



# HACETTEPE ÜNİVERSİTESİ EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Eğitim Bilimleri Ana Bilim Dalı  
Eğitimde Ölçme ve Değerlendirme Programı

TIMSS 2015 FEN DUYUŞSAL ÖZELLİKLERİNİN CİNSİYET  
VE BÖLGELERE GÖRE İNCELENMESİ

Mehmet Can DEMİR

Yüksek Lisans Tezi

Ankara, 2020

Liderlik, arařtırma, inovasyon, kaliteli eęitim ve deęiřim ile

*Daha ileriye ... En İyiyeye ...*



**HACETTEPE ÜNİVERSİTESİ**  
**EĞİTİM BİLİMLERİ ENSTİTÜSÜ**

Eğitim Bilimleri Ana Bilim Dalı  
Eğitimde Ölçme ve Değerlendirme Programı

TIMSS 2015 FEN DUYUŞSAL ÖZELLİKLERİNİN CİNSİYET  
VE BÖLGELERE GÖRE İNCELENMESİ

AN EXAMINATION OF TIMSS 2015 SCIENCE AFFECTIVE FACTORS  
WITH REGARD TO GENDER AND REGIONS

Mehmet Can DEMİR

Yüksek Lisans Tezi

Ankara, 2020

## Öz

Dünya çapında, eğitim programlarının değerlendirilmesi amacıyla birçok uluslararası ve ulusal sınav uygulanmaktadır. Bu sınavlardan biri olan Uluslararası Matematik ve Fen Eğilimleri Araştırması (TIMSS) ile yaklaşık 250.000 öğrencinin bilgi ve becerileri test edilmektedir. Test sonuçları ile uluslararası ve ulusal düzeyde birçok karşılaştırma yapılmaktadır. Fakat, bu kadar büyük bir kitlenin sonuçlarının anlamlı bir şekilde karşılaştırılabilmesi için sınav sonuçlarının başarıdan farklı bir değişkenden ötürü farklılık göstermemesi, aynı ölçüm sonuçlarını yansıtması gereklidir. Bundan hareketle, araştırma kapsamında TIMSS 2015 Türkiye uygulamasına katılan 8. sınıf öğrencilerinin fen duyuşsal özellik değişkenlerinden öğrenci anketine verilen yanıtlar kullanılarak Yapısal Eşitlik Modellemesi (YEM) aracılığıyla bir model oluşturulmuştur. Daha sonra bu modelin cinsiyet ve İstatistiki Bölge Birimleri Sınıflaması (İBBS) Düzey 1'e göre Türkiye'de yer alan bölgelere göre ölçme değişmezliği, çok gruplu doğrulayıcı faktör analizi (ÇGDFA) ile incelenmiştir. Araştırmanın örneklemini TIMSS 2015 Türkiye uygulamasına katılan 5344 öğrenciden oluşmaktadır. Analizler SPSS 23.0 ve AMOS 23.0 programları ile yürütülmüştür. Çalışma sonucunda, oluşturulan modelin cinsiyetler arasında katı değişmezlik koşulunu ancak bölgeler arasında ölçek değişmezliği koşulunu sağladığı gözlenmiştir.

**Anahtar sözcükler:** TIMSS, fen duyuşsal özellikleri, yapısal eşitlik modellemesi, ölçme değişmezliği ve çok gruplu doğrulayıcı faktör analizi.

## **Abstract**

Various international and national tests are applied to assess curricula across the world. About 250.000 students' knowledge and skills are tested via Trends in International Mathematics and Science Study (TIMSS). Many international and national comparisons are made with the test results. However, to be able to compare the test results of such a big crowd of people, the test results must not differ because of any other variable but achievement; in short, they must represent the same results of measurement. In the light of this information, the purpose of the study is to examine the measurement invariance of the model of the science affective factors of the 8th graders participated in TIMSS 2015 Turkey which is constructed via Structural Equation Modelling (SEM), among Nomenclature of Territorial Units for Statistics – Level 1 (NUTS-1) regions and genders in Turkey, through Multi-Group Confirmatory Factor Analysis (MG-CFA). The sample of this research consists of 5344 students and the analyses was made by using SPSS 23.0 and AMOS 23.0 softwares. As a result of this study, the model met strict invariance for genders and scalar invariance for regions.

**Keywords:** TIMSS, science affective factors, structural equation modeling, measurement invariance, and multi-group confirmatory factor analysis.

## Teşekkür

Gerek ders dönemi gerekse tez dönemi boyunca hiçbir sorumu cevapsız bırakmayan, karşılaştığım sorunlara çözüm üretmek için elinden gelen tüm yardımı sunan ve desteğini hep hissettiğim değerli hocam ve akademik danışmanım Prof. Dr. Selahattin GELBAL'a,

Tez savunma jürimdeki yapıcı önerileri ile tezime yaptıkları önemli katkılar için değerli hocalarım Prof. Dr. Cem Oktay GÜZELLER ve Prof. Dr. Nuri DOĞAN'a,

Derslerinde öğrettikleri bilgiler ve kazandırdıkları tecrübeler için değerli hocalarım Prof. Dr. Hülya KELECİOĞLU ve Doç. Dr. Burcu ATAR'a,

Tez çalışmam için gerekli verileri paylaşarak araştırmamı yürütmemi sağlayan Millî Eğitim Bakanlığı Ölçme Değerlendirme Sınav Hizmetleri Genel Müdürlüğü Veri Analizi İzleme ve Değerlendirme Daire Başkanlığı'na,

Kısa süre önce tanışmış olsak da ilk günden itibaren beni aileden biri olarak kabul ederek desteklerini sunan değerli çalışma arkadaşlarım Dr. Öğr. Üyesi Sema SULAK, Dr. Öğr. Üyesi Beyza AKSU DÜNYA ve Arş. Gör. Eda AKDOĞDU'ya,

Tek tek isimlerini yazmaya kalksam birilerinin eksik kalacağını bildiğim ancak bu yolda koşullar ne olursa olsun bana yardım eden, inanan ve güç veren tüm dostlarıma,

Bugünlere gelmem için sayısız fedakarlıkta bulunan, her zaman yanımda olan ve desteklerini hiçbir zaman esirgemeyen sevgili aileme

teşekkürlerimi sunarım.

## İçindekiler

Öz.....	ii
Abstract.....	iii
Teşekkür.....	iv
Tablolar Dizini.....	vii
Şekiller Dizini.....	viii
Simgeler ve Kısaltmalar Dizini.....	ix
Bölüm 1 Giriş.....	1
Problem Durumu.....	1
Araştırmanın Amacı ve Önemi.....	3
Araştırma Problemi.....	5
Sınırlılıklar.....	6
Tanımlar.....	6
Bölüm 2 Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar.....	7
Ölçme Değişmezliği.....	7
Yapısal Eşitlik Modellemesi.....	9
Açımlayıcı Faktör Analizi.....	13
Doğrulayıcı Faktör Analizi.....	14
Çok Gruplu Doğrulayıcı Faktör Analizi.....	15
İlgili Araştırmalar.....	21
Bölüm 3 Yöntem.....	26
Araştırmanın Yöntemi.....	26
Araştırmanın Evreni ve Örnekleme.....	26
Veri Toplama Araçları ve Süreci.....	27
Verilerin Analizi.....	28
Bölüm 4 Bulgular ve Yorumlar.....	41
Bölüm 5 Sonuçlar ve Öneriler.....	49

Sonuçlar .....	49
Öneriler .....	51
Kaynaklar .....	53
EK-A: Etik Komisyonu Onay Bildirimi .....	60
EK-B: Veri Talebi Dilekçesi.....	61
EK-C: Veri Teslim Yazısı .....	62
EK-Ç: Etik Beyanı.....	63
EK-D: Yüksek Lisans/Doktora Tez Çalışması Orijinallik Raporu .....	64
EK-E: Thesis/Dissertation Originality Report.....	65
EK-F: Yayımlama ve Fikrî Mülkiyet Hakları Beyanı .....	66



## Tablolar Dizini

Tablo 1 <i>Uyum İndeksleri</i> .....	12
Tablo 2 <i>Ölçme Değişmezliği İncelemesinde Kullanılan Parametreler</i> .....	20
Tablo 3 <i>Bölgelere Göre Öğrenci Sayıları</i> .....	27
Tablo 4 <i>Boyut Matrisi</i> .....	35
Tablo 5 <i>Duyuşsal Özellik Modeli Uyum Değerleri</i> .....	36
Tablo 6 <i>Düzenlenen Modelin Uyum Değerleri</i> .....	39
Tablo 7 <i>Cinsiyete Göre Model Uyum Değerleri</i> .....	41
Tablo 8 <i>Cinsiyete Göre Ölçme Değişmezliği Uyum Değerleri</i> .....	42
Tablo 9 <i>Bölgelere Göre Model Uyum Değerleri</i> .....	45
Tablo 10 <i>Bölgelere Göre Ölçme Değişmezliği Uyum Değerleri</i> .....	46

## Şekiller Dizini

Şekil 1. TIMSS 2015 Türkiye 8. sınıf fen bilimleri bölge ortalamaları.....	4
Şekil 2. Yapısal eşitlik modelinin oluşturulması .....	13
Şekil 3. Yapısal değişmezlik .....	16
Şekil 4. Metrik değişmezlik .....	18
Şekil 5. Ölçek değişmezliği.....	19
Şekil 6. Katı değişmezlik .....	20
Şekil 7. Yamaç-Birikinti Grafiği .....	34
Şekil 8. Standartlaştırılmamış faktör yükleri ve hatalar.....	38
Şekil 9. Düzenlenen modele ilişkin standartlaştırılmış faktör yükleri.....	40

## **Simgeler ve Kısaltmalar Dizini**

- AERA:** American Educational Research Education
- AFA:** Açımlayıcı Faktör Analizi
- APA:** American Psychological Association
- CFI:** Comparative Fit Index
- ÇGDFA:** Çok Gruplu Doğrulayıcı Faktör Analizi
- DFA:** Doğrulayıcı Faktör Analizi
- IEA:** International Association for the Evaluation of Education Achievement
- IFI:** Incremental Fit Index
- KTK:** Klasik Test Kuramı
- MEB:** Millî Eğitim Bakanlığı
- MTK:** Madde Tepki Kuramı
- NCME:** National Council on Measurement in Evaluation
- OECD:** Organization for Economic Co-operation and Development
- PIRLS:** Progress in International Reading Literacy Study
- PISA:** Programme for International Student Assessment
- RMSEA:** The Root Mean Square Error of Approximation
- SRMR:** Standardized Root Mean Squared Residual
- TIMSS:** Trends in International Mathematics and Science Study
- TLI:** Tucker Lewis Index
- YEM:** Yapısal Eşitlik Modellemesi

## **Bölüm 1**

### **Giriş**

Bu bölümde önce araştırmanın problem durumu belirtilmiş, sonra araştırmanın amacı ve önemi açıklanmış, daha sonra araştırmanın problem cümlesi ile alt problemleri ifade edilmiş ve son olarak da araştırmanın sayıtları ve sınırlılıkları verilmiştir.

#### **Problem Durumu**

Eğitimdeki başarı, uluslararası rekabetteki konumu belirlemede ve toplumsal kalkınmada kritik öneme sahiptir. Geçmişten bugüne dek gerçekleşen her bilimsel ve teknolojik ilerleme, yaşamımızı ekonomik ve sosyal açıdan kritik bir şekilde değiştirmiştir. Gün geçtikçe teknolojiye bağlı değişimler artmaktadır. Buna paralel bir şekilde değişimlerin hayatımıza etkisi de artmakta ve bu etkiler daha somut hale gelmektedir. Teknolojik ilerlemeyi fenden bağımsız olarak düşünmek mümkün olmadığı için çoğu ülke fen bilimleri okuryazarlığının ülkenin ekonomik iyileşmesi üzerinde ne kadar önemli olduğunu anlamıştır. Bu nedenle ülkeler, son yıllarda eğitimde izlemiş oldukları politikaları sürekli olarak değerlendirmekte ve değerlendirme sonuçlarına göre geliştirmektedir.

Teknolojik gelişmelerin kolaylaştırıcı etkisiyle bugün eğitim alanında, ulusal ve uluslararası karşılaştırmalar içeren birçok çalışma (ÖBBS, ABİDE, TIMSS, PISA, PIRLS, vb.) yapılmaktadır. Bu çalışmaların sonuçları neticesinde ülkeler, kendilerinin ve diğer ülkelerin eğitimdeki başarı düzeyleri hakkında bilgi sahibi olmaktadır. Ülkeler, aynı zamanda kendi içerisindeki çeşitli özellikler açısından farklı grupların (sosyoekonomik düzey, kültür, cinsiyet vb.) başarı durumları hakkındaki bilgilere de ulaşmakta ve bu bilgilerle eğitim sistemlerini farklı açılardan iyileştirme/geliştirme imkânı bulmaktadır.

IEA (The International Association for the Evaluation of Educational Achievement) tarafından düzenlenen TIMSS (The Trends in International Mathematics and Science Study: Uluslararası Matematik ve Fen Eğilimleri Araştırması), 4. ve 8. sınıfa devam öğrencilerinin fen bilimleri ve matematik alanlarındaki başarılarının değerlendirildiği uluslararası bir tarama araştırmasıdır (Mullis ve Martin, 2013). TIMSS aracılığıyla, uluslararası düzeyde, 4. ve 8. sınıf

düzeyindeki öğrencilerin matematik ve fen bilimleri derslerinde kazanılan bilgi ve becerileri değerlendirilmekte ve öğrencilerin okulda öğrendiklerini gündelik yaşamda ne ölçüde kullanabildikleri ile üst düzey zihinsel becerilere sahip olma durumlarının belirlenmesi amaçlanmaktadır. TIMSS döngüleri kapsamında 1995 yılından itibaren her dört yılda bir veri toplanmıştır. Türkiye, 1995 ve 2003 yıllarındaki uygulamalara katılmamış; 1999 ve 2007 yıllarındaki uygulamalara 8. sınıf düzeyinde; 2011 ve 2015 yıllarındaki uygulamalara ise hem 4. hem de 8. sınıf düzeyinde katılmıştır (MEB, 2016).

TIMSS yalnızca öğrencilerin matematik ve fen başarı testlerinden oluşmamaktadır, çalışma kapsamında öğrenci anketleri de yer almaktadır. Öğrenci anketleri ile öğrencilerin sosyoekonomik düzeyleri, ev ve okuldaki eğitim olanakları, okul iklimi, vb. hakkında birçok detaylı bilgi elde edilmektedir. TIMSS'te kullanılan öğrenci anketinde ilgi, motivasyon, tutum, okula aidiyet, akran zorbalığı, derse verilen değer gibi öğrenci başarısı üzerinde etkisi olan duyuşsal özelliklere ilişkin maddeler de yer almaktadır (Mullis, Martin, Foy & Hooper, 2016).

Sosyal bilimler ve eğitim alanında yapılan araştırmalarda, ölçülen özelliğin (örn: başarı) örneklem içerisindeki farklı gruplar arası (örn: cinsiyet, anne baba eğitim düzeyi veya sosyoekonomik düzey) karşılaştırmalarına çokça yer verilmektedir (Cronbach ve Meehl, 1955). TIMSS 2015'te yer alan anketlerden elde edilen bilgilerle, farklı değişkenler açısından ulusal ve uluslararası düzeyde birçok başarı karşılaştırması ve değerlendirmesi yapılmıştır.

TIMSS 2015 Türkiye uygulamasının sonuçları incelendiğinde, İstatistiki Bölge Birimleri Sınıflandırması (İBBS) Düzey 1'de yer alan 12 bölge arasında başarı puanları açısından farklar olduğu görülmüştür. Aynı şekilde, cinsiyete göre karşılaştırılma yapıldığında kız ve erkek öğrencilerin ortalama puanları arasında istatistiksel olarak anlamlı fark olduğu görülmüştür. Bireyler farklı özelliklere sahip olduğu için ölçme sonuçlarının da farklı çıkması doğal bir sonuçtur. Fakat başarı puanları arasındaki farklılığın tek sebebinin bireylerin farklı olması ve yalnızca bununla açıklanması her zaman doğru değildir (Başusta, 2010). Çünkü, farklılığın sebebi ölçülen özelliğin kendisi olabileceği gibi ölçme aracının kendisi de olabilir.

Bir ölçme aracı farklı gruplara uygulanmak üzere geliştirildiyse, gözlenen göstergelerin psikometrik özelliklerinin tüm gruplarda denk olduğunun kanıtlanması gerekmektedir. Yapılan karşılaştırmaların sağlıklı olabilmesi

örneklemedeki farklı alt gruplardaki yer alan ve ölçümlenen değişken açısından aynı yetenek düzeyindeki bireylerin eşit puanlar almasıyla mümkündür (Schmitt ve Kuljanin, 2008). Bu denklik sağlandığı zaman, ölçeklerden alınan puanların gruplar arasında karşılaştırılabilir olduğu yani ölçme değişmezliğinin sağlandığı sonucuna ulaşılır. Ölçme değişmezliği, gizil değişkenlerin göstergeleriyle yani gözlenen değişkenlerle ilişkilerinin tüm gruplar arasında aynı şekilde olmasıdır (Widaman ve Reise, 1997). Ölçme değişmezliği ölçümün elde edildiği bireylerin özellikleriyle değil ölçme aracının kendisiyle ilgilidir. Ölçme değişmezliğinin hangi düzeyde olduğunun kanıtlanmamış olduğu gruplar arasında yapılan karşılaştırmalar her zaman doğru sonuçlar vermeyebilir. Dolayısıyla ölçme değişmezliği sağlanamadığında, gruplar arasında anlamlı karşılaştırma ve/veya değerlendirmeler yapmak mümkün değildir (Öğretmen, 2006).

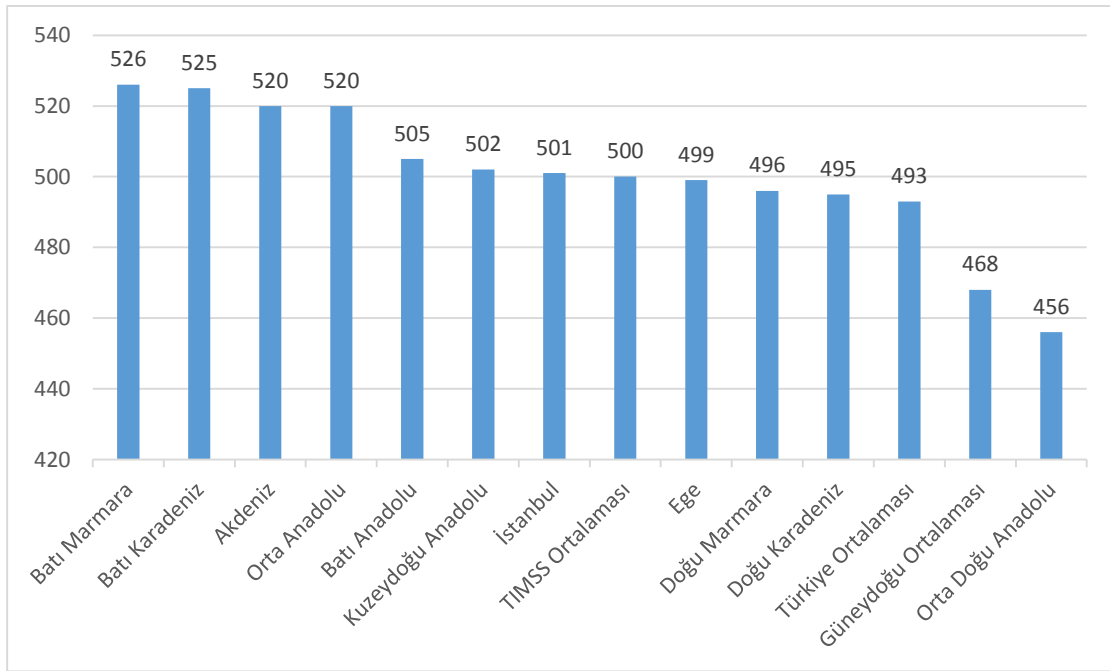
Ölçme değişmezliği, geçerlik kanıtı toplama yöntemlerinden birisi olan yapı geçerliği ile ilgili bir kavramdır. Yapı geçerliği, var olduğu kabul edilen ve ölçülmeye çalışılan özelliğin yani yapının sınırlarının belirlenmesiyle ve gözlenebilir hale getirilebilmesiyle ilgilidir. Ölçme değişmezliğinin sağlanmış olması yapı geçerliği hakkında kanıt sağlanması anlamına gelmektedir. Ölçme değişmezliğinin varlığında, söz konusu yapı tüm alt gruplarda benzer olacağı için ölçme değişmezliği yapı geçerliğinde kritik öneme sahiptir. Ölçme değişmezliğinin sağlanmaması durumunda ölçme sonuçlarının geçerliği ispat edilememiş olacaktır. Bu durumda elde edilen sonuçlarla yapılacak tüm karşılaştırmalar ve yorumların doğruluğu şüpheli olacaktır. Dolayısıyla ölçme değişmezliği çalışmaları, söz konusu araştırmalardan elde edilen sonuçların geçerliği hakkında önemli kanıtlar sağlamasından ötürü büyük bir öneme sahiptir.

### **Araştırmanın Amacı ve Önemi**

Bugüne kadar yapılan farklı birçok ulusal ve uluslararası çalışmada öğrenci başarısını etkileyen faktörler araştırılmıştır. Daha sonra bu faktörler (sosyoekonomik düzey, öğrenci özellikleri, öğretmen özellikleri, vb.) açısından öğrenci başarısı karşılaştırmalarına yer verilmiştir. TIMSS ile asıl amaçlanan fen ve matematik öğretiminin geliştirilmesini sağlamak olsa da araştırmaya ilişkin raporlarda öğrenci başarısı ve duyuşsal özellikleri açısından karşılaştırmalar yapıldığı görülmektedir (Mullis, Martin, Foy ve Hooper, 2016). Ancak,

karşılaştırmaların yapılabilir olup olmadığını incelemeksizin karşılaştırmalar yapıp sonuçları olduğu gibi kabul etmek hatalı kararlara yol açabilir. Karşılaştırmalar yapılmadan önce çalışmaya ilişkin ölçme değişmezliği incelemesi yapılması gerekir. Çünkü ölçme değişmezliği çalışması sonucunda hem karşılaştırma için kullanılacak faktörler hakkında hem de karşılaştırmaların ne derece yapılabilir olduğu hakkında kesin bilgiler elde edilmiş olur.

