



## TIMSS 2015 Matematik Başarısı İle İlgili Bazı Değişkenlerin Cinsiyete Göre Ölçme Değişmezliğinin İncelenmesi \*

### The Investigation of Measurement Invariance of the Variables Related to TIMSS 2015 Mathematics Achievement in terms of Gender

Zafer ERTÜRK\*\*

Oya ERDİNÇ-AKAN\*\*\*

Received: 04 April 2018

Research Article

Accepted: 03 December 2018

**ABSTRACT:** TIMSS is an international study which aims to follow the changes and developments in student achievement and identify the differences between the education systems of the participating countries. In this study, it is aimed to test whether the variables which are effective in mathematics achievement of fourth grade students provide the measurement invariance conditions or not. It is a correlational study and its sample consists of 6456 fourth grade students participated in TIMSS 2015 from Turkey. The Multigroup Confirmatory Factor Analysis (MG-CFA) was used for testing the measurement invariance of the factors related to the student achievement. The configural, metric, scalar and strict invariance tests in terms of genders were done incrementally for liking mathematics, interest in mathematics and confidence in mathematics latent variables in the research. CFI, comparative fit indices, differences between the unconstrained condition (configural invariance) and the conditions in which more restrictions were added incrementally were assessed in the analyses. As a result, it was seen that the three of the latent variables tested provided the configural invariance condition, however only liking mathematics latent variable provided the strict invariance condition in gender groups.

**Keywords:** TIMSS, measurement instability, structural equation model, gender.

**ÖZ:** TIMSS, öğrenci başarısındaki değişim ve gelişimin izlenmesi ve ülkeler arası eğitim sistemi farklılıklarının belirlenmesi amacıyla yapılan uluslararası bir çalışmadır. Bu çalışmada 4. sınıf öğrencilerinin matematik başarısında etkili olan değişkenlerin ölçme değişmezliğinin incelenmesi amaçlanmaktadır. Araştırma ilişkisel bir araştırma olup, örneklemini TIMSS 2015'e Türkiye'den katılan 6456 4. sınıf öğrencisi oluşturmaktadır. Öğrencilerin matematik başarısı ile ilişkili olan değişkenlerin ölçme değişmezliklerinin test edilmesinde Çok Gruplu Doğrulayıcı Faktör Analizi (ÇG-DFA) kullanılmıştır. Araştırmada yer alan matematiği sevme, matematiğe olan ilgi ve matematiğe ilişkin özgüven gizil değişkenleri için cinsiyet açısından aşamalı bir şekilde biçimsel, metrik, skalar ve katı değişmezlik testleri yapılmıştır. Analizlerde hiçbir sınırlamanın yapılmadığı biçimsel değişmezlik durumu ile aşamalı olarak daha fazla sınırlamanın eklendiği diğer değişmezlik testleri arasındaki CFI karşılaştırmalı uyum iyiliği indeks farkları değerlendirilmiştir. Araştırmada test edilen üç gizil değişkenin biçimsel değişmezlik koşulunu yerine getirdiği, ancak cinsiyet gruplarında katı değişmezlik koşulunu sadece "matematiği sevme" gizil değişkeninin sağladığı bulunmuştur.

**Anahtar kelimeler:** TIMSS, ölçme değişmezliği, yapısal eşitlik modeli, cinsiyet.

\*This research was presented at the International Congress on Science and Education (ICSE) organized in Afyonkarahisar on March 23-25, 2018.

\*\*Corresponding Author: Res. Asst., Gazi University, Ankara, Turkey, [zerturk35@gmail.com](mailto:zerturk35@gmail.com)

\*\*\*Res. Asst., Gazi University, Ankara, Turkey, [oyaerdincakan@gmail.com](mailto:oyaerdincakan@gmail.com)

#### Citation Information

Ertürk, Z., & Erdinç-Akan, O. (2018). TIMSS 2015 matematik başarısı ile ilgili bazı değişkenlerin cinsiyete göre ölçme değişmezliğinin incelenmesi. *Kuramsal Eğitim Bilim Dergisi [Journal of Theoretical Educational Science]*, *UBEK-2018*, 204-226

## Giriş

Toplumun yapısını oluşturan kurumların her biri toplumsal ilerlemenin, değişimin, bütünlüğün ya da gereksinimlerin sağlanmasında önemli görevleri yerine getirmektedir. Toplumsal değerlerin yürütülmesinde yer alan kurumlardan birisi de eğitim ve öğretim faaliyetlerinin yapıldığı okullardır. Okulların toplumsal amaçları yerine getirirken bazı dönemlerde sorgulanması, nitelik ve verimlilik yönünden belirlenen birtakım ölçütlere göre çok yönlü incelenmesi ve karşılaştırılması gerekmektedir. Uluslararası düzeyde yapılan sınavlar aracılığıyla bu karşılaştırmalar yapılabilmektedir.

Uluslararası düzeyde yapılan sınavlardan birisi de TIMSS (Uluslararası Matematik ve Fen Eğilimleri Araştırması)'dir. TIMSS, çalışmaya katılan ülkelerdeki öğrencilerin matematik ve fen alanlarında kazandıkları bilgi ve becerilerinin değerlendirildiği bir tarama çalışmasıdır. TIMSS, Uluslararası Eğitim Başarılarını Değerlendirme Kuruluşu (International Association for the Evaluation of Educational Achievement) IEA'nın bir projesidir. TIMSS dört yılda bir yapılmakta olup dördüncü ve sekizinci sınıf düzeyindeki öğrencilere uygulanmaktadır. TIMSS'in amacı öğrenci başarılarındaki değişim ve gelişimleri izlemek ve ülkelerin eğitim sistemleri arasındaki farklılıkları değerlendirmektir.

Türkiye, ilk olarak 1999 yılında yapılan TIMSS-R (The Third International Mathematics and Science Study–Repeat) olarak bilinen Üçüncü Uluslararası Matematik ve Fen Bilimleri Çalışması'na sadece sekizinci sınıf düzeyinde katılmıştır ve matematik başarı testlerinde 429 ortalama puan elde etmiştir (Millî Eğitim Bakanlığı Eğitimi Geliştirme ve Araştırma Dairesi Başkanlığı, 2003). TIMSS 2003'e katılmayan Türkiye TIMSS 2007'ye sekizinci sınıf düzeyinde katılım göstermiştir. Türkiye, ilk kez dördüncü sınıf düzeyinde TIMSS 2011'e katılmıştır. TIMSS'in altıncısı 2015 yılında yapılmış ve Türkiye sekizinci sınıf düzeyinde dördüncü kez katılırken, dördüncü sınıf düzeyinde ikinci kez bu çalışmaya katılmıştır. Dört yılda bir düzenlenen TIMSS'e Türkiye ile birlikte dünyanın birçok bölgesinden ülkede katılmaktadır. Yapılan birçok araştırmada uluslararası sınavlara katılan ülkelerin sınavlardan elde ettikleri sonuçlar çeşitli değişkenlere göre karşılaştırılmış ve başarıya etki eden faktörler belirlenmeye çalışılmıştır (Eurydice, 2010; Türkan, Üner, & Alıcı, 2015; Yılmaz & Aztekin, 2012).

Karşılaştırmaya dayalı yapılan birçok çalışmada elde edilen sonuçlar bireylerin cinsiyet, sosyo-ekonomik düzey, sınıf düzeyi gibi farklı özellikler taşıması nedeni ile anlamlı farklılıklar oluşturacaktır. Oluşan farklılığın sadece bireylerin sahip olduğu özelliklere bağlanması ve bu şekilde açıklanması her zaman doğru değildir. Ortaya çıkan farklılık çalışmada kullanılan ölçme aracının kendisinden de kaynaklanıyor olabilir. Farklı kültürlerde uygulanan bir matematik tutum ölçeğinden elde edilen puanlar, farklı kültür gruplarında yer alan bireylerin maddeleri yorumlama biçimlerine göre farklılık gösterecektir. Aslında ölçme aracı oluşturulurken her grupta aynı özelliği ölçebileceği varsayımı ile hazırlanmaktadır. Dolayısıyla bu varsayım test edilmeden yapılan cinsiyet, sosyo-ekonomik düzey ve kültürler arası karşılaştırma çalışmalarında yanlış sonuçlar ortaya çıkabilecektir.

Bir araştırmada kullanılan ölçeğin faktör yapısının, ölçeğin uygulandığı farklı gruplarda aynı yapıya sahip olması ölçek maddelerinin faktör yüklerinin, faktörler arası korelasyonlarının ve hata varyanslarının aynı olması anlamına gelir (Byrne, 1998; Jöreskog & Sörbom, 1993). Ölçeğin faktör yapısı için yapılan değişmezlik çalışmaları

ile ölçek yeniden geliştirilmez. Bollen'e (1989) göre, ölçümlerin değişmezliği gruplar arası karşılaştırma çalışmalarında yapılan karşılaştırmanın anlamlılığının bir önkoşuludur (aktaran, Cheung & Rensvold, 2000). Yani ölçme değişmezliği ölçme aracından elde edilen ölçümlerde aynı yapının ortaya çıkıp çıkmadığını belirlemeye çalışır. Dolayısıyla ölçme değişmezliği bireylere ilişkin özelliklerle değil ölçme aracının kendisiyle ilgili bir durumdur. Yapılan karşılaştırmalar sonucu tüm parametrelerin eşitliği söz konusuysa modelin gruplar için değişmez olduğu söylenebilir. Belirli bir ölçme aracının kullanıldığı çalışmalarda, gruplar arasındaki farklılıkların ölçülmek istenen psikolojik yapı açısından gruplar arası "gerçek" farklılık olduğu kabul edilmektedir (Mark & Wan, 2005).

