487 öğrenci ile çalışma gerçekleştirilmiştir. Farklı yıllarda dördüncü sınıf TIMSS matematik alt testinde yer alan çoktan seçmeli maddelerden MEB tarafından erişime açık olanlar, veri toplama aracını hazırlamak amacıyla havuza alınmıştır. Pilot çalışma kapsamında geçerlik ve güvenirlik analizlerinin tamamlanmasıyla veri toplama aracının nihai hali oluşturulmuş ve ana uygulama gerçekleştirilmiştir.

Bulgular. Çalışma kapsamında öğrenciler en yüksek performansı veri, daha sonra geometri ve ölçme öğrenme alanında göstermişleridir. En düşük başarı ise sayılar ve işlemler öğrenme alanında gözlenmiştir. Öğrenme alanı ve bilişsel seviye birlikte değerlendirildiğinde elde edilen verilere göre öğrencilerin en yüksek performansı, her üç öğrenme alanının da en üst bilişsel seviyesi olan akıl yürütme seviyesinde sergiledikleri gözlenmiştir. Cinsiyete göre elde edilen veriler incelendiğinde ise sayılar ve işlemler öğrenme alanının uygulama düzeyi ile geometri öğrenme alanının bilgi ve uygulama düzeylerinde erkek öğrencilerin anlamlı bir şekilde kız öğrencilerden daha yüksek bir ortalamaya sahip oldukları gözlenmiştir.

Tartışma Sonuç ve Öneriler. Öğrencilerin en yüksek performansı veri, daha sonra geometri ve ölçme öğrenme alanında; en düşük başarıyı ise sayılar ve işlemler öğrenme alanında sergilemiş olmaları alan yazında bazı çalışmalarla benzerlik gösterirken (Büyüköztürk, Çakan, Tan, & Atar, 2014; Kılıç, Aslan-

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

Tutak, & Ertaş (2014). bazıları ile zıtlık sergilemektedir (Dindyal, 2008) . Bununla birlikte öğrenme alanı ve bilişsel seviye birlikte değerlendirildiğinde elde edilen verilere göre öğrencilerin en yüksek performansı, her üç öğrenme alanının da en üst bilişsel seviyesi olan akıl yürütme seviyesinde sergiledikleri gözlenmiştir. Elde edilen bu bulgu için alan yazında hem benzer hem tersi sonuçlar mevcuttur. Cinsiyete göre elde edilen sonuçlar sayılar ve işlemler öğrenme alanının uygulama düzeyi ile geometri öğrenme alanının bilgi ve uygulama düzeylerinde erkek öğrencilerin anlamlı bir şekilde kız öğrencilerden daha yüksek bir ortalamaya sahip olduklarını göstermektedir. Elde edilen bu sonuç, alan yazında var olan bulguların bazıları ile örtüşürken, bazıları ile farklılık göstermektedir. İlerleyen çalışmalarda da uluslararası düzeyde gerçekleştirilen testlerin, farklı iller ya da ilçeler bazında uygulanarak sonuçların raporlanması önerilebilir. Böylece söz konusu bölgelerin durumlarını daha detaylı şekilde görüp bu sonuçlara göre il ve ilçe milli eğitim müdürlükleri ile okul yönetimleri gereken düzenlemeleri gerçekleştirebilir.

Özyıldırım Gümüş, F. ve Atılgan, O. (2023). The reflection of TIMSS mathematics subtest performance in a middle-sized city. *Batı Anadolu Eğitim Bilimleri Dergisi, 14*(2), 949-968.
DOI. 10.51460/baebd.1273731

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

Introduction

In today's world, there are many large-scale tests that measure mathematics achievement at the national or international levels. In mathematics tests that are included as subtests of tests such as YKS (Yükseköğretim Kurumları Sınavı- [Higher Education Institution Exam]), KPSS (Kamu Personeli Seçme Sınavı- [Public Personnel Selection Examination]), ALES (Akademik Personel ve Lisansüstü Eğitimi Giriş Sınavı- [Academic Personnel and Postgraduate Education Entrance Exam]), or centralized exams for transition to secondary education institutions at the national level, test items are generally based on the objectives included in the curriculum or the knowledge and skills that the relevant educational level aims to provide. With these tests, a measurement process is carried out, and usually, a placement is made as a result of this measurement. On the other hand, with the data obtained from the mathematics subtests of large-scale tests such as TIMSS (Trends in International Mathematics and Science Study) and PISA (Programme for International Student Assessment), which are carried out at the international level, not only are the countries' achievements compared but also different variables that may be related to this achievement level are examined. Moreover, with these data, countries can structure their educational policies by seeing their strengths and weaknesses. Therefore, it can be said that participating in these exams at the international level is important not only in terms of providing a global perspective but also in identifying the deficiencies in the education process.

In international large-scale exams that include a mathematics subtest, such as PISA, students around the age of 15 usually participate, while TIMSS allows participation from two different grade levels, fourth and eighth grade. Therefore, if a change is made in education policies, using TIMSS data may be more effective to see the impact of this change. In other words, if the group that participated in TIMSS fourth grade also participates in the eighth-grade application in later years, it would be more meaningful to measure the effectiveness of education practices that took place during this time. In this context, due to the opportunity to work with a younger age group, it has been decided to use TIMSS items as a data collection tool in this study.

Conceptual framework about TIMSS mathematics cognitive area skills and mathematics curriculum learning areas

TIMSS stands for Trends in International Mathematics and Science Study, which is the world's most comprehensive study project conducted every four years at the international level, as noted above, focusing on the cognitive and affective characteristics of students in mathematics and science at the fourth and eighth grade levels (Işlak & Altıntaş, 2022). It aims to measure not only basic mathematical concepts and computational skills, but also problem-solving and reasoning skills based on routine or non-routine problems in mathematics (Karaca, 2018; Yıldırım, Yıldırım, & Ceylan, 2017). According to Kılıç, Aslan-Tutak, and Ertaş (2014), TIMSS is an exam that aims to measure not only subject knowledge, but also cognitive skills. So, the items in the TIMSS math subtest are prepared according to cognitive levels related to knowing, applying, and reasoning (Mullis, Martin, Ruddock, O'Sullivan, & Preuschoff, 2009).