Şekil 1'de görüldüğü üzere, açıklanan TIMSS 2015 Ulusal Ön Raporu'nda 8. sınıf fen bilimleri başarısına ilişkin bölüm incelendiğinde en yüksek başarı ortalamasına sahip olan Batı Marmara Bölgesi ile en düşük başarı ortalamasına sahip olan Ortadoğu Anadolu Bölgesi arasında 70 puanlık bir fark mevcuttur (MEB, 2016). Erkek ve kız öğrencilerin yeterlik düzeyi puanları incelendiğinde erkek öğrencilerin ortalamasının 484, kız öğrencilerin ortalamasının 503 ve aradaki farkın istatistiksel olarak anlamlı olduğu görülmüştür.



Şekil 1. TIMSS 2015 Türkiye 8. sınıf fen bilimleri bölge ortalamaları

Öğrencilerin ders başarılarını duyuşsal özelliklerinden bağımsız olarak değerlendirmek mümkün değildir. Öğrencilerin derslere duygusal anlamdaki eğilimleri öğrencilerin duyuşsal özelliklerini oluşturmaktadır (Bloom, 2012). Duyuşsal özellikler ilgi, tutum, kaygı, motivasyon, benlik değeri, vb. değişkenlerin bir bileşimi şeklinde yorumlanabilir. Bir derse ilişkin duyuşsal özelliklerin o dersin başarısı üzerinde etkisi olduğu daha önce yapılan araştırmalar ile ortaya

konulmuştur. Bloom (2012)'ye göre öğrenci başarısındaki farklılıkların yaklaşık %25'i duyuşsal özelliklerden kaynaklanmaktadır. Duyuşsal özelliklerin ders başarısı farkı üzerindeki önemli etkisinden ötürü başarı karşılaştırmalarının yapıldığı durumlarda duyuşsal özellikler açısından farklılaşmanın etkisinin de hesaba katılması gereklidir.

Alanyazın incelendiğinde TIMSS ile diğer ulusal ve uluslararası değerlendirme çalışmalarında elde edilen test ve anket puanlarına yönelik ölçme değişmezliği çalışmalarının yapıldığı görülmektedir. Bu çalışmada, fen başarısı üzerinde önemli etkisi olan fen duyuşsal özelliklerinden oluşturulacak bir modelin değişmezliği incelenmesi planlanmıştır. Duyuşsal özelliklerin ders başarısı üzerindeki etkisi ile bölge ve cinsiyetler arasındaki puan farkları göz önüne alındığında, fen duyuşsal özelliklerinin cinsiyete ve bölgelere göre ölçme değişmezliğinin araştırılmasının alanyazına katkı sağlaması beklenmektedir.

Cinsiyet ve bölge ortalama puanları arasındaki farklılıktan yola çıkılarak bu araştırmada, TIMSS 2015 Türkiye uygulamasındaki öğrenci anketi maddeleri kullanılarak oluşturulan fen duyuşsal özellik modelinin cinsiyet ve bölgeler arasındaki ölçme değişmezliği test edilmiştir. Yürütülen araştırmanın amacı 8. sınıf fen duyuşsal özellik modelinin cinsiyetler ve bölgeler arasındaki ölçme değişmezliğini incelemek ve puanların geçerliği hakkında kanıt sağlamaktır.

### **Araştırma Problemi**

TIMSS 2015 verilerine göre öğrenci anketiyle oluşturulan fen bilimleri duyuşsal özellik modeli cinsiyetler ve Türkiye İBBS Düzey 1 bölgeleri arasında ölçme değişmezliği sağlamakta mıdır?

#### **Alt problemler.**

1. Fen bilimleri duyuşsal özellik modeli, erkek ve kız öğrenciler arasında ölçme değişmezliği aşamalarından;

- a. şekil değişmezlik koşulunu sağlamakta mıdır?
- b. metrik değişmezlik koşulunu sağlamakta mıdır?
- c. ölçek değişmezliği koşulunu sağlamakta mıdır?
- ç. katı değişmezlik koşulunu sağlamakta mıdır?



2. Fen bilimleri duyuşsal özellik modeli, Türkiye İBBS Düzey 1 bölgeleri arasında ölçme değışmezliđi aşamalarından;

- a. şekil değışmezlik koşulunu sağlamakta mıdır?
- b. metrik değışmezlik koşulunu sağlamakta mıdır?
- c. ölçek değışmezliđi koşulunu sağlamakta mıdır?
- ç. katı değışmezlik koşulunu sağlamakta mıdır?

### **Sınırlılıklar**

Araştırma, TIMSS 2015'te yer alan 8. sınıf öğrenci anketindeki maddelere verilen yanıtlarla sınırlıdır.

### **Tanımlar**

İstatistiki Bölge Birimleri Sınıflandırması (İBBS): Avrupa Birliđi (AB) İstatistik Bürosu (Eurostat) 1988 yılında geliştirilen İstatistiksel Bölge Birimleri Sınıflandırması (İBBS) – Nomenclature of Territorial Units for Stats (NUTS), bölgelere ilişkin istatistikleri bir yapı içerisinde sınıflandırmak ile ekonomik ve toplumsal problemlerin çözümü için bölgesel politikalar oluşturabilmek amacıyla oluşturulmuştur. Türkiye'nin AB adaylıđı kapsamında 2001 yılında Türkiye'deki istatistiksel bölgeler belirlenmiş ve 2002 yılında yürürlüğe alınmıştır (Öztürk, 2009). Yapılan sınıflandırma cođrafi özellikler, temel istatistiksel göstergeler ve sosyoekonomik gelişmişlik düzeyleri dikkate alınarak oluşturulmuştur ve Türkiye'deki istatistiki bölge birimleri üç düzeyden oluşmaktadır (Taş, 2006). Ülkemizde uygulanan ulusal ve uluslararası eğitim araştırmalarındaki karşılaştırmalarda en sık kullanılan kategori olan Düzey 1'de 12 bölge bulunmaktadır: İstanbul, Batı Marmara, Ege, Dođu Marmara, Batı Anadolu, Akdeniz, Orta Anadolu, Batı Karadeniz, Dođu Karadeniz, Kuzeydođu Anadolu, Ortadođu Anadolu ve Güneydođu Anadolu. Düzey 2'de, bir önceki düzeyde yer alan bölgeler içerisindeki şehirlere göre sosyoekonomik açıdan daha gelişmiş durumda olan İstanbul, İzmir, Ankara, vd. 26 bölge yer almaktadır. Bu düzeydeki bölgeler -çođunlukla- Büyükşehir statüsünde olan iller ve bu illere komşu olan diđer illerin bir araya gelmesinden oluşmaktadır. Düzey 3'te ise Türkiye Mülki İdare Haritası'nda yer alan 81 şehir bulunmaktadır.

## Bölüm 2

### Araştırmanın Kuramsal Temeli ve İlgili Araştırmalar

Bu kısımda araştırmaya ilişkin kuramsal temel ve ölçme değişmezliği ile ilgili yürütülmüş araştırmalar sunulmuştur. Araştırmanın kuramsal temeli; ölçme değişmezliği, yapısal eşitlik modellemesi, açımlayıcı faktör analizi, doğrulayıcı faktör analizi ve çok gruplu doğrulayıcı faktör analizi başlıklarından oluşmaktadır.

#### Ölçme Değişmezliği

Her ölçme aracının öngörülen bir amacı vardır. Önceden belirlenmiş amaç doğrultusunda ölçme aracıyla puanlar elde edilir ve elde edilen puanlar ile yorumlamalar yapılır (örn: başarılı, başarısız). Geçerlik, testlerin öngörülen kullanımları için test puanlarının yorumlarının teori ve kanıtlarla desteklenme derecesi olarak tanımlanmıştır (AERA, APA ve NCME, 2014). Dolayısıyla, bir çalışmada kullanılan ölçme araçlarından elde edilen puanlar ile yapılacak yorumlar ancak puan yorumlarının geçerliğine ilişkin kanıtlar bulunduğu durumlarda uygundur.

Geçerlik, ölçme aracı kullanılarak elde edilen puanlara göre değişiklik gösterir. Geçerlik kanıtı elde etmek amacıyla hesaplanan teste ve maddelere ilişkin istatistikler, gruptaki bireylerin özelliklerine göre şekillenmektedir (Crocker ve Algina, 1986). Dolayısıyla, ölçme aracının psikometrik özellikleri verilerin toplandığı grupların özelliklerine göre farklılık gösterebilir. Ölçme araçlarından elde edilen puanlar ile farklı gruplar arasında karşılaştırma yapılacaksa ölçme aracının teorik yapısı farklı gruplarda aynı olmalıdır. Başka bir deyişle, böyle bir karşılaştırmanın geçerli olması için kullanılan ölçeğin, farklı gruplarda aynı modele sahip özdeş yapıları ölçmesi gerekir (van de Schoot, Lugtig ve Hox, 2012). Yani, puanların karşılaştırılabilir nitelikte olabilmesi için ölçeğin gruplar arasında ölçme değişmezliği sağlanmalıdır (Kıbrıslıoğlu, 2015).

Ölçme değişmezliği, farklı gruplarda yer alan ve belirli bir gizil yapı açısından denk olan bireylerin ilgili gözlenen değişken açısından da denk olmasıdır. Başka bir ifadeyle, belli bir özelliğe aynı düzeyde sahip olan bireylerin test maddelerini aynı şekilde idrak etmesi ve yorumlamasıdır (Bryne ve Watkins, 2003). Dolayısıyla, ölçme değişmezliği çalışmalarında gizil değişkenler ile gözlenen değişkenlerin ilişkisinin karşılaştırma yapılan tüm gruplarda aynı olup

olmadığı incelenmektedir. Yapılan incelemelerin amacı, karşılaştırmaların ve puan yorumlarının geçerliğine ilişkin kanıt sağlamaktır. Gruplar arası ölçme değişmezliği sağlanamadığında, puanların karşılaştırılmasına dayalı olarak yapılan yorumların geçerliği düşer (Şekercioğlu, 2018).

Ölçme değişmezliği incelemesi çalışmalarında gözlenen ve gizil değişkenler arasında bir ölçme modeli kurulur daha sonra modelin aralarında karşılaştırma yapılan farklı gruplarda bu modelin aynı yapıda olup olmadığı incelenir. Modelin farklı gruplarda aynı yapıya sahip olması ile ölçme aracındaki maddelerin faktör yüklerinin, faktörlerin kendi aralarındaki korelasyonlarının ve faktörlere ilişkin hata varyanslarının eşitliği ifade edilmektedir (Jöreskog ve Sörbom, 1993). Ölçme değişmezliği için bugüne kadar yapılan çalışmalar incelendiğinde iki farklı yaklaşım görülmektedir: Madde Tepki Kuramı (MTK) temelli ve yapısal eşitlik modellemesi (YEM) temelli yaklaşımlar (Reise, Widaman ve Pugh, 1993; Raju, Laffitte ve Byrne, 2002). İki yaklaşımda da kuramsal bir ölçme modeli oluşturulur ancak, model oluşturmadaki yöntemsel farklılıklardan ötürü yaklaşımlar kendi içlerinde avantaj ve dezavantajlara sahiptir (Meade ve Lautenschlager, 2004).

MTK temelli ölçme değişmezliği çalışmalarında log lineer bir model oluşturulur ve modelin gruplar arası karşılaştırması çok boyutludur. Değişmezlik, modelin madde ve test düzeyindeki işlev farkları üzerinden incelenmektedir. Madde düzeyinde değişmezlik sağlanmamış ise Değişen Madde Fonksiyonu (DMF), test düzeyinde sağlanmamış ise Değişen Test Fonksiyonu (DTF) olduğu şeklinde yorum yapılır (Osterlind ve Everson, 2009). Değişmezliğin sağlanmadığı durumlarda model ve veri arasındaki uyumun sağlanması koşuluyla alt grupların her biri için model parametreleri kestirilip ölçeklendikten sonra madde karakteristik eğrileri karşılaştırılarak değişmezlik kaynağı olan madde/maddeler belirlenir (Karakoç Alatlı, 2016).

YEM'e dayalı çalışmalarda çok gruplu doğrulayıcı faktör analizi (ÇGDFA) ve ortalama ve kovaryans yapılarının değişmezliği (OKYD) ile kovaryans yapılarının eşitliği dikkate alınır (Yandı, Köse, Uysal ve Oğul, 2017). Ancak, yapılan çalışmalara bakıldığında ÇGDFA yönteminin ölçme değişmezliği çalışmalarında daha sık kullanıldığı görülmektedir. Bu çalışmada da ÇGDFA aracılığıyla YEM yaklaşımı kullanılmıştır. Bu nedenle, açılımlayıcı faktör analizi (AFA), YEM ve ÇGDFA aşağıda ayrıntılı bir şekilde ele alınmıştır.

## Yapısal Eşitlik Modellemesi

YEM'in tanımı ve YEM'de kullanılan teknikler ile ilgili tek bir görüş olmasa da araştırmacılar, YEM'in bir istatistiksel teknik değil birden fazla istatistiksel tekniği kapsayan bir araç olduğu konusunda hemfikirdir (Çokluk, Şekercioğlu ve Büyüköztürk, 2016). YEM ile bir ya da daha fazla sürekli veya kesikli değişken ile bir ya da daha fazla sürekli veya kesikli değişken arasındaki ilişki çeşitli istatistiksel teknikler aracılığıyla modellenir (Kıbrıslıoğlu, 2015). Bu nedenle, nedensel modelleme, eş zamanlı eşitlik modellemesi, kovaryans yapılarının analizi veya doğrulayıcı faktör analizi olarak da adlandırılmaktadır (Tabachnick ve Fidell, 2007). Fakat yukarıda belirtildiği gibi, YEM bir analiz tekniği değildir; farklı analizleri genişleterek ve birleştirerek yorumlamayı sağlayan bir araçtır. YEM ile çoklu regresyon ve faktör analizinden elde edilen sonuçlar birleştirilerek gizil ve gözlenen değişkenlerin ilişkisi yalnızca tek bir modelde belirlenir ve değerlendirilir (Hoyle, 2012). YEM'deki istatistiksel hesaplamalar temel olarak sahip olunan modeldeki değişkenler arasındaki korelasyonel ilişkilerin örüntüsünü belirleme ve değişkenlerin varyansının mümkün olan en büyük kısmını bu model ile açıklama amacıyla yapılır (Önen, 2009). YEM, diğer yöntemlerden farklı olarak gözlenen değişkenlerin ölçme hatalarının etkisini de istatistiksel hesaplama dâhil ettiği için sıklıkla kullanılmaktadır (Raykov ve Marcoulides, 2006).

Kullanılan birçok istatistiksel yöntemde olduğu gibi YEM ile ilgili analizler öncesinde de verilerin yonteme ilişkin varsayımların ve gerekliliklerin sağladığının doğrulanması gereklidir. Çokluk ve diğerleri (2016) YEM'e ilişkin doğrulanması gereken hususları şu şekilde sıralamıştır:

- Örneklem büyüklüğü
- Kayıp veriler
- Normallik
- Doğrusallık
- Çoklu doğrusallık ve tekliklik
- Uç değerler

Yukarıdaki varsayımlar ve gereklilikler sağlandıktan sonra YEM'e ilişkin analizler yürütülmeye başlanır. Kline (2011)'e göre analizler genel olarak altı aşamada yapılır. Birinci adımda araştırmacının hipotezleri doğrultusunda model oluşturulur. İkinci adımda model tanımlanır. Buradaki amaç modelin örneklem ve verinin etkisinden arınmış bir şekilde kuramsal olarak tanımlanıp tanımlanamadığının kontrol edilmesidir. Üçüncü adımda modele ilişkin veriler toplanır ve toplanan veriler düzenlenir. Dördüncü adımda uyum değerlendirmesi, parametre kestirimleri yorumlaması ve denk modeller değerlendirmesi ile modelin kestirimi yapılır. Veri ile model arasında uyumun sağlanamadığı durumlarda model yeniden belirlenerek veri model uyumu sağlanana kadar aynı adımlar tekrar edilir. Son adımda ise elde edilen sonuçlar raporlaştırılır. Yukarıda bahsedilen adımlar aşağıda detaylı olarak açıklanmıştır.

**Modelin oluşturulması ve tanımlanması.** Bu aşamada, araştırmadaki gözlenen ve gizil değişkenler arasındaki ilişkiler, kuramsal ve deneysel bilgiler temel alınarak belirlenir ve şekilsel olarak ifade edilir. Daha sonraki analizler bu aşamada oluşturulacak modele göre yapılacağı için YEM'in en kritik adımıdır (Kline, 2011). Modellerin şekilsel gösteriminde yaygın kullanılan üç yöntem vardır: Yol diyagramı, denklem kümesi ve matrisler (Hoyle, 2012). Bu araştırmada model oluşturulurken yol diyagramı kullanılmıştır.

Model oluşturulduktan sonra, modelin tanımlanabilirliği kontrol edilir. Modelin tanımlanabilmesi için modele ilişkin parametrelerin her biri için tek bir değer elde edilmesi gerekir (Hoyle, 2012). Bu adımda elde edilen verilerden yola çıkarak modele ilişkin parametrelerin kestirilip kestirilemediği kontrol edilir (Schumacker ve Lomax, 2010). Bu kontrol sırasında aşağıdaki bilgiler göz önüne alınır (Kline, 2011):

1. Modelin serbestlik derecesi sıfıra eşit veya sıfırdan büyük ise model tanımlanabilir ( $sd \geq 0$ ).

2. Modeldeki tüm gizil değişkenler; faktör yüklerinden biri ya da gizil değişkenlerden birinin varyansı "1"e eşitlenerek ölçeklenir.

3. Model tek boyutlu ise bu boyutun minimum üç, model iki veya daha fazla boyutlu ise her boyutun minimum iki gözlenen değişkene sahip olmalıdır.

**Modelin uyumunun kestirilmesi.** Bu aşamada, modeldeki gözlenen ve kestirilen değişkenlere ilişkin kovaryans matrislerinin farkı minimum düzeye indirilerek örneklere ait parametreler kestirilir (Ülkü, 2019). Model kestirimi yapılırken En Çok Olabilirlik (Maximum Likelihood, ML), En Küçük Kareler ve Bayes kestirimleri gibi kullanılacak farklı yöntem grupları mevcuttur (Kıbrıslıoğlu, 2015). ML yöntemlerinin sayıltıları çoklu normallik ve değişkenlerin sürekli olmasıdır. Farklı ML yöntemleri eş zamanlı olarak analizi mümkün kıldığı için diğer yöntemlere göre daha fazla bilgi sağlamaktadır. Buna ek olarak, istatistiksel uygunluk sağlandığı ve model doğru şekilde oluşturulduğu zaman ML, büyük örneklemlerde asimptotik olarak yansız, tutarlı ve etkili kestirimler yapmaktadır (Kline, 2011). Bu araştırmanın verileri çoklu normallik ve süreklilik sayıltılarını sağladığından ML yöntemiyle kestirim yapılmıştır.

**Uyum iyiliği istatistiklerinin değerlendirilmesi.** Gözlenen ve kestirilen değişkenlerin kovaryans matrisleri birbirlerinden istatistiksel olarak anlamlı düzeyde farklı değilse model ile veri arasında genel anlamda bir uyum var demektir (Wang ve Wang, 2012). Uyum istatistikleri incelenerek bu uyumun düzeyi değerlendirilir. Uyum istatistikleri; model test istatistikleri ve yaklaşım uyum indeksleri şeklinde iki başlık altında incelenir (Kline, 2011).

**Model test istatistikleri.** Gizil ve gözlenen değişkenlerin kovaryans matrisleri arasında istatistiksel açıdan anlamlı fark olup olmadığı ki-kare ( $\chi^2$ ) istatistiği ile incelenir. İki matrisin  $\chi^2$  puanları arasında fark varsa ( $p < 0,05$ ) model ve veri arasında iyi uyum olmadığı yorum yapılır. Farkın nedeni verilerin normal dağılmıyor olması gibi gözlenen değişkenler arasındaki korelasyonun ve örneklem büyüklüğünün etkisi de olabilir (Kline, 2011). Bu nedenle  $\chi^2$ 'ye ek olarak, örneklemin etkisini minimize etmek amacıyla normlaştırılmış ki-kare ( $\chi^2/sd$ ) istatistiği hesaplanır (West, Taylor ve Wu, 2012).

**Yaklaşık uyum indeksleri.** Oluşturulan modelin toplanan veriyle ne derece uyum sağladığını belirlemek için kullanılan standart değerlerdir (Erşan, 2016). Daha önce belirtilmiş farklı uyum indeksleri vardır fakat en çok kullanılmakta olan indeksler RMSEA bu indekslerin uyum aralıkları Tablo 1'de belirtilmiştir.