Yapısal eşitlik modeline dayalı yapılan gruplar arası karşılaştırma çalışmalarında klasik karşılaştırma çalışmalarında olduğu gibi gözlenen değişkenlere ilişkin yapılan ölçmelerden ziyade örtük değişkenlerin içinde bulunduğu hipotezler test edilmektedir. Örtük değişken ortalamalarının test edilmesi klasik yöntemlere göre daha duyarlı bir tekniktir ve iki grup arasındaki farklılığı daha kesin bir şekilde ortaya koyar (Hong, Malik, & Lee, 2003). Tyson'a (2004) göre ölçme değişmezliği, karşılaştırma yapılacak gruplar ile ilgili psikolojik ölçümlerin geçerlik ve güvenirlik düzeylerinin birbirine benzer olma derecesidir. Aslında ölçme değişmezliğinin test edilmesindeki amaç, karşılaştırma grupları için ilgili ölçümlerin geçerliğini sağlamaktır. Ölçme değişmezliği ile ilgili ayrıntılı bilgiler aşağıda verilmiştir.

## Kavramsal Çerçeve

### Ölçme Değişmezliği

Ölçme değişmezliği testleri bir tür kovaryans yapı analizi olup farklı gruplarda belli bir yapının ölçülmesi üzerine tasarlanmıştır. İstatistiksel olarak ölçme değişmezliği test edilirken en çok ve yaygın kullanılan yöntem Çok Gruplu Doğrulamalı Faktör Analizidir (ÇG-DFA) (Jöreskog & Sörbom, 1999). ÇG-DFA hiçbir sınırlandırmanın koyulmadığı modelden en sınırlı modele doğru karşılaştırma yaparak gruplar arası parametrelerin eşdeğerliğini inceleyen aşamalı bir süreçtir (Horn & McArdle, 1983).

Ölçme değişmezliği için test edilmesi gereken dört aşama bulunmaktadır (Meredith, 1993). Bu aşamalar sırasıyla şu şekildedir:

1. Biçimsel (Yapısal) değişmezlik (Configural invariance)
2. Metrik (Zayıf) değişmezlik (Weak factorial invariance= Metric invariance)
3. Skalar değişmezlik (Strong invariance= Scalar invariance)
4. Katı (Tam) değişmezlik (Strict invariance).

**Biçimsel (Yapısal) değişmezlik.** Biçimsel değişmezlik hiçbir sınırlandırmanın konulmadığı en temel modeldir (McArdle & Mason, 1983; aktaran, Chen, Sausa, & West, 2005). Bu aşamada temel model yapısı gruplar için değişmezse biçimsel değişmezliğin sağlandığı söylenebilir. Bu modelde hiçbir sınırlandırma yapılmadığı için, grupların farklı parametre değerleri almasına izin verilir. Böylece, ölçme değişmezliği aşamasında yer alan tüm modeller ile arasında bir karşılaştırma yapma imkânı oluşur. Biçimsel değişmezlik modeli kritik bir öneme sahiptir çünkü biçimsel değişmezliğin sağlanamaması durumunda, uygulanan ölçme aracının gruplar arası aynı

psikolojik yapıyı ölçmediğine işaret etmektedir. Biçimsel değişmezliğin sağlanması halinde aşağıdaki model eşitliği sağlanabilir

$$\Lambda_{1.Grup} = \Lambda_{2.Grup}$$

**Metrik (Zayıf) değişmezlik.** Metrik değişmezlik aşamasında, farklı gruplardaki bireylerin maddeleri aynı şekilde cevapladığı ve böylece her bir gruptan gelen puanların karşılaştırmasının da anlamlı olduğu sayıtları bulunmaktadır (Steenkamp & Baumgartner, 1998). Metrik değişmezlikte şekilsel değişmezlikteki değişmez faktör örüntüsünün tanımlanmasına ek olarak bu örüntü içerisindeki maddelerin faktör yük değerlerinin de gruplar arası eşit olduğu sınırlaması bulunmaktadır (Bollen, 1989). Metrik değişmezliğin sağlanması durumunda aşağıdaki denklem eşitliği kurulabilir.

$$\lambda_{ij}^g = \lambda_{ij}^{g'}$$

**Skalar değişmezlik.** Ölçme değişmezliğinin güçlü bir seviyesi olan bu aşamada, “madde eğimlerine ek olarak madde sabitlerinin gruplar arasında eşit olması gerekmektedir” (Tucker, Ozer, Lyubomirsk, & Boehm, 2006, s. 344). Yani aynı örtük yapı üzerindeki değerlerin gözlenen değer üzerinde de eşit olduğu sayıtları bulunmaktadır. Bu da “gözlenen maddeler üzerindeki grup farklılıkları örtük yapıların ortalamalarından kaynaklanır” anlamına gelmektedir. Dolayısıyla örtük değişkenlerin ortalamalarının karşılaştırılması bu model için anlamlı olacaktır (Başusta, 2010).

**Katı (Tam) değişmezlik.** Bu aşamada ölçme aracını oluşturan maddelere ilişkin hata terimlerinin karşılaştırma grupları için eşit olduğu şeklindeki bir hipotez test edilmektedir:  $\varepsilon_g = \varepsilon_{g'}$  (Önen, 2009). Gözlenen değişken ortalamalarındaki ve dağılım matrisindeki grup farklılıkları hesaplanır.

Ölçme değişmezliğinin test edilmesinde her bir aşamada kurulan model için bir hipotez test edilmektedir. Aşamalı bir sürecin izlendiği ölçme değişmezliğinde her bir model, kendisinden önceki aşamada yer alan modele belirli parametre sınırlandırmaların koyulması ile oluşturulmaktadır. Bu nedenle bir aşamadaki ölçme değişmezliği testi, o aşamadaki model ile bir önceki aşamadaki modelin veriye uyum düzeylerinin karşılaştırılması yoluyla yapılmaktadır.

Türkiye’de ölçme değişmezliği konusu ile ilgili çeşitli çalışmalar bulunmaktadır. Önen’in (2009), Epistemolojik İnançlar Envanterinin Türkçe’ye uyarlamasını yaptığı çalışmada ölçek Türkçe’ye uyarlandıktan sonra ölçeğin cinsiyet değişkeni açısından ölçme değişmezliği test edilmiştir. Analizler sonucunda ölçeğin Türkçe formu için cinsiyet grupları arasında şekilsel değişmezliğe, tam metrik değişmezliğe, kısmi ölçme değişmezliğine, kısmi değişmez özgüllüğe ve faktör varyanslarının değişmezliğine ilişkin kanıtlar elde edilmiştir. Uzun ve Öğretmen (2010) yapmış oldukları çalışmada TIMSS 1999 Türkiye verilerini kullanarak fen başarısına etki eden duyuşsal faktörler belirleyerek, bu faktörlerin ölçme değişmezliğini test etmişlerdir. Yapılan analizler sonucunda çalışmada yer alan öz yeterlik, önem, tutum ve sınıf içi öğrenci etkinlikleri gizil değişkenlerinin metrik değişmezlik koşulunu sağladığı fakat cinsiyet gruplarında katı değişmezlik koşulunu hiçbirinin sağlayamadığı görülmüştür. Akyıldız (2009) yapmış olduğu çalışmada, PIRLS 2001 başarı testinin 35 ülkede yapılmış olan uygulamasından elde edilen faktör yapılarının birbirine denkliğini çok gruplu doğrulayıcı faktör analizi ile incelemiştir. PIRLS 2001 başarı testinin faktör yapısının 35

ülkede değişmezliği tüm ülkelerde uygulanmış olan dört ayrı alt ölçek bağlamında ayrı ayrı test edilmiştir. Yapılan testler sonucunda PIRLS 2001 testinin faktör yapısının uygulamaya katılan tüm ülkelerde birbirine orta düzeyde denk sayılabileceği sonucuna ulaşılmıştır. Bahadır'ın (2012) yapmış olduğu bir çalışmada ise PISA 2009 Türkiye verisi kullanılarak öğrencilerin okuma becerilerini etkileyen değişkenler modellenmiş ve bu modeldeki değişkenler için bölgeler arası ölçme değişmezliği testleri yapılmıştır.

TIMSS, öğrencilerin fen ve matematik başarısını ve bu başarıyı etkileyebilecek çeşitli özellikler hakkında birey, okul ve ülke bazında bilgi sağlayan uluslararası kapsamlı bir çalışmadır. TIMSS'te öğrencilerin matematik başarılarını etkileyebilecek öğrenci, öğretmen ve okul düzeyinde değişkenler bulunmaktadır. Öğrenci düzeyinde matematik başarısının incelendiği çalışmalara bakıldığında, matematik başarısını en çok etkileyen faktörlerin başında matematiğe karşı olan tutum, matematik ile ilgili beklentiler, bireylerin kendilerine olan özgüvenleri ve matematiği sevmeleri gelmektedir (Akyüz, 2014; Çavdar, 2015; Doğan & Barış, 2010; Leder & Forgasz, 2006; Yayan, 2003). Bu çalışmada da matematik başarısında öğrenci düzeyinde etkili olduğu düşünülen matematiği sevmeye, matematiğe olan ilgi ve matematiğe ilişkin özgüven gizil değişkenleri ile kurulan modelin cinsiyet grupları açısından ölçme değişmezliklerinin test edilmesi amaçlanmaktadır. Bu amaç doğrultusunda çalışmanın problem cümlesi aşağıdaki şekilde ifade edilmiştir:

TIMSS 2015'e katılan Türkiye dördüncü sınıf öğrencilerinin matematik başarıları üzerinde etkili olduğu düşünülen matematiğe olan ilgi, matematiği sevmeye ve matematiğe ilişkin özgüven gizil değişkenlerinin kız ve erkek öğrenci gruplarında ölçme değişmezlik düzeyleri nasıldır?

### **Araştırmanın Amacı ve Önemi**

Çalışmada TIMSS-R (Third International Mathematics and Science Study-Repeat) verileri kullanılmıştır. TIMSS 2015 Türkiye verilerine göre dördüncü sınıf öğrencilerinin matematik başarılarını etkileyen gizil değişkenler aracılığı ile oluşturulan modelin cinsiyet grupları açısından ölçme değişmezliği testleri yapılmıştır. Bu çalışmanın temel amacı matematik başarısını etkileyen matematiği sevmeye, matematiğe olan ilgi ve matematiğe ilişkin özgüven gizil değişkenleri ile oluşturulan ölçme modelinin kız ve erkek grupları için aynı yapıda olup olmadığını test etmektir.