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

According to the report published by the Ministry of National Education (MEB, 2020), TIMSS mathematics cognitive domains include the ability to express a problem situation mathematically using graphs and symbols, to propose claims for a strategy to be used in problem-solving by modeling problem situations, and to utilize tools such as ruler and calculator. Each learning area consists of questions designed for each of the three cognitive domains. The distribution of questions follows a pattern of 40% for knowing, 40% for applying, and 20% for reasoning (Mullis & Martin, 2017). The knowing level encompasses the knowledge, concepts, and processes that students are expected to possess. The applying level involves the application of acquired knowledge in appropriate contexts. The highest level, reasoning, entails the logical thinking processes students engage in for non-routine problem situations (MEB, 2020).TIMSS mathematics subtest consists of the following content areas that form the cognitive domains: for the knowing domain, it includes recall, recognition/discrimination, classification/ordering, computation, retrieving/reading information, and measurement; for the applying domain, it includes identification/decision making, presentation/modeling, and application; and for the reasoning domain, it includes analysis, synthesis, evaluation, inference, generalization, and verification (MEB, 2020). Detailed information is presented in Table 1 below.

Table 1.
Content areas forming the cognitive domains in the TIMSS mathematics subtest (MEB, 2020)

Cognative Domain	**Subject Fields**	**Content**
Knowing	Recall	Remembering definitions, terminology, number properties, units of measurement, geometric properties, and formulas (e.g., axb=ab, a+a+a=3a).
	Recognition/Discrimination	Distinguishing numbers, expressions, quantities, and shapes, distinguishing mathematically equivalent entities (e.g., equivalent fractions, decimal numbers and percentages, different positions of simple geometric shapes).
	Classification/Ordering	Classifying numbers, expressions, quantities, and shapes based on their common characteristics.
	Computation	Using algorithmic methods for addition, subtraction, multiplication, and division, as well as their combinations with natural numbers, fractions, decimal numbers, and integers. Applying basic algebraic processes.
	Retrieving/Reading Information	Understanding information from graphs, tables, texts, and other sources.
	Measurement	Using measurement tools and selecting appropriate units of measurement.
Applying	Determination/Decision Making	Determining effective/appropriate operations, strategies, and tools for solving common problem-solving methods.
	Presentation/Modeling	Representing data with tables or graphs, creating diagrams for equations, inequalities, geometric shapes, and problem

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

		situations, and generating equivalent representations of mathematical relationships.
	Application	Applying strategies to solve problems that involve mathematical concepts and procedures.
Reasoning	Analysis	Determining, defining, and utilizing the relationships between numbers, expressions, quantities, and shapes.
	Synthesis	Making connections between information, relevant representations, and different elements of procedures to solve problems.
	Evaluation	Evaluating alternative problem-solving strategies and solutions.
	Drawing Conclusions	Making valid inferences based on information and evidence.
	Generalization	Establishing statements that demonstrate relationships in more general and broad applicable conditions.
	Verification	Presenting mathematical claims to support a strategy or a solution.

In this context, TIMSS has indicated that 40% of the items in the mathematics subtests for the fourth-grade level are at the knowing level, 40% are at the applying level, and 20% are at the reasoning level (MEB, 2020). Additionally, 50% of the items in the related tests are from the number learning area (25% natural numbers - 15% expressions, simple equations, and relationships - 10% fractions and decimal representations), 30% are from the measurement and geometry learning area (15% geometry - 15% measurement), and 20% are from the data learning area (15% data reading, interpretation, and representation - 5% using data to solve problems) (MEB, 2020).

Turkey first participated in TIMSS in 1999. Table 2 below presents information on which grade levels Turkey participated in TIMSS in which years.

Table 2.
Years Turkey Participated in TIMSS (MEB, TIMSS 2019 Turkey Report, 2020)

TIMSS application years	**Grade Levels Attended**	
	fourth grade level	**eighth grade level**
1995		
1999		✓
2003		
2007		✓
2011	✓	✓
2015	✓	✓
2019	✓ (fifth grade)	✓

As can be seen from the table above (Table 2), Turkey participated in TIMSS at both grade levels in the last three application. However, for the 2019 application, the Ministry of Education (MEB) decided that due to age averages and compatibility with the framework of the exam, it was more

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

appropriate for TIMSS to be applied to fifth and eighth grades for Turkey instead of fourth and eighth grades . Therefore, Turkey participated in the TIMSS 2019 administration with fifth and eighth grade levels. Because the Ministry of National Education (MEB, 2020) has decided that it is more appropriate to apply TIMSS in 5th and 8th grades instead of 4th and 8th grades in terms of both average age and compliance with the framework of the exam.