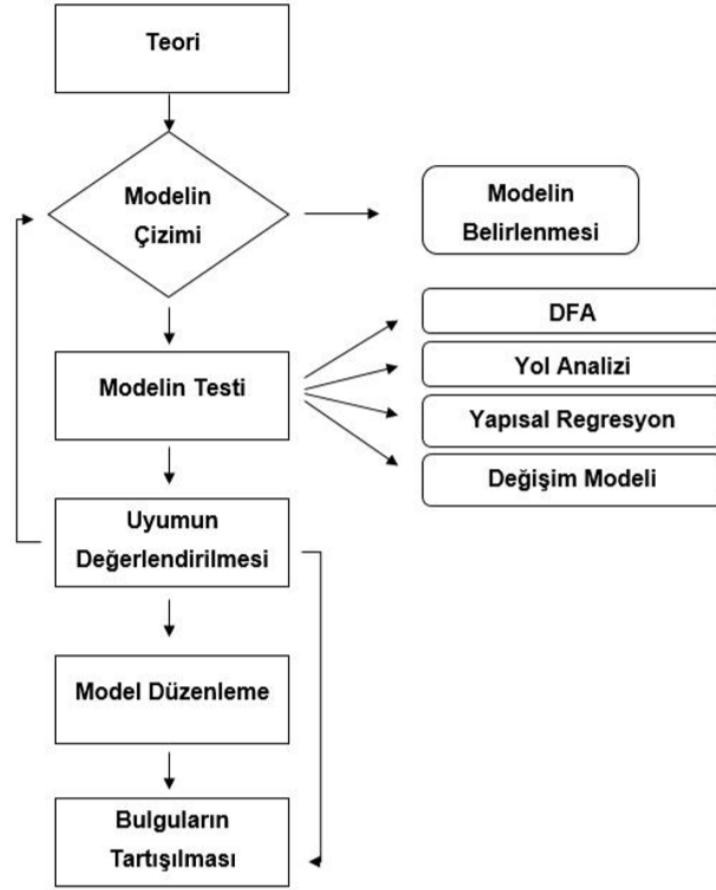
Tablo 1  
*Uyum İndeksleri*

İndeks	İyi Uyum	Kabul Edilebilir Uyum
$\chi^2$	$p > 0.05$	$p > 0.05$
$\chi^2/sd$	$0 \leq \chi^2/sd \leq 3$	$3 < \chi^2/sd \leq 5$
RMSEA	$0 \leq RMSEA \leq 0.05$	$0.05 < RMSEA \leq 0.08$
SRMR	$0 \leq SRMR \leq 0.05$	$0.05 < SRMR \leq 0.08$
GFI	$0.95 \leq GFI \leq 1.00$	$0.90 \leq GFI < 0.95$
AGFI	$0.95 \leq AGFI \leq 1.00$	$0.90 \leq AGFI < 0.95$
CFI	$0.95 \leq CFI \leq 1.00$	$0.90 \leq CFI < 0.95$
IFI	$0.95 \leq IFI \leq 1.00$	$0.90 \leq IFI < 0.95$
TLI	$0.95 \leq TLI \leq 1.00$	$0.90 \leq TLI < 0.95$

Kaynak: Hu & Bentler, 1999.

**Modelin yeniden betimlenmesi.** Bu aşama tüm çalışmalar için zorunlu değildir. Model-veri uyumu sağlanmış ise bir sonraki aşamaya geçilir. Model-veri uyumunun sağlanamadığı durumlarda ise model; kuramsal bilgiler ve modifikasyon indeksleri çerçevesinde yeniden tanımlanarak model-veri uyumu iyileştirilmelidir. Ancak, yapılan modelde yapılan tüm değişiklikler kuramsal bir dayanağa sahip olmalıdır (Çokluk ve diğerleri, 2016).

**Sonuçların raporlanması.** Bu aşamada analiz sonuçları ve sonuçlar üzerinden yapılan yorumlar raporlanır (Çokluk ve diğerleri, 2016). YEM oluşturmanın aşamaları aşağıda yer alan Şekil 2 ile özet olarak verilmiştir.



Şekil 2. Yapısal eşitlik modelinin oluşturulması

### Açımlayıcı Faktör Analizi

Açımlayıcı faktör analizi (AFA)'nın amacı gözlenen değişkenler ile ilişkili olduğu düşünülen faktörlerin ve faktörlerin kendi aralarındaki ilişki yapısını somut veriler üzerinden ortaya çıkarmaktır. AFA ile elde edilen gözlenen değişkenlerin faktörleşmeye uygun olup olmadığı ve hangi şekilde faktörleşebildiği test edilir. Bir gözlenen değişkenin uygun bir şekilde faktörleşebilmesi için değişkenin ilişkili olduğu boyuttaki faktör yükünün minimum 0.32 ve diğer boyutlardaki faktör yükünün görece daha düşük olması istenir (Tabachnick ve Fidell, 2013). Ancak faktör yükü koşulu, madde sayısı az olduğu zaman 0.30 şeklinde kullanılabilir (Büyüköztürk, 2005, s. 124). Bu araştırmada alt sınır 0,32 olarak belirlenmiştir. AFA aracılığıyla faktörler ve faktörler altındaki değişkenler ortaya çıkarıldıktan sonra bu faktörler alanyazından yola çıkılarak adlandırılır (Tabachnick ve Fidell, 2013).



AFA yapılmadan önce bazı varsayımların kontrol edilmesi gerekir: Örneklemin büyüklüğü, kayıp veriler, uç değerler, normallik ve eş doğrusallık. Verilerin normalliği basıklık ve çarpıklıkla ilgili katsayılar ile kontrol edilebilir. Kaiser-Meyer-Olkin (KMO) Testi ile örneklemdaki veri sayısının faktör analizi için ne düzeyde uygun olduğu, Bartlett Küresellik Testi ile eldeki verilerin dağılımının çoklu normalliği sağlayıp sağlamadığı veya başka bir deyişle verilerin faktörleşmeye uygun olup olmadığı test edilir (Tavşancıl, 2005).

AFA için farklı bilgisayar yazılımları mevcuttur (SPSS, R, jamovi, vb.). Yapılan analizler faktörlere ilişkin yükler, yamaç-birikinti grafiği ve öz-değerler aracılığıyla yorumlanır. Yamaç-birikinti grafiği ve özdeğerler kontrol edilerek faktör sayısı; maddelere ilişkin faktör yükleri kontrol edilerek maddelerin hangi boyutta olacağı hakkında karar verilir.

### **Doğrulayıcı Faktör Analizi**

Doğrulayıcı Faktör Analizi (DFA), gözlenen değişkenler ile gizil değişkenler arasındaki ilişkiyi gösteren ölçme modelinin test edildiği istatistiksel tekniktir (Brown, 2015). DFA'daki temel amaç, model içerisinde yer alan değişkenlerin ilişkileriyle ilgili kurulmuş hipotezleri test ederek modelin doğrulanıp doğrulanmadığını belirlemektir (Maruyama, 1998). Kurulan hipotezlerden yola çıkarak, DFA'da gözlenen değişkenler ile boyutlar ve boyutların kendi arasındaki ilişkileri test edilir (Çokluk ve diğerleri, 2016).

DFA, test geliştirme ve uyarılama ile aynı zamanda YEM çalışmalarında sıklıkla kullanılmaktadır. YEM'de gözlenen değişkenler ve faktörlerin ilişkileri ve değişkenlere ilişkin hata hataların ilişkiler üzerine kurulan ölçme modelinin verilerle uyumu DFA ile test edilir (Brown, 2015). DFA'nın uygulama aşamaları YEM ile birbirine çok benzemektedir. YEM'deki gibi DFA'da da model betimlenir, tanımlanır, kestirilir, değerlendirilir ve gerekli durumlarda tekrar kestirilir.

AFA sonucunda gözlenen değişkenler arasında ortaya çıkan ve kuramsal bilgiyle desteklenmiş olan faktör yapısının testi DFA ile yapılır. DFA sonucunda gözlenen değişkenler ile faktör yapısının ne derece temsil edildiği belirlenir (Thompson, 2004). DFA'da gözlenen değişkenler ile gizil değişkenlerin ilişkileri gizil değişkenlerden gözlenen değişkenlere doğru çizilen yollar (path) aracılığıyla yol diyagramında şekilsel olarak ifade edilir (Şencan, 2005). Daha sonra da

AFA'daki faktör yüklerine benzeyen lambda katsayıları ( $\lambda$ ) ile gizil değişkenin bir birimlik değişiminin gözlenen değişken üzerindeki değişim miktarı ifade edilir.  $\lambda$  katsayısının büyük çıkması gözlenen ile gizil değişkenin ilişkisinin güçlü olduğunu belirtir (Çokluk ve diğerleri, 2016).

### Çok Gruplu Doğrulamalı Faktör Analizi

Çok gruplu doğrulamalı faktör analizi (ÇGDFA), DFA'nın iki veya daha çok grupta aynı anda uygulanmasını ifade etmektedir. ÇGDFA ile, oluşturulan modelin örneklem içerisindeki alt gruplarda aynı yapıyı temsil edip etmediği sınıranır (Tabachnick ve Fidell, 2013).

Vandenberg ve Lance (2000)'e göre ÇGDFA aracılığıyla ölçme değişmezliği şöyle incelenmektedir;  $j$  ölçülen değişkeni için  $k$  grubundaki  $i$  bireyinin puanı  $X_{ijk}$  olarak kabul edilirse,  $X_{ijk}$  için faktör modeli şu şekildedir:

$$X_{ijk} = \tau_{jk} + \gamma_{jk}W_{jk} + u_{jk} \quad (1)$$

Denklemdaki  $\tau_{jk}$  ifadesi gözlenen değişkenler ile örtük yapılar arasındaki katsayı vektörünü,  $\gamma_{jk}$  ifadesi  $r$  madde sayısı olmak üzere  $r \times 1$  desenindeki faktör yüklerine ilişkin matrisi,  $W_{jk}$  ifadesi  $i$  bireyinin  $r \times 1$  desenindeki ortak faktör yükleri vektör matrisini ve  $u_{jk}$  ifadesi de bağımsız gözlenen değişkenlere ilişkin hata vektörünü temsil etmektedir. Buna ek olarak, ölçmeye ilişkin hataların kendi aralarında ve ortak faktöre ilişkin puanlarla ilişkisinin sıfıra eşit olduğu ( $E(W_{jk}, u_{jk}) = 0$ ) varsayımı ile kovaryans denklemi şu şekildedir:

$$cov(X_{ijk}) = \Sigma_k = \Lambda_k \Phi_k \Lambda_k' + \theta_k \quad (2)$$

Denklemdaki  $\Lambda_k$  ifadesi  $\gamma_{jk}$ 'den oluşan  $p$  boyut olmak üzere  $p \times r$  deseninde olan matristir.  $\Phi_k$  ifadesi,  $\gamma_{jk}$ 'daki varyans ve kovaryanslardan oluşur ve  $\theta_k$  ifadesi ölçmeye ilişkin hataların köşegen matrisini temsil eder. Aynı şekilde,  $X_{ik}$  değerlerine ait ortalamaya ilişkin vektör de şu şekilde gösterilebilir:

$$E(X_{ik}) = \mu_k = \tau_k + \Lambda_k K_k \quad (3)$$

Denklemdaki  $\tau_k$  ifadesi ölçme hatalarına ilişkin vektördür. Yukarıda sıralanan denklemlerden hareketle ölçme değişmezliği çalışmalarında  $[\tau_k, \Lambda_k, \theta_k]$  parametrelerinin örneklem içerisindeki  $k$  sayıda alt grupta eşitliği sınıranır

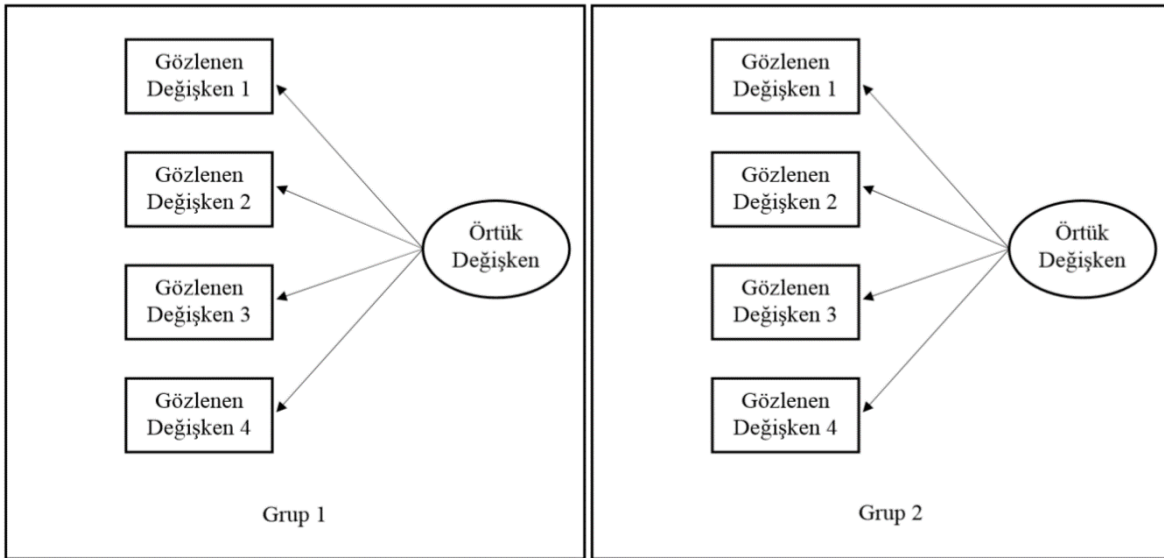
(Vandenberg ve Lance, 2000, s. 10; Jöreskog ve Sörborm, 2001; Millsap ve Olivera-Aguilar, 2012, s.381).

ÇGDFA aracılığıyla ölçme değişmezliği testi, hiyerarşik dört modelin sırayla sınanmasıyla yapılmaktadır. Ölçme değişmezliğinin dört aşaması ya da başka bir deyişle sırasıyla test edilen modeller yapısal değişmezlik, metrik değişmezlik, ölçek değişmezliği ve katı değişmezliktir (Bialosiewicz, Murphy ve Berry, 2013).

**Yapısal değişmezlik.** Bu aşamada, oluşturulan modelin gruplarda aynı olup olmadığı test edilir (Kline, 2011). Bu test,  $\Lambda_k$  matrisindeki değerlerin yani sabitlerin ve serbest faktör yüklerinin gruplarda eşit olup olmaması üzerinden yapılır. (Widaman ve Reise, 1997).

$$\Lambda_k^{(1)} = \Lambda_k^{(2)} \quad (4)$$

Yukarıdaki eşitlik sağlanıyorsa gruplarda faktör yapısının aynı olduğu ve maddelerin aynı yapıyı ölçtüğü yorumu yapılır. Ancak bu eşitlik sağlanamamış ise faktör yapısının gruplarda aynı olmadığı yani maddelerin farklı gruplarda aynı yapıyı ölçmediği anlamı çıkar dolayısıyla gruplar arasında karşılaştırma yapmaya ve değişmezliğin bir sonraki adımına geçmeye gerek yoktur (Vandenberg ve Lance, 2000). Şekilsel değişmezlik sağlanamıyorsa diğer üç değişmezlik de sağlanamaz (Kline, 2011). Çünkü ölçülen yapının aynı olmadığı alt gruplarda, daha hassas olan diğer üç kriterin sağlanması beklenemez.



Şekil 3. Yapısal değişmezlik

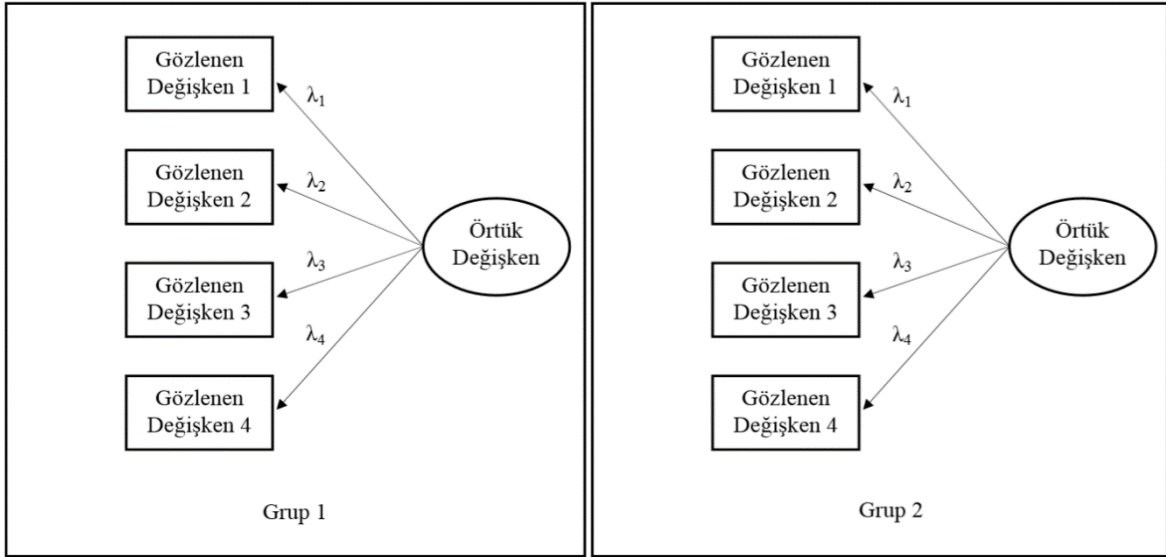
Yukarıda Şekil 3 ile yapısal değişmezlik modeli gösterilmiştir. Şekilde yer alan yapısal değişmezlik modelinde, faktör sayıları ile faktör örüntüsünün alt gruplarda aynı olup olmadığı test edilmektedir. Eğer Şekil 3'teki gibi gözlenen değişkenlerin sayısı ve örüntüsü, tüm gruplarda aynıysa yapısal değişmezlik sağlanmış demektir.

**Metrik değişmezlik.** Bu aşamada, gruplarda yapıların eşitliğine ek olarak madde faktör yüklerinin eşit olup olmadığı test edilir (Millsap ve Olivera-Aguilar, 2012). Gizil değişkenler madde faktör yükleri bağlantısıyla gözlenen değişkenler üzerinde etkiye sahiptirler (Bollen, 1989). Dolayısıyla gizil değişkenin gözlenen değişken üzerindeki etki derecesini gösteren faktör yüklerinin aynı olmadığı gruplarda yapı değişmez bir şekilde ölçülemez.

$$\gamma_{jk}^{(1)} = \gamma_{jk}^{(2)} \quad (5)$$

Yukarıdaki eşitliğin testi yapısal model ile bu adımdaki model arasında yapılan ki-kare fark testi ve uyum indeksleri kullanılarak yapılır. Ki-kare değerleri açısından anlamlı fark yoksa ve indeksler kabul aralığında ise metrik değişmezliğin sağlandığı kabul edilir. Buna dayanarak faktör yüklerinin gruplarda değişmediği yani gruplarda yer alan bireyler maddeleri benzer şekilde yorumladığı yorumu yapılır. Gruplar arasındaki faktör yükleri eşit olduğu için grupların faktör varyans ve kovaryansları karşılaştırma yapmak için uygundur fakat grupların ortalamalarındaki farklılığın kaynağı halen tam olarak belli değildir. (Millsap ve Olivera-Aguilar, 2012). Faktör yüklerinin gruplarda eşit olmaması maddelerin bireyler tarafından aynı şekilde yorumlanmadığını ifade etmektedir (Bialosiewicz, Murphy ve Berry, 2013). Maddelerin farklı gruplarda aynı şekilde yorumlanmaması da bir ya da birkaç faktörün gruplarda aynı anlamı ifade etmediği ya da yanlılığa sahip olduğunu işaret eder (Gregorich, 2006).

Aşağıda Şekil 4 ile metrik değişmezlik modeli gösterilmiştir. Bu modelde, önceki aşamaya ek olarak faktör yüklerinin alt gruplarda aynı olup olmadığı test edilmektedir.  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3$  ve  $\lambda_4$  ile ifade edilen faktör yüklerinin her biri tüm gruplarda eşit ise metrik değişmezlik sağlanmış demektir.



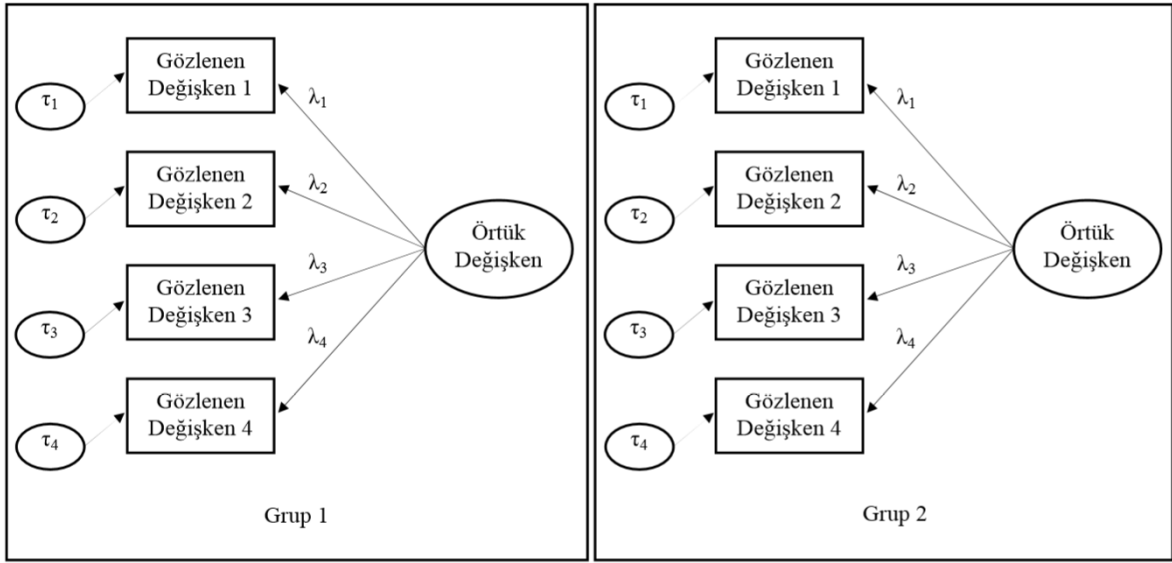
Şekil 4. Metrik değişmezlik

**Ölçek değişmezliği.** Bu aşamada madde faktör yüklerinin eşitliğine ek olarak maddelere ilişkin sabitlerin gruplarda eşit olup olmadığı test edilir (Millsap ve Olivera-Aguliar, 2012).

$$\tau_{jk}^{(1)} = \tau_{jk}^{(2)} \quad (6)$$

Yukarıdaki eşitliğin, testi yapısal model ile bu adımdaki model arasında yapılan ki-kare fark testi ve uyum indeksleri kullanılarak yapılır. Ki-kare değerleri açısından anlamlı fark yoksa ve indeksler kabul aralığında ise ölçek değişmezliğinin sağlandığı kabul edilir. Bu eşitliğin sağlanması, faktör ve gözlenen değişkenlere ilişkin ortalamaların karşılaştırılabilir olduğu anlamına gelmektedir (Gregorich, 2006). Yani maddeler, faktörleri farklı gruplarda aynı düzeyde açıklamaktadır. Aynı şekilde, ölçek değişmezliği sağlandığı için grupların ortalamaları arasındaki farkın kaynağının gizil değişken dışındaki başka bir değişken değil gizil değişkenin kendisi olduğu yorumu yapılabilir (Başusta ve Gelbal, 2015). Bu koşul sağlanamadığında ise ortalamalara ilişkin karşılaştırmalara dayalı yorumlar hataya yol açar (Kıbrıslıoğlu, 2015). Çünkü madde sabiti tüm gruplarda eşit olmadığı için bir veya daha fazla sayıda madde gruplardan en az birinde farklı işlemektedir yani yanlılık vardır.