Matematik başarısını etkileyen yapıların belirlendiği birçok çalışma bulunmasına rağmen, PISA ve TIMSS gibi uluslararası karşılaştırmaları konu alan çalışmalarda duyuşsal özelliklere ilişkin bilgi toplamayı amaçlayan anketlerin eşdeğerlikleri ile çok fazla çalışmanın yapılmadığı görülmektedir (Schulz, 2003, 2005, 2008). Türkiye'de matematik başarısının gruplar arasında ölçme değişmezliğinin incelendiği çalışmalar birkaç örnek verilebilir. Örneğin, Gülleroğlu (2017) yaptığı çalışmada, PISA 2012 Matematik uygulamasına katılan Türk öğrencilerinin duyuşsal özelliklerinin cinsiyete göre ölçme değişmezliğini incelemiştir. Uzun ve Öğretmen (2010) fen başarısına etki eden duyuşsal özelliklerin TIMSS Türkiye örneğinde cinsiyete göre ölçme değişmezliğini değerlendirmişlerdir. Başka bir çalışmada ise Karakoç-Alatlı, Ayan, Polat-Demir ve Uzun (2016), TIMSS 2011 dördüncü sınıf matematik başarı testinin kültürler arası ölçme değişmezliğini incelemiştir. Araştırmalarda ölçme değişmezliği çalışmalarının yapılması grupların karşılaştırılmasında daha doğru ve güvenilir sonuçlar vermektedir. Bu çalışmanın TIMSS 2015 dördüncü sınıf öğrenci anketinde yer alan

değişkenlerin cinsiyete göre yapılan karşılaştırma çalışmalarının anlamlılığı konusunda daha geçerli bilgiler sağlayabileceği düşünülmektedir. Bunun dışında araştırma bulgularının Türkiye'deki eğitim sisteminin değişim ve gelişimi için yapılan deneysel çalışmalara da yardımcı bilgiler sağlayabileceği umulmaktadır.

### Yöntem

Bu bölümde araştırmanın modeli, evren ve örneklem, veri toplama yöntemi, verilerin analizi ile ilgili bilgilere yer verilmiştir.

#### Araştırmanın Modeli

Bu araştırma, TIMSS 2015'e katılan dördüncü sınıf Türkiye örnekleminde elde edilen verilerle öğrencilerin matematik başarısını açıkladığı düşünülen bir model ortaya konularak, modelin cinsiyet gruplarına göre ölçme değişmezliğinin test edildiği ilişkisel bir araştırmadır. İlişkisel araştırmalar değişkenler arasındaki ilişkilerin ve bağlantıların incelendiği araştırmalardır. İlişkisel araştırmalar ile değişkenler arasındaki olası bağlantılar derinlemesine araştırılarak, olgular ve kavramlar daha iyi anlaşılabilir (Büyüköztürk, Çakmak, Akgün, Karadeniz, & Demirel, 2014).

#### Evren ve Örneklem

Araştırmanın evrenini Türkiye'de öğrenim görmekte olan dördüncü sınıf öğrencileri oluşturmaktadır. Araştırmanın örneklemini ise TIMSS 2015'e Türkiye'den katılan dördüncü sınıf öğrencileri oluşturmaktadır. TIMSS 2015 Türkiye örneklemini belirlemek için iki aşamalı tabakalı örnekleme yöntemi kullanılmıştır. İlk aşamada okullar, ikinci aşamada ise bu okullardaki sınıflar rastgele seçilmiştir. Eğitimi Araştırma ve Geliştirme Dairesinin (EARGED) oluşturduğu ilköğretim okulları listesinden, TIMSS. 2015 çalışmasına katılacak okullar IEA tarafından seçkisiz yöntemle belirlenmektedir. Okullar belirlendikten sonra sınıfların seçimi EARGED tarafından yine seçkisiz örnekleme yöntemi ile belirlenmektedir. TIMSS 2015 Türkiye örnekleminde toplam 6456 dördüncü sınıf öğrencisi bulunmaktadır. Ölçme değişmezliği testine geçmeden önce kayıp verilerin ve uç değerlerin belirlenmesi için analizler gerçekleştirilmiştir. Kayıp veri analizi için Little's MCAR (Missing Completely at Random) testi yapılmıştır. Test sonuçlarına göre dördüncü sınıf öğrenci anketinden elde edilen veride %1.5 kayıp veri olduğu ve test sonucunda ( $p=0.000$ ) kayıp verinin sistematik bir örüntüden ziyade rastgele örüntüler içerdiği sonucuna ulaşılmıştır (Garson, 2015). Kayıp verinin %5'in altında olduğu ve tamamlanmamış verinin sistematik bir örüntü oluşturmadığı görüldükten sonra kayıp verilerin yerine ortalama değerler atanmıştır. Çok değişkenli uç değerlerin tespit edilmesi amacıyla değişkenlerin Cook uzaklıkları hesaplanmıştır. Cook uzaklığı için ölçüt değer  $4/n$ 'dir. Örneklem büyüklüğünün 6456 olmasından dolayı  $4/6456=0.00619$ 'dur. Cook uzaklığı bu değerden büyük olan 310 veri uç değer olarak değerlendirilip analiz dışı bırakılmıştır (Cook, 1977). Dolayısıyla analizler 6146 kişi üzerinden gerçekleştirilmiştir. Çalışmaya dördüncü sınıf düzeyinde öğrenim görmekte olan 3049 (%49.6) kız ve 3097 (%50.4) erkek öğrenci katılmıştır.

#### Veri Toplama Araçları

Araştırmada yer alan gizil değişkenler, TIMSS 2015'de uygulanan öğrenci anketinden elde edilmiştir. Öğrenci anketinden matematiği sevme, matematiğe ilişkin

özgüven ve matematiğe olan ilgi değişkenlerinin ölçüldüğü maddelere ilişkin cevaplar seçilmiştir. Araştırma kapsamında kullanılan bu veriler TIMSS verilerinin yer aldığı web sayfasından (<http://timssandpirls.bc.edu/timss2015/>) elde edilmiştir.

**Öğrenci anketleri.** TIMSS 2015’de öğrencilere uygulanan anket maddeleri incelenmiş ve bu araştırmada incelenen matematiği sevme, matematiğe olan ilgi ve matematiğe ilişkin özgüven gizil değişkenlerini ölçen maddeler belirlenmiştir. Analizlerde bu maddelerden elde edilen puanlar kullanılmıştır.

Öğrenci anketinde yer alan “matematiği sevme”, “matematiğe olan ilgi”, ve “matematiğe ilişkin özgüven” değişkenlerini ölçmeye yönelik olan maddelerin tepki kategorileri “çok katılıyorum, az katılıyorum, katılmıyorum, hiç katılmıyorum” şeklindedir. Her bir alt ölçekten elde edilen toplam puanların yüksek olması, öğrencilerin sırasıyla matematiği sevmelerinin, matematiğe olan ilgilerinin ve matematiğe ilişkin özgüvenlerinin yüksek olduğunu göstermektedir. Ölçeklerde yer alan maddeler ve kodları Tablo 1’de yer almaktadır.

Tablo 1

*Ölçekler Maddeleri ve Kodları*

<b>Ölçek</b>	
<b>Matematiği Sevme</b>	<b>Kod</b>
Matematik öğrenirken eğleniyorum.	ASBM01A
Keşke matematik çalışmak zorunda olmasam.	ASBM01B*
Matematik sıkıcıdır.	ASBM01C*
Matematik dersinde ilginç şeyler öğrenirim.	ASBM01D
Matematiği seviyorum.	ASBM01E
Okul işleri sayıları içerir.	ASBM01F
Matematik problemlerini severim.	ASBM01G
Matematik derslerini dört gözle beklerim.	ASBM01H
Favori dersim matematiktir.	ASBM01I
<b>Matematiğe olan ilgi</b>	<b>Kod</b>
Öğretmenimin benden ne istediğini biliyorum.	ASBM02A
Öğretmenimin anlattıklarını kolay anlıyorum.	ASBM02B
Öğretmenimin söyledikleri ile ilgileniyorum.	ASBM02C
Öğretmenim yapmam için ilginç şeyler verir.	ASBM02D
Öğretmenime net cevaplar veririm.	ASBM02E
Öğrenim iyi bir şekilde açıklar.	ASBM02F
Öğretmenime öğrendiklerimi gösteririm.	ASBM02G
Öğretmenim çeşitli şekilde öğretir.	ASBM02H
Daha iyi olmaya çalışırım.	ASBM02I
Öğretmenimi dinlerim.	ASBM02J

Matematiğe ilişkin özgüven	Kod
Genellikle matematikte iyiyim.	ASBM03A
Benim için matematik diğer derslerden daha zor.	ASBM03B*
Sadece matematikte iyi değilim.	ASBM03C*
Matematik dersinde bir şeyleri çabuk öğrenirim.	ASBM03D
Matematik beni sınırlendirir.	ASBM03E*
Matematik problemlerini çözmede iyiyim.	ASBM03F
Matematikte iyiyim.	ASBM03G
Matematik dersleri benim için daha zordur.	ASBM03H*
Matematik benim kafamı karıştırır.	ASBM03I*

\* Madde ters kodlanmıştır.

Araştırma kapsamında ele alınan duyuşsal özelliklere ilişkin puanların geçerlikleri için AFA yapılmış ve güvenilirlikleri için de Cronbach alfa iç tutarlılık katsayıları hesaplanmıştır.