There are various studies in the literature that focus on TIMSS mathematics cognitive domain skills and levels. For instance, Gündüz and Çakan (2020) examined the learning outcomes in the middle school mathematics curriculum based on TIMSS cognitive domain skills, considering both grade levels and learning outcomes. The findings revealed that the cognitive levels of the learning outcomes varied across grades, with the highest proportion of knowing level outcomes observed in the 5th grade. Delil, Özcan, and Işlak (2020) examined the learning outcomes in the primary school mathematics curriculum from the year 2018, based on TIMSS 2019 cognitive domain skills, considering grade levels and learning domains. According to the results, out of a total of 229 learning outcomes across all grade levels, 58% were at the knowing level, 32% were at the applying level, and 10% were at the reasoning level. However, when comparing the distribution of cognitive levels in the TIMSS mathematics items, there was inconsistency as the learning outcomes in the 4th-grade mathematics curriculum comprised 53% knowing, 34% applying, and 13% reasoning level outcomes. In addition to studies comparing mathematics curriculum with TIMSS cognitive domain skills, there are also studies that analyze exam questions from teacher-made tests or large-scale national exams in the context of TIMSS cognitive domain skills. For example, Taştekinoğlu and Aydın (2014) aimed to compare the cognitive level distributions of teacher-made mathematics exam questions at the 4th-grade level with the mathematics curriculum and TIMSS 2011 questions. Their study found that the cognitive level distributions of the TIMSS questions were not consistent with the cognitive level distributions of the teacher-made mathematics exam questions. Specifically, while 40% of the TIMSS 2011 exam questions were at the knowing level, 40% were at the applying level, and 20% were at the reasoning level, the teacher-made test questions included in the study were observed to be 67% at the knowing level, 18% at the applying level, and 15% at the reasoning level. Yolcu, Tetik, and Delil (2015), in their study, examined the mathematics questions from the central measurement and evaluation exams administered by the Ministry of National Education (MEB) between 1998 and 2015 based on the TIMSS 2015 cognitive level categories. The results of their study indicated that in the central measurement and evaluation exams administered by the MEB for 8th-grade students between 1998 and 2015, the highest number of questions were at the applying level, followed by knowing and reasoning levels. However, it was noted that there was no example of a reasoning level question in 2008. Based on these findings, it can be observed that both the teacher-made mathematics tests and the large-scale tests administered by the MEB had question distributions in the mathematics subtests that were inconsistent with the cognitive level distributions specified by TIMSS. This inconsistency can be interpreted as a possible reason for the low achievement of our students in TIMSS.

Interesting results have also been obtained in comparative studies conducted among countries. For example, in the TIMSS 2011 implementation, Singapore achieved the highest performance in mathematics at the fourth-grade level, while Yemen exhibited the lowest performance. It was determined that both countries shared the highest achievement in the domain of number learning (Büyüköztürk, Çakan, Tan, & Atar, 2014). Similarly, according to the TIMSS 2015

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

results, fourth-grade Turkish students showed the highest average performance in the domain of numbers (Yıldırım, Özgürlük, Parlak, Gönen, & Polat, 2016). However, in the TIMSS 2019 data, this pattern showed some variability. According to the 2019 data, Turkish students at the fourth-grade level were more successful in the domains of measurement and geometry, as well as in numbers, compared to the domain of data (MEB, 2020).

When comparing cognitive level performances, it has been observed that Turkish fourth-grade students obtained the highest average score at the knowing level, followed by the applying level, and finally the reasoning level in the TIMSS 2011 and 2015 mathematics subtests (Büyüköztürk, Çakan, Tan, & Atar, 2014; Yıldırım, Özgürlük, Parlak, Gönen, & Polat, 2016). On the other hand, Yemeni students, who demonstrated the lowest mathematics achievement in TIMSS 2011, showed higher performance in the reasoning level for all three learning areas (Büyüköztürk, Çakan, Tan, & Atar, 2014). In the TIMSS 2019 implementation, fifth-grade Turkish students exhibited a higher average performance in the applying level items compared to others (MEB, 2020). Shanmugam (2015) compared the scores obtained by eighth-grade students in the TIMSS 2011 mathematics subtest in Indonesia, Malaysia, Singapore, and Thailand based on the cognitive levels of the items. The study found that the majority of students in Indonesia and Thailand performed better in items requiring higher-level thinking (reasoning level) compared to items requiring lower-level thinking (knowing and applying levels).

Studies that examine gender, cognitive domains, and learning domains together have reached different results. For example, in the Turkish data of the TIMSS 2011, it was found that male students were more successful in the numbers learning area, while female students were more successful in the geometric shapes and measurement, and data representation areas (Büyüköztürk, Çakan, Tan, & Atar, 2014). When the same data were examined based on cognitive levels, it was found that female students were more successful at the knowing and reasoning levels, while male students were more successful at the applying level (Büyüköztürk, Çakan, Tan, & Atar, 2014). Additionally, according to the TIMSS 2015 and 2019 results, Turkish fourth-grade students did not show a significant difference in total scores based on gender (MEB, 2020; Yıldırım, Özgürlük, Parlak, Gönen, & Polat, 2016). However, in a study by Dindyal (2008) on TIMSS 2003 data from eighth-grade students in Chinese Taipei, Hong Kong-SAR, Indonesia, Japan, Republic of Korea, Malaysia, Singapore, and the Philippines, it was noted that females exhibited higher performance in mathematics.

Importance of the study

For many years, TIMSS, which has been conducted in very large samples, has naturally been the subject of various research studies. In this section, brief mention is made of some of the research studies focused on the mathematics subtest. For example, when the international literature is examined, it is observed that studies have been conducted to examine the relationship between academic achievement in school mathematics and achievement in the TIMSS mathematics subtest (Wiberg, 2019) and the factors that affect students' achievement in the TIMSS mathematics subtest (Dodeen, Abdelfattah, Shumrani, & Hilal, 2012). Additionally, there are studies that investigate the gender factor in TIMSS mathematics subtest achievement (Cheng & Seng, 2001) and whether the mathematics curricula implemented in countries are compatible with the content of the TIMSS

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

mathematics subtest and students' achievement in TIMSS (Anamuah-Mensah & Mereku, 2005; Kuiper, Bos, & Plomp, 1999).

When examining the researches conducted on the TIMSS mathematics subtest in our country, studies examining the compatibility of examples in math textbooks or the achievements in math curricula with the learning areas and cognitive levels in TIMSS mathematics tests are among the leading studies (Delil, Özcan, & Işlak, 2020; Gündüz & Çakan, 2020; Güner, 2015; İncikabı, Mercimek, Ayanoğlu, Aliustaoğlu, & Tekin, 2016; Kılıç, Aslan Tutak, & Ertaş, 2014; Taşpınar-Şener & Bulut, 2022; Toptaş, Elkatmış, & Karaca, 2012). In addition, studies comparing math exam questions (Taştekinoğlu & Aydın, 2014) or questions in the math subtests of large-scale national exams with questions in the TIMSS mathematics test (Yolcu Tetik & Delil, 2015) are also among these studies. In addition, there are studies in the literature that compare Turkish students' scores on the TIMSS mathematics test with those of students from other countries (Abazoğlu, Yatağan, Yıldızhan, Arifoğlu, & Umurhan, 2015) and examine the factors that affect students' TIMSS mathematics achievement (Akyüz, 2014; Ertürk & Erdinç Akan, 2018; Ölçüoğlu & Çetin, 2016; Sarı, Arıkan, & Yıldızlı, 2017). Moreover, Okudan and Yeşilyurt (2021) administered an achievement test consisting of items from the TIMSS mathematics subtests to a sample of eighth grade students from the Erzurum region and found low academic achievement in mathematics. However, the learning areas and cognitive level distributions of the items in the achievement test used in that study to examine students' academic achievement did not take into account the proportions recommended by TIMSS.