Aşağıda Şekil 5 ile ölçek değişmezliği modeli gösterilmiştir. Bu modelde, önceki aşamaya ek olarak madde sabitlerinin alt gruplarda aynı olup olmadığı test edilmektedir.  $\tau_1$ ,  $\tau_2$ ,  $\tau_3$  ve  $\tau_4$  ile ifade edilen madde sabitlerinin her biri eşit ise ölçek değişmezliği sağlanmış demektir.



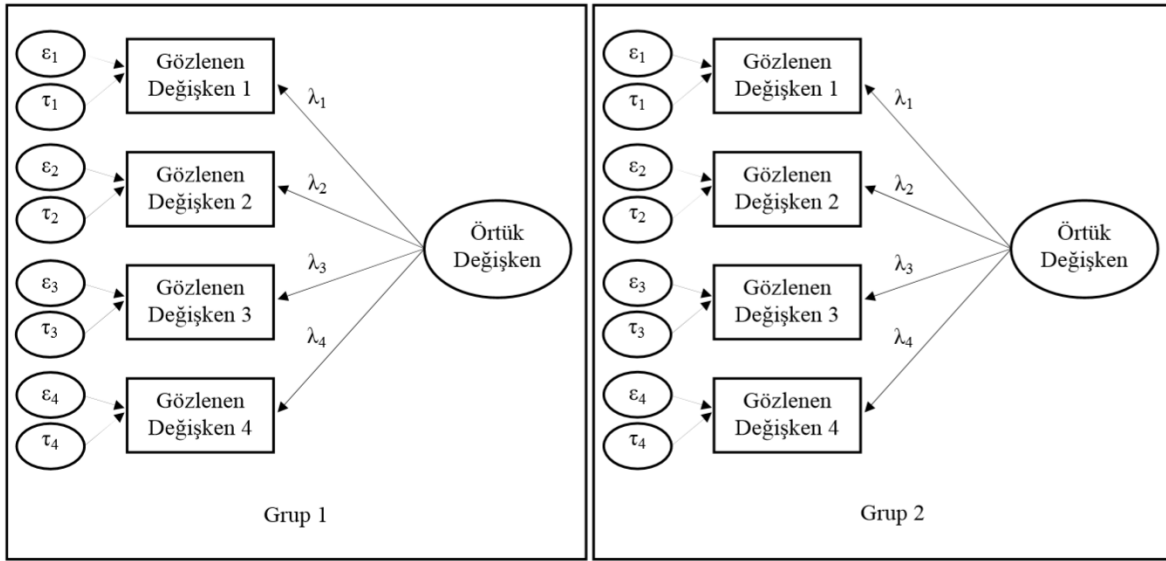
Şekil 5. Ölçek değişmezliği

**Katı değişmezlik.** Bu aşamada madde sabitlerinin eşitliğine ek olarak maddelere ilişkin hata varyanslarının gruplarda eşit olup olmadığı test edilir (Widaman ve Reise, 1997).

$$\varepsilon_k^{(1)} = \varepsilon_k^{(2)} \quad (7)$$

Yukarıdaki eşitliğin testi, yapısal model ile bu adımdaki model arasında yapılan ki-kare fark testi ve uyum indeksleri kullanılarak yapılır. Ki-kare değerleri açısından anlamlı fark yoksa ve indeksler kabul aralığında ise katı değişmezliğin sağlandığı kabul edilir. Katı değişmezliğin sağlanması gözlenen değişkenlerin varyansları ve kovaryanslarının gruplar arasında karşılaştırılabilir olduğu anlamına gelmektedir (Gregorich, 2006). Ancak, bu model madde hata varyanslarının eşitliği koşuluna dayandığı için değişmezliği pratikte karşılamak kolay değildir. Çünkü gizil değişkene ilişkin varyansın artması madde hata varyanslarının artmasına neden olmaktadır (Widaman ve Reise, 1997).

Aşağıda Şekil 6 ile katı değişmezlik modeli gösterilmiştir. Bu modelde, önceki aşamaya ek olarak madde hata varyanslarının alt gruplarda aynı olup olmadığı test edilmektedir.  $\varepsilon_1, \varepsilon_2, \varepsilon_3$  ve  $\varepsilon_4$  ile ifade edilen hata varyanslarının her biri eşit ise katı değişmezlik sağlanmış demektir.



Şekil 6. Katı değişmezlik

Yukarıda sıralanan paragraflarda, ölçme değişmezliğinin aşamaları detaylı olarak anlatılmıştır. Aşağıda yer alan Tablo 2’de ise, ölçme değişmezliği incelenirken kullanılan hiyerarşik dört model ile her değişmezlik modelinde sabit tutulan ve eşitliği test edilen parametreler gösterilmiştir.

Tablo 2

*Ölçme Değişmezliği İncelemesinde Kullanılan Parametreler*

Değişmezlik Modeli	Sabit Tutulan Parametreler	Test Edilen Parametreler
Yapısal değişmezlik	-	Madde/faktör grupları
Metrik değişmezlik	Faktör varyansları ve kovaryansları	Faktör yükleri
Ölçek değişmezliği	Faktör gözlenen değişken ortalamaları	Madde sabitleri
Katı değişmezlik	Gözlenen varyans ve kovaryanslar	Madde hata varyansları

Kaynak: Gregorich, 2006

Tablo 2’de gösterimi yapılan ikinci sütündeki parametreler sabit tutularak üçüncü sütündeki parametreler test edilir ve bu şekilde ölçme değişmezliği incelenir. Her değişmezlik aşamasında, bir önceki aşamada test edilen parametrelere yeni bir parametre dahil edilir. Buna göre, yapısal değişmezlikte madde/faktör grupları; metrik değişmezlikte madde/faktör grupları ve faktör yükleri; ölçek değişmezliğinde madde/faktör grupları, faktör yükleri ve madde sabitleri; katı değişmezlikte madde/faktör grupları, faktör yükleri, madde sabitleri ve madde hata varyansları serbestçe kestirilmektedir. Parametre kestirimleri yapılırken paralel

şekilde, her aşamada birer parametre de sabitlenmektedir. Yapısal değişmezlikte herhangi bir parametre sabit tutulmazken, metrik değişmezlikte faktör varyansları ve faktör kovaryansları; ölçek değişmezliğinde faktör varyansları, faktör kovaryansları ile faktör ve gözlenen değişken ortalamaları; katı değişmezlikte ise faktör varyansları, faktör kovaryansları, faktör ve gözlenen değişken ortalamaları ile gözlenen varyans ve kovaryanslar sabitlenmektedir.

### **İlgili Araştırmalar**

Bu bölümde, uluslararası değerlendirme sınavlarıyla ilgili Türkiye’de ve dünyada yürütülen ölçme değişmezliği çalışmalarına yer verilmiştir.

Ercikan ve Koh (2005), araştırmalarında TIMSS 1995 uygulamasında yer alana matematik ve fen bilimleri alt testlerinin Fransızca ve İngilizce formları arasındaki değişmezliğini incelemiştir. Çalışma kapsamında İngilizce ve Fransızca formları için ÇGDFA yöntemi ile ölçme değişmezliği, MTK ile değişen madde fonksiyonu incelemesi yürütülmüştür. Analizler sonucunda her iki alt testte yer alan maddelerin de iki grup arasında hem değişen madde fonksiyonu gösterdiği hem de ölçme değişmezliği sağlamadığı ortaya koyulmuştur. Bundan hareketle testin iki farklı dildeki formunun denkliğinin sağlanmadığı dolayısıyla yapılacak karşılaştırmalara dayalı yorumların geçerliğinin düşük olduğu yorumunda bulunulmuştur.

Wu, Li ve Zumbo (2007), TIMSS 1999 uygulamasındaki matematik başarı testinden elde edilen puanlara ilişkin ölçme değişmezliğini ülkelere göre incelemiştir. Araştırma kapsamında, benzer kültürel özellikler gösteren Yeni Zelanda, Avustralya, Kanada ve ABD bir grupta, Japonya, Kore ve Tayvan diğer grupta toplanarak hem benzer kültürler hem de farklı kültürler arasında ölçme değişmezliği test edilmiştir. Bulgular incelendiğinde, benzer kültürler arasında katı değişmezliğinin sağlandığı ancak farklı kültürler arasında değişmezliğin en fazla metrik ve yapısal düzeyde sağlandığı görülmüştür.

Uzun (2008), TIMSS-R Türkiye uygulamasına katılan 7841 öğrencinin fen bilimleri başarı testi ve öğrenci anketi verileri aracılığıyla fen başarısını etkileyen değişkenleri kullanarak bir model oluşturmuştur. Daha sonra bu modelde yer alan her bir boyutun ölçme değişmezliğini cinsiyet açısından incelemiştir. Araştırma kapsamında yapısal eşitlik modellemesi ve çok gruplu doğrulayıcı faktör analizi



yöntemleri kullanılmıştır. Analiz sonuçlarına göre modelde yer alan öz yeterlik, tutum ve sınıf içi öğrenci etkinlikleri boyutlarının metrik değişmezlik, tutum boyutunun ise ölçek değişmezliği koşulunu sağladığı belirtilmiştir.

Akyıldız (2009), PIRLS 2001 uygulamasında yer alan dört alt teste ilişkin verileri kullanarak başarı testlerinde yer alan faktörlerin yapısının uygulamaya katılan 35 ülke arasındaki ölçme değişmezliğini incelemiştir. ÇGDFA yöntemi kullanılarak yapılan analizler sonucunda, dört alt testin karşılaştırma yapılan ülkeler arasında sadece yapısal değişmezlik sağladığı yani yalnızca aynı faktör yapısına sahip olduğu görülmüştür.

Uyar (2011), PISA 2009 Türkiye uygulamasının öğrenci anketlerindeki öğrenmeyi öğrenme kısmında yer alan maddelerden yola çıkarak öğrenme stratejilerine ilişkin bir model oluşturmuştur. Daha sonra bu modelin cinsiyetler, okul türleri ve istatistiksel bölge birimleri arasındaki ölçme değişmezliğini incelemiştir. Uygulamaya katılan 4340 öğrenciye ilişkin verilerle yürütülen analizler sonucunda modelin, cinsiyet ve okul türüne göre en fazla metrik değişmezlik, istatistiksel bölge birimine göre ise katı değişmezlik koşullarını sağladığı görülmüştür. Bundan hareketle, oluşturulan modelin bölgeler arasında yapılacak karşılaştırmalar için uygun olduğu yorumunda bulunulmuştur.

Bahadır (2012), PISA 2009 Türkiye uygulamasında yer alan veriler aracılığıyla okuma becerilerine ilişkin bir model oluşturmuştur. Daha sonra bu modelde değişkenlerin ölçme değişmezliğini Türkiye'deki bölgeler arasında incelemiştir. Açımlayıcı faktör analizi sonucunda aile, evdeki kültürel varlıklar ve okumaya karşı tutum değişkenlerinde oluşan bir model ortaya konulmuştur. Bu modelin ölçme değişmezliği incelendiğinde, modelin metrik değişmezliği sağladığı ancak ölçek değişmezliğini sağlamadığı görülmüştür. Kısmi değişmezliğin sağlanmış olmasından ötürü değişmezlik aşamaları test edilmiş ve bölgeler arasındaki en fazla farkın anne-baba eğitim düzeyi ile öğrenciye sağlanan olanaklar açısından olduğu sonucuna varılmıştır.

Ölçüoğlu (2015), TIMSS 2011 Türkiye uygulamasına katılan 8. sınıf öğrencilerinin matematik başarısını etkileyen değişkenleri modellemiş, daha sonra bu modeli Türkiye'deki yedi coğrafi bölge için test ederek modelin tüm bölgelerde geçerli olup olmadığını incelemiştir. Modelin tüm bölgelerde geçerli olduğu

doğrulandıktan sonra çok gruplu doğrulayıcı faktör analizi ile bu modelin coğrafi bölgeler arasında ölçme değişmezliği sağlayıp sağlamadığı incelenmiştir. Analizler sonucunda modelin tüm bölgelerde ölçek değişmezliği ve katı değişmezlik koşulunu yani tam eşdeğerliği sağlamadığı görülmüştür. Dolayısıyla puan farklılıklarının bölgelerden mi ölçme aracından mı kaynaklandığı kesin olarak belirlenemediği için bölgeler arasında yapılacak karşılaştırmaların yorumlanmasında dikkatli olunması gerektiği sonucuna ulaşılmıştır.

Kıbrıslıoğlu (2015), PISA 2012 uygulamasında yer alan matematik öğrenme anketinden elde edilen verilerle matematik öğrenme algısına ilişkin bir model oluşturmuştur. Daha sonra bu modelin değişmezliğini çok gruplu doğrulayıcı faktör analizi aracılığıyla cinsiyetler ve Türkiye, Çin-Şangay ve Endonezya örneklemi arasında incelemiştir. Sonuçlar, matematik öğrenme algısına ilişkin modelin cinsiyete göre katı değişmezlik, ülkelere arasında ise sadece şekilsel değişmezlik koşulunu sağladığını göstermiştir. Buna bağlı olarak cinsiyet puanları arasındaki farkın ölçme aracından kaynaklanmadığı ancak bölgeler arasındaki farkın kaynağının tam olarak belirlenemediği sonucuna varılmıştır.

Karakoç Alatlı (2016), PISA 2012'de yer alan fen ve matematik okuryazarlığı ile okuma becerileri başarı testlerinin farklı kültürler arasında dil açısından ölçme değişmezliğini incelemiştir. Karşılaştırmalar Avusturalya, Fransa, Şangay-Çin ve Türkiye'de eğitim görmekte olan ve ortak kullanılan kitapçıktaki soruları cevaplayarak teste katılmış 2149 öğrenciye ilişkin verilerle yürütülen analizlerde ÇGDFA yöntemi kullanılarak yapılmıştır. Analizler sonucunda üç alt testin de dil açısından en fazla yapısal değişmezlik koşulunu sağladığı belirlenmiştir.

Yandı, Köse ve Uysal (2017), çalışmalarında PISA 2012 uygulamasındaki alt ölçeklerden birisi olan Problem Çözmeye Açıklık Ölçeği'nin Türkiye ve Finlandiya örneklemi arasındaki ölçme değişmezliğini farklı yöntemler kullanarak incelemiş ve sonuçları karşılaştırmışlardır. Ölçme değişmezliği incelenirken, YEM temelli yöntemlerden biri olan ortalama ve kovaryans yapılarının değişmezliği ile örtük sınıf analizi temelli yöntemlerden biri olan çoklu grup örtük sınıf analizi yöntemleri kullanılmıştır. Çalışma kapsamında ayrıca, verilerin normal dağılım göstermediği koşullarda normallik varsayımı gerektiren ve normallik varsayımı gerektirmeyen yöntemlerin sonuçları arasında fark olup

olmadığı da incelenmiştir. Bulgular incelendiğinde normallik varsayımını sağlamayan verilerin, normal dağılım gerektiren ortalama ve kovaryans yapılarının değişmezliği yöntemiyle test edildiğinde katı değişmezlik koşulunu sağlandığı ancak normal dağılım varsayımına dayalı olmayan çoklu grup örtük sınıf analizi yöntemiyle test edildiğinde kısmi homojen modelin kabul edildiği yani verilerin zayıf değişmezlik koşulunu sağladığı görülmüştür. Bu bilgilerden hareketle ölçme değişmezliği incelemesi yapılırken varsayımların test edilmesi ve kullanılacak yöntemin verilere uygun olmasına dikkat edilmesi gerektiği sonucuna ulaşılmıştır.

Gülleroğlu (2017), PISA 2015 matematik alt testine katılan öğrencilerin derse yönelik duyuşsal özellikleriyle oluşturduğu modelin ölçme değişmezliğini ÇGDFA aracılığıyla cinsiyet değişkenine göre incelemiştir. Bu kapsamda matematiğe yönelik ilgi, matematik kaygısı, matematiğe yönelik benlik algısı ve matematik öz yeterliliği boyutları kullanılarak cinsiyetler arasında model karşılaştırması yapılmıştır. Sonuçlar incelendiğinde matematiğe yönelik ilgi boyutunun ölçek değişmezliği, matematik kaygısı boyutunun metrik değişmezlik, matematik benlik algısı boyutunun yapısal değişmezlik koşulunu sağladığı; matematik öz yeterliliği boyutunun ise hiçbir değişmezlik koşulunu sağlamadığı görülmüştür.

Polat (2019), TIMSS 2015 uygulamasındaki 8. sınıf öğrenci anketlerindeki verileri kullanarak matematik ve fen bilimleri derslerine ilişkin duyuşsal özelliklere ilişkin birer model oluşturmuştur. Daha sonra bu modellerin ölçme değişmezliği çok gruplu doğrulayıcı faktör analizi yöntemini kullanarak cinsiyete, bölgelere ve kültürlere (Singapur, Suudi Arabistan ve Türkiye) göre incelenmiştir. Yürütülen analizler sonucunda, fen ve matematik derslerine ilişkin oluşturulan modellerin kültürler ve bölgeler arasında ölçek değişmezliği, cinsiyetler arasında ise katı değişmezlik koşullarını sağladığı görülmüştür. Dolayısıyla cinsiyete göre yapılacak karşılaştırmaların anlamlı olacağı ancak bölgeler ve kültürler arasında yapılacak karşılaştırmalarda ortaya çıkacak olası farkların herhangi bir yanlılıktan kaynaklanmadığı sonucuna ulaşılmıştır.

Güngör (2019), PISA 2015 uygulamasındaki öğrenci anketlerinde yer alan Fen Öğreniminde Araçsal Güdülenme ve Fen Özyeterliliği ölçeklerinden elde edilen puanlar aracılığıyla oluşturulan modelin ölçme değişmezliği, cinsiyetler ve Türkiye'deki istatistikî bölgeler arasında incelenmiştir. YEM tekniğine dayalı olarak

yapılan analizler sonucunda modelin cinsiyetler arasında şekilsel, bölgeler arasında ise zayıf değişmezlik koşulunu sağladığı görülmüştür. Analiz sonuçlarına göre, fen bilimleri başarı testi puanlarıyla cinsiyet ve bölgelere göre yapılacak karşılaştırmaların yorumları konusunda dikkatli olunması gerektiği önerilmektedir.

Ülkü (2019), ABİDE 2016 uygulamasında yer alan öğretmen anketi verilerini kullanarak öğrencilerin Fen Bilimleri ve Türkçe testlerindeki başarılarının ölçme değişmezliğini öğretmenlerin eğitim düzeyi ve mesleki deneyimine göre incelemiştir. Yapılan analizler sonucunda her iki derse ilişkin oluşturulan modelin de ölçme değişmezliğinin tüm aşamalarını sağladığı belirlenmiştir. Dolayısıyla öğretmenlerin eğitim düzeyi ve mesleki deneyimine göre ortalamalar açısından yapılacak karşılaştırmalar ve yorumların geçerli olacağı sonucuna varılmıştır.

Tekin (2019), PISA 2015 Xandar alt testine ait verileri kullanarak iş birlikli problem çözme becerilerine ilişkin bir model oluşturmuştur. Modeli doğrulanmış ve bu modelin kültürler arasındaki ölçme değişmezliğini Türkiye, Norveç ve Singapur'dan elde edilen 2990 kişiye ilişkin veri üzerinde ÇGDFA aracılığıyla incelemiştir. Yapılan analizlerle modelin kültürler arasında yapısal düzeyde değişmezlik gösterdiği ve buna bağlı olarak üç ülke arasındaki karşılaştırmaların anlamlı olmadığı sonucuna ulaşılmıştır.

Son yıllarda yürütülen araştırmalarda, uluslararası karşılaştırmalar yapmak amacıyla yapılmış değerlendirme çalışmalarının ölçme değişmezliği sıkça incelenmiştir. Ölçme değişmezliği incelemelerinin benzer ve farklı ülkeler/kültürler, cinsiyetler, bölgeler ve yaş grupları arasında, öğrenci başarısını etkileyen değişkenler ile oluşturulmuş modellerle yapıldığı görülmüştür. Değişmezlik incelemelerinde çoğunlukla yapısal eşitlik modellemesi ve çok gruplu doğrulayıcı faktör analizi yöntemleri kullanılmıştır. Çalışmaların sonuçlarında, değişmezliği incelenen testlerin katı değişmezlik koşulunu nadiren sağladığı görülmüştür. Bu durumun çalışmaların geçerliğini olumsuz etkilediğini ve test puanlarıyla anlamlı karşılaştırmalar yapılabilmesi için testlerdeki maddelerin düzenlenmesi gerektiği belirtilmiştir. Dolayısıyla, yapı geçerliği hakkında kanıt, karşılaştırmalar için dayanak ve daha sonra yapılacak çalışmalar için bilgi sağlaması açısından TIMSS 2015 sekizinci sınıf fen öğrenci anketi puanlarına yönelik ölçme değişmezliği incelemesinin alana katkı sunması beklenmektedir.