### Verinin Analizi

Çalışmada temel olarak iki istatistiksel analiz yönteminden yararlanılmıştır. İlk olarak matematiği sevme, matematiğe olan ilgi ve matematiğe ilişkin özgüven ölçeklerinin faktör yapısını belirlemek amacıyla her bir ölçek için açımlayıcı faktör analizi (AFA) yapılmıştır. Daha sonra matematik başarısı bağımlı değişken olarak ele alınıp yapısal eşitlik modeli (YEM) gerçekleştirilmiştir. Yapısal Eşitlik Modeli (YEM) ile yapısal modelde yer alan değişkenler arasındaki yapısal ilişkiler ve faktör analizindeki faktörler ile maddeler arasındaki ilişkiler tek bir analizde birleştirilmektedir (Başusta & Gelbal, 2015). Ölçme aracının cinsiyet grupları açısından ölçme değişmezliğinin test edilmesi amacıyla çok gruplu doğrulayıcı faktör analizi (ÇG-DFA) uygulanmıştır. Grup karşılaştırmalarında çok gruplu doğrulayıcı faktör analizi grup parametrelerinin eşit kalarak, faktör ortalamalarının karşılaştırılmasını sağlamaktadır. Gözlenen değişkene dair ölçümlerden ziyade yapısal eşitlik modeli kullanılarak yapılan grup karşılaştırmasında örtük değişkenlerin içinde bulunduğu hipotezlerin test edilme olanağı bulunmaktadır (Başusta & Gelbal, 2015). Çoklu grup YEM analizlerinde karşılaştırmalı uyum katsayıları (CFI) arasındaki farklar kullanılarak modeller karşılaştırılmaktadır. Çalışmamızda da temel model çerçevesinde sınırlandırılmayan modelden en sınırlı modele doğru giderken modeller arasındaki CFI farkları dikkate alınmıştır. Veri ile uyum gösteren ÇG-DFA değişkenleri için  $-0.01 \leq \Delta CFI \leq 0.01$  olan değişmezlik koşulunun sağlanıp sağlanmadığı test edilmiştir.

Ayrıca değişmezlik testinin her aşamasında kurulan ölçme modelinin model veri uyumunu sağlayıp sağlamadığını belirlemek için  $\chi^2/df$ , GFI, CFI, NFI, NNFI RMSEA indeksleri rapor edilmiştir. Model veri uyumunun sağlanabilmesi için hesaplanan  $\chi^2/df$  oranının 5'ten küçük olması, GFI değerinin 0.90'dan yüksek olması, RMSEA değerinin ise 0.05'ten düşük çıkması ölçüt değerler olarak kabul edilmektedir. (Jöreskog & Sörbom, 1993; Marsh & Hocevar, 1988). Bununla birlikte, GFI'nin 0.85'ten büyük olması, RMSEA değerinin 0.10'dan düşük çıkması, model veri uyumu için kabul



edilebilir alt sınırlar olarak değerlendirilmektedir (Anderson & Gerbing, 1984; Cole, 1987; Marsh, Balla, & McDonald, 1988). Karşılaştırmalı uyum indeksleri olan CFI, NFI ve NNFI değerlerinin ise 0.95'in üzerinde olması çok iyi bir uyumu, 0.90-0.95 olması ise kabul edilebilir bir uyumu göstermektedir (Sümer, 2000).

### Bulgular

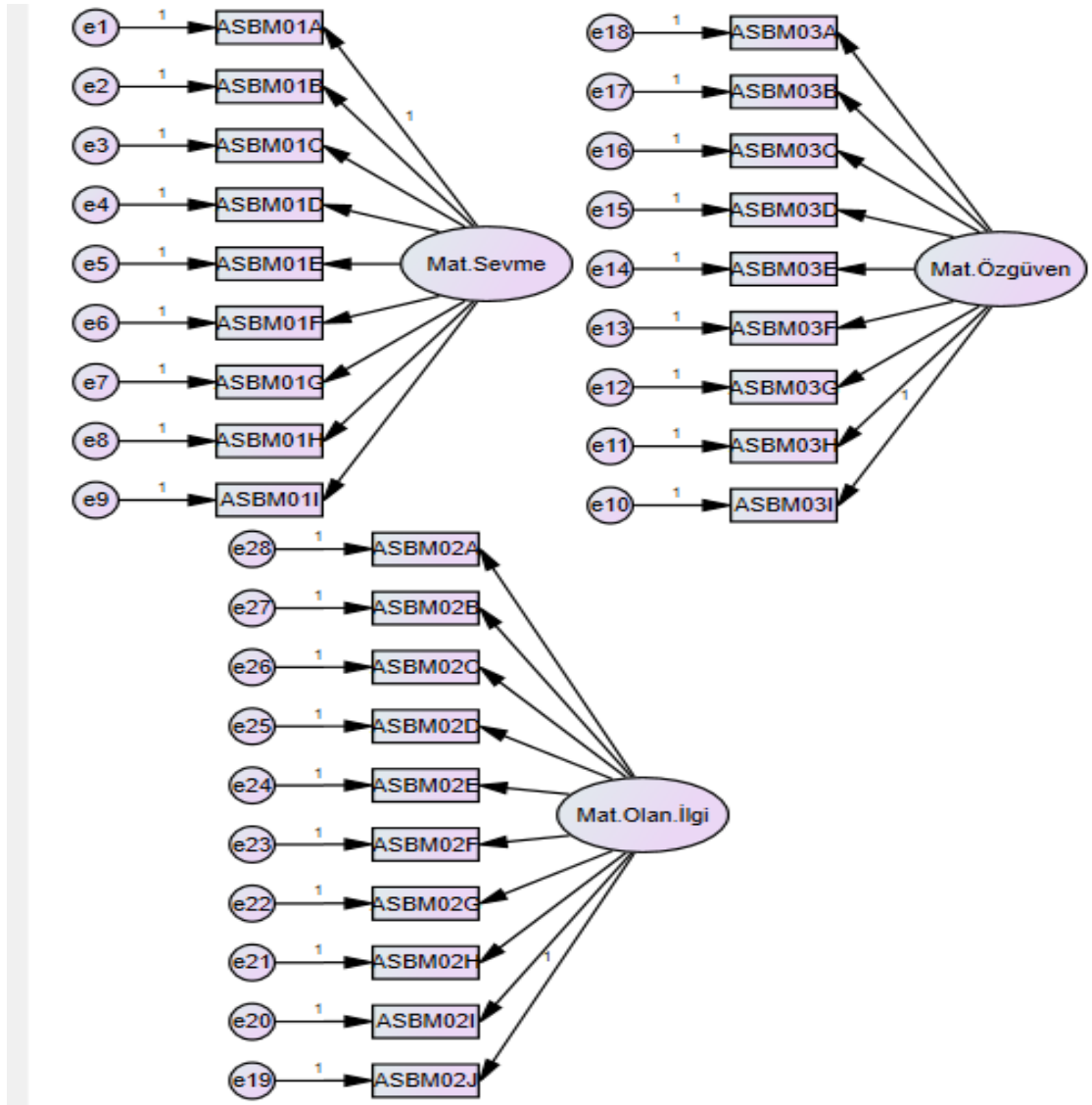
TIMSS öğrenci anketinin matematik başarısı ile ilgili olan matematiği sevme, matematiğe olan ilgi ve matematiğe ilişkin özgüven gizil değişkenlerinin her birine ilişkin ölçümlere dayalı olarak ayrı ayrı AFA yapılmıştır. Yapılan AFA' larda faktör sayısına sınırlama konulmadan temel bileşenler analizi uygulanmış ve herhangi bir döndürme işlemi gerçekleştirilmemiştir. Veri setinin AFA'ya uygunluğunu değerlendirmek amacıyla Kaiser – Meyer – Olkin (KMO) katsayısı ve Barlett küresellik testi sonuçları incelenmiştir. Ayrıca her bir ölçek için Cronbach alfa iç tutarlılık katsayısı 0.70 ve üzeri ise “yüksek” güvenilirlik ve 0.70 ile 0.60 arasında ise “kabul edilebilir” ve 0.60 ile 0.40 arasında ise “zayıf” güvenilirlik yorumu yapılabilir (George & Mallery, 2003). Her bir duyuşsal özellik için yapılan AFA sonuçları ile Cronbach alfa değerleri aşağıda sıra ile verilmiştir.

**Matematiği sevme.** Matematiği sevme ölçeği için yapılan AFA sonucunda elde edilen KMO katsayısı (0.90) ve Barlett Küresellik testi (18073.438;  $p=0.00$ ) değerleri, örneklemin faktör analizi yapabilmek için yeterli olduğu ve verilerin faktör analizine uygun olduğunu göstermektedir (Büyüköztürk, 2015). AFA sonucunda, ilk olarak öz değerleri 3.40 ve 1.53 olan ve varyansın %59.16'sını açıklayan iki faktör elde edilmiştir. Fakat alan yazında bir faktörün en az 3 maddeden oluşması ve faktör yüklerinin de yüksek olması gerektiği (Ford, MacCallum, & Tait, 1986; Hogarty, Hines, Kromrey, Ferron, & Mumford, 2005) göz önüne alınarak analizler yeniden yapılmıştır. Yeniden yapılan AFA sonucunda öz değeri 3.93 olan ve varyansın %49.08'ini açıklayan tek bir faktör oluşmuştur. Maddelerin faktör yük değerleri ise  $\lambda=0.38$  ila  $\lambda=0.84$  arasında değişmektedir. Dolayısıyla toplam sekiz madde ve tek faktörden oluşan “Matematiği sevme” ölçeği elde edilmiştir. Ayrıca matematiği sevme ölçeğinin Cronbach alfa iç tutarlılık katsayısı 0.65 olarak hesaplanmış ve ölçeğin “kabul edilebilir” düzeyde güvenilir olduğu söylenebilir.

**Matematiğe olan ilgi.** Matematiğe olan ilgi ölçeğine ilişkin elde edilen AFA sonucunda KMO katsayısı (0.81) ve Barlett Küresellik testi (5453.666;  $p=0.00$ ) olarak bulunmuştur. Buna göre veri yapısı AFA yapabilmek için yeterli düzeydedir (Büyüköztürk, 2015). AFA sonucunda, ilk olarak öz değerleri 2.95 ve 1.43 olan ve varyansın %45.12'sini açıklayan iki faktör elde edilmiştir. Ancak ikinci faktör altında yer alan dört madde de aynı özelliği ölçmek amacı ile yazılmış oldukları için birlikte başka bir faktör oluşturmaktadır. Dolayısıyla maddelerden faktör yük değerleri en düşük olan üç madde sırası ile analiz dışı bırakılarak yeniden AFA yapılmıştır. Yeniden yapılan AFA sonucunda ise öz değeri 2.51 olan ve varyansın %35.83'ünü açıklayan tek bir faktör elde edilmiştir. Maddelerin faktör yük değerleri ise  $\lambda=0.38$  ila  $\lambda=0.68$  arasında değişmektedir. Toplam yedi madde ve tek faktörden oluşan “Matematiğe olan ilgi” ölçeğinin faktör yapısı elde edilmiştir. Matematiğe olan ilgi ölçeğinin Cronbach alfa iç tutarlılık katsayısı 0.63 olarak hesaplanmış ve ölçeğin “kabul edilebilir” düzeyde güvenilir olduğu söylenebilir.