In this study, an achievement test consisting of items from TIMSS mathematics subtests was used, and this achievement test was administered to a sample of fifth grade students (due to the fact that the 2019 TIMSS application in our country was conducted with fifth graders instead of fourth graders) attending a medium-sized provincial center in the Central Anatolia Region. Since TIMSS is not administered in every school, it may not be possible to see situations related to smaller areas in such large-scale tests. Therefore, this study provides an opportunity to examine the reflection of an internationally conducted test in a medium-sized city using data obtained from samples selected from schools in the location where the research is conducted. In addition to that, one of the important factors that make this study valuable is that the proportions prescribed by TIMSS for both the learning areas and the cognitive level distributions of the items included in the achievement test were taken into consideration. It can be said that examining the compatibility of the information obtained by applying a data collection tool with these characteristics in a certain location with the data obtained at the national level also contributes to the literature. In addition, the fact that the students' math performances were examined in terms of both cognitive level and learning area in this study increases its importance, as no similar research has been found in the literature. Based on this, the research question of this study is "Do the scores of fifth grade students attending state schools in a medium-sized city center in the Central Anatolia region on an achievement test prepared from the TIMSS mathematics subtest items differ according to the cognitive levels, learning areas, and genders of the students?" Two sub-problems were addressed in the scope of this research problem, which are presented below:

- In which learning areas and at which cognitive levels are students more successful?
- Does scores obtained from different cognitive levels and learning areas differ according to gender?

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

Method

Research design

The research design adopted for this study is a quantitative research method called correlational design, which is used to investigate relationships between multiple variables (Can, 2013). In this study, the scores of students on a TIMSS mathematics achievement test, the learning areas and cognitive levels of the test items, and the students' gender were examined using correlational design.

Population and sample

The population of the study consists of fifth grade students in state schools located in the center of a medium-sized city in the Central Anatolian region. According to Karasar (2007), cluster sampling method, which is an unbiased sampling method, is used to select clusters from the population to conduct the study. In this context, each of the state middle schools in the related city center was determined as a cluster, and a sufficient number of students were randomly selected from these clusters to obtain the required sample size. In this regard, a total of 487 students participated in the study, with 154 students in the pilot study and 333 students in the main study for the test and item statistics of the data collection tool.

Data collection tool

In the study, multiple-choice items from the fourth grade TIMSS mathematics subtests, which were publicly available by the Ministry of Education, were initially included in the pool to prepare the data collection tool. It was stated that there were 20 to 25 items in each TIMSS subtest for fourth grade students, and 36 minutes were given for each subtest (MEB, 2012). Therefore, it was aimed to have around 20 items in the final version of the data collection tool to be used in this study, and a draft form consisting of 30 items was prepared for the initial version of the data collection tool in line with this aim. In order to ensure the content validity of the data collection tool, the distribution of learning areas and items published by MEB (2020) for fifth graders in the 2019 TIMSS application was considered. With the inclusion of cognitive levels, it was decided that 6 out of the 15 items in the numbers and operations learning area should be at the knowing level, 6 at the applying level, and 3 at the reasoning level. For the geometry and measurement learning area, it was considered that 3 out of the 8 items should be at the knowing level, 3 at the applying level, and 2 at the reasoning level. Finally, it was decided that 3 out of the 7 items selected for the data learning area should be at the knowing level, 3

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

at the applying level, and 1 at the reasoning level, and the distribution of the items in terms of learning areas, and cognitive levels in the draft form is presented in Table 3.

Table 3.
Distribution of items included in the draft form by learning areas, sub-areas and cognitive levels

Learning area	**Cognative Level**			**Total**
	Knowing	**Applying**	**Reasoning**	
Numbers and Operations	6	6	3	15
Geometry and Measurement	3	3	2	8
Data	3	3	1	7
Total	12	12	6	30

On the other hand, since the learning areas and cognitive levels of TIMSS items made available by MEB are already disclosed by MEB, expert opinion was not consulted in this scope. According to Büyüköztürk (2012), validity is related to the ability of the items in the test to measure the intended behavior both quantitatively and qualitatively. In this context, expert opinion was consulted on the suitability of the draft data collection tool as a qualitative aspect and necessary arrangements were made (Tavşancıl, 2002).

A pilot study was conducted with 154 fifth grade students for the draft data collection tool that have 30 items. According to Tavşancıl (2002), it is appropriate to work with a sample size that is five times the number of items for item analysis. In the pilot application conducted within one lesson hour with 154 students, coding was performed by giving 1 point to each correct answer and 0 points to each incorrect answer. After coding, item statistics were examined and items with discrimination index below 0.30 were discarded. According to Büyüköztürk (2012), if an item's discrimination index is below 0.30, that item needs to be revised or removed from the data collection tool. Moreover, Hasançelebi, Terzi, and Küçük (2020) emphasized that the average difficulty level of a test should be around 0.50. Therefore, items that were found to disrupt this balance after the pilot study were excluded. The distribution of the items considered suitable for inclusion in the final data collection tool, based on the pilot study's item analysis, according to the learning areas and cognitive levels, is presented in Table 4.