## **Bölüm 3**

### **Yöntem**

Bu bölümde araştırmanın yöntemi, örnekleme, veri toplama araçları, veri toplama ve veri analizi süreci detaylı olarak açıklanmıştır.

#### **Araştırmanın Yöntemi**

Çalışma, TIMSS 2015 değerlendirmesinin geçerlik düzeyini belirlemek, değerlendirmeler yapmak ve olaylar arasındaki olası ilişkileri ortaya çıkarmak amaçlarından dolayı betimsel bir araştırmadır (Büyüköztürk ve diğerleri, 2012). Bununla birlikte, araştırma problemlerine nicel veriler kullanılarak yanıt arandığı için nicel araştırma yöntemleri benimsenmiştir.

#### **Araştırmanın Evreni ve Örnekleme**

TIMSS'in hedef kitlesi dördüncü ve sekizinci sınıf düzeyinde örgün eğitime kayıtlı olan tüm öğrencilerdir. TIMSS 2015 uygulamasına 60 ülkeden yaklaşık 600.000 öğrenci katılmıştır (TIMSS, 2016). TIMSS'te okul örnekleme, tabakalı seçkisiz örnekleme yöntemiyle belirlenmektedir. TIMSS 2015 uygulaması için ilk aşamada İBBS Düzey 1, okul türü, okulların buldukları yer ve okulların idari biçimleri tabakaları kullanılarak okullar tabakalı seçkisiz örnekleme yöntemiyle belirlenmiştir, ikinci aşamada ise bu okullarda uygulamaya katılacak olan öğrenciler seçkisiz olarak belirlenmiştir.

TIMSS 2015 Türkiye uygulamasının araştırma evrenini bu iki düzeyde Türkiye'de öğrenim gören tüm öğrenciler oluşturmaktadır. Ancak, bu araştırma 8. sınıf düzeyindeki öğrenciler ile sınırlandırıldığı için bu araştırmanın evreni 1.187.893 öğrenciden oluşmaktadır (MEB, 2016). TIMSS 2015 uygulamasına İBBS Düzey 1'e göre yer alan 12 bölge ve kırsal kesimde yer alan 8. sınıf düzeyindeki öğrencileri temsil eden 6079 öğrenci katılmıştır. Örnekleme yer alan kız öğrenci sayısı 2943 (%48,4) iken, erkek öğrenci sayısı 3136 (%51,6)'dır. Öğrenci sayılarına ilişkin bilgiler Tablo 3'te gösterilmiştir.

Tablo 3

*Bölgelere Göre Öğrenci Sayıları*

Bölge	Öğrenci Sayısı	Yüzde
İstanbul (TR1)	913	17,08
Batı Marmara (TR2)	214	4,00
Ege (TR3)	607	11,36
Doğu Marmara (TR4)	577	10,80
Batı Anadolu (TR5)	341	6,38
Akdeniz (TR6)	570	10,67
Orta Anadolu (TR7)	274	5,13
Batı Karadeniz (TR8)	195	3,65
Doğu Karadeniz (TR9)	252	4,72
Kuzeydoğu Anadolu (TRA)	208	3,89
Orta Doğu Anadolu (TRB)	327	6,12
Güneydoğu Anadolu (TRC)	866	16,21
Toplam	5344	100

Tablo 3'e göre uygulamaya katılan öğrenci sayısı bölgelerin nüfuslarıyla orantılı olarak İstanbul'da en yüksek, Batı Karadeniz'de en düşüktür. Kırsal bölgeye ilişkin verilerin tüm Türkiye'deki kırsal bölgeleri kapsamından ötürü, çalışma kırsal bölge haricindeki bölgelerde yer alan öğrencilerin verileriyle yürütülmüştür. Buna göre, analizler 12 bölgede yer alan 5344 öğrenci verisiyle yürütülmüştür.

**Veri Toplama Araçları ve Süreci**

TIMSS 2015 uygulaması temel olarak matematik ve fen bilimleri başarı testleri ve bu derslere yönelik duyuşsal özellik anketleriyle yürütülmüştür. Başarı testlerinde bulunan maddeler TIMSS Matematik ve Fen Değerlendirme Çerçevesi kapsamında kazanımlara yönelik olarak hazırlanmıştır. Öğrenci anketlerinde ailenin sosyoekonomik düzeylerine, evde ve okulda bulunan eğitim kaynaklarına, matematik ve fen bilimleri derslerine yönelik tutumlarına ve okul iklimine yönelik maddeler bulunmaktadır. Başarı testleri 45'er dakikalık iki oturumda, anketler ise toplam 30 dakikalık tek oturumda uygulanmıştır. TIMSS'te öğrenci testi ve anketlerine ek olarak uygulamaya katılan okulların idarecileri ile sınıfların fen

bilimleri ve matematik dersi öğretmenleri için hazırlanan anketler de yer almaktadır.

Bu araştırmada, TIMSS 2015 Türkiye sekizinci sınıf uygulamasında yer alan fen öğrenci anketine verilmiş olan yanıtlar veri olarak kullanılmıştır. Araştırmaya başlanmadan önce Hacettepe Üniversitesi Etik Komisyonundan Ek A'da yer alan araştırma izni alınmıştır. Bölge kodlarına ilişkin veriler, Milli Eğitim Bakanlığı Ölçme Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü'nün Ek C'de yer alan 06.02.2020 tarihli yazısı ile elden teslim alınmıştır. Bölge kodları haricindeki tüm veriler IEA'nın resmi internet sitesinden (<https://www.iea.nl/data-tools/repository/timss>) alınmıştır.

## Verilerin Analizi

Veri analizi üç adımda tamamlanmıştır. İlk olarak, verilerin analizlere uygunluğu test edilmiş ve veriler analizlere uygun hale getirilmiştir. Daha sonra, açımlayıcı faktör analizi ile fen bilimleri duyuşsal özellik modeli oluşturulmuş ve bu model doğrulayıcı faktör analizi ile doğrulanmıştır. Son olarak modelin cinsiyete ve bölgelere göre ölçme değışmezliđi çok gruplu doğrulayıcı faktör analizi ile test edilmiştir. Analiz sürecindeki tüm adımlar aşağıda detaylı olarak açıklanmıştır.

**Sayıtların incelenmesi.** Bu aşamada normallik, doğrusallık, çoklu bağlantı, tekillik ve eşvaryanslılık varsayımları ile örneklem büyüklüğü, kayıp veriler ve uç değerler incelenmiştir.

**Kayıp veriler.** Faktör analizi yöntemi korelasyona dayalı bir yöntem olduğu ve korelasyon değeri veri sayısından etkilendiđi için, veri seti içerisindeki kayıp değerlerin korelasyon değerleri üzerindeki etkisi kontrol edilmelidir (Şencan, 2005). Araştırmacıların eldeki verinin tamamıyla çalışması tavsiye edilmektedir ancak bunun mümkün olmadığı durumlarda kayıp veriyle baş etmek amacıyla üç farklı yaklaşım bulunmaktadır. Bu yaklaşımlar kayıp veri tespit edilen değışken veya katılımcıların silinmesi, kayıp verilerin yerine yaklaşık değeri atanması ve kayıp verilerin ötelemeye dayalı yöntemler ile kestirilmesidir. Yaklaşımlardan hangisinin kullanılacağına karar verme aşamasında eksik verilerin tesadüfi olarak dağılıp dağılmadığı, tahmin ya da silme durumunda oluşacak örneklemin büyüklüğü gibi durumlar dikkate alınmalıdır (Çokluk ve diđerleri, 2016). Tabachnick ve Fidell (2013)'e göre kayıp veriler rastlantısal olarak dağılıyorsa ve kayıp veri oranı

%5'ten daha düşük ise bu kayıp büyük örneklem için ciddi bir problem oluşturmayacak ve kayıp veri ile baş etme yöntemleriyle elde edilen sonuçlarla benzer sonuçlar elde edilecektir. Ancak, kayıp veri sayısı çok fazlaysa ve kayıp verilerin dağılımı rastlantısal değilse veri silme işlemi çok fazla veri kaybına neden olmaktadır. Kayıp veriler yerine yaklaşık değer atama yöntemi kullanıldığında ise varyans olduğundan daha düşük hesaplanır ve bu durum verilerin normalde olduğundan daha uyumlu olarak görünmesine neden olabilir. Silme ve yaklaşık değer atama yöntemlerinin verilerin yanlı hale gelmesine neden olabileceği görülmektedir.

Örneklem içerisinde yer alan 208 öğrenciye ilişkin veriye çeşitli nedenlerden dolayı ulaşılamamıştır. Cinsiyet, yaş gibi demografik bilgilere ilişkin sorular da dahil olmak üzere çok fazla soruya yanıt vermeyen 29 öğrencinin verisi araştırmadan çıkarılmıştır. Araştırmada yer alan değişkenlere ilişkin kayıp veri oranı %0,2 - %2,0 aralığındadır. Değişkenlerin her biri ve tüm kombinasyonlarına göre kayıp veri oranı %5'ten düşük olduğu için kayıp veriler, ötelemeye dayalı yöntemlerden Beklenti Maksimizasyonu (BM) yöntemi kullanılarak kestirilmiştir.

**Uç değerler.** Veri setindeki beklenmedik veya aşırı değerlere sahip olan veriler uç değer olarak adlandırılmaktadır (Schumacker ve Lomax, 2010). Beklenmedik ya da aşırı olmalarından ötürü uç değerler, doğrusallığı ve normalliği dolayısıyla elde edilen analiz sonuçlarını olumsuz yönde etkilemektedir. Uç değerlerin olası nedenleri hatalı veri girişi, katılımcının örneklemin yapıldığı evrenin üyesi olmaması ve katılımcının örnekleme yer alan diğer katılımcılardan farklı olmasıdır. Uç değerler, tek değişkenli olabileceği gibi çok değişkenli de olabilir (Tabachnick ve Fidell, 2013). Tek değişkenli uç değer ile ifade edilmek istenen söz konusu katılımcının verisinin tek bir değişken açısından aşırı olmasıdır. Çok değişkenli uç değer ise katılımcının iki veya daha fazla değişkene ilişkin puanlarının kombinasyonlarının aşırı olmasıdır. Tek değişkenli uç değer analizinde veriler Z puanına dönüştürülür ve eğer geniş bir örneklem ( $n > 100$ ) ile çalışılıyorsa  $\pm 4$  aralığında olmayan değerler uç değer olarak kabul edilir (Çokluk ve diğerleri, 2016). Çok değişkenli uç değerlerin analizinde, verilere ilişkin Mahalanobis uzaklıkları hesaplanır ve kritik  $\chi^2$  değerinden daha büyük değerler uç değer olarak belirlenir (Ho, 2013). Bu kriterlere göre uç olarak kabul edilen değerler veri setinden temizlenir.



Kayıp verilerle ilgili işlemler gerçekleştirildikten sonra uç değerlerin analizi yapılmıştır. Tek değişkenli uç değer analizi sonucunda z puanları  $\pm 4$  aralığında olmayan 93 veri temizlenmiştir. Daha sonra yapılan çok değişkenli uç değer analizi sonucunda, hesaplanan Mahalanobis uzaklıkları kritik  $\chi^2$  değerinden yüksek olan 557 veri temizlenmiştir. Uç değerlerin temizlenmesinin ardından 4457 öğrenciye ait veri üzerinden çalışılmaya devam edilmiştir. Veri setinin son halindeki cinsiyete ve bölgelere göre öğrenci sayısı dağılımı, kayıp verilerin temizlenmemiş olduğu haliyle benzerdir.

**Normallik ve Doğrusallık.** Yeterli örneklem büyüklüğü sağlandığı ve değişkenler arasındaki ilişkiler verilere uygun yöntemlerle betimsel açıdan yorumlandığı sürece normallik varsayımının sağlanması güç değildir. Normalliğin test edilmesi sırasında en sık kullanılan test istatistikleri çarpıklık ve basıklık katsayılarıdır. Çarpıklık katsayısı ile verilerin dağılımının normal dağılımdan uzaklaşmasının derecesi, basıklık katsayısı ile verilerin dağılımının genişliği belirlenir. Her iki katsayı için 0 ideal olmakla birlikte  $\pm 1$  aralığının dışında kalan değerler anlamlı düzeyde çarpıklık ve basıklığı belirtmektedir (Harrington, 2009).

Açımlayıcı faktör analizinin kullanıldığı durumlarda tek değişkenli normalliğe ek olarak çok değişkenli normallik varsayımının da doğrulanması gerekir. Çok değişkenli normallik ile ifade edilen, analizlerde yer alan değişkenlerin ve değişkenlerin doğrusal kombinasyonlarının tümünün normal dağılıyor olmasıdır (Tabachnick ve Fidell, 2013). Çok değişkenli normallik varsayımı Bartlett Küresellik Testi ile kontrol edilir.  $\chi^2$  istatistiği ile hesaplanan test sonuçlarının anlamlılığı üzerinden veri setinde yer alan kovaryans matrisinin birim matristen farklı olup olmadığı yani eldeki verilerden faktör çıkarılıp çıkarılamayacağı hakkında yorum yapılır. Hesaplanan anlamlılık değerinin 0,05'ten küçük olması faktör analizi yapılabileceği, 0,05'ten büyük olması ise faktör analizi yapılamayacağı anlamına gelmektedir (Şencan, 2005). Ancak,  $\chi^2$  değeri örneklemin büyüklüğü kullanılarak hesaplandığı için görece büyük örneklerde  $\chi^2$  değerinin manidar çıkma olasılığı da örneklem büyüklüğüne bağlı olarak artacaktır. Dolayısıyla, büyük örneklerde Bartlett Küresellik Testi sonuçları yorumlanırken dikkatli olunmalıdır (Çokluk ve diğerleri, 2016).

Çok değişkenli normallik varsayımına bağlı olarak veri setinde yer alan değişken çiftleri arasındaki ilişkilerin doğrusal olması gerekmektedir. Doğrusallık

varsayımının testi, deęişken çiftleri arasında hesaplanan korelasyon katsayılarının yorumlanması üzerinden yapılır. Korelasyon katsayıları üzerinden doğrusallık ile ilgili yorum yapılırken, korelasyon katsayısının örneklem büyüklüğünden etkilendięi dikkate alınarak yorumlanması önerilmektedir. Ayrıca, veriler içerisinde kategorik ölçümlere sahip deęişkenler yer alıyorsa doğrusallık varsayımı göz ardı edilmiş olur. Dolayısıyla bu verilerle yapılan analiz sonuçları yanıltıcı nitelikte olabilir (Çokluk ve dięerleri, 2016).

Tek deęişkenli normallik incelemesinde bazı maddelerin çarpıklık ve basıklık katsayılarının  $\pm 1$  aralığında olmadığı görülmüştür. Ancak çarpıklık ve basıklık katsayılarının 0'dan anlamlı düzeyde farklı olması, büyük örneklerde normalliği önemli düzeyde bozacak bir şekilde fark yaratmaz (Tabachnick ve Fidell, 2013). Örneklem büyüklüğü 4457 olduğu için tek deęişkenli normalliğin sağlandığı kabul edilmiştir. Bartlett Küresellik Testi sonucu istatistiksel olarak manidar çıkmıştır ( $\chi^2=53630,707$ ,  $p<0,05$ ) ancak örneklem büyüdükçe test sonucunun manidar olma ihtimalinin artması ve bu araştırmanın örneklem büyüklüğünün 4457 olması göz önüne alınarak çok deęişkenli normalliğin sağlandığı kabul edilmiştir. Saçılım diyagram matrisleri incelendiğinde saçılımların eliptik bir biçimde olduğu yani çok deęişkenli normallik ve doğrusallığa ilişkin herhangi bir problem olmadığı görülmüştür. Korelasyon katsayılarının ideal aralıkta olması da doğrusallık varsayımına ilişkin problem olmadığına işaret etmektedir.

**Çoklu bağlantı ve teklik.** Çoklu bağlantı, doğrusal ilişkiye sahip olması beklenen deęişken ikililerinin arasında yüksek ilişki bulunması yani deęişkenler arasındaki korelasyon katsayılarının 0,90 değerinden büyük olması anlamına gelmektedir. Teklik ise bu korelasyon katsayısının 1,00 değerine eşit olması durumudur (Şencan, 2005). Çoklu bağlantı; tolerans, varyans şişkinlik faktörü (VIF) ve koşul indeksi (CI) değerleri aracılığıyla da incelenmektedir. Tolerans değerinin 0,01'den büyük, VIF değerinin 10'dan küçük ve CI değerinin 30'dan küçük olması çoklu bağlantı sorunu olmadığını göstermektedir. Çoklu doğrusallık ve teklik, faktör analizi yöntemleri için problem yaratmaktadır ve bu duruma sahip olan deęişkenlerin verilerden silinmesi yoluna gidilebilir (Tabachnick ve Fidell, 2013).

Maddeler arasındaki korelasyon katsayıları incelendiğinde 0.90 değerinden yüksek bir ilişkiye rastlanmamıştır. Tüm maddelere ilişkin tolerans değerlerinin

0.01'den büyük, VIF değerlerinin 10'dan küçük ve CI değerlerinin 30'dan küçük olduğu görülmüştür. Bu sonuçlar, veri setinde çoklu bağlantı ve tekillik problemi olmadığını göstermektedir.

**Eşvaryanslılık (Homojenlik).** Eşvaryanslılık, bir sürekli değişkene ilişkin puanlardaki değişimin, diğer değişken puanlarında da paralel şekilde gözlenmesidir. Eşvaryanslılık, çok değişkenli normallik eşvaryanslılığı gerektirdiği için normallik ile bağlantılıdır (Tabachnick ve Fidell, 2013). Çok değişkenli istatistikler için eşvaryanslılık, Box M Testi ile sınanabilir. Bu test ile kovaryans matrislerinin eşit olduğunu ifade eden  $H_0$  hipotezi test edilir. Test sonucunun manidar olması eşvaryanslılık sayılımasının sağlanmadığını gösterir. Stevens (1992) ise, Box M Testi'nin çok değişkenli normallikten etkilendiğini, çok değişkenli normalliğin sağlanmadığı durumlarda test sonucunun manidar olmasının her zaman eşvaryanslılığın sağlanmadığı şeklinde yorumlanmaması gerektiğini ifade etmektedir (Akt. Mertler ve Vannatta, 2005).

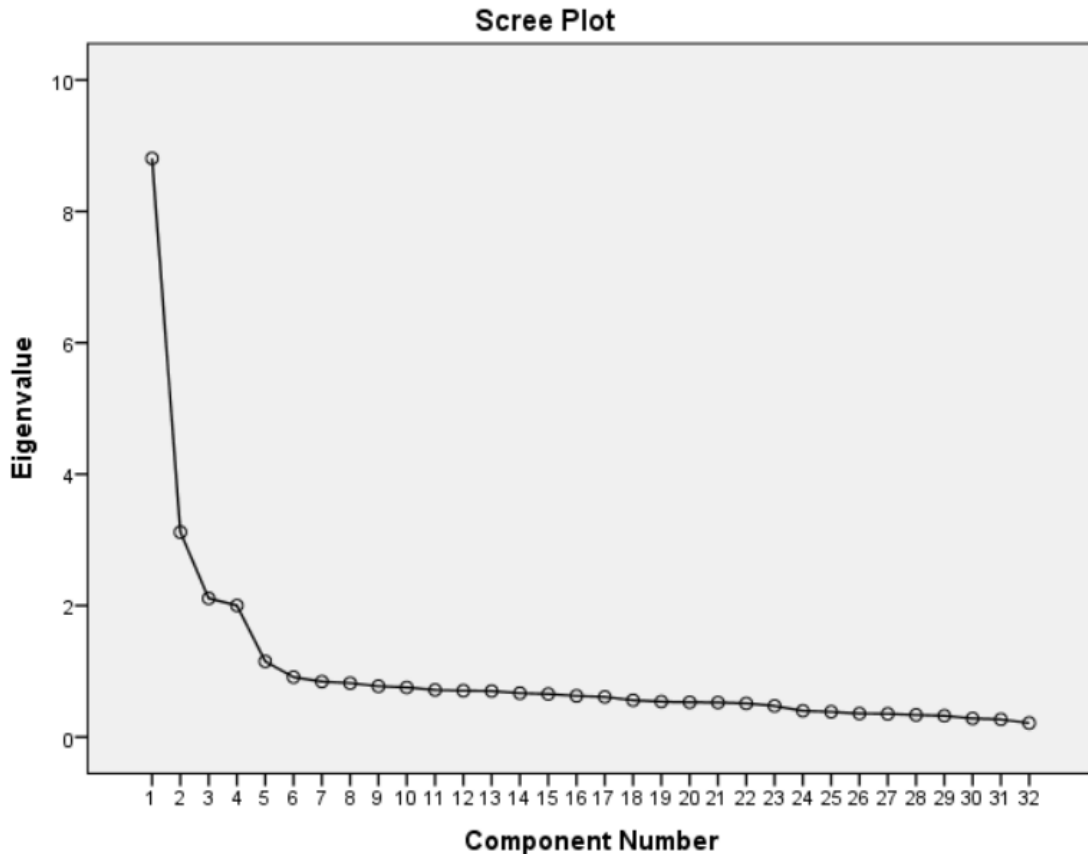
Eşvaryanslılık varsayımının kontrol edilmesi için gerçekleştirilen Box M testi sonuçlarına göre verilerde eşvaryanslılığa ilişkin bir problem yoktur ( $p>0,01$ ).