**Matematiğe ilişkin özgüven.** Matematiğe ilişkin özgüven ölçeğinin faktör yapısını belirlemek amacıyla yapılan AFA sonucunda KMO katsayısı 0.84 ve Barlett Küresellik testi (12374.035;  $p=0.00$ ) olarak bulunmuştur. Buna göre veri AFA yapabilmek için yeterli düzeydedir (Büyüköztürk, 2015). AFA sonucunda, ilk olarak öz değerleri 3.74 ve 1.36 olan ve varyansın %53.11’ni açıklayan iki faktör elde edilmiştir. İkinci faktör altında iki madde yer almıştır. Fakat alan yazında faktör yükleri bir faktörün en az 3 maddeden oluşması ve faktör yüklerinin de yüksek olması gerektiği (Ford ve diğerleri, 1986; Hogarty ve diğerleri, 2005) göz önüne alınarak faktör sayısı bire sabitlenip analizler yeniden yapılmıştır. Yeniden yapılan AFA sonucunda ise öz değeri 3.29 olan ve varyansın %41.13’nü açıklayan tek bir faktör oluşmuştur. Maddelerin faktör yük değerleri ise  $\lambda=0.57$  ile  $\lambda=0.69$  arasında değişmektedir. Böylece toplam sekiz madde ve tek faktörden oluşan “Matematiğe İlişkin Özgüven” ölçeğinin faktör yapısı belirlenmiş oldu. Ayrıca ölçeğin hesaplanan Cronbach alfa iç tutarlılık katsayısı 0.43’dür ve güvenilirliğin “zayıf” düzeyde olduğu söylenebilir. Yapılan AFA sonucunda ölçekler ve ölçeklerde yer alan maddelere ilişkin elde edilen yol diyagramı Şekil 1’de gösterilmektedir.

Şekil 1. Ölçeklere İlişkin Yol Diyagramları



Matematik başarısını açıklayan üç gizil değişkene ilişkin kurulan modelin ölçme değişmezliğini test amacıyla her bir aşamadaki CFI katsayı farkları kullanılarak anlamlılık testi yapılmıştır. Her bir gizil değişkenin ölçme değişmezliğinin test edilmesinde biçimsel değişmezlik ile metrik değişmezlik, skalar değişmezlik ve katı değişmezlik arasındaki CFI fark değeri hesaplanmış ve  $-0.01 \leq \Delta CFI \leq 0.01$  kriter değeri dikkate alınmıştır. İlk aşama olarak biçimsel değişmezliğin test edilmesinde her bir gizil değişken için erkek ve kız gruplarına ait yapılarıdaki hiçbir parametreye sınırlandırma getirilmemiştir. Her iki grup için tüm parametreler (faktör yükleri, faktörler arası korelasyon, hata varyansları) serbestçe kestirilmiş ve farklı değerler almalarına izin verilmiştir. Biçimsel değişmezlik için her bir ölçüğe ilişkin elde edilen model veri uyum istatistikleri Tablo 2’de gösterilmektedir.

Tablo 2

*Biçimsel Değişmezlik Aşamasına İlişkin Uyum İstatistikleri*

Ölçekler	$\chi^2/df$	RMSEA	CFI	GFI	NFI	NNFI
Matematiği Sevme	27.43	0.093	0.91	0.96	0.91	0.86
Matematiğe Olan İlgi	21.59	0.008	0.92	0.96	0.92	0.88
Matematiğe İlişkin Özgüven	33.85	0.010	0.88	0.93	0.88	0.83

Tablo 2’de yer alan değerler incelendiğinde  $\chi^2/df$  değerleri büyük çıkmıştır. Bunun sebebi ise  $\chi^2$  değerinin örneklem büyüklüğünden etkilenmesidir. Dolayısıyla model-veri uyumu için diğer indeksler göz önüne alınmıştır. RMSEA değerleri incelendiğinde sadece “Matematiği sevme” değişkeni için RMSEA değeri 0.09 olduğu için zayıf model-veri vardır yorumu yapılabilir. “Matematiğe olan ilgi” ve “Matematiğe ilişkin özgüven” değişkenleri için RMSEA değerleri ise 0’a oldukça yakındır (Tabachnick & Fidell, 2001). CFI, GFI ve NFI ve NNFI değerleri ise üç gizil değişken içinde kabul edilir aralıklar içerisindedir (Meredith, 1993). Dolayısıyla elde edilen değerler için alt gruplarda bu model veri ile uyum sergilemektedir. Böylece üç gizil değişkende hem kız hem de erkekler için aynı yapıdadır. Yani gizil değişkenler kız ve erkek gruplarında değişmezdir. Bu bulgular neticesinde cinsiyet grupları için biçimsel değişmezliğin matematiği sevme, matematiğe olan ilgi ve matematiğe ilişkin özgüven ölçekleri için sağlandığı yorumu yapılabilir.

Hangi değişkenlerin metrik değişmezliği sağladığına karar vermek için biçimsel değişmezlikteki sınırlandırmalara ek olarak gruplar arasındaki faktör yüklerinin de aynı olması sınırlılığına yer verilmektedir. Tablo 3’de faktör yüklerinin eşit olduğu koşulu ile hazırlanan metrik değişmezliğe ilişkin uyum katsayıları yer almaktadır.

Tablo 3

*Metrik Değişmezlik Aşamasına İlişkin Uyum İstatistikleri*

Ölçekler	$\chi^2/df$	RMSEA	CFI	GFI	NFI	NNFI
Matematiği Sevme	20.56	0.080	0.90	0.96	0.90	0.90
Matematiğe Olan İlgisi	16.61	0.071	0.93	0.97	0.92	0.91
Matematiğe İlişkin Özgüven	22.97	0.085	0.91	0.96	0.90	0.89

Tablo 3 incelendiğinde metrik değişmezlik için model-veri uyumunun üç gizil değişken içinde sağlandığı görülmektedir. Elde edilen bu uyum değerlerine göre her bir değişkene ilişkin ölçme modelinin alt gruplar için anlamlı olduğu söylenebilir. Değişmezlik koşulunu sağlayan değişkenlere karar vermek için metrik model ile biçimsel model kıyaslanarak bu iki model arasındaki ( $\Delta CFI$ ) uyum katsayılarının fark değerleri incelenmiş ve elde edilen sonuçlar Tablo 4'te gösterilmiştir.

Tablo 4

*Biçimsel Değişmezlik ve Metrik Değişmezlik İçin Fark Testleri Tablosu Uyum İstatistikleri*

Ölçekler	$\Delta CFI (CFI_{BİÇİMSEL} - CFI_{METRİK})$
Matematiği Sevme	0.01
Matematiğe Olan İlgisi	-0.01
Matematiğe İlişkin Özgüven	-0.03

Metrik değişmezlik için fark değerinin  $-0.01 \leq \Delta CFI \leq 0.01$  aralığında olması gerekmektedir. Buna göre Tablo 4 incelendiğinde “Matematiği sevme” ve “Matematiğe olan ilgi” değişkenlerinin metrik değişmezliği sağladığı, “Matematiğe ilişkin özgüven” değişkeninin ise metrik değişmezliği sağlamadığı görülmektedir. Dolayısıyla “Matematiği sevme” ve “Matematiğe olan ilgi” gizil değişkenleri için grupların maddelere aynı biçimde cevap verdiğini ve böylece iki gruptan elde edilen puanları karşılaştırmada anlamlı olduğu sonucuna ulaşılmıştır. Buna karşın “Matematiğe ilişkin özgüven” değişkeni için metrik değişmezlik sağlanmadığı için bu gruplar arası fark olduğu anlamına gelmektedir. Yani gözlenen maddeler üzerindeki grup farklılıkları bakımından cinsiyetler arası bir karşılaştırma anlamlı olmayacaktır.

Değişmezliğin üçüncü aşaması olan skalar değişmezlik, metrik değişmezlikte koyulan sınırlandırmalara ek olarak benzer maddelerin gruplar arası eş olduğu sınırlandırılması da yapılır. Böylece erkek ve kız gruplarında hem faktör yükleri hem de faktörler arası korelasyonların eşit olma sınırlılığı getirilmiş olur. Matematiğe ilişkin özgüven değişkeni metrik değişmezliği sağlamadığı için skalar değişmezlik testi de yapılmamıştır. Metrik bakımdan değişmez olan “Matematiği sevme” ve “Matematiğe olan ilgi” değişkenlerine skalar değişmezlik sınırlandırılması da eklenerek test edilmiştir. Skalar değişmezlik için elde edilen model-veri uyum indeks değerleri Tablo 5’de gösterilmiştir.

Tablo 5

*Skalar Değişmezlik Aşamasına İlişkin Uyum İstatistikleri*

Ölçekler	$\chi^2/df$	RMSEA	CFI	GFI	NFI	NNFI
Matematiği Sevme	20.56	0.080	0.90	0.96	0.90	0.90
Matematiğe Olan İlgisi	11.65	0.059	0.94	0.99	0.94	0.94

Tablo 5 incelendiğinde RMSEA, CFI, GFI, NFI ve NNFI kabul edilir uyum sınırları içerisinde bulunmaktadır (Meredith, 1993). Elde edilen bu değerlere göre “alt gruplarda bu model veri ile mükemmel uyum sergilemektedir” yorumu yapılabilir. Bu uyum istatistiklerinden her bir değişkene ilişkin ölçme modelinin alt gruplar için anlamlı olduğu sonucu çıkar. Yani gizil değişkenler hem kız hem de erkekler için aynı yapıdadır. Başka bir deyişle, tüm gizil değişkenler gruplar için değişmezdir.