Table 4.
Distribution of items in the final version of the data collection tool

Learning area	**Cognative Level**			**Total**
	Knowing	**Applying**	**Reasoning**	
Numbers	4	4	2	10
Geometry and Measurement	2	2	1	5
Data	1	2	1	4
Total	7	8	4	19

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

The total of 19 items in the final data collection tool had an average difficulty index of 0.61 and an average discrimination index of 0.60. The KR-20 reliability coefficient of the test was calculated as 0.86. Based on these values, it was concluded that the data collection tool was valid, reliable, and suitable for use.

Data analysis

After the pilot study, in the actual study conducted with 333 students, total scores were obtained for each student by giving 1 point for each correct answer and 0 points for each incorrect answer. Calculation of total scores was performed separately for each cognitive level and learning area group.

For the first sub-problem of the study, first, the arithmetic mean of the scores obtained by the students from the items in each learning area and cognitive level group was calculated. As there are different numbers of items in each learning area and cognitive level group, the maximum possible scores also differ. This situation prevents a healthy comparison. Therefore, the total scores obtained from each learning area and cognitive level group were transformed so that they are equal to a maximum of 100, making the averages comparable.

The normality of the scores obtained by the students from each cognitive level and learning area was examined for the second sub-problem of the research. In this context, the Shapiro-Wilk W Test was used for normality analyses, and the corresponding values are presented in Table 5.

Table 5.
Results of normality analysis for scores in learning areas and cognitive levels

Learning area	**Cognative Level**	***df***	***sd***	***skewness***	***kurtosis***	***p***
Numbers and Operations	Knowing	333	32.73	-0.45	-0.91	0.00
	Applying		32.28	-0.51	-0.87	0.00
	Reasoning		36.80	-0.78	-0.74	0.00
Geometry and Measurement	Knowing	333	35.96	0.40	-0.99	0.00
	Applying		30.67	-0.14	-0.50	0.00
	Reasoning		39.46	-1.56	0.46	0.00
Data	Knowing	333	49.79	0.21	-1.96	0.00
	Applying		34.61	-0.78	-0.59	0.00
	Reasoning		32.20	-2.39	3.74	0.00

When the values in Table 5 are examined, although some of the total scores obtained for each learning area and cognitive level group have acceptable skewness and kurtosis values, it is clear that all of the obtained total scores do not follow a normal distribution ($p=0.00<0.05$). Since the scores do

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

not show a normal distribution, Mann-Whitney U Test was used for the comparison of means according to gender variable for the second sub-problem of the research (Can, 2013).

Findings

In which learning areas and cognitive levels are students more successful?

The data collection tool used in the study includes a total of 10 items in the numbers and operations learning area, with 4 items at the knowing level, 4 items at the applying level, and 2 items at the reasoning level. There is a total of 5 items in the geometry learning area, with 2 items at the knowing level, 2 items at the applying level, and 1 item at the reasoning level. Finally, there is a total of 4 items in the data processing learning area, with 1 item at the knowing level, 2 items at the applying level, and 1 item at the reasoning level. Since the number of items in each learning area and cognitive level is different, the averages for each level were first converted into a hundred-point system in order to compare the possible scores. Descriptive statistics obtained after the conversion are presented in Table 6.

Table 6.
Descriptive statistics of scores for learning areas and cognitive levels

Learning area	***n***	***X***	***Median***	**Cognative Level**	***X***	***Median***
Numbers and operations	333	34.04	36.84	Knowing	62.53	75
				Applying	64.18	75
				Reasoning	69.97	100
Geometry and Measurement	333	55.19	60	Knowing	37.68	50
				Applying	59.90	50
				Reasoning	80.78	100
Data	333	68.78	75	Knowing	44.74	0
				Applying	71.02	100
				Reasoning	88.28	100

As seen in Table 6, the average scores increase as the cognitive level of the items in each learning area increases. In other words, students have a lower average score in the most basic cognitive level, which is the knowing level, and a higher average score in the next cognitive level, which is the applying level, compared to the knowing level items. Similarly, students have a lower average score in the applying level items that are at a lower cognitive level than the reasoning level items.

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

However, as stated in the data analysis section, the scores obtained by cognitive level and learning area do not show a normal distribution. In such cases, it is suggested that the median is more appropriate than the arithmetic mean for comparing success (Sevgi & Çağlıköse, 2020). From this perspective, when data were evaluated according to learning areas, students showed the highest performance in data, followed by geometry and measurement learning areas. The lowest achievement was observed in the numbers and operations learning area. When performance was examined by cognitive level along with learning areas, it was observed that students' achievement in the knowledge and applying levels in the numbers and operations learning area was lower than their achievement at the reasoning level. A similar situation is valid for the geometry learning area. As for the data processing learning area, the lowest achievement was observed at the knowing level, while a higher and equal achievement was observed at the application and reasoning levels.

Does scores obtained from different cognitive levels and learning areas show differences according to gender?

In order to compare the relevant data, Mann-Whitney U test was performed to see if there is a difference by gender. The results of the analysis are presented in Table 7.

Table 7.
Results of comparison of students' scores by gender

Learning area	**Cognative Level**	**Gender**	***n***	***X***	***sd***	***p***
Numbers and operations	Knowing	F	165	63.03	32.38	0.81
		M	168	62.05	33.16	
	Applying	F	165	61.06	30.90	0.03*
		M	168	67.26	33.39	
	Reasoning	F	165	69.09	37.23	0.67
		M	168	70.83	36.46	
Geometry and Measurement	Knowing	F	165	30.90	33.34	0.00*
		M	168	44.34	37.28	
	Applying	F	165	56.36	29.81	0.03*
		M	168	63.39	31.18	
	Reasoning	F	165	83.63	37.10	0.19
		M	168	77.97	41.56	
Data	Knowing	F	165	44.24	49.81	0.85
		M	168	45.23	49.92	
	Applying	F	165	71.81	33.26	0.84
		M	168	70.23	35.97	
	Reasoning	F	165	91.51	27.95	0.07

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

	M	168	85.11	35.69	

* p<.05

The values in Table 7 show that there are significant differences between male and female students in the scores of the applying level of the numbers and operations learning area ($p=0.03<0.05$), the knowing level of the geometry learning area ($p=0.00<0.05$), and the applying level of the geometry learning area ($p=0.03<0.05$). In this context, it was observed that male students were more successful than female students in the three score groups where significant differences were observed. However, no significant differences were found between male and female students in the scores of other learning areas and cognitive level groups.