**Örneklem büyüklüğü.** Comrey ve Lee (1992), örneklem büyüklüğü olarak 100'ün zayıf, 200'ün orta, 300'ün iyi, 500'ün çok iyi ve 1000'in ise mükemmel olduğunu ifade etmiştir. Bryman ve Cramer (2001), örneklem büyüklüğünün madde sayısının 5 veya 10 katı olmasının yeterli olduğunu belirtmişlerdir. Kline (2011) ise, örneklem büyüklüğünün 200 olmasının yeterli olacağını ifade etmiştir. Bu ölçütlere ek olarak, örneklem büyüklüğüne bağlı olarak verilerin yapısının faktör analizine ne düzeyde uygun olduğunu yorumlamak için kullanılan Kaiser-Meyer-Olkin (KMO) Örnekleme Yeterliği Ölçümü de kullanılmaktadır. KMO Testi aracılığıyla gözlenen korelasyon katsayılarının ve kısmi korelasyon katsayılarının büyüklüğü karşılaştırılır (Kalaycı, 2017). Hesaplanan değere bağlı olarak ölçekte yer alan değişkenlerin birbirleri tarafından ne düzeyde tahmin edilebileceği hakkında yorum yapılır. Örneklem büyüklüğü konusunda başka araştırmacılar tarafından ortaya atılmış farklı değerlerin de olduğu, dolayısıyla kesin bir kriterin olmadığı görülmektedir. Bu duruma çözüm olarak, alan yazında önerilen ölçütlerin en az ikisinin karşılandığı bir örneklem büyüklüğüne ulaşılması önerilmektedir (Çokluk ve diğerleri, 2016).

Analizlerde kullanılan veri setinde 32 madde ve 4457 katılımcı bulunmaktadır. Örneklem büyüklüğünün faktör analizi için uygun olup olmadığını sınamak amacıyla gerçekleştirilen KMO Testi'nin sonucu 0,93 olarak hesaplanmıştır. Bu istatistiğe göre, örneklem büyüklüğünün faktör analizi için mükemmel düzeyde olduğu sonucuna ulaşılmıştır (Şencan, 2005). Bu istatistikler ile yukarıda bahsedilen koşulların sağlandığı yani örneklem büyüklüğünün ilgili analizler için uygun olduğu sonucuna varılmıştır.

**Modelin oluşturulması.** Bu çalışma kapsamında TIMSS 2015 sekizinci sınıf öğrenci anketinde yer alan fen bilimleri dersi ile ilgili maddeler incelenmiştir. Model oluşturmak için, önce fen bilimleri ile ilgili duyuşsal özelliklere yönelik maddelere açımlayıcı faktör analizi yapılmış sonra bu model doğrulayıcı faktör analizi yapılarak doğrulanmıştır.

**Açımlayıcı faktör analizi.** KMO ve Bartlett Küresellik Testi sonuçları ile verilerin açımlayıcı faktör analizi için uygun olduğuna karar verilmiştir. Açımlayıcı faktör analizinde, faktör çıkarma yöntemi olarak Temel Bileşenler Analizi ve döndürme yöntemi olarak dik döndürme yöntemlerinden varimax (en fazla değişkenlik) kullanılmıştır. Faktör yük değerlerinin minimum kabul düzeyi 0,32 olarak belirlenmiştir.



### Şekil 7. Yamaç-Birikinti Grafiği

Şekil 7’de yer alan açımlayıcı faktör analizine ilişkin yamaç-birikinti grafiği incelendiğinde özdeğeri 1’den büyük olan 4 bileşen olduğu görülmüştür. İlgili grafik daha dikkatli incelendiğinde dördüncü boyuttan sonra bileşenlerin özdeğerleri daha küçük ve benzer düzeye gelmiştir. Boyut matrisi de modelin 4 boyutlu olduğunu doğrulamaktadır. Analize dahil edilen tüm maddelerin faktör yükleri 0,32’den yüksektir. Birinci boyutta 10 madde yer almaktadır. Bu boyutta yer alan bir madde (BSBS22C) ikinci boyutta da faktör yükü vermiştir ancak iki boyut arasındaki faktör yükleri farkı 0,1’den yüksek olduğu ve görece büyük olan faktör yükü bu boyutta olduğu için maddeyle ilgili işlem yapılmamıştır. Bu boyutta yer alan 10 maddenin faktör yükleri 0,48 ile 0,80 arasında değişmektedir. İkinci boyutta 9 madde yer almaktadır ancak üç madde (BSBG21G, BSBG21D ve BSBG21H) birinci boyutta da faktör yükü vermiştir. Söz konusu maddeler fen dersine ilişkin görüşlerle ilgili olduğu düşünülen maddeler olsa da öğretmene ilişkin görüşler boyutuyla binişiklik göstermiştir. Bu maddelerin boyutlar arasındaki faktör yükleri farkı 0,1’den daha düşük olduğu için farkı en az olandan en çok olana doğru sırayla analizden çıkarılmıştır. İkinci boyutta geriye kalan 6 maddenin faktör yükleri 0,62 ile 0,78 arasında değişmektedir. Üçüncü boyutta 7 madde yer almaktadır ve bu maddelerin faktör yükleri 0,38 ile 0,73 arasında değişmektedir. Dördüncü boyutta ise 6 madde yer almaktadır ve bu maddelerin faktör yükleri 0,57 ile 0,72 arasında değişmektedir. Maddelerin silinmesi sonucunda oluşan boyut matrisi ve madde faktör yükleri Tablo 4’te gösterilmiştir. Boyutlarda yer alan maddeler dikkate alındığında bu boyutlar sırasıyla öğretmene ilişkin görüşler (ÖG), fen dersine ilişkin görüşler (FG), okula ilişkin görüşler (OG) ve akran zorbalığı (AZ) şeklinde isimlendirilmiştir. 29 madde ve 4 boyuttan oluşan modelin toplam varyans açıklama oranı %52,11’dir.

Tablo 4

*Boyut Matrisi*

Madde	Boyut			
	1	2	3	4
BSBS22E	Öğretmenim sorularıma anlaşılır cevaplar verir	0,803		
BSBS22F	Öğretmenim feni açıklamada iyidir	0,802		
BSBS22I	Öğretmenim, hata yaptığımda hatamı nasıl düzeltereğimi söyler	0,794		
BSBS22J	Öğretmenim söylemek istediklerimi dinler	0,783		
BSBS22G	Öğretmenim öğrendiklerimi göstermeme izin verir	0,758		
BSBS22B	Öğretmenimi anlamak kolaydır	0,758		
BSBS22H	Öğretmenim öğrenmemizi sağlamak için farklı şeyler yapar	0,695		
BSBS22C	Öğretmenimin anlattıklarıyla ilgileniyorum	0,671	0,377	
BSBS22A	Öğretmenimin benden ne beklediğini bilirim	0,548		
BSBS22D	Öğretmenim bana yapılacak ilginç şeyler verir	0,484		
BSBS21E	Feni severim		0,776	
BSBS21I	Fen en sevdiğim derslerden birisidir		0,770	
BSBS21A	Fen öğrenmeyi severim		0,734	
BSBS21F	Okulda fen öğrenmeyi dört gözle beklerim		0,704	
BSBS21C	Fen sıkıcıdır		0,727	
BSBS21B	Keşke fen çalışmak zorunda kalmasam		0,619	
BSBS21G	Fen bana dünyadaki şeylerin nasıl çalıştığı öğretir	0,409	0,485	
BSBS21D	Fen dersinde birçok ilginç şey öğrenirim	0,369	0,442	
BSBS21H	Fen deneyleri yürütmeyi severim	0,325	0,421	
BSBG15F	Bu okula gitmekten gurur duyarım			0,729
BSBG15C	Bu okula aitmişim gibi hissederim			0,689
BSBG15B	Okuldayken güvende hissederim			0,680
BSBG15A	Okulda olmayı severim			0,678
BSBG15G	Okulda çok şey öğrenirim			0,598
BSBG15D	Okuldaki arkadaşlarımı görmeyi severim			0,492
BSBG15E	Okulumdaki öğretmenler bana karşı adil davranır			0,382
BSBG16C	Hakkımda yalan söylendi			0,719
BSBG16G	Hakkımda utanç verici bilgiler paylaşıldı			0,700
BSBG16E	Fiziksel şiddete maruz kaldım ya da yaralandım			0,667
BSBG16B	Oyunları ya da aktivitelere alınmadım			0,628
BSBG16A	Benimle dalga geçildi ya da bana lakap takıldı			0,622
BSBG16D	Bana ait bir eşya çalındı			0,574

**Doğrulayıcı faktör analizi.** Boyutlar isimlendirildikten sonra doğrulayıcı faktör analizine geçilmiştir. Doğrulayıcı faktör analiziyle, açıklayıcı faktör analizi sırasında oluşturulan model doğrulanmaya çalışılmıştır. Bu adımda, her boyuttaki en büyük faktör yüküne sahip gözlenen değişken referans değişken olarak belirlenmiştir ve diğer gözlenen değişkenler serbestçe kestirilmiştir. Öğretmene ilişkin görüşler boyutu için BSBS22E, fen dersine ilişkin görüşler boyutu için BSBS21E, okula ilişkin görüşler boyutu için BSBSG15F ve akran zorbalığı boyutu için BSBG16C referans değişken olarak seçilmiştir. Veriler sürekli düzeyde ve normal dağılıma sahip olduğu için model kestirimi için Maximum Likelihood (En Çok Olabilirlik) Yöntemi kullanılmıştır. DFA, AMOS programı kullanılarak yapılmıştır. Modeldeki değişkenlerin faktör yükleri ve hata varyansları ile modelin uyum indeksleri aşağıda yorumlanmıştır.

DFA sonucunda hesaplanan uyum istatistiklerine bakıldığında RMSEA ve SRMR değerleri iyi; GFI, AGFI, CFI ve TLI değerleri kabul edilebilir düzeydedir. Ancak,  $\chi^2$  değerinin anlamlı ( $p < 0,05$ ) ve  $\chi^2/sd$  değerinin 5'ten büyük olduğu görülmüştür. Bu durum,  $\chi^2$  değerinin örneklem büyüklüğünden etkilenmesi ve büyük örneklerde daha anlamlı sonuçlar vermeye yatkın olması ile açıklanabilir (Kline, 2016). Analizlere dahil edilen kişi sayısının 4457 olduğu göz önünde bulundurulduğunda, bu sonuçlar  $\chi^2$  ve  $\chi^2/sd$  değerleriyle ilgili problem olmadığı şeklinde yorumlanmıştır. Tüm uyum iyiliği istatistiklerine bakılarak uyum iyiliğinin yeterli düzeyde olduğu sonucuna varılmıştır. Modele ilişkin hesaplanan uyum indeksi değerleri Tablo 5'te sunulmuştur.

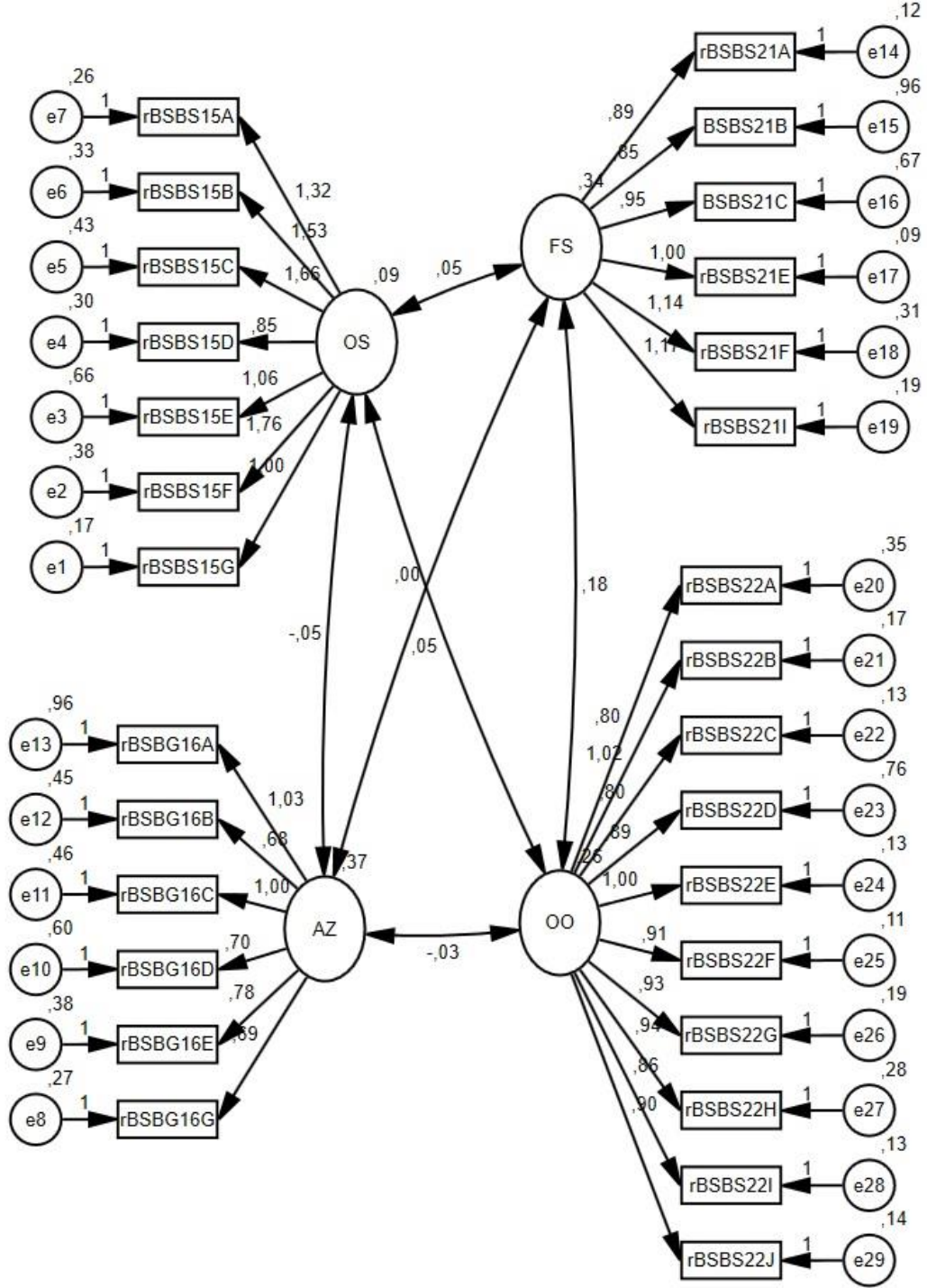
Tablo 5

*Duyuşsal Özellik Modeli Uyum Değerleri*

Uyum İndeksi	Hesaplanan Değer
$\chi^2$	$p < 0,05$
$\chi^2/sd$	12,304
RMSEA	0,050
SRMR	0,029
GFI	0,932
AGFI	0,921
CFI	0,921
TLI	0,914

Fen bilimleri duyuşsal özellik modelinin yol diyagramı incelendiğinde modelde yer alan 29 gözlenen deęişkenin 4 gizil deęişken altında toplandıęı görölmektedir. Gözlenen deęişkenlerin standartlaştırılmış faktör yükleri 0,36 ile 0,89 arasında deęişmektedir ve tüm faktör yükleri  $p=0,05$  düzeyinde anlamlıdır. Faktörlere ilişkin hata varyansları ise 0,09 ile 0,96 arasında deęişmektedir. Modelde yer alan standartlaştırılmamış faktör yüklerini ve hataları içeren yol diyagramı Şekil 8'de, standartlaştırılmış faktör yüklerini içeren yol diyagramı Şekil 9'da gösterilmiştir.





Şekil 8. Standartlaştırılmamış faktör yükleri ve hatalar

0,90'ın üzerindeki hata varyanslarının problem oluşturma ihtimali de düşünülerek modifikasyon indeksleri incelenmiştir. 0,90'ın üzere hata varyansı olan iki maddeden birisi olan BSBG16A maddesi incelendiğinde aynı faktör altında yer alan başka bir madde ile modifikasyon önerisi olmadığı için bu maddeyle ilgili

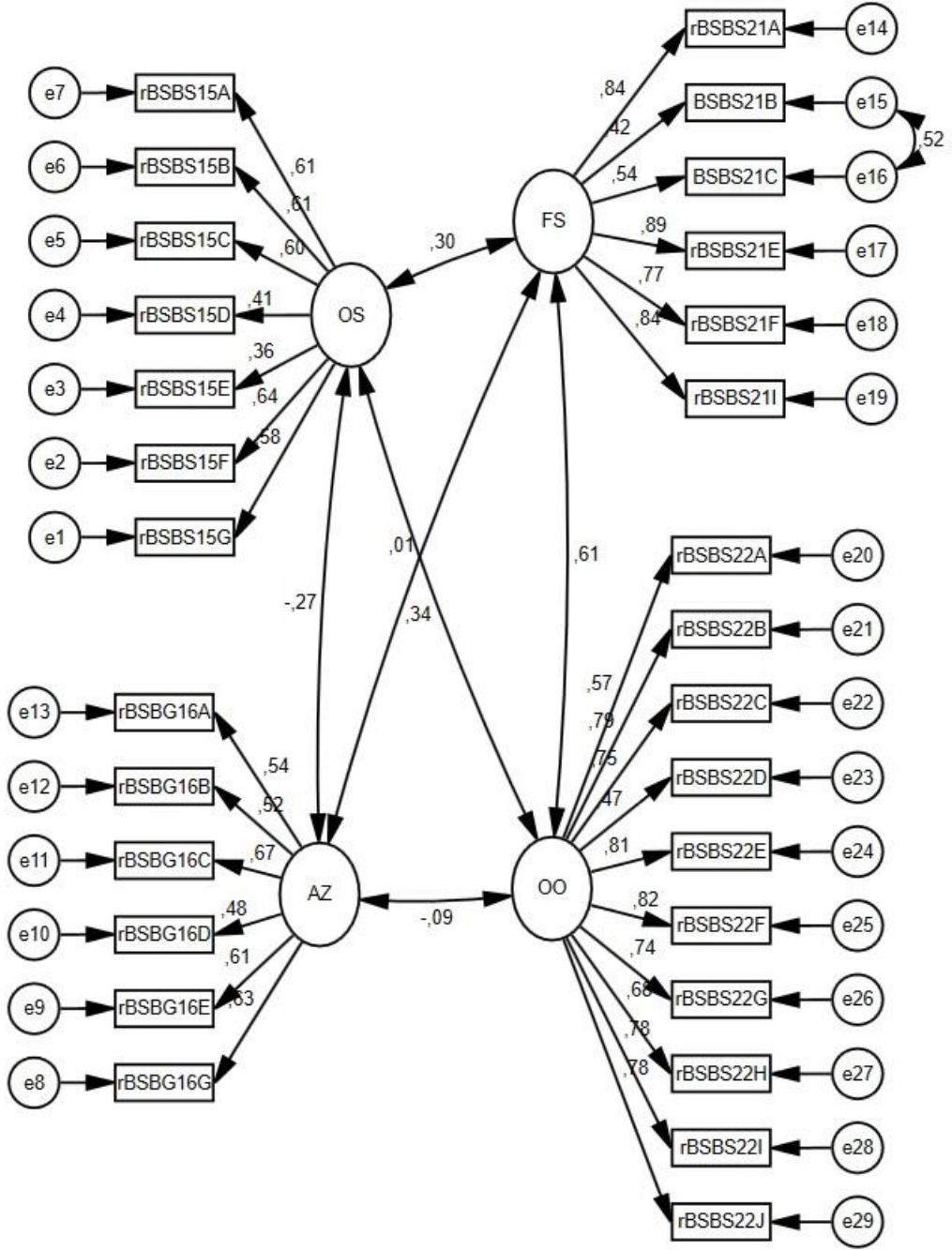
herhangi bir işlem yapılmamıştır. Ancak BSBS21B maddesi incelendiğinde, bu madde ile BSBS21C arasında kovaryans oluşturulması program tarafından önerilmiştir. Anket formunda yer alan maddeler incelendiğinde, BSBS21B maddesinin “Keşke fen çalışmak zorunda olmasam” ve BSBS21C maddesinin “Fen sıkıcıdır” şeklinde olduğu görülmüştür. Bu maddelerin benzer şekilde algılanabileceği ve maddelere benzer şekilde cevap verilmiş olunabileceğinden yola çıkılarak modifikasyon uygulanmış ve model yeniden oluşturulmuştur.

Modifikasyon sonrasında hesaplanan RMSEA ve SRMR değerlerinde iyileşme olduğu,  $\chi^2/sd$  değerinin 12,304’ten 8,725’e düştüğü ve diğer indeks değerlerinin (GFI, AGFI, CFI ve TLI) 0,94’e yükselerek iyi uyuma yaklaştığı görülmüştür. Yapılan modifikasyon sonucunda model veri uyumunun iyileştiği ifade edilebilir. Düzenlenen modelin uyum değerleri Tablo 6’da gösterilmiştir.

Tablo 6

*Düzenlenen Modelin Uyum Değerleri*

Uyum İndeksi	Hesaplanan Değer
$\chi^2$	p < 0,05
$\chi^2/sd$	8,725
RMSEA	0,042
SRMR	0,022
GFI	0,949
AGFI	0,941
CFI	0,946
TLI	0,941



Şekil 9. Düzenlenen modele ilişkin standartlaştırılmış faktör yükleri

Düzenlenen modele ilişkin standartlaştırılmamış faktör yükleri Şekil 10'da gösterilmiştir. Yapılan analizler sonucunda TIMSS 2015 Türkiye sekizinci sınıf fen duyuşsal özellik modeli doğrulanmıştır. Model doğrulandıktan sonra modelin cinsiyet ve bölge grupları arasındaki ölçme değişmezliğini test etmek için AMOS programı aracılığıyla çok ÇGDFA yapılmıştır. ÇGDFA ile ilgili sonuçlara bulgular ve yorumlar bölümünde detaylı olarak yer verilmiştir.