“Matematiği sevme” ve “Matematiğe olan ilgi” değişkenlerinin skalar değişmezliği sağlayıp sağlamadığını belirlemek için biçimsel değişmezlik ile skalar değişmezlik arasındaki CFI fark değerine bakılmıştır. Biçimsel değişmezliğe ilişkin uyum istatistikleri incelenerek iki model arasındaki CFI fark testine ait sonuçlar Tablo 6’da gösterilmiştir.

Tablo 6

*Biçimsel Değişmezlik ve Skalar Değişmezlik için Fark Testleri Tablosu Uyum İstatistikleri*

Ölçekler	$\Delta CFI (CFI_{BİÇİMSEL} - CFI_{SKALAR})$
Matematiği Sevme	0.01
Matematiğe Olan İlgisi	-0.02

Tablo 6 incelendiğinde “Matematiğe olan ilgi” değişkenine ilişkin fark değerinin  $-0.01 \leq \Delta CFI \leq 0.01$  aralığında olmadığı görülmektedir. Yani kız ve erkek öğrenci grupları için “Matematiğe olan ilgi” gizil değişkeni için skalar değişmezlik sağlanamamıştır. Bu bulgu skalar değişmezliğin sağlanamadığı değişkenlerde, gruplar arası fark olduğu anlamına gelmektedir. Bunun yanı sıra “Matematiği sevme” gizil değişkeni için skalar değişmezlik sağlanmıştır. Yani gözlenen maddeler üzerindeki grup farklılıkları bakımından cinsiyetler arası bir karşılaştırma anlamlı olabilecektir.

Ölçme değişmezliğinin son aşaması olan katı değişmezliği için bir önceki aşama olan skalar değişmezlikteki sınırlamalara ek olarak hata varyanslarının gruplar arası eşit olma sınırlaması da eklenmiştir. Skalar değişmezlik sadece “Matematiği sevme” ölçeği için sağlandığından CFI fark testi sadece bu ölçek için yapılacaktır. Ölçme değişmezliği aşamalı bir süreç olduğundan bir aşama sağlanmadığı zaman diğer aşamaya geçiş söz konusu olmaz. Matematiği sevme değişkeni için yapılan katı değişmezlik testi model-veri uyum sonuçları Tablo 7’de gösterilmektedir.

Tablo 7

*Katı Değişmezlik Aşamasına İlişkin Uyum İstatistikleri*

Ölçekler	$\chi^2/df$	RMSEA	CFI	GFI	NFI	NNFI
Matematiğe Sevme	20.76	0.080	0.90	0.96	0.90	0.90

Tablo 7 incelendiğinde “Matematiği sevme” ölçeğinin veri ile uyum gösterdiği görülmektedir. Katı değişmezliğin sağlanıp sağlanmadığına ilişkin yapılan CFI fark sonuçları ise Tablo 8’de gösterilmektedir.

Tablo 8

*Biçimsel Değişmezlik ve Katı Değişmezlik İçin Fark Testleri Tablosu Uyum İstatistikleri*

Ölçekler	$\Delta CFI (CFI_{BİÇİMSEL} - CFI_{KATI})$
Matematiği Sevme	0.01

Tablo 8’deki bilgilere göre “Matematiği sevme” gizil değişkeni için katı değişmezlik sağlanmıştır. Dolayısıyla “Matematiği sevme” gizil değişkeni için gözlenen maddeler üzerindeki grup farklılıkları bakımından cinsiyetler arası bir karşılaştırma yapılması anlamlı olabilecektir. Değişkenler üzerinden yapılacak olan grup karşılaştırmalarının anlamlılığı değişkenlerin değişmezlik aşamalarını sağlayıp sağlayamadığı ile ilgilidir.

Üç ölçek içinde biçimsel değişmezlik sağlanmaktadır. Matematiğe ilişkin özgüven ölçeği için metrik değişmezlik sağlanmazken, matematiği sevme ve matematiğe olan ilgi ölçekleri için metrik değişmezlik sağlanmaktadır. Dolayısıyla bu iki ölçek için cinsiyetler arası yapılacak karşılaştırma çalışmalarının anlamlı olacağı ve alt gruplarda meydana gelen farklılığın ölçme aracından kaynaklanmadığı yorumu yapılabilecektir. İlgili ölçeği için skalar değişmezlik sağlanmamaktadır ve dolayısıyla bir sonraki aşama olan katı değişmezlik testlerine de geçilmemiştir. Buna karşın, matematiği sevme ölçeği skalar değişmezlik ve skalar değişmezlikten sonraki aşama olan katı değişmezlik testlerini de yerine getirmektedir. Dolayısıyla, matematiği sevme ölçeğinde kızlar ve erkekler için sağlanmış olan değişmezlik testleri ile maddelerin faktör yük değerleri eşittir ve örtük değişkenlerin hata varyanslarına ait değerler de erkeklerde ve kızlarda farklılık göstermemektedir. Buna göre araştırmada yer alan üç gizil değişken kullanılarak oluşturulan bir modelde (kız öğrenciler ve erkek öğrenciler için biçimsel değişmezlik dosyasına ait standardize edilmiş değerler) madde puanlarına ilişkin yapılacak karşılaştırmalar anlamlı olabilecektir. Çünkü her üç gizil değişkeninde karşıladığı değişmezlik koşulu biçimsel değişmezliktir. Ancak bu üç değişkenin bulunduğu faktör yapısına göre yapılan karşılaştırmaların anlamlılığı anlamına gelmemelidir. Bunun yanında tüm değişmezlik testlerinin (biçimsel, metrik, skalar ve katı) sağlandığı gizil değişken olan matematiği sevme değişkeni için yapılacak olan tüm

karşılaştırmaların anlamlı olacağı sonucuna ulaşılabilir. Böylece matematiği sevmeye gizil değişkeninin kızlarda ve erkeklerde matematik başarısını açıklayan değişken olduğu yorumu yapılabilir.

### Sonuç ve Tartışma

Bu çalışmada, TIMSS 2015 matematik uygulamasına Türkiye’den katılan öğrencilerin matematiği sevmeye, matematiğe olan ilgi ve matematiğe ilişkin özgüven değişkenlerinin cinsiyete göre ölçme değişmezliği test edilmiştir. Alanyazın incelendiği zaman matematiğe ilişkin özgüven değişkeninin bazı çalışmalarda başarıyı pozitif yönde etkilediği (Doğan & Barış, 2010; Pajares & Kranzler, 1995) görülürken, bazı çalışmalarda ise matematik başarısını negatif yönde (Çavdar, 2015; Ölçüoğlu, 2015; Stankov & Lee, 2014) yordadığı bulunmuştur. Ayrıca matematiği sevmeye ve matematiğe olan ilgi de matematik başarısını yordamada önemli değişkenler olarak karşımıza çıkmaktadır (Akyüz, 2014; Khine, Al-Mutawah, & Afari, 2015; Oral & McGivney, 2013; Önen, 2018). Birçok araştırmada matematik ile ilgili duyuşsal özelliklerin incelendiği göz önüne alınarak, cinsiyet açısından bu değişkenlerin ölçme değişmezliğini sağlayıp sağlamadığı önemli görülmektedir. Yapılan bu çalışmada üç duyuşsal özelliğin, cinsiyete göre ölçme değişmezlik koşullarının hepsini yerine getirmediği görülmüştür.

Ölçme değişmezliğinin ilk aşaması olan biçimsel değişmezlikte modele ilişkin kavramsal yapı test edilmektedir. Matematiği sevmeye, matematiğe olan ilgi ve matematiğe ilişkin özgüven gizil değişkenleri için kurulan modellerin veri uyumu sağlanmaktadır. ÇG-DFA ile analizi yapılan bu üç değişkenin de biçimsel değişmezlik aşamasını sağladığı bulunmuştur. Böylece kız ve erkek grupların almış oldukları testteki sevmeye, ilgi ve özgüven ölçeklerine ilişkin kavramsal yapıları değişmezdir. Gülleroğlu (2017) tarafından, PISA 2012 verileri kullanılarak matematiğe ilişkin duyuşsal özelliklerin cinsiyete göre ölçme değişmezliğinin incelendiği çalışmada matematiğe yönelik öz yeterlik değişkeni biçimsel değişmezlik modelini sağlamamıştır. Buna karşılık, Uzun’un (2008) TIMSS 1999 Türkiye örneklemindeki fen başarısını etkileyen değişkenlerin cinsiyete göre ölçme değişmezliğini incelediği çalışmasında ise duyuşsal özelliklerden özgüven değişkeninin biçimsel değişmezlik aşamasını sağladığı görülmüştür. Bu araştırmadan elde edilen sonuçlar, Uzun’un (2008) yaptığı çalışmanın sonuçları ile benzerlik göstermektedir.

Metrik model ile biçimsel değişmezlikte yer alan sınırlandırılmamış model arasında yapılan CFI fark testi sonucunda modeldeki “Matematiği sevmeye” ve “Matematiğe olan ilgi” gizil değişkenleri metrik değişmezlikteki sınırlamaları karşılamıştır. Böylece bu iki değişken metrik bakımdan erkek ve kızlar için değişmez bulunmuştur. Buna karşılık “Matematiğe ilişkin özgüven” değişkeni ise metrik değişmezliği karşılayamamıştır. Yani, “Matematiğe ilişkin özgüven” değişkeni metrik model bakımdan cinsiyet gruplarında değişmez değildir. Matematiğe ilişkin özgüven değişkeninin metrik model bakımından cinsiyet grupları arasında değişmezliği sağlayamamasının nedeni, kız ve erkek öğrencilerin matematiğe yönelik öz yeterliliklerinin yapısal olarak farklı olmasından kaynaklanabilmektedir (Gülleroğlu, 2017).

Skalar değişmezlik için yapılan sınırlamanın karşılandığı tek gizil değişken “Matematiği sevmeye” değişkenidir. Dolayısıyla modele alınan değişkenlerden skalar

değişmezliği sağlamayan iki değişken olan matematiğe ilişkin özgüven ve matematiğe olan ilgi değişkenleri için yapılacak sınırlamanın değişmezlik çalışmalarında aşamalılık ilkesinden ötürü anlamsız olduğu belirtilmiştir.