Discussion, Conclusion, and Suggestions

When the data obtained within the framework of the first sub-problem is examined on a learning area basis, it can be stated that the students performed the best in the data, followed by geometry and measurement learning areas. The lowest success was observed in the number and operations learning area. This finding is similar to the TIMSS 2011 results for Turkey but does not match the TIMSS 2015 findings. When the TIMSS 2011 reports are examined, it can be seen that the highest value of the fourth-grade Turkish students' mathematics subtest achievement averages is in the data learning area with an average of 478, followed by the numbers and operations learning area with an average of 477, and the lowest average of 447 is in the geometry and measurement learning area (Büyüköztürk, Çakan, Tan, & Atar, 2014). In other words, according to the TIMSS 2011 reports, fourth grade Turkish students have the highest and very close averages in the data learning area and numbers and operations learning area. On the other hand, the results obtained from this study contradict the situation of students in Singapore and Yemen under TIMSS 2011. In the TIMSS 2011 exam, it was observed that students in Singapore, who showed the highest achievement in mathematics at the fourth-grade level, had the highest achievement in the numbers and operations learning area, while students in Yemen, who showed the lowest achievement, had the highest achievement in the numbers and operations learning area. (Büyüköztürk, Çakan, Tan, & Atar, 2014). However, this result also contradicts the results of the TIMSS 2015 exam. According to the TIMSS 2015 results, fourth grade Turkish students, who had an average score of 489 in the numbers field, 475 in the geometry learning field, and 476 in the data representation learning field, showed the highest average in the numbers and operations learning area (Yıldırım, Özgürlük, Parlak, Gönen, & Polat, 2016). In the TIMSS 2019 data, this situation has shown some variability. According to the 2019 data, Turkish students at the fourth-grade level scored 525 in the numbers field, 527 in the measurement and geometry field, and 510 in the data field, thus being more successful in the measurement and geometry field and the numbers field than in the data field (MEB, 2020).

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

According to the data obtained, when learning areas and cognitive levels were evaluated together, students showed the highest performance at the highest cognitive level, which is reasoning level, in all three learning areas. There are both similar and opposite results in the literature for this finding. For example, according to TIMSS 2011 and 2015 reports, Turkish fourth grade students received the highest average scores for the knowing level, followed by the applying level, and finally the reasoning level for the items in the math subtest (Büyüköztürk, Çakan, Tan, & Atar, 2014; Yıldırım, Özgürlük, Parlak, Gönen, & Polat, 2016). On the other hand, for Yemen, which showed the lowest math achievement in TIMSS 2011, the situation is different, and Yemeni students have shown more success at the reasoning level in all three learning areas (Büyüköztürk, Çakan, Tan, & Atar, 2014). The results obtained in this study are similar to the situation in Yemen. In the TIMSS 2019 application, fifth grade Turkish students showed a higher average in items requiring applying level compared to others, with 514 points in the knowing level, 531 points in the applying level, and 509 points in the reasoning level (MEB, 2020). Shanmugam (2015) compared the scores obtained based on the cognitive levels of items in the TIMSS 2011 mathematics subtest in four Southeast Asian countries (Indonesia, Malaysia, Singapore, and Thailand) and stated that items at the knowing and applying levels require lower-order thinking skills (LOTS) while items at the reasoning level require higher-order thinking skills (HOTS) in his study conducted with data from eighth grade students. The findings from his study indicate that the majority of students in Indonesia and Thailand showed a higher performance in HOTS items than in LOTS items. In Malaysia, students generally showed higher performance in LOTS items. On the other hand, in Singapore, students in the bottom 25% performed better in HOTS items, while others performed better in LOTS items.

In addition, within the scope of this study, it was concluded that students' achievement levels in the knowing level for all three learning areas were lower, compared to other levels. The fact that students showed the lowest achievement in the knowing level is similar to the findings of Kılıç, Aslan-Tutak, and Ertaş (2014). Researchers who examined the eighth grade TIMSS mathematics subtest for 2011 stated that Turkish students showed their lowest achievement in the items related to the knowing level of seventh grade objectives. Furthermore, when the data obtained within this study is evaluated according to the learning area, it is observed that students' achievements at the knowing and applying levels in the numbers and operations domain are lower than their achievements at the reasoning level. While this finding is similar to the findings obtained by Kilic, Aslan-Tutak, and Ertas (2014) for the numbers and operations learning area in some respects, it differs in some other respects. They stated that out of the math items with low mean scores, 8 were at the knowing level, 3 were at the applying level, and 3 were at the reasoning level. In this context, in the study, students showed the lowest average at the knowing level in the numbers and operations learning area, while they showed slightly higher averages at the applying and reasoning levels. Having the lowest average at the knowing level and a high average at the reasoning level in terms of achievement shows similarity between both studies, while the success at the applying level differs between the two studies.

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

According to the findings obtained for the second sub-problem of the study, male students have significantly higher averages than female students in the applying level of the numbers and operations learning area and in the knowing and applying levels of the geometry and measurement learning area. This result overlaps with some of the findings in the literature, while differing from others. For example, in the Turkey data of the TIMSS 2011 application, it was found that male students had higher averages than female students in the numbers and operations learning area, while female students were more successful than male students in other areas such as geometry and measurements, and data (Büyüköztürk, Çakan, Tan, & Atar, 2014). When the same data was examined in terms of cognitive levels, it was found that female students were more successful than male students at the knowing and reasoning levels, while male students were more successful at the applying level (Büyüköztürk, Çakan, Tan, & Atar, 2014). However, according to both TIMSS 2015 and TIMSS 2019 results, although the difference in total score at the fourth-grade level is not statistically significant, male students have exhibited a higher mathematics performance than female students (MEB, 2020; Yıldırım, Özgürlük, Parlak, Gönen, & Polat, 2016). This situation contradicts the findings of Dindyal (2008). Dindyal (2008) examined the mathematics performance of eighth grade students by gender in eight countries (Chinese Taipei, Hong Kong SAR, Indonesia, Japan, Republic of Korea, Malaysia, Singapore, and the Philippines) in the TIMSS 2003 application and stated that the mathematics performance of the eight countries in the study reflected the international average by gender (except Japan and Republic of Korea), and that girls performed better.