## Bölüm 4

### Bulgular ve Yorumlar

Bu bölümde, araştırmadaki ölçme değişmezliği analizinin bulgularına ve yorumlarına yer verilmiştir. Bu kapsamda 8. sınıf fen duyuşsal özellik modelinin cinsiyetler ve İBBS Düzey 1 bölge birimleri arasındaki ölçme değişmezliği incelemesinin bulguları sunulmuştur. ÇGDFA yöntemi ile yapılan ölçme değişmezliği incelenmesi aşamaları hiyerarşik olarak yapısal, metrik ve ölçek ve katı değişmezlik şeklinde sıralanarak yorumlanmıştır.

#### Araştırmanın 1. Alt Problemine İlişkin Bulgular

“TIMSS 2015 8. sınıf fen duyuşsal özellik modeli cinsiyetler arasında ölçme değişmezliği sağlamakta mıdır?” problemini yanıtlamak amacıyla erkek ve kız öğrencilerin verileri eş zamanlı şekilde kullanılarak ÇGDFA yapılmıştır. Ölçme değişmezliğinin incelenmesine başlamadan önce tüm verilerle doğrulanmış model, her iki cinsiyet grubu için doğrulanmıştır. Bu amaçla kız ve erkek öğrenciler için ayrı ayrı DFA yapılmıştır. Cinsiyetlere ilişkin model uyum istatistikleri Tablo 7’de sunulmuştur.

Tablo 7

#### *Cinsiyete Göre Model Uyum Değerleri*

Cinsiyet	$\chi^2$	sd	$\chi^2/sd$	RMSEA	CFI	IFI	TLI
Kız	1839,322	370	4,97	0,042	0,942	0,942	0,936
Erkek	1848,288	370	4,99	0,042	0,947	0,947	0,942

Tablo 7’de yer alan uyum iyiliğine ilişkin sonuçlar incelendiğinde  $\chi^2/sd$  değerlerinin 4,97 ve 4,99; RMSEA değerlerinin 0,042; CFI değerlerinin 0,942 ve 0,947; IFI değerlerinin 0,942 ve 0,947; TLI değerlerinin 0,936 ve 0,942 olduğu görülmüştür. Buna göre  $\chi^2/sd$ , CFI, IFI ve TLI değerleri kabul edilebilir, RMSEA değeri ise iyi uyum aralığında yer almaktadır. Tüm uyum indeksi değerlerinin ideal uyum aralığında olmasından ötürü modelin kız ve erkekler öğrenciler için uyumlu olduğu kabul edilmiş ve model doğrulanmıştır. Modelin doğrulanmasının ardından ÇGDFA yapılmıştır. Cinsiyetler arası ÇGDFA sonucunda hesaplanan uyum istatistiklerine Tablo 8’de yer verilmiştir.

Tablo 8

*Cinsiyete Göre Ölçme Değişmezliği Uyum Değerleri*

Değişmezlik Aşamaları	$\chi^2$	sd	$\chi^2/sd$	RMSEA	CFI	IFI	TLI	$\Delta CFI$
Yapısal Değişmezlik	3687,610	740	4,98	0,030	0,945	0,945	0,939	-
Metrik Değişmezlik	4071,076	769	5,29	0,031	0,938	0,938	0,934	0,007
Ölçek Değişmezliği	4523,402	798	5,67	0,032	0,930	0,930	0,929	0,008
Katı Değişmezlik	5164,267	834	6,19	0,034	0,919	0,918	0,921	0,011

**Yapısal değişmezlik.** Ölçme değişmezliğinin birinci aşaması olan yapısal değişmezlik aşamasında yapının tüm gruplar için benzer olup olmadığı test edilir. Tablo 8’de gösterilen hesaplanmış uyum indeksleri incelendiğinde RMSEA (0,030) değerinin iyi uyum;  $\chi^2/sd$  (4,98), CFI (0,945), IFI (0,945), TLI (0,939) değerlerinin ise kabul edilebilir uyum aralığında olduğu görülmüştür. Bununla birlikte CFI, IFI ve TLI değerlerinin iyi uyum aralığı alt sınırına çok yakın seyrettiği gözlenmiştir. Tüm değerlerin en az kabul edilebilir düzeyde olmasından dolayı bu sonuç, yapının tüm gruplarda benzer olduğu yani modelin cinsiyete göre yapısal değişmezlik gösterdiği şeklinde yorumlanmıştır. Yapısal değişmezliğin sağlanmasının ardından bir sonraki aşama olan metrik değişmezlik aşamasına geçilmiştir.

**Metrik değişmezlik.** Ölçme değişmezliğinin ikinci aşaması olan metrik değişmezlik aşamasında yapıya ek olarak madde faktör yüklerinin tüm gruplarda benzer olup olmadığı test edilir. Tablo 8’de gösterilen hesaplanmış uyum indeksleri incelendiğinde  $\chi^2/sd$  (5,29), RMSEA (0,031), CFI (0,938), IFI (0,938) ve TLI (0,934) değerleri görülmüştür.  $\chi^2/sd$  değeri kabul edilebilir aralığın hemen dışında olmakla birlikte bu sonuçlar, modelin yeterli düzeyde uyum gösterdiğini belirtmektedir. Ancak  $\chi^2$  istatistiklerinin büyük örneklerde yanlış sonuç verme olasılığından ötürü bu değerlere ek olarak, bir önceki değişmezlik aşamasındaki CFI değeriyle olan farkı gösteren  $\Delta CFI$  hesaplanmıştır. Hesaplanan  $\Delta CFI=0,007$  değerinin uyum için kabul edilen sınırlar olan  $\pm 0,01$  aralığında olduğu görülmüştür. Bu sonuçlarla, fen bilimleri duyuşsal özellik modelindeki yapıya ek olarak modeldeki madde faktör yüklerinin kız ve erkek öğrenciler için aynı olduğu yani metrik değişmezliğin sağlandığı kabul edilmiştir. Metrik değişmezliğin

sağlanmasının ardından bir sonraki aşama olan ölçek değişmezliği aşamasına geçilmiştir.

**Ölçek değişmezliği.** Ölçme değişmezliğinin üçüncü aşaması olan ölçek değişmezliği aşamasında faktör yüklerine ek olarak maddelere ilişkin sabitlerin tüm gruplarda benzer olup olmadığı test edilir. Tablo 8'de gösterilen hesaplanmış uyum indeksleri incelendiğinde  $\chi^2/sd$  (5,67), RMSEA (0,032), CFI (0,930), IFI (0,930) ve TLI (0,929) değerleri görülmüştür. Sonuçlar, modelin bu aşamada da yeterli düzeyde uyum gösterdiğini belirtmektedir. Metrik değişmezlik aşamasında hesaplanan CFI değeriyle farkı belirten  $\Delta CFI=0,008$  değerinin uyum için kabul edilen sınırlar olan  $\pm 0,01$  aralığında olduğu görülmüştür. Bu sonuçlarla, fen bilimleri duyuşsal özellik modelindeki yapı ve madde faktör yüklerine ek olarak madde sabitlerinin de kız ve erkek öğrenciler için aynı olduğu yani metrik değişmezliğin sağlandığı kabul edilmiştir. Madde sabitlerinin benzer olması puanlardaki cinsiyetler arası farklılığın gizil değişken dışında başka bir kaynağı olmadığını yani madde düzeyinde yanlılık bulunmadığını gösterir. Bu eşitliğin sağlanmış olması cinsiyet ortalama puanlarının karşılaştırılmasını anlamlı hale getirir.

**Katı değişmezlik.** Ölçme değişmezliğinin son aşaması olan katı değişmezlik aşamasında madde sabitlerine ek olarak madde hata varyanslarının tüm gruplarda benzer olup olmadığı test edilir. Tablo 8'de gösterilen hesaplanmış uyum indeksleri incelendiğinde  $\chi^2/sd$  (6,19), RMSEA (0,035), CFI (0,919), IFI (0,918) ve TLI (0,921) değerleri görülmüştür.  $\chi^2/sd$  ideal aralığın hemen dışında; RMSEA değeri iyi uyum aralığında; diğer uyum indeks değerleri kabul edilebilir uyum aralığındadır. Ancak, ölçek değişmezliği aşamasında hesaplanan CFI değeriyle farkı belirten  $\Delta CFI=0,011$  değerinin uyum için kabul edilen sınırlar olan  $\pm 0,01$  aralığında olmadığı görülmüştür. Elde edilen 0,011 sonucu kabul aralığının hemen üstündedir. RMSEA değerinin iyi uyum; diğer indeks değerlerinin ise kabul edilebilir uyum aralığında olduğu göz önünde bulundurulursa modelin katı değişmezlik koşulunu sağladığı ifade edilebilir.

Bu sonuçlar, sekizinci sınıf fen duyuşsal özellik modelinin cinsiyete göre değişmez olduğunu göstermektedir. Katı değişmezlik koşulu sağlandığına göre kız ve erkek öğrencilerin duyuşsal özellik modeli için madde hata varyansları da dahil olmak üzere yapılacak tüm karşılaştırmalarda bir problem yoktur. Geçerliğin

kullanılan ölçme aracının ölçmek üzere planlandığı özelliği başka bir özelliğin etkisi olmadan ölçebilme derecesi olduğu göz önünde bulundurulduğunda, katı değişmezliğin sağlanmış olması elde edilen puanların geçerliği hakkında kanıt sağlamıştır. Eldeki duruma göre, duyuşsal özellik modelinde yer alan değişken puanları kız ve erkek öğrenciler arasında karşılaştırılabilir çünkü modelin teorik yapısı her iki grupta da aynıdır. Başka bir ifadeyle, gizil değişken puanları eşit olan kız ve erkek öğrencilerin gözlenen değişken puanları da eşittir.

Uzun (2008), TIMSS-R Türkiye uygulamasındaki fen başarısını etkileyen derse yönelik tutum, derse verilen önem, özyeterlik ve sınıf için öğrenci etkinlikleri değişkenleriyle bir model oluşturmuş ve bu modelin cinsiyet açısından ölçme değişmezliğini test etmiştir. Sonuçlara göre tutum boyutu ölçek değişmezliğini sağlarken diğer üç boyut metrik değişmezlik koşulunu sağlamıştır. İki araştırmanın sonucu birbiriyle kısmen benzerlik gösterse de farklı sonuçlar elde edilmesi modellerde yer alan değişkenlerin farklı olmasıyla açıklanabilir.

## **Araştırmanın 2. Alt Problemine İlişkin Bulgular**

“TIMSS 2015 8. sınıf fen duyuşsal özellik modeli bölgeler arasında ölçme değişmezliği sağlamakta mıdır?” problemini yanıtlamak amacıyla tüm bölgelerin verileri eş zamanlı şekilde kullanılarak ÇGDFA yapılmıştır. Bölgeler arasında ölçme değişmezliğinin incelenmesine başlamadan önce tüm verilerle doğrulanan model bölgelerin her biri için doğrulanmıştır. Bu amaçla her bölge için ayrı ayrı DFA yapılmıştır. Bölgelere ilişkin model uyum istatistikleri Tablo 9’da sunulmuştur.

Tablo 9

*Bölgelere Göre Model Uyum Değerleri*

Bölgeler	$\chi^2$	sd	$\chi^2/sd$	RMSEA	CFI	IFI	TLI
İstanbul	880,585	370	2,38	0,045	0,940	0,941	0,935
Batı Marmara	603,018	370	1,63	0,063	0,888	0,890	0,887
Ege	693,544	370	1,87	0,039	0,958	0,958	0,954
Doğu Marmara	770,502	370	2,08	0,047	0,928	0,929	0,921
Batı Anadolu	791,125	370	2,14	0,059	0,899	0,900	0,890
Akdeniz	872,937	370	2,36	0,052	0,924	0,924	0,916
Orta Anadolu	552,487	370	1,49	0,048	0,905	0,907	0,896
Batı Karadeniz	541,922	370	1,47	0,051	0,915	0,916	0,906
Doğu Karadeniz	635,954	370	1,72	0,060	0,913	0,914	0,904
Kuzeydoğu Anadolu	539,438	370	1,46	0,044	0,921	0,923	0,914
Ortadoğu Anadolu	671,826	370	1,82	0,052	0,925	0,926	0,918
Güneydoğu Anadolu	874,085	370	2,36	0,048	0,923	0,923	0,915

Tablo 9'da yer alan uyum iyiliğine ilişkin sonuçlar incelendiğinde  $\chi^2/sd$  değerlerinin 1,49 - 2,38; RMSEA değerlerinin 0,044 – 0,063; CFI değerlerinin 0,888 – 0,958; IFI değerlerinin 0,890 – 0,958; TLI değerlerinin 0,887 – 0,954 aralığında oldukları görülmüştür. Buna göre, tüm bölgelerin  $\chi^2/sd$  ve RMSEA değerleri iyi veya kabul edilebilir değerlere sahiptir. Batı Marmara, Batı Anadolu ve Orta Anadolu bölgeleri dışındaki tüm bölgelerin CFI, IFI ve TLI değerleri kabul edilebilir veya iyi uyum aralığında yer almaktadır. Bununla birlikte, Batı Marmara bölgesinin CFI (0,888), IFI (0,890) ve TLI (0,887); Batı Anadolu bölgesinin CFI (0,899) ve TLI (0,890); Orta Anadolu bölgesinin TLI (0,896) değerleri kabul edilebilir alt sınır olan 0,90'dan düşüktür. Ancak, hem  $\chi^2/sd$  ve RMSEA değerlerinin yeterli uyuma sahip olması hem de kabul edilebilir aralığın dışında



olan CFI, IFI ve TLI değerlerin kabul sınırının çok az altında olmasından ötürü modelin bölgelerin her birinde iyi uyum gösterdiği kabul edilmiştir ve model doğrulanmıştır. Modelin doğrulanmasının ardından ÇGDFA yapılmıştır. Bölgeler arası ÇGDFA sonucunda hesaplanan uyum istatistiklerine Tablo 10'da yer verilmiştir.

Tablo 10

*Bölgelere Göre Ölçme Değişmezliği Uyum Değerleri*

Değişmezlik Aşamaları	$\chi^2$	sd	$\chi^2/sd$	RMSEA	CFI	IFI	TLI	$\Delta CFI$
Yapısal Değişmezlik	8431,570	4440	1,90	0,014	0,927	0,927	0,919	-
Metrik Değişmezlik	9190,014	4759	1,93	0,014	0,918	0,919	0,917	0,009
Ölçek Değişmezliği	9764,563	5078	1,92	0,014	0,914	0,913	0,917	0,004
Katı Değişmezlik	11193,307	5474	2,05	0,015	0,895	0,894	0,906	0,029

**Yapısal değişmezlik.** Tablo 8'de gösterilen hesaplanmış uyum indeksleri incelendiğinde  $\chi^2/sd$  (1,90) ve RMSEA (0,014) değerlerinin iyi; CFI (0,927), IFI (0,927) ve TLI (0,919) değerlerinin kabul edilebilir uyum aralığında olduğu gözlenmiştir. Tüm değerlerin en az kabul edilebilir düzeyde olmasından dolayı yapının tüm gruplarda benzer olduğu yani fen bilimleri duyuşsal özellik modelinin bölgeler arasında yapısal değişmezlik gösterdiği doğrulanmıştır. Yapısal değişmezliğin sağlanmasının ardından bir sonraki aşama olan metrik değişmezlik aşamasına geçilmiştir.

**Metrik değişmezlik.** Tablo 8'de gösterilen hesaplanmış uyum indeksleri incelendiğinde  $\chi^2/sd$  (1,93), RMSEA (0,014), CFI (0,918), IFI (0,919) ve TLI (0,917) değerleri görülmüştür. Bu sonuçlar, modelin yeterli düzeyde uyum gösterdiğini belirtmektedir. Ancak  $\chi^2$  istatistiklerinin büyük örneklemelerde yanlış sonuç verme olasılığından ötürü bu değerlere ek olarak, bir önceki değişmezlik aşamasındaki CFI değeriyle olan farkı gösteren  $\Delta CFI$  hesaplanmıştır. Hesaplanan  $\Delta CFI=0,009$  değerinin uyum için kabul edilen sınırlar olan  $\pm 0,01$  aralığında olduğu görülmüştür. Bu sonuçlarla, fen bilimleri duyuşsal özellik modelindeki yapıya ek olarak modeldeki madde faktör yüklerinin tüm bölgelerde aynı olduğu yani metrik

değişmezliğin sağlandığı kabul edilmiştir. Metrik değişmezliğin sağlanmasının ardından bir sonraki aşama olan ölçek değişmezliği aşamasına geçilmiştir.

**Ölçek değişmezliği.** Tablo 8'de gösterilen hesaplanmış uyum indeksleri incelendiğinde  $\chi^2/sd$  (1,92), RMSEA (0,014), CFI (0,914), IFI (0,913) ve TLI (0,917) değerleri görülmüştür. Bu sonuçlar, modelin bu aşamada da yeterli düzeyde uyum gösterdiğini belirtmektedir. Metrik değişmezlik aşamasında hesaplanan CFI değeriyle farkı belirten  $\Delta CFI=0,004$  değerinin uyum için kabul edilen sınırlar olan  $\pm 0,01$  aralığında olduğu görülmüştür. Bu sonuçlarla, fen bilimleri duyuşsal özellik modelindeki yapı ve madde faktör yüklerine ek olarak madde sabitlerinin de tüm bölgelerde aynı olduğu yani metrik değişmezliğin sağlandığı kabul edilmiştir. Madde sabitlerinin benzer olması puanlardaki bölgeler arası farklılığın gizil değişken dışında başka bir kaynağı olmadığını yani madde düzeyinde yanlılık bulunmadığını gösterir. Bu eşitliğin sağlanmış olması bölgelerin ortalama puanlarının karşılaştırılmasını anlamlı hale getirir.

**Katı değişmezlik.** Tablo 8'de gösterilen hesaplanmış uyum indeksleri incelendiğinde  $\chi^2/sd$  (2,05), RMSEA (0,015), CFI (0,895), IFI (0,894) ve TLI (0,906) değerleri görülmüştür. Bu sonuçların bir kısmı iyi düzeyde, bir kısmı kabul edilebilir sınırlara çok yakın seyretse de ölçek değişmezliği aşamasında hesaplanan CFI değeriyle farkı belirten  $\Delta CFI=0,029$  değerinin uyum için kabul edilen sınırlar olan  $\pm 0,01$  aralığında olmadığı görülmüştür. Hesaplanan 0,029 değeri kabul aralığının çok üzerindedir ve buna ek olarak iki indeks değeri kabul edilebilir aralığın dışında yer almaktadır. Bu durum, madde hata varyanslarının tüm gruplarda benzer olmadığını yani fen duyuşsal özellik modelinin bölgeler arasında katı değişmezliği sağlamadığını göstermektedir.

Elde edilen sonuçlar fen duyuşsal özellik modelinin bölgelere göre tamamen değişmez olmadığını göstermektedir. Ölçek değişmezliği sağlandığı için bölgelere göre duyuşsal özellik puanları karşılaştırıldığında doğru yorumlamalar yapılabilir ancak tam bir değişmezlikten söz edilemediği için madde hata varyanslarının karşılaştırılması doğru olmayacaktır. Ölçme değişmezliği incelemeleri, çalışmaların geçerlik düzeyini belirlemeye yönelik araştırmalardır ve tam değişmezliğinin sağlanamaması elde edilen puanlarının geçerliğini olumsuz yönde etkilemektedir.

Katı deęişmezlięin saęlanamamış olması, katı deęişmezlięin saęlanması zor bir koşul olmasıyla birlikte bu çalışmada yer alan bölgeler arasındaki farklılıklarla da açıklanabilir. Söz konusu karşılaştırmada yer alan bölgeler gelişmişlik düzeyi, nüfus yoğunluğu, coęrafi özellikleri; bölgelerde ikamet eden öğrencilerin anne-baba eğitim düzeyi, sosyoekonomik statü, vb. deęişkenler açısından belirgin düzeyde farklıdır. Bölgelerin birbirlerinden belirgin düzeyde farklı olması duyuşsal özellik puanları arasındaki farkı dolaylı yoldan etkilemektedir.

Ölçüoęlu (2015), TIMSS 2011 verileriyle Türkiye sekizinci sınıf matematik başarısı matematik başarısı etkileyen üç boyutu bir model oluşturmuş ve bu modelin Türkiye’de yer alan 7 coęrafi bölgeye göre ölçme deęişmezlięini test etmiştir. Analizler sonucunda modelin ölçek deęişmezlięini saęladığını ancak katı deęişmezlięi saęlamadığını belirtilmiştir.

Polat (2019), TIMSS 2015 verileriyle Türkiye sekizinci sınıf fen bilimleri başarısı etkileyen dört boyutlu bir model oluşturmuş ve bu modelin Türkiye İBBS Düzey 1’de yer alan bölgelere göre ölçme deęişmezlięini incelemiştir. Oluşturulan modelin bölgeler arasında ölçek deęişmezlięi saęladığını ancak katı deęişmezlik koşulunu saęlamadığını görülmüştür. Bu çalışmadan elde edilen sonuçlar, alanyazındaki araştırmaların sonuçlarıyla benzerlik göstermektedir.

## **Bölüm 5**

### **Sonuçlar ve Öneriler**

Bu bölümde araştırmanın alt problemlerine ilişkin sonuçlar özetlenmiş ve öneriler sunulmuştur.

#### **Sonuçlar**

Bu araştırmada, TIMSS 2015 sekizinci sınıf fen bilimleri öğrenci anketinde yer alan maddeler kullanılarak oluşturulan fen bilimleri duyuşsal özellik modelinin cinsiyete ve Türkiye İBBS Düzey 1'deki bölgelere (İstanbul, Batı Marmara, Ege, Doğu Marmara, Batı Anadolu, Akdeniz, Orta Anadolu, Batı Karadeniz, Doğu Karadeniz, Kuzeydoğu Anadolu, Ortadoğu Anadolu ve Güneydoğu Anadolu) göre ölçme değişmezliği test edilmiştir. Yapılan analizler sonucunda oluşturulan modelin cinsiyet için tüm değişmezlik koşullarını sağladığı, bölgeler için ölçek değişmezliği koşulunu sağladığı ancak katı değişmezlik koşulunu sağlamadığı görülmüştür.