Katı değişmezlik aşamasında test edilebilen değişken “Matematiği sevme” değişkenidir. Matematiği sevme değişkeni ölçme değişmezliğinin dördüncü ve son aşaması olan en sınırlayıcı modelin kurulduğu katı değişmezliği de sağlamıştır. Dolayısıyla ölçme aracından elde edilen matematiği sevme gizil değişkenine ilişkin puanlara dayalı olarak kız ve erkek öğrencilerin belirtilen ölçme modeli çerçevesinde karşılaştırılmasının uygun olacağı söylenebilir. Yapılan testler sonucunda tüm gizil değişkenler ile kurulan bir modelin kız ve erkek gruplar arası karşılaştırılmasının anlamlı olacağı aşama biçimsel değişmezlik aşamasıdır.

Sonuç olarak, bu çalışmanın asıl odağı olan “TIMSS 2015’e katılan Türkiye dördüncü sınıf öğrencilerinin matematik başarıları üzerinde etkili olduğu düşünülen matematiğe olan ilgi, matematiği sevme ve matematiğe ilişkin özgüven gizil değişkenlerinin kız ve erkek öğrenci gruplarında ölçme değişmezlik düzeyleri nasıldır?” sorusu şu biçimde cevaplanabilir: Tüm gizil değişkenler en azından biçimsel değişmezliğe sahiptir; ölçme değişmezliğinde katı değişmezliğe kadar sağlanması gereken tüm koşulları sağlayabilen tek değişken “Matematiği sevme” gizil değişkenidir. Bu değişken üzerinden yapılacak olan karşılaştırmalar anlamlı olacaktır. Ölçüoğlu (2015) yaptığı çalışmasında matematik başarıları ile ilişkili olan duyuşsal özellikler (matematiği sevme, matematiğe olan ilgi ve matematiğe ilişkin özgüven) ile ev ortamı ve okul ortamı değişkenlerinin Türkiye’deki yedi coğrafi bölgeye göre ölçme değişmezliğini incelemiş ve tüm gizil değişkenlerin metrik değişmezlik modelini sağladığı fakat üçüncü aşama olan skalar değişmezlik aşamasını sağlamadığı sonucuna ulaşmıştır. Uzun ve Öğretmen’in (2010) yaptıkları çalışmada öğrencilerin fen başarısına etki eden duyuşsal özelliklerin cinsiyete göre ölçme değişmezliğini sağlayıp sağlamadığını test etmiştir. Araştırmadan elde edilen sonuçlar, öz yeterlik, tutum, önem ve sınıf içi öğrenme etkinlikleri değişkenlerinin metrik değişmezlik koşulunu sağladığını, fakat cinsiyet grupları arasında katı değişmezlik koşulunu hiçbirinin sağlamadığını göstermiştir.

Bu çalışmada modele dâhil edilmeyen değişkenlerin bağımsız değişkenle ilişkileri nedensellik ilişkisini etkilemiş olabilir. Bu nedenle dikkate alınmayan değişkenler açısından da değişmezlik testleri ile karşılaştırmalar yapılmasına ihtiyaç olduğu düşünülmektedir. Matematik başarılarını yordayan değişkenlerin cinsiyete göre ölçme değişmezliğinin incelendiği bu çalışmada, matematiğe ilişkin özgüven ve matematiğe olan ilgi değişkenlerinin skalar ve katı değişmezlik aşamalarını neden sağlayamadıklarına ilişkin ek çalışmalar yapılabilir. Gelecek çalışmalarda öğretmen ve okullara ilişkin anketler kullanılarak benzer araştırmalar yapılması ve değişmezlik testleri ile zenginleştirilmesi, mesleki eğitim ve bölge gelişimi açısından önemli bilgiler sağlayabilir. Matematik başarıları ile ilişkili olan duyuşsal özelliklerin cinsiyete göre ölçme değişmezliği Türkiye örneğinde bütün değişkenler için sağlanamamıştır. Türkiye ile sosyal ve coğrafi yönden benzerlik gösteren başka ülkelerde de bu çalışmada yer alan duyuşsal özelliklerin ölçme değişmezliği çalışmaları yapılarak, ülkeler arası karşılaştırmalar yapılabilir. Ülkemizin bütününe yönelik anlamlı genellemeler yapılabilmesi için değişmezlik çalışması ile farklı il ve bölgeler çapında karşılaştırmalar yapılması, sosyal yapının ve konu alanına yönelik durumun açıklanmasında yararlı



olacaktır. Matematik başarısına etki eden duyuşsal özelliklerin ölçme deęişmezlięinin araştırıldığı bu çalışma, fen başarısına etki eden duyuşsal özellikler ile de yapılabilir.

## Summary

**Purpose and Significance:** The main purpose of this study is to examine whether the measurement model set up with the latent variables; liking of mathematics, interest in mathematics and confidence in mathematics and the items under these latent variables have the same structure for both girls and boys. Although there are many studies in which the structures that effect mathematics achievement are identified, there are few studies dealing with the measurement invariance of these structures for different groups. This study will provide more valid information about the meaningfulness of the comparison studies conducted with the variables in TIMSS 2015 fourth grade student questionnaires in terms of gender.

**Methods:** This study is comprised of testing measurement invariance of a model, which is thought to explain students' achievement with the data gathered from TIMSS 2015 Turkey sample, according to genders. Hence this study is a correlational study because it examines the relationships between the variables. The population of the study is comprised of fourth grade students in Turkey, and the sample of the study consists of the fourth grade students participated in TIMSS 2015 from Turkey. There are totally 6456 fourth grade students in TIMSS 2015 Turkey sample. The analyses were carried out to identify the missing value and the outliers before the measurement invariance test was done. The measurement invariance test was conducted with 6146 students after doing missing value and multivariate outlier analyses. 3049 (49.6%) girl and 3097 (50.4%) boy students in fourth grade participated in the study.

First, the explanatory factor analysis (EFA) was performed to determine the factor structure of each scale; liking mathematics, interest in mathematics and confidence in mathematics using SPSS packaged software. Then, the Structural Equation Model (SEM) analysis was conducted using LISREL. Multigroup Confirmatory Factor Analysis (MGCFA) technique was used with the purpose of testing the measurement invariance of the measurement instrument in terms of the gender groups. In group comparisons, MGCFA enables comparing the factor means leaving the group parameters equal. Group comparisons drew with SEM prepare the way for testing the hypothesis containing latent variables rather than the investigations with observed variables (Başusta & Gelbal, 2015). The models are compared using the differences between the comparative fit indices (CFI) in MGSEM analysis. In this study, the CFI differences were compared between the models from the unconstrained to the most constrained pursuant to the basic model. It is tested whether the invariance conditions,  $-0.01 \leq \Delta CFI \leq 0.01$ , are provided for the MG-CFA variables fitting with the data, or not.  $\chi^2/df$ , GFI, CFI, NFI, NNFI and RMSEA indices were reported for determining if the measurement models set in each level of the invariance test provided the model fit or not. The ratio of  $\chi^2/df$  should be fewer than 5, GFI should be more than 0.90, RMSEA should be fewer than 0.10, and CFI, NFI and NNFI indices should be between 0.90 and 0.95 to provide model fit.

**Results:** Explanatory factor analysis was done on the measurements taken by each of the latent variables relating to student achievement in TIMSS student questionnaires; liking mathematics, interest in mathematics and confidence in mathematics. Principle

components analysis was done without limiting the factor number, and no factor rotation was performed. The Kaiser-Meyer-Olkin coefficient and the results from Barlett's test of sphericity were used to assess the fit of the data in with factor analysis. According to EFA results, liking mathematics scale has totally 8 items and a single factor, interest in mathematics scale has 7 items and a single factor, and confidence in mathematics scale has 8 items in total and a single factor.

In testing for measurement invariance, CFI differences between configural invariance and metric invariance, and scalar invariance and strict invariance for each latent variable were calculated and interpreted.

$\chi^2/df$ , GFI, CFI, NFI, NNFI and RMSEA indices of the models set up for liking mathematics, interest in mathematics and confidence in mathematics latent variables are acceptable in configural invariance, the first step of measurement invariance. Accordingly, the model-fits of these three latent variables were provided. The difference between CFI values in metric model and configural model was examined for metric invariance. The difference for the metric invariance should be between  $-0.01 \leq \Delta CFI \leq 0.01$ . When CFI values were analyzed, it was seen that liking mathematics and interest in mathematics variables provided metric invariance, but confidence in mathematics didn't.

In scalar invariance, the third step of the measurement invariance, the restriction that items have the same factor loads are identic between the groups is introduced in addition to the restrictions in metric invariance.

Liking mathematics and interest in mathematics variables which were metrically invariant were tested adding the scalar invariance restriction. The model-fit indices of RMSEA, CFI, GFI, NFI and NNFI values for liking mathematics and interest in mathematics are in the acceptable limits. The CFI value between the configural invariance and scalar invariance was checked whether liking mathematics and interest in mathematics variables provided the scalar invariance. It was seen that the difference value of interest in mathematics variable isn't between  $-0.01 \leq \Delta CFI \leq 0.01$ . In other words, the scalar invariance for interest in mathematics latent variable between girl and boy student groups isn't provided. However, the scalar invariance for liking mathematics latent variable is provided. Namely, a comparison between gender groups in point of group differences on observed items will be meaningful.

For the strict invariance, the last step of measurement invariance, the restriction of the equivalence of the error variances between the groups was added to the working file in which the scalar invariance was tested earlier. Because the scalar invariance was provided just for liking mathematics variable, the difference test was done only for this variable. Liking mathematics latent variable provided the model-fit. The strict invariance for this variable was also provided. Thus, a comparison between gender groups in point of group differences on observed items will be meaningful.

**Discussion and Conclusions:** As a result, the main focus of this research, whether the comparison of TIMSS variables between girl and boy groups is meaningful, can be answered as: all the latent variables have configural invariance at least, and the only variable which provides all the conditions up to strict invariance in measurement invariance is liking mathematics. The comparisons on this variable will be meaningful.