In conclusion, the findings obtained in this study, as well as those in the literature, show both parallel and opposite results. The reasons for this difference may be many factors, such as the culture, socioeconomic level of the region where the data was collected, students' affective characteristics towards mathematics, and teachers' approaches to the curriculum. Being able to describe the performance displayed by students in a small location in a success test prepared by compiling items from the math subtest of international exams such as TIMSS is important in terms of being able to see the reflection of the situation at the national and international level. Therefore, studies to be conducted with data collected in different locations, such as this study, can make it possible to make comparisons over a wider range. Because it is known that the results obtained from international exams such as TIMSS affect educational policies (Aydın, 2017; MEB, 2016). In addition, given the low performance of students in the knowledge-level items of all learning areas in this study, it is recommended that the content related to knowing level be enriched in both textbooks and in-class applications.

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

References

Abazoğlu, İ., Yatağan, M., Yıldızhan, Y., Arifoğlu, A., & Umurhan, H. (2015). Öğrencilerin matematik başarısının uluslararası fen ve matematik eğilimleri araştırması sonuçlarına göre değerlendirilmesi. *Turkish Studies, 10*(7), 33-50.

Akyüz, G. (2014). TIMSS 2011'de öğrenci ve okul faktörlerinin matematik başarısına etkisi. *Eğitim ve Bilim, 39*(172), 150-162.

Anamuah-Mensah, J., & Mereku, D. K. (2005). Ghanaian JSS2 students' abysmal mathematics achievement in TIMSS 2003: A consequence of the basic school mathematics curriculum. *Mathematics Connection, 5*(1), 1-13.

Aydın, M. (2017). TIMSS 2011 öğretmen ölçeklerinin faktör yapısının ve psikometrik özelliklerinin incelenmesi. Marmara Üniversitesi Atatürk Eğitim Fakültesi Eğitim Bilimleri Dergisi, 46(46), 21-36.

Büyüköztürk, Ş. (2012). *Sosyal bilimler için veri analizi el kitabı.* Ankara: Pegem Akademi.

Büyüköztürk, Ş., Çakan, M., Tan, Ş., & Atar, H. Y. (2014). *TIMSS uluslararası matematik ve fen eğilimleri araştırması. TIMSS 2011 ulusal matematik ve fen raporu 4. sınıflar*. Ankara: Milli Eğitim Bakanlığı.

Can, A. (2013). *SPSS ile bilimsel araştırma sürecinde nicel veri analizi.* Ankara: Pegem Akademi Yayınları.

Cheng, S. K., & Seng, Q. K. (2001). Sex differences in TIMSS mathematics achievement of four Asian nations. *Studies in Educational Evaluation, 27*(4), 331-40.

Delil, A., Özcan, B. N., & Işlak, O. (2020). İlkokul matematik dersi öğretim programı kazanımlarının TIMSS-2019 değerlendirme çerçevesine göre analizi. *Celal Bayar Üniversitesi Sosyal Bilimler Dergisi, 18*(1), 280-282. https://doi.org/10.18026/cbayarsos.669086

Dindyal, J. (2008). An overview of the gender factor in mathematics in TIMSS-2003 for the Asia-Pacific region. *ZDM, 40*, 993-1005. doi:10.1007/s11858-008-0111-2

Dodeen, H., Abdelfattah, F., Shumrani, S., & Hilal, M. A. (2012). The effects of teachers' qualifications, practices, and perceptions on student achievement in TIMSS mathematics: A comparison of two countries. *International Journal of Testing, 12*(1), 61-77. https://doi.org/10.1080/15305058.2011.621568

Ertürk, Z., & Erdinç Akan, O. (2018). TIMSS 2015 matematik başarısını etkileyen değişkenlerin yapısal eşitlik modeli ile incelenmesi. *Ulusal Eğitim Akademisi Dergisi, 2*(2), 14-34. https://doi.org/10.32960/uead.407078

Gündüz, T., & Çakan, M. (2020). The classification of students based on mathematics achievement test in TIMSS 2015 Turkey sample by cognitive diagnosis models. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi, 15*(4), 1149-1163. doi: 10.17522/balikesirnef.741401

Güner, N. (2015). 6.-8. sınıf matematik ders kitaplarındaki geometri, veri ve olasılık sorularının TIMSS bilişsel düzeylerine göre sınıflandırılması. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 77-90.

Hasançelebi, B., Terzi, Y., & Küçük, Z. (2020). Madde güçlük indeksi ve madde ayırt edicilik indeksine dayalı çeldirici analizi. *Gümüşhane Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 10*(1), 224-240. https://doi.org/10.17714/gumusfenbil.615465

İncikabı, L., Mercimek, O., Ayanoğlu, P., Aliustaoğlu, F., & Tekin, N. (2016). Ortaokul matematik dersi öğretim programı kazanımlarının TIMSS bilişsel alanlarına göre değerlendirilmesi. *İlköğretim Online, 15*(4), s. 1149-1163. https://doi.org/10.17051/io.2016.54792

Işlak, O., & Altıntaş, G. (2022). İlkokul 4. sınıf ders kitaplarındaki soruların TIMSS 2019 bilişsel alanlarına göre analizi. *Manisa Celal Bayar Üniversitesi Sosyal Bilimler Dergisi, 20*(1), 172-184. https://doi.org/10.18026/cbayarsos.1016386

Karaca, F. (2018). *Sekizinci sınıf öğrencilerinin TIMSS matematik başarılarının bazı değişkenler açısından incelenmesi: Eskişehir ili örneği* (*Unpublished master's thesis*). Eskişehir: Eskişehir Osmangazi Üniversitesi.