#### **Cinsiyete göre ölçme değişmezliği sonuçları.**

Yapısal değişmezliğin sağlanmış olması, modelin iki grupta da aynı olduğunu göstermektedir yani model kız ve erkek öğrenci gruplarında için aynı yapıyı ölçmektedir. Bu sonuçlar, iki grubun yapısının karşılaştırmaya uygun nitelikte olduğunu ifade etmektedir.

Metrik değişmezliğin sağlanması maddelerin faktör yüklerinin kız ve erkek öğrenciler için eşit olduğunu göstermektedir. Faktör yüklerinin eşitliği, kız ve erkek öğrencilerin modelde yer alan maddeleri birbirlerinden farklı bir şekilde yorumlamadığı anlamına gelmektedir.

Ölçek değişmezliğinin karşılanması modeldeki madde sabitlerinin kız ve erkek öğrenciler için eşit olması, maddelerin kız veya erkek öğrenciler lehine yanlılık göstermediğini ve maddelere verilen yanıtlardaki farklılığın kaynağının örtük değişken puanlarındaki farklılık olduğunu gösterir. Bu durumda, kız ve erkek öğrencilerin gözlenen değişken ortalamaları anlamlı bir şekilde karşılaştırılabilir.

Katı değişmezlik aşamasında hata varyanslarının gruplar arasında aynı olup olmadığı incelenmektedir ancak, gizil değişken varyansları arttıkça gözlenen değişkenlerin hata varyansları da arttığı için katı değişmezlik, pratikte karşılanması

zor bir kořuldur. Cinsiyete gre katı deęiřmezlięin saęlanmıř olması modelde yer alan madde hata varyanslarının kız ve erkek ęrenciler arasında benzer olduęunu gstermektedir. Gzlenen deęiřken puanlarının varyans ve kovaryansları cinsiyete gre karřılařtırmaya uygundur.

### **Blgelere gre lme deęiřmezlięi sonuları.**

Yapısal deęiřmezlik saęlandıęı iin model tm blgelerde aynı yapıyı temsil etmektedir. Buna gre, modeldeki faktr yapısı tm blgeler iin aynıdır ve yapı, blgeler arasında karřılařtırılabilir.

Metrik deęiřmezlik kořulunun karřılanması modelin madde faktr yklerinin tm blgeler iin eřit olduęu anlamına gelmektedir. Bu sonuca gre, tm blgelerdeki ęrenciler maddeleri benzer řekilde yorumlamıřtır.

lek deęiřmezlięinin saęlanmıř olması, madde sabitlerinin tm blgelerde eřittir ve hibir blge lehine yanlılık olmadıęını gstermektedir. Yani duyuřsal zelliklerin blgeler arasındaki farkını sebebi rtk deęiřkenin kendisidir. Bu sonuca gre, modelde yer alan maddelerin ortalamalarının blgelere gre karřılařtırılması herhangi bir hatalı yoruma yol amaz.

Katı deęiřmezlik kořulu, gruplar arasında adil ve yansız bir karřılařtırma iin gerekli olsa da hata varyanslarının gizil deęiřkenle doęru orantılı olmasından tr saęlanması gtr. Blgeler arasında katı deęiřmezlięin karřılanamaması modeldeki hata varyanslarının tm blgelerde eřit olmadıęı anlamına gelmektedir. Hata varyanslarının eřit olmamasından dolayı, blgelere gre gzlenen deęiřken varyans ve kovaryanslarını anlamlı bir řekilde karřılařtırmak mmkn deęildir.

## **Öneriler**

Bu çalışmada ölçme değişmezliği incelemesi, TIMSS 2015 Türkiye uygulamasındaki 8. sınıf öğrenci anketinde yer alan fen bilimleri dersi başarısını etkilediği düşünülen maddelerin verileriyle ÇGDFA yöntemi kullanılarak cinsiyetler ve bölgeler arası düzeyde yürütülmüştür. Elde edilen sonuçlardan hareketle uygulamaya ve gelecek araştırmalara yönelik öneriler aşağıda sıralanmıştır.

### **Uygulamaya yönelik öneriler.**

Model, cinsiyetler arasında tam değişmez iken bölgeler arasında değildir. Bölgeler arasında tam değişmezliğin sağlanmaması geçerliği olumsuz yönde etkilemektedir. Bu durumda bölgelerin duyuşsal özellik puanlarına ilişkin yorumlar için kanıt sağlanmamıştır. Duyuşsal özelliklerin ders başarısı üzerindeki etkisi dikkate alınırsa bölgeler arasında başarı açısından bir karşılaştırma yapılacağı zaman bu sonuç göz önünde bulundurulmalıdır.

Araştırmada kurulan modeldeki öğretmene ilişkin görüşler boyutunun pozitif yönde ve en büyük etkiye sahip olduğu gözlenmiştir. Bu boyutta en büyük faktör yükü veren maddelere bakıldığında öğretmenin ders anlatımındaki ve hataları açıklamadaki ifade gücüyle ilgili maddeler görülmektedir. Öğretmenleri ifade becerisinin iyileşmesi öğrencilerin fen başarısı üzerinde olumlu etkiler yaratacaktır. Fen başarısını artırmak için yapılacak olası çalışmalarda öğretmenlerin ifade gücünü geliştirmeye yönelik çalışmaların olması önemli kazanımlar sağlayacaktır.

### **Gelecek araştırmalara yönelik öneriler.**

Bölgeler arasında tam değişmezliğin sağlanmaması istenmedik bir durumdur. Daha sonra yapılacak araştırmalarda çalışmanın yürütüleceği bireylerin özellikleri göz önünde bulundurulmalıdır. Geçerlik açısından ortaya çıkabilecek problemlere en başından çözüm üretmek amacıyla asıl uygulama yapılmadan değişmezlik çalışmaları yürütülmelidir.

8. sınıfa ek olarak 4. sınıf verileri de kullanılarak iki sınıf düzeyi arasında karşılaştırmalar yapılabilir.

Fen bilimleri dersine ek olarak matematik dersi için de bir model oluşturularak iki model arasında karşılaştırılmalar yapılabilir.

Bu alıřmadaki modelde kullanılan deęiřkenlere ek olarak bařka deęiřkenler de kullanılarak daha kapsamlı bir model aracılıęıyla lme deęiřmezlięi test edilebilir.

đrenci anketi verilerine ek olarak đretmen ve/veya okul yneticisi anketleri de kullanılarak her bir anket iin ayrı yapısal eřitlik modeli kurulabileceęi gibi ok dzeyli yapısal eřitlik modeliyle tm anket verileri aynı anda analiz edilerek daha kapsamlı bir arařtırma modeli kurulabilir.

GDFA yntemine ek olarak lme deęiřmezlięinin incelendięi bařka yntem veya yntemler kullanılarak farklı yntem sonularını karřılařtıran alıřmalar yrtlebilir.

## Kaynaklar

- AERA, APA, & NCME (2014). *The standards for educational psychological testing*. Washington, US: American Educational Research Association.
- Akyıldız, M. (2009). PIRLS 2001 testinin yapı geçerliliğinin ülkelerarası karşılaştırılması. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 6(1), 18-47.
- Bahadır, E. (2012). *Uluslararası öğrenci değerlendirme programına (PISA 2009) göre Türkiye'deki öğrencilerin okuma becerilerini etkileyen değişkenlerin bölgelere göre incelenmesi* (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Başusta, N. B. (2010). Ölçme eşdeğerliği. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 1(2), 58-64.
- Başusta, N. B. & Gelbal, S. (2015). Gruplar arası karşılaştırmalarda ölçme değişmezliğinin test edilmesi: PISA öğrenci anketi örneği. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 30(4), 80-90.
- Bialosiewicz, S., Murphy, K., & Berry, T. (2013, Haziran). An introduction to measurement invariance testing: Resource packet for participants. <http://comm.eval.org/HigherLogic/System/DownloadDocumentFile.ashx?DocumentFileKey=63758fed-a490-43f2-8862-2de0217a08b8> adresinden elde edildi.
- Bloom, B. (2012). *İnsan nitelikleri ve okulda öğrenme*. (Çev. D. A. Özçelik). Ankara: MEB Yayınları. (Orijinal yayın tarihi, 1979).
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley-Interscience Publication.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (Second Edition). New York: The Guilford Press.
- Bryman, A. & Cramer, D. (2001). *Quantitative data analysis with SPSS release 10 for windows: A guide for social scientists*. Philadelphia: Routledge.



- Bryne, B. M., & Watkins, D. (2003). The issue of measurement invariance revisited. *Journal of Cross-Cultural Psychology, 34*(2), 155-175.
- Büyüköztürk, Ş., Çakmak, E. K., Akgün, Ö. E., Karadeniz, Ş. ve Demirel, F. (2012). *Bilimsel araştırma yöntemleri* (12. Baskı). Ankara: Pegem Akademi Yayıncılık.
- Büyüköztürk, Ş. (2005). *Sosyal bilimler için veri analizi el kitabı: İstatistik, araştırma deseni, SPSS uygulamaları ve yorum*. Ankara: Pegem Akademi Yayıncılık.
- Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). New Jersey: Lawrence Erlbaum.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Philadelphia: Harcourt Brace Jovanovich College Publishers.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*(4), 281-302.
- Çokluk, Ö., Şekercioğlu, G. & Büyüköztürk, Ş. (2016). *Sosyal bilimler için çok değişkenli istatistik SPSS ve LISREL uygulamaları*. Ankara: Pegem Akademi.
- Ercikan, K., & Koh, K. (2005). Examining the construct comparability of the English and French version of TIMSS. *International Journal of Testing, 5*(1), 23-35.
- Erşan, Ö. (2016). *TIMSS 2011 sekizinci sınıf öğrencilerinin matematik başarılarını etkileyen faktörlerin çok düzeyli yapısal eşitlik modeliyle incelenmesi* (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups?: Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care, 44*(1), 78-94.
- Gülleroğlu, H. D. (2017). PISA 2012 Matematik uygulamasına katılan Türk öğrencilerin duyuşsal özelliklerinin cinsiyete göre ölçme değişmezliğinin incelenmesi. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi, 37*(1), 151-175.

- Güngör, M. (2019). *Fen motivasyonu ve öz yeterliği modelinin ölçme değişmezliğinin incelenmesi: PISA 2015 Türkiye örneği* (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Harrington, D. (2009). *Confirmatory factor analysis*. New York: Oxford University Press.
- Hoyle, R. H. (Ed.). (2012). *Handbook of structural equation modeling*. New York, US: The Guilford Press.
- Ho, R. (2013). *Handbook of univariate and multivariate data analysis with IBM SPSS* (2nd ed). Northwestern, US: CRC Press.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.
- Jöreskog, K. G., & Sörbom, D. (1993). *Lisrel 8: Structural equation modeling with the simplis command language*. Lincolnwood: Scientific Software International, Inc.
- Jöreskog, K., & Sörbom, D. (2001). *Lisrel 8: User's reference guide*. Scientific Software International Inc.
- Kalaycı, Ş. (2017). *SPSS uygulamalı çok değişkenli istatistik teknikleri* (7. bs.). Ankara: Asil Yayınevi.
- Karakoç Alatlı, B. (2016). *Uluslararası Öğrenci Değerlendirme Programı (PISA-2012) okuryazarlık testlerinin ölçme değişmezliğinin incelenmesi* (Yayımlanmamış doktora tezi). Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Kıbrısoğlu, N. (2015). *PISA 2012 matematik öğrenme modelinin kültürlere ve cinsiyete göre ölçme değişmezliğinin incelenmesi: Türkiye – Çin (Şangay) – Endonezya örneği* (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Kline, R. B. (2011). *Principles and practices of structural equation modeling*. New York: The Guilford Press.

- Maruyama, G. (1997). *Basics of structural equation modeling*. CA: Sage Publications.
- Meade, A. W., & Lautenschlager, G. J. (2004). A comparison of item response theory and confirmatory factor analytic methodologies for establishing measurement equivalence/invariance. *Organizational Research Methods*, 7(4), 361-388.
- MEB. (2016). *TIMSS 2015 Ulusal Matematik ve Fen Ön Raporu: 4. ve 8. Sınıflar*. Ankara.
- Meydan, C. H. ve Şeşen, H. (2015). *Yapısal eşitlik modellemesi ve AMOS uygulamaları*. Ankara: Detay Yayıncılık.
- Millsap, R. E., & Olivera-Aguilar, M. (2012). *Investigating measurement invariance using confirmatory factor analysis*. R. H. Hoyle (Ed.), Handbook of structural equation modeling, içinde (s. 380-392). New York, US: Guilford Press.
- Mullis, I. V. S., & Martin, M. O. (Eds.) (2013). *TIMSS 2015 Assessment Frameworks*. Boston, US.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS 2015 international results in mathematics*. Boston, US.
- Osterlind, S. J., & Everson, H. T. (2009). *Differential item functioning*. (2<sup>nd</sup> Edition). California, US: SAGE Publications.
- Öğretmen, T. (2006). *Uluslararası okuma becerilerinde gelişim projesi (PIRLS) 2001 testinin psikometrik özelliklerinin incelenmesi: Türkiye-Amerika Birleşik Devletleri örneği* (Yayımlanmamış doktora tezi). Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Ölçüoğlu, R. (2015). *TIMSS 2011 Türkiye sekizinci sınıf matematik başarısını etkileyen değişkenlerin bölgelere göre incelenmesi* (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Önen, E. (2009). *Ölçme değişmezliğinin yapısal eşitlik modelleme teknikleri ile incelenmesi* (Yayımlanmamış doktora tezi). Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.

- Öztürk, A. (2009). *Homojen ve fonksiyonel bölgelerin tespiti ve Türkiye için istatistikî bölge birimleri önerisi* (Devlet Planlama Teşkilatı planlama uzmanlığı tezi). T.C. Başbakanlık Devlet Planlama Teşkilatı Müsteşarlığı, Bölgesel Gelişme ve Yapısal Uyum Genel Müdürlüğü, Ankara.
- Polat, M. (2019). *TIMSS 2015 matematik ve fen duyuşsal özellik modellerinin kültürlere, cinsiyete ve bölgelere göre ölçme değışmezliđinin incelenmesi* (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü.
- Raju, N. S., Laffitte, L. J., & Byrne, B. M. (2002). Measurement equivalence: A comparison of methods based on confirmatory factor analysis and item response theory. *Journal of Applied Psychology, 87*(3), 517-529.
- Raykov, T., & Marcoulides, G. A. (2008). *An introduction to applied multivariate analysis*. NY: Taylor & Francis Group.
- Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin, 114*(3), 552-566.
- Schumacker, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modeling* (3<sup>rd</sup> Edition). New York: Taylor & Francis Group.
- Schmitt, N., & Kuljanin, G. (2008). Measurement invariance: Review of practice and implication. *Human Resources Management Review, 18*(4), 210-222.
- Somer, O., Korkmaz, M., Dural, S. ve Can, S. (2009). Ölçme eşdeğerliğinin yapısal eşitlik modellenmesi ve madde cevap kuramı kapsamında incelenmesi. *Türk Psikoloji Dergisi, 24*(64), 61-75.
- Şekerciođlu, G. (2018). Measurement invariance: Concept and implementation. *International Online Journal of Education and Teaching, 5*(3), 609-634.
- Şencan, H. (2005). *Sosyal ve davranışsal ölçümlerde güvenirlilik ve geçerlik*. Ankara: Seçkin Yayınları.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (5<sup>th</sup> Edition). Boston, US: Pearson Education.

- Tabachnick, B. G., & Fidell, L. S. (2007). *Experimental design using ANOVA*. Belmont, CA: Duxbury.
- Taş, B. (2006). AB uyum sürecinde Türkiye için yeni bir bölge kavramı: İstatistiki bölge birimleri sınıflandırması (İBBS). *Afyon Kocatepe Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 8(2), 185-198.
- Tavşancıl, E. (2005). *Tutumların ölçülmesi ve SPSS ile veri analizi*. Ankara: Nobel Basımevi.
- Tekin, Y. T. (2019). *2015 PISA işbirlikli problem çözme becerilerinin ülkelere göre ölçme değişmezliğinin incelenmesi: Türkiye, Norveç, Singapur* (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington: American Psychological Association.
- Uyar, Ş. (2011). *PISA 2009 Türkiye örnekleminde öğrenme stratejileri modelinin farklı gruplarda ölçme değişmezliğinin incelenmesi* (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Uzun, N. B. (2008). *TIMSS-R Türkiye örnekleminde fen başarısını etkileyen değişkenlerin cinsiyetler arası değişmezliğinin değerlendirilmesi* (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Ülkü, S. (2019). *ABİDE 2016 Türkçe ve Fen bilimleri alt-testlerinin öğretmen özelliklerine göre ölçme değişmezliğinin incelenmesi* (Yayımlanmamış yüksek lisans tezi). Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, 9(4), 486-492.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices and

- recommendations for organizational research. *Organizational Research Methods*, 3(1), 4-70.
- Wang, J. & Wang, X. (2012). *Structural equation modelling: Applications using Mplus* (1<sup>st</sup> Edition). UK: Wiley Publications.
- West, S. G., Taylor, A. B., & Wu, W. (2012). *Model fit and model selection in structural equation modeling*. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling*, (s. 209-231). New York, US: The Guilford Press.
- Widaman, K. F., & Reise, S. P. (1997). *Exploring the measurement invariance of psychological instruments: Applications in the substance use domain*. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (p. 281–324).
- Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: a demonstration with TIMSS data. *Practical Assessment, Research & Evaluation*, 12(3).
- Yandı, A., Köse, İ. A. ve Uysal, Ö. (2017). Farklı yöntemlerle ölçme değişmezliğinin incelenmesi: PISA 2012 örneği. *Mersin Üniversitesi Eğitim Bilimleri Dergisi*, 13(1), 243-253.
- Yandı, A., Köse, İ. A., Uysal, Ö. ve Oğul, G. V. (2017). *PISA 2015 öğrenci anketinin (St094Q01na - St094Q05na) ölçme değişmezliğinin farklı yöntemlerle incelenmesi*. Demirel, Ö. ve Dinçer, S. (Ed.), *Küresel Dünyada Eğitim, içinde* (s. 333-344). Ankara: Pegem Akademi.

## EK-A: Etik Komisyonu Onay Bildirimi



T.C.  
HACETTEPE ÜNİVERSİTESİ  
Rektörlük

Tarih: 26.06.2018 16:19  
Sayı: 35853172-300-E.00000113154  
  
E.00000113154

Sayı : 35853172-300  
Konu : Mehmet Can DEMİR Hk. (Etik Komisyon İzni Hk)

### EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi : 12.06.2018 tarihli ve 51944218-300/00000096483 sayılı yazınız.

Enstitünüz Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı tezli yüksek lisans programı öğrencilerinden **Mehmet Can DEMİR**'in **Prof. Dr. Selehattin GELBAL** danışmanlığında yürüttüğü "**TIMSS Fen Bilimleri Başarısının Bölgelere Göre İncelenmesi**" başlıklı tez çalışması, Üniversitemiz Senatosu Etik Komisyonunun **18 Haziran 2018** tarihinde yapmış olduğu toplantıda incelenmiş olup, etik açıdan uygun bulunmuştur.

Bilgilerinizi ve gereğini saygılarımla rica ederim.

e-imzalıdır  
Prof. Dr. Rahime Meral NOHUTCU  
Rektör Yardımcısı

## EK-C: Veri Teslim Yazısı



T.C.  
MİLLÎ EĞİTİM BAKANLIĞI  
Ölçme, Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü

Sayı : 57750415-605.01-E.2711746  
Konu : Veri Talebi

06.02.2020

Sayın Mehmet Can DEMİR  
(Bartın Üniversitesi Eğitim Fakültesi Eğitim Bilimleri Bölümü Merkez / BARTIN)

İlgi : 08.01.2020 tarihli dilekçeniz.

"TIMSS 2015 Fen Bilimleri Başarısının Bölgelere Göre İncelenmesi" konulu tez çalışmanız kapsamında Genel Müdürlüğümüzden TIMSS 2015 Fen Bilimleri Sınavı'na ait İstatistiki Bölge Kodları ilgili dilekçe ile talep edilmektedir.

İlgi dilekçe ve ekleri Genel Müdürlüğümüzce incelenmiş olup talep edilen kodların yukarıda anılan tez çalışmasında kullanılması Genel Müdürlüğümüzce uygun görülmüştür. Ancak söz konusu tez çalışmasının Teşekkür Sayfasında Genel Müdürlüğümüz Veri Analizi İzleme ve Değerlendirme Daire Başkanlığına yer verilmesi ve tezin bir nüshasının Genel Müdürlüğümüze teslim edilmesi hususları tarafınızca taahhüt edilmesi hâlinde talep edilen kodlar teslim edilecektir.

Bilgilerinizi ve gereğini rica ederim.

Dr. Sadri ŞENSOY  
Bakan a.  
Genel Müdür

Ek: Taahhütname (1 sayfa)

Güvenli Elektronik İmza  
Aslı ile Aynıdır.  
: / /20

Teknikokullar/ANKARA  
Elektronik Ağ: www.meb.gov.tr  
Kep: meb@hs01.kep.tr

Ayrıntılı bilgi için: Mustafa CENGİZ / Şef  
Tel: (0 312) 413 32 76  
e-posta: mustafa.cengiz@meb.gov.tr

Bu evrak güvenli elektronik imza ile imzalanmıştır. <https://evraksorgu.meb.gov.tr> adresinden d3b5-e3e0-3d79-a740-7dad kodu ile teyit edilebilir.