### Kaynakça

- Akyıldız, M. (2009). PIRLS 2001 testinin yapı geçerliliğinin ülkelerarası karşılaştırılması. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 6(1), 18-47.
- Akyüz, G. (2014). The effects of student and school factors on mathematics achievement in TIMSS 2011. *Education and Science*, 39(172), 150-162.
- Anderson, J. C., & Gerbing, D.W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*, 49, 155-173.
- Bahadır, E. (2012). *Uluslararası Öğrenci Değerlendirme Programı'na (PISA 2009) göre Türkiye'deki öğrencilerin okuma becerilerini etkileyen değişkenlerin bölgelere göre incelenmesi* (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi, Ankara.
- Başusta, N. B. U. (2010). Ölçme eşdeğerliği. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 1(2), 58-64.
- Başusta, N. B., & Gelbal, S. (2015). Gruplararası karşılaştırmalarda ölçme değişmezliğinin test edilmesi: PISA öğrenci anketi örneği. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 30(4), 80-90.
- Bollen, K.A. (1989). *Structural equations with latent variables*. New York: John Wiley & Sons, Inc.
- Büyüköztürk, Ş. (2015). *Sosyal bilimler için veri analizi el kitabı: İstatistik, araştırma deseni, SPSS uygulamaları ve yorum* (21.Baskı). Ankara: Pegem Akademi.
- Büyüköztürk, Ş., Çakmak, E. K., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2014). *Bilimsel araştırma yöntemleri* (16. Baskı). Ankara: Pegem Akademi.
- Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. New Jersey, NY: Lawrence Erlbaum.
- Chen, F. F., Sousa, K. H., & West, S. G. (2005). Testing measurement invariance of second-order factor models. *Structural Equation Modeling*, 12(3), 471-492.
- Cheung, G. W., & Rensvold, R. B. (2000). Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling. *Journal of Cross-Cultural Psychology*, 31(2), 187-212.
- Cole, D. A. (1987). Utility of confirmatory factor analysis in test validation research. *Journal of Consulting and Clinical Psychology*, 55(4), 1019-1031.
- Cook, R. D. (1977). Detection of influential observation in linear regression. *Technometrics*, 19(1), 15-18.
- Çavdar, D. (2015). *TIMSS 2011 matematik başarısının öğrenci ve öğretmen özellikleri ile ilişkisi* (Yayımlanmamış yüksek lisans tezi). Gazi Üniversitesi, Ankara.
- Doğan, N., & Barış, F. (2010). Tutum, değer ve özyeterlik değişkenlerinin TIMSS-1999 ve TIMSS-2007 sınavlarında öğrencilerin matematik başarılarını yordama düzeyleri. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 1(1), 44-50.
- Eurydice. (2010). *Eğitim çıktılarında cinsiyet farklılıkları: Avrupa'da alınan tedbirler ve mevcut durum*. Eurydice Raporu, Brüksel.

[http://eacea.ec.europa.eu/education/eurydice/documents/thematic\\_reports/120TR.pdf](http://eacea.ec.europa.eu/education/eurydice/documents/thematic_reports/120TR.pdf)

- Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The applications of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*, 39, 291-314.
- Garson, G. D. (2015). *Missing value analysis and data imputation*. Ashebora, NC: Statistical.
- George, D., & Mallery P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.)*. Boston: Allyn & Bacon.
- Gülleroğlu, H. D. (2017). PISA 2012 matematik uygulamasına katılan türk öğrencilerin duyuşsal özelliklerinin cinsiyete göre ölçme değışmezliğinin incelenmesi. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 37(1), 151-175.
- Hogarty, K. Y., Hines, C. V., Kromrey, J. D., Ferron, J. M., & Mumford, K. R. (2005). The quality of factor solutions in exploratory factor analysis: The influence of sample size, communality and over determination. *Educational and Psychological Measurement*, 65(2), 202-226.
- Hong, S., Malik, M. L., & Lee, M. K. (2003). Testing configural, metric, scalar, and latent mean invariance across genders in sociotropy and autonomy using a non-Western sample. *Educational and Psychological Measurement*, 63(4), 636-654.
- Horn, J. L., McArdle, J.J., & Mason, R. (1983). When is invariance not invariant: A practical scientist's look at the ethereal concept of factor invariance. *The Southern Psychologist*, 1(4), 179-188.
- Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Scientific Software International.
- Jöreskog, K. G., & Sörbom, D. (1999). *LISREL 8.30 and PRELIS 2.30*. Chicago,IL: Scientific Software International.
- Karakoc-Alatlı, B., Ayan, C., Polat-Demir, B., & Uzun, G. (2016). Examination of the TIMSS 2011 fourth grade mathematics test in terms of cross-cultural measurement invariance. *Eurasian Journal of Educational Research*, 66, 389-406.
- Khine, M. S., Al-Mutawah, M., & Afari, E. (2015). Determinants of affective factors in mathematics achievement: Structural equation modeling approach. *Journal of Studies in Education*, 5(2), 199-211.
- Leder, G. C., & Forgasz, H. J. (2006). Affect and mathematics education: PME perspectives. In A. Gutiérrez & P. Boero (Eds.), *Handbook of research on the psychology of mathematics education: Past, present and future* (1st ed., pp. 403-427). Rotterdam, The Netherlands: Sense Publishers.
- Mark, B. A., & Wan, T.T.H (2005). Testing measurement equivalence in a patient satisfaction instrument. *Western Journal of Nursing Research*, 27(6), 772-787.
- Marsh, H.W., & Hocevar, D. (1988). A new more powerful approach to multitrait multi method analyses: Application of second-order confirmatory factor analysis. *Journal of Applied Psychology*, 73(1), 107-117.  
<http://psycnet.apa.org/index.cfm?fa=buy.optionToBuy&id=1989-07244-001.pdf>.

- Marsh, H.W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, 103(3), 391-410.
- Meredith, W. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika*, 58(4), 525-543.
- Millî Eğitim Bakanlığı Eğitimi Geliştirme ve Araştırma Dairesi Başkanlığı, (2003). TIMSS 1999 Ulusal Raporu. [http://timss.meb.gov.tr/?page\\_id=25](http://timss.meb.gov.tr/?page_id=25).
- Oral, I., & McGivney, E. (2013). *Türkiye’de matematik ve fen bilimleri alanlarında öğrenci performansı ve başarının belirleyicileri TIMSS 2011 analizi*. İstanbul: Eğitim Reformu Girişimi Raporu.
- Ölçüoğlu, R. (2015). *TIMSS 2011 Türkiye sekizinci sınıf matematik başarısını etkileyen değişkenlerin bölgelere göre incelenmesi* (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi, Ankara.
- Önen, E. (2009). *Ölçme değişmezliğinin yapısal eşitlik modelleme teknikleri ile incelenmesi* (Yayımlanmamış doktora tezi). Ankara Üniversitesi, Ankara.
- Önen, E. (2018). Öğrenci, öğretmen ve öğretimsel nitelikler açısından TIMSS-2015’e dayalı olarak öğrencilerin sınıflandırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 9(1), 64-84.
- Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving. *Contemporary Educational Psychology*, 20(4), 426-443.
- Schulz, W. (2003). *Validating questionnaire constructs in international studies. Two examples from PISA 2000*. Paper Presented at the Annual Meetings of the American Educational Research Association (AERA) in Chicago, 21-25 April.
- Schulz, W. (2005). *Testing parameter in variance for questionnaire indices using confirmatory factor analysis and item response theory*. Paper Presented at the Annual Meetings of the American Educational Research Association (AERA) in San Francisco, 7-11 April.
- Schulz, W. (2008). *Questionnaire construct validation in the international civic and citizens hip education study*. Paper presented to the 3rd IEA International Research Conference in Taipei, September.
- Stankov, L., Morony, S., & Lee, Y. P. (2014). Confidence: The best non-cognitive predictor of academic achievement? *Educational Psychology*, 34(1), 9-28.
- Steenkamp, J. B. E., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, 25(1), 78-90.
- Sümer, N. (2000). Yapısal eşitlik modelleri: Temel kavramlar ve örnek uygulamalar. *Türk Psikoloji Yazıları*, 3(6), 49-74.
- Tabachnick, G. B., & Fidell, L. S. (2001). *Using multivariate statistics*. Needham Heights, CA: Allyn & Bacon, Inc.
- Tucker, K. L., Ozer, D. J., Lyubomirsk, S., & Boehm, J. K. (2006). Testing for measurement invariance in the satisfaction with life scale: A comparison of Russians and North Americans. *Social Indicators Research*, 78(2), 341-360.
- Turkan, A., Uner, S., & Alci, B. (2015). 2012 PISA matematik testi puanlarının bazı değişkenler açısından incelenmesi. *Ege Eğitim Dergisi*, 16(2), 358-372.

- Tyson, H. E. (2004). Ethnic differences using behavior rating scales to assess the mental health of children: a conceptual and psychometric critique. *Child Psychiatry and Human Development*, 34(3), 167-201.
- Uzun, N. B. (2008). *TIMSS-R Türkiye örnekleminde fen başarısını etkileyen değişkenlerin cinsiyetler arası değişmezliğinin değerlendirilmesi* (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi, Ankara.
- Uzun, B., & Öğretmen, T. (2010). Fen başarısı ile ilgili bazı değişkenlerin TIMSS-R Türkiye örnekleminde cinsiyete göre ölçme değişmezliğinin değerlendirilmesi. *Eğitim ve Bilim*, 35(155), 26-35.
- Yayan, B. (2003). *Üçüncü uluslararası matematik ve fen çalışması-tekrar (TIMSS-R) daki matematik başarısının kültürler arası karşılaştırması* (Yayımlanmamış yüksek lisans tezi). Orta Doğu Teknik Üniversitesi, Ankara.
- Yılmaz, H. B., & Aztekin, S. (2012). *Türkiye'deki 15 yaş grubu öğrencilerin matematik okuryazarlığı başarılarını etkileyen bazı faktörlerin okul ve öğrenci düzeyine göre incelenmesi*. X. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi, Niğde.



This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0). For further information, you can refer to <https://creativecommons.org/licenses/by-nc-sa/4.0/>