Karasar, N. (2007). *Bilimsel araştırma yöntemi: Kavramlar, ilkeler, teknikler.* Ankara: Nobel Yayın Dağıtım.

Kılıç, H., Aslan Tutak, F., & Ertaş, G. (2014). TIMSS merceğiyle ortaokul matematik öğretim programındaki değişiklikler. *Mersin Üniversitesi Eğitim Fakültesi Dergisi, 10*(2), 129-141.

Batı Anadolu Eğitim Bilimleri Dergisi, (2023), 14 (2), 949-968.
Western Anatolia Journal of Educational Sciences, (2023), 14 (2), 949-968.
Araştırma Makalesi / Research Paper

Kuiper, W., Bos, K., & Plomp, T. (1999). Mathematics achievement in the Netherlands and appropriateness of the TIMSS mathematics test. *Educational Research and Evaluation, 5*(2), 85-104.

MEB. (2012). *TIMSS 2011 tanıtım kitapçığı* . Ankara: EARGED. http://yegitek.meb.gov.tr/pdf/TIMSS_2011_kitapcigi.pdf adresinden alındı.

MEB (2016).TIMSS 2015 Ulusal Matematik ve Fen Bilimleri Ön Raporu 4. ve 8. Sınıflar. http://timss.meb.gov.tr/wpcontent/uploads/TIMSS_2015_Ulusal_Rapor.pdf adresinden alındı.

MEB. (2020). *TIMSS 2019 Türkiye Raporu. Eğitim analiz ve değerlendirme raporları serisi No: 15.* Ankara: Ölçme, Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü. https://odsgm.meb.gov.tr/meb_iys_dosyalar/2020_12/10175514_TIMSS_2019_Turkiye_On_Raporu_.pdf adresinden alındı

Mullis, I. V.S., & Martin, M. O. (2017). *TIMSS 2019 assessment frameworks*. TIMSS and PIRLS International Study Center. Chestnut Hill, MA: Lynch School of Education, Boston College.

Mullis, I. V., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., & Preuschoff, C. (2009). *TIMSS 2011 assessment frameworks.* Boston: TIMSS & PIRLS International Study Center.

Okudan, Ü., & Yeşilyurt, E. (2021). Sekizinci sınıf öğrencilerinin matematik dersi akademik başarısının değerlendirilmesi: TIMSS soruları örneği. *Disiplinlerarası Eğitim Araştırmaları Dergisi. 5*(10), 139-153.

Ölçüoğlu, R., & Çetin, S. (2016). TIMSS 2011 sekizinci sınıf öğrencilerinin matematik başarısını etkileyen değişkenlerinin bölgelere göre incelenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 7*(1), 202-220. https://doi.org/10.21031/epod.34424

Sarı, M. H., Arıkan, S., & Yıldızlı, H. (2017). 8. sınıf matematik akademik başarısını yordayan faktörler-TIMSS 2015. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi, 8*(3), 246-265.

Sevgi, S., & Çağlıköse, M. (2020). Altıncı sınıf öğrencilerinin üstbiliş becerilerinin bazı değişkenler açısından incelenmesi. *Cumhuriyet Uluslararası Eğitim Dergisi, 9*(1), 139-157.

Shanmugam, S. K. (2015). Eight grade students' mathematics achievement in TIMSS 2011 cognitive domains-a comparison across four Southeast Asian countries. *Southeast Asian Mathematics Education Journal, 5*(1), 3-30.

Taşpınar-Şener, Z., & Bulut, A. S. (2022). 4. ve 8. sınıf matematik ders kitaplarının TIMSS bilişsel alanlarına göre analizi. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi, 23*, 46-83. doi: 10.29299/kefad.999519

Taştekinoğlu, E., & Aydın, G. (2014). 4. sınıf matematik sınav sorularının TİMSS 2011 bilişsel alanları ve öğretim programlarıyla karşılaştırılması. *III. Türkiye Lisansüstü Çalışmalar Kongresi*, Sakarya.

Tavşancıl, E. (2002). *Tutumların ölçülmesi ve spss ile veri analizi.* Ankara: Nobel Yayıncılık.

Toptaş, V., Elkatmış, M., & Karaca, T. (2012). İlköğretim 4. sınıf matematik programının öğrenme alanları ile matematik öğrenci çalışma kitabındaki soruların zihinsel alanlarının TIMSS'e göre incelenmesi. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi, 13*(1), 17-29.

Wiberg, M. (2019). The relationship between TIMSS mathematics achievements, grades, and national test scores. *Education Inquiry, 10*(4), 328-343. https://doi.org/10.1080/20004508.2019.1579626

Yıldırım, A., Özgürlük, B., Parlak, B., Gönen, E., & Polat, M. (2016). *TIMSS uluslararası matematik ve fen eğilimleri araştırması TIMSS 2015 ulusal matematik ve fen bilimleri ön raporu 4. ve 8. sınıflar.* Ankara: MEB Ölçme, Değerlendirme Ve Sınav Hizmetleri.

Yıldırım, H. H., Yıldırım, S., & Ceylan, E. (2017). *Türkiye perspektifinden TIMSS 2015 sonuçları (TEDMEM Analiz Dizisi 4).* Ankara: Türk Eğitim Derneği Yayınları.

Yilmaz, N., Ay, Z., & Aydin, Ş. (2021). An investigation of tasks in the mathematics textbooks and objectives in mathematics curriculum from 4th to 8th grade related with data content domain according to TIMSS 2019 cognitive domains. *Cukurova University Faculty of Education Journal, 50*(2), 1397-1436.

Yolcu Tetik, B., & Delil, A. (2015). 8. Sınıf merkezi sınavlardaki matematik sorularının TIMSS-2015 bilişsel alanlarına göre analizi. *Celal Bayar Üniversitesi Sosyal Bilimler Dergisi, 13*(4), 166-184. doi: 10.18026/cbusos.87